Analysis of Environmental Treaty Design:
A Data Science Approach

Martina Kunz

Darwin College

Centre for Environment, Energy and Natural Resource Governance

Department of Land Economy

University of Cambridge

———

July 2023

This thesis is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Acknowledgements and specified in the text.

I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

This thesis does not exceed the regulation length of 80'000 words including footnotes, references and appendices.

Martina Kunz

Summary

Analysis of Environmental Treaty Design: A Data Science Approach
Martina Kunz

There are hundreds if not thousands of international agreements governing all sorts of environmental problems, from endangered species and pollution to stratospheric ozone depletion and climate change. Analysing and describing the provisions of all these treaties using the traditional 'reading and writing' approach has become all but impossible. The main proposals for solving this epistemic challenge involve either time-consuming manual approaches to building datasets, or use statistical natural language processing (NLP) for a different kind of content analysis. This thesis proposes an intermediate approach, leveraging rule-based NLP for dataset construction and employing statistics and machine learning only for downstream analysis. Traditional legal research can thus be supported and complemented while taking advantage of data science and automation. The approach is developed with a set of about 120 open multilateral environmental agreements and about 50 treaty design variables. Regular expression pattern matching is found to be well suited for accurate and precise extraction of information from common treaty provisions such as those on entry into force, amendment, supplementary agreements, treaty organs, withdrawal, termination and dispute settlement. Implementation-related provisions, including national reporting, international verification of compliance, treaty progress review, non-compliance procedures and sanctions are more difficult to capture and compare across treaties, but this difficulty itself is of interest for the analysis of treaty design. The variables, their distribution and associations are described and the speed of entry into force is predicted using various techniques including linear regression and neural networks.

Regarding the larger epistemic challenge, the scalability of the approach is assessed and limitations of existing treaty databases and research practices are identified. Drawing from achievements of the bioinformatics and linked open data communities, I argue that a collaborative, incrementally expanding database, or findable, accessible, interoperable and reusable (FAIR) datasets would make the approach scalable. This relies on a standardised vocabulary or formal ontology for data integration. Accordingly, the thesis builds a proof-of-concept Public International Law Ontology and an NLP pipeline to populate the ontology with data gathered from treaty texts and participation records. Output formats and interfaces are designed for wide accessibility, without requiring programming skills. All software and data accompanying this thesis are available under a free and open source licence.

To my parents

Acknowledgements

First, I would like to gratefully acknowledge the guidance, support and encouragement of my PhD supervisor, Prof. Jorge E. Viñuales, and warmly thank my first year, third year and final examiners, Professors Douglas Crawford-Brown, Markus Gehring, Shaun Larcom and Sergio Puig, for their helpful insights and gratifying praise. My appreciation extends to four anonymous reviewers from the Swiss National Science Foundation for their feedback and support of my Doc.Mobility Fellowship application. Without this funding I would have had to work much longer on the side, although especially the research and teaching assistance experiences were stimulating and valuable in their own right. I am also indebted to the Leverhulme Centre for the Future of Intelligence and my supervisor there, Dr. Seán Ó hÉigeartaigh, for letting me exploit synergies between my doctoral research and my projects at the Centre during the three years I was employed there.

Other kinds of support and equally inspiring environments were provided by the Centre for Environment, Energy and Natural Resource Governance of the Department of Land Economy, the Cambridge Conservation Initiative at the David Attenborough Building where I was lucky enough to have a desk for a couple of years, and Darwin College, where I had influential conversations which set me on a course to study data science, natural language processing and machine learning. Events at the Lauterpacht Centre for International Law made me feel at home when I overstretched my intellectual comfort zone.

Interdisciplinary doctoral research is bound to be solitary. I am grateful to Haicun, Bérénice, Laura, Klara, Andy, Wolfgang, Andrea, Jan, Douwe, Pablo, José, Victoria, Sam, Miriam, Liliana, Shahar, Huw, Mo, Nati, Roberto, Armeen, Tanja, Anija and Jürgen for their patient listening, insightful comments and enthusiastic support of my venture down this rabbit hole. Last but not least, I could not have done this without my parents and I wish my brother had lived to see me complete this project.

Contents

List of treaties

List of tables

List of figures

List of abbreviations and acronyms

Treaty-related

AustLII	Australasian Legal Information Institute
COP	Conference of the Parties
EIF	Entry into force
FAO	Food and Agriculture Organisation
IAEA	International Atomic Energy Agency
ICJ	International Court of Justice
IEAdb	International Environmental Agreements Database
ILC	International Law Commission
ILO	International Labour Organisation
IRENA	International Renewable Energy Agency
MEAs	Multilateral Environmental Agreements
NCP	Non-compliance procedure
UN	United Nations
UNEP	UN Environment Programme
UNGA	UN General Assembly
UNSC	UN Security Council
UNSG	UN Secretary-General
UNTS	UN Treaty Series
VCLT	Vienna Convention on the Law of Treaties
WTO	World Trade Organisation

Data science

AI	Artificial intelligence
BFO	Basic Formal Ontology
CL	Computational linguistics
DL	Deep learning
ML	Machine learning
MLP	Multilayer perceptron (neural network)
NLP	Natural language processing
OLS	Ordinary least squares (regression)
OWL	Web Ontology Language
RDF	Resource Description Framework (format)
Regex(p)	Regular expression
UNDO	UN System Document Ontology

1   Introduction

Not a day goes by without news on some environmental risk or disaster and commentary on causes or solutions. The stakes range from mild nuisance to existential threats for communities, nations and perhaps our species as a whole, not to mention many other species.[1] How to prioritise different concerns, especially between human and nonhuman survival and wellbeing, between the hardships of today and the risks of tomorrow, and between national and international interests, is an inherently political and moral question. There may be irreconcilable differences between, to take an extreme example, those in favour of space colonisation and advocates of antinatalism or voluntary extinctionism. A less extreme example is the value we attribute to the survival and wellbeing of certain animal species as compared to the survival and wellbeing of certain human communities, as in the case of large predators near human dwellings, recreational hunting, and farm animals kept in abject conditions.

Law both reflects and shapes different political and moral perspectives and standpoints. At the international level, divergent views collide and crystallise in treaties, which are also the main legal instrument for addressing the environmental problems we face. Apart from treaties adopted with the intention of tackling an environmental problem head on, there are also many international agreements that have a different primary focus, e.g. human health, labour standards, human rights, crime, trade, investment, energy, the conduct of war, etc. but that are of high relevance for environmental problems in that they govern the causes, contributors or solutions to such problems. As problem definitions and our understanding of contributing factors evolve, so does the range of relevant treaties to examine. Analysing the content, scope and language of agreements is thus useful both from an applied/problem-solving point of view and from a more meta-level interest in how states have defined, described, understood, and purported to address environmental risks and challenges.

Environmental treaty databases suggest that the number of multilateral agreements addressing environmental problems has grown from a few dozen to over seven

---

[1]   Gerardo Ceballos and others, 'Accelerated Modern Human-induced Species Losses: Entering the Sixth Mass Extinction' (2015) 1(5) Science Advances e1400253.

hundred in less than a century.[2] The most comprehensive general contemporary treaty database, the United Nations Treaty Series (UNTS), contains over 7000 multilateral agreements and over 50'000 bilateral treaties, including amendments and supplementary agreements.[3] Some treaties have lengthy appendices which are regularly amended in light of new developments, while others abstain from formal amendments but hold frequent meetings of treaty organs who adopt decisions, recommendations, or 'soft law'. Legislative, executive and adjudicative bodies from the local to the global level in turn contribute to implementing treaties in line with interpretative guidance given by treaty organs.

Contemplating all the texts produced by these governance processes, one can see that it has become impossible to gain a comprehensive overview and keep track of legal developments with traditional analytical tools. A lifetime would not be enough to read all the primary sources, let alone secondary ones. Even with regard to a smaller subset of treaties, e.g. in the hundreds instead of thousands, there are certain cognitive limitations affecting all of us: human memory is fallible and we are subject to cognitive biases hindering accurate assessments.[4] Hence, researchers aiming to grasp and describe all but the smallest number of agreements are faced with a considerable epistemic challenge. This challenge provides the starting point for the present thesis.

Note that this is by no means a new challenge. Legal informatics proponents started warning about an 'information crisis' or 'information avalanche' in the 1960s,[5] and the phenomenon of a perceived 'information explosion' across the sciences and a battle for control over information resources is said to have emerged after World War II, in the context of the Cold War.[6] With 'challenge', I deliberately opt for less alarmist and militaristic terminology to invite calmer reflection.[7]

An optimistic legal positivist objection to this premise contends that no single

---

[2]   The exact number depends on how one defines 'environmental' treaties, and whether amendments and supplementary agreements are counted separately or not, more on this in Chapter 2.

[3]   Own calculation based on a custom-built web scraper, which probably missed about 20% of the records. See Appendix B.1 for more information.

[4]   Amos Tversky and Daniel Kahneman, 'Judgment under Uncertainty: Heuristics and Biases' (1974) 185(4157) Science 1124.

[5]   See e.g. Spiros Simitis, 'Die Informationskrise des internationalen Rechts und die Datenverarbeitung' (1969) 9 Zeitschrift für Rechtsvergleichung 276; Freie Universität Berlin Arbeitsgruppe EDV und Recht, 'Juradat kämpft gegen die Informationslawine' (1970) 3(4) Kritische Justiz 463 (about a legal informatics company founded in 1969).

[6]   See Mark D Bowles, 'The Information Wars: Two Cultures and the Conflict in Information Retrieval, 1945-1999' (1999).

[7]   A striking example of the opposite approach is Ashley Deeks, 'High-Tech International Law' (2020) 88 George Washington Law Review 574, arguing for urgent uptake of 'machine learning and big data' in international law on the basis that "states such as China will begin to deploy these tools in power-enhancing ways" (p. 574).

researcher needs to read and comprehend all the relevant primary sources, because we build on each others' work as a community. I agree with this to some extent but think that there are certain drawbacks inherent in current research and publication practices which impede progress and impact of the field of international law. While there are a whole host of reasons for this, including insufficient training in international law at some universities, language barriers, and paywalled publications inaccessible to international law researchers at less privileged institutions, firms and organisations, the focus here lies on the above-mentioned epistemic challenge which remains even in the absence of socioeconomic and linguistic barriers. To illustrate the challenge, consider a sceptical entrant tasked with doing research on some type of provision common to a large number of treaties. The specific research questions could for instance be about its prevalence, variations (subtypes) and/or co-occurrence with other types of provisions. Many legal publications mention this type of provision and describe some aspects of it, but each only includes the smallest number of references to primary sources which are necessary to make a well-supported argument. After all, publications have a word limit and superfluous references or excerpts can make it seem like the author is only capable of rote copy-pasting. Even if there are supplementary materials in an appendix, which is rare, if all they do is list additional treaty references, then our entrant would still need to look up all those provisions to determine whether they adequately support the claims made in the publication. Moreover, the sceptical researcher is Popperian[8] and would not be satisfied with supporting evidence alone. Instead, she would search for uncited provisions which could falsify the author's claims. Doing all this is nearly as time-consuming as the original research that is being scrutinised, and there may be several relevant publications illuminating different aspects of this kind of provision, or making different arguments about it. Mere replication would not be rewarded (in academia at least), so our researcher would be expected to refute, refine, combine, review or extend the claims found in the literature.[9]

This kind of comprehensive verification and contribution is impracticable in most circumstances due to the sheer amount of primary and secondary sources to consider. Therefore, entrants learn which publication outlets, scholars, and styles of writing to trust. Specialising in a particular area or topic of international law also helps to distinguish strong arguments from more tenuous or flawed ones, because counter-examples may readily come to mind. Still, relying on intuitive trust and background

---

[8]   Karl R Popper, *The Logic of Scientific Discovery* (Basic Books 1959).

[9]   See e.g. Jorge E Viñuales, 'On Legal Inquiry' in Denis Alland and others (eds), *Unity and Diversity of International Law: Essays in Honour of Professor Pierre-Marie Dupuy* (Nijhoff 2014).

knowledge only works within a community of experienced and like-minded international lawyers. This is problematic because interest in treaties (justifiably) exceeds the legal discipline.

An inadequate map of the legal landscape also entails an inability to accurately assess the effect of the rules therein, and it is the latter which has attracted the most interdisciplinary interest with regard to environmental problem-solving. Indeed, scholars of different disciplinary backgrounds have devoted considerable efforts to the assessment of compliance, implementation, effectiveness and impact of international environmental treaties, using a wide range of theoretical and empirical approaches. International lawyers typically conduct in-depth treaty or country case studies and are interested in how to increase compliance and implementation.[10] Economists generally prefer quantitative methods and game-theoretical reasoning and find the question of compliance not worthy of enquiry because it might be coincidental; they instead seek to build a convincing 'no-regime counterfactual' to compare with the actual outcome in order to establish whether the treaty has made any difference at all (about which they are often sceptical),[11] or they develop theories about how to move from the Nash equilibrium to the social optimum with regard to the environmental problem under consideration by choosing the right institutional design.[12]

---

[10]   Peter H Sand (ed), *The Effectiveness of International Environmental Agreements: A Survey of Existing Legal Instruments* (Grotius Publications 1992); Abram Chayes and Antonia Handler Chayes, *The New Sovereignty: Compliance with International Regulatory Agreements* (Harvard University Press 1995); Edith Brown Weiss, 'Understanding Compliance with International Environmental Agreements: The Baker's Dozen Myths' (1999) 32 University of Richmond Law Review 1555; Edith Brown Weiss and Harold K Jacobson (eds), *Engaging Countries: Strengthening Compliance with International Environmental Accords* (MIT Press 1998); Rüdiger Wolfrum, 'Means of Ensuring Compliance with and Enforcement of International Environmental Law' English [1998] Recueil des cours de l'Académie de droit international de La Haye = Collected courses of the Hague Academy of International Law 9; Claude Impériali and Alexandre Charles Kiss, *L'effectivité du droit international de l'environnement: contrôle de la mise en oeuvre des conventions internationales* (Economica 1998); Ulrich Beyerlin (ed), *Ensuring Compliance with Multilateral Environmental Agreements: A Dialogue between Practitioners and Academia* (Studies on the Law of Treaties vol. 2, Martinus Nijhoff 2006); Tullio Treves, Attila Tanzi, and Laura Pineschi (eds), *Non-Compliance Procedures and Mechanisms and the Effectiveness of International Environmental Agreements* (Asser Press 2009); Sandrine Maljean-Dubois and Lavanya Rajamani, *La mise en oeuvre du droit international de l'environnement* (Martinus Nijhoff 2011).

[11]   Such as James C Murdoch and Todd Sandler, 'The Voluntary Provision of a Pure Public Good: The Case of Reduced CFC Emissions and the Montreal Protocol' English (1997) 63(03) Journal of Public Economics 331; but see the response by Robin Mason and Timothy Swanson, 'A Kuznets Curve Analysis of Ozone-Depleting Substances and the Impact of the Montreal Protocol' English (2003) 55(1) Oxford Economic Papers 1; as well as Ulrich J Wagner, 'The Voluntary Provision of a Pure Public Good? Another look at CFC Emissions and the Montreal Protocol' (2009) 61(1) Oxford Economic Papers 183.

[12]   E.g. Frank Stähler, *On the Economics of International Environmental Agreements* (Springer 1998); Timothy M Swanson and Sam Johnston, *Global Environmental Problems and International Environmental Agreements: The Economics of International Institution Build-*

Political scientists/international relations scholars are more eclectic in their methods, and their focus often lies on highly context-specific factors of regime influence such as personal leadership, coalitions and other networks, problem characteristics, power structures, hybrid governance, and institutional interplay.[13] Ecologists, geographers and environmental scientists usually focus on in-depth case studies and only occasionally engage in large-scale analysis.[14] The picture that emerges is not uniform, and this is merely a small snapshot of the disciplinary approaches involved in assessing MEAs and their effects.

Thus, similar to how the definition and prioritisation of environmental problems and solutions are political, so are the boundaries and labelling of the object of study and the most appropriate methods of analysis.[15] Regardless, including 'environmental' in the title of the present thesis is not intended as a political statement against 'sustainable development law'[16] or 'earth system law',[17] but merely to give an idea of what agreements the thesis is about in a maximally succinct way. After all, "design of open multilateral treaties relevant for tackling global environmental problems" is considerably longer than 'environmental treaty design', but it would have been more accurate. Transparency and clarity seem like the best approach for dealing with fuzzy and contested boundaries, hence the relatively detailed description of the

ing (Edward Elgar: Published in association with UNCTAD 1999); Amitrajeet A Batabyal (ed), *The Economics of International Environmental Agreements* (Ashgate 1999); Carsten Schmidt, *Designing International Environmental Agreements: Incentive Compatible Strategies for Cost-Effective Cooperation* (New horizons in environmental economics, Edward Elgar 2000); Carsten Helm, *Economic Theories of International Environmental Cooperation* (Edward Elgar 2000); Michael Finus, *Game Theory and International Environmental Cooperation* (Edward Elgar 2001); Scott Barrett, *Environment and Statecraft: The Strategy of Environmental Treaty-Making* (Oxford University Press 2005). For an overview see Scott Barrett, 'An Economic Theory of International Environmental Law' in Daniel Bodansky, Jutta Brunnée, and Ellen Hey (eds), *The Oxford Handbook of International Environmental Law* (Oxford University Press 2007).

[13] Peter M Haas, Robert O Keohane, and Marc A Levy, *Institutions for the Earth: Sources of Effective International Environmental Protection* (MIT Press 1993); David G Victor, Kal Raustiala, and Eugene B Skolnikoff, *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice* (International Institute for Applied Systems Analysis; MIT Press 1998); Oran R Young, *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms* (MIT Press 1999); Jørgen Wettestad, *Designing Effective Environmental Regimes: the Key Conditions* (Edward Elgar 1999); Edward L Miles, *Environmental Regime Effectiveness: Confronting Theory with Evidence* (MIT Press 2002); Helmut Breitmeier, Oran R Young, and Michael Zürn, *Analyzing International Environmental Regimes: from Case Study to Database* (MIT Press 2006).

[14] See e.g. M Hoffmann and others, 'The Impact of Conservation on the Status of the World's Vertebrates' en (2010) 330(6010) Science 1503.

[15] See Martti Koskenniemi, 'The Fate of Public International Law: Between Technique and Politics' (2007) 70(1) Modern Law Review 1, for an incisive critique of supposedly apolitical managerial governance.

[16] Marie-Claire Cordonier Segger and Ashfaq Khalfan, *Sustainable Development Law: Principles, Practices, and Prospects* (Oxford University Press 2004).

[17] Louis J Kotzé, 'Earth System Law for the Anthropocene' (2019) 11(23) Sustainability.

delimitation of the treaty set in Chapter 2 and the somewhat arduous restraint in generalising findings beyond the treaty set in subsequent chapters.

'Treaty' is used in the sense it has in international law,[18] and is the core focus of this thesis.

As for 'design', the point of this addition is to underline that the scope and content of international agreements is the product of a deliberate choice by the negotiators involved, however contested and constrained. These choices are what the proposed data science approach purports to analyse. Before outlining the approach, a brief overview of existing ways to deal with the above-mentioned epistemic challenge is in order.

In political science, economics and other disciplines interested in treaties, collecting structured data on agreement content has gained in popularity. Data is published as an in-text table, supplementary materials, online database, or in some cases only summarised in the text.[19] By structured data I mean information about treaties that is expressed in a set of defined variables.[20] A natural language summary of the text alone would not count, whereas a set of keywords based on a shared data model used across all treaties would. For instance, Scott Barrett's structured data on the treaties he analysed for his book on MEAs comprises a categorical variable assigning one of twelve themes to each treaty, two numeric variables (minimum ratifications and number of signatories), and three binary variables (openness, trade restrictions and administrative organisation).[21] A larger set of treaties and variables is included in Ronald Mitchell's International Environmental Agreements Database (IEAdb).[22]

While these efforts are commendable for their ambition, they tend to suffer from three main issues that arguably make them unsuitable for legal research: (i) lack of precision (in typology and evidence), (ii) low accuracy (in texts, metadata, treaty and content classifications), and (iii) insufficiently holistic/systemic (failure

---

[18]   Vienna Convention on the Law of Treaties (adopted 23 May 1969, entered into force 27 January 1980) 1155 UNTS 331, art 2(a) ('"treaty" means an international agreement concluded between States in written form and governed by international law, whether embodied in a single instrument or in two or more related instruments and whatever its particular designation.'); see also Jan Klabbers, *The Concept of Treaty in International Law* (Kluwer Law International 1996).

[19]   An early example of this approach is James F Keeley, 'Coding Treaties: An Example from Nuclear Cooperation' (1985) 29(1) International Studies Quarterly 103.

[20]   Text, audio and video is unstructured data, as this term is commonly used in computer science and information technology.

[21]   Barrett, *Environment and Statecraft* (n 12), p.165-194 ("Appendix 6.1. Multilateral Environmental Agreements"). More on these variables in Chapter 2.

[22]   Ronald B Mitchell *International Environmental Agreements Database Project* version 2020.1 ⟨http://iea.uoregon.edu/⟩, described in Ronald B Mitchell and others, 'What We Know (and Could Know) About International Environmental Agreements' (2020) 20(1) Global Environmental Politics 103.

to recognise the legal context of a treaty and the system of international law it is a part of). As a result, such data misrepresent international law and any findings based on it should be taken with a grain of salt. These approaches also typically do not follow best practices in data science and reproducible research.

While non-lawyers gloss over crucial details, some lawyers have honed in on the details of the epistemic challenge while arguably losing sight of the objective of solving it. The fact that methodological fields like artificial intelligence and law, legal informatics, and computational law/legal studies have emerged is one indication for this turn to methods development as an objective in and of itself. To clarify relevant terms as I will use them here, 'artificial intelligence' (AI)[23] is an umbrella term including both rule-based, symbolic AI, such as expert systems and theorem provers, as well as more data-driven approaches typically called 'machine learning' (ML)[24] of which the subfield of 'deep learning',[25] using neural networks with multiple layers of artificial neurons to learn complex functions, achieved significant leaps in performance across a number of tasks in the first two decades of the 21st century thanks to better hardware and larger datasets becoming available. The aim of AI research is to build intelligent machines, but a much larger scholarly community is more interested in what these data and hardware trends mean for science.[26] Buzzwords seem to be replaced or redefined every decade or so, hence there is little use in attempting a definitive classification. Still, Blei and Smyth's perspective of data science as the "child of statistics and computer science"[27] going back to Tukey's view of data analysis[28] is convincing. In particular, their emphasis that the practice of data science encompasses the whole process of "data preprocessing, exploration, selection, transformation, analysis, interpretation, and communication" with due regard for scientific reproducibility is one that I think reflects a general understanding of this term and makes it a better moniker than the previously more popular 'data mining' or 'big data analytics' which have a narrower focus. It also appropriately subordinates the choice of analysis tool or method to the research questions and data at hand or to be collected.

---

[23] Stuart J Russell and Peter Norvig, *Artificial Intelligence : A Modern Approach* (Pearson Education 2022).

[24] Michael I Jordan and Tom M Mitchell, 'Machine learning: Trends, perspectives, and prospects' (2015) 349(6245) Science 255.

[25] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, 'Deep Learning' (2015) 521(7553) Nature 436; for an introductory textbook see Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (http://www.deeplearningbook.org, MIT Press 2016).

[26] David M Blei and Padhraic Smyth, 'Science and Data Science' (2017) 114(33) Proceedings of the National Academy of Sciences 8689.

[27] Ibid 8689.

[28] John W Tukey, 'The Future of Data Analysis' (1962) 33(1) The Annals of Mathematical Statistics 1.

As for text data (such as treaty texts), the labels used also differ in focus, scope and objective. 'Natural language processing' (NLP), at its broadest, is "any kind of computer manipulation of natural language"[29] with 'natural language' referring to naturally evolved languages used for human communication, as opposed to deliberately designed programming languages or mathematical notations. As an academic field, it is sometimes equated with 'computational linguistics' (CL), but NLP is about solving language-based tasks like machine translation and text classification while CL is concerned with the study of language, using a computational approach.[30] Similar to data mining, 'text mining' started to emerge as a buzzword in the 1990s for knowledge discovery from large text collections (also called corpora in linguistics) in the context of applications in business, surveillance and suchlike.[31] A related but more recent and less widely known label is 'text-as-data', mostly used to refer to NLP methods applied to political texts,[32] but seeing some uptake in other social sciences and even international law.[33] In my view, introducing new names for old methods is only justified if the new label is more informative or accurate than adapting an existing one. For instance, 'computational social science'[34] is a good name because it evokes the application of computational methods and tools to a given academic discipline, similar to 'computational linguistics', 'computational neuroscience', and 'computational biology'. 'Text mining' and 'text-as-data' neither add information as to methods nor as to the application domain, and are thus in my view inferior as a new label for a field, approach or framework.[35]

---

[29]   Steven Bird, Ewan Klein, and Edward Loper, *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit* (O'Reilly 2009), p.ix.

[30]   Yorick Wilks, 'Computational Linguistics: History' in Keith Brown (ed), *Encyclopedia of Language & Linguistics* (2nd Edition, Elsevier 2006), p.762 ("NLP does require a task: it is not in itself a program of scientific investigation, which is what CL normally claims to be".). By contrast, see Julia Hirschberg and Christopher D Manning, 'Advances in Natural Language Processing' (2015) 349(6245) Science 261, presenting CL as a subfield of computer science, rather than an approach to linguistics or an interdisciplinary field.

[31]   For an overview of emergence and positioning with regard to existing fields, see Marti A Hearst, 'Untangling text data mining' (1999).

[32]   For a widely cited introduction, see Justin Grimmer and Brandon M Stewart, 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts' (2013) 21(3) Political Analysis 267.

[33]   In particular by Wolfgang Alschner and collaborators, e.g. in Wolfgang Alschner and Dmitriy Skougarevskiy, 'Mapping the Universe of International Investment Agreements' (2016) 19(3) Journal of International Economic Law 561; and Wolfgang Alschner, Julia Seiermann, and Dmitriy Skougarevskiy, 'Text of Trade Agreements (ToTA)—A Structured Corpus for the Text-as-Data Analysis of Preferential Trade Agreements' (2018) 15(3) Journal of Empirical Legal Studies 648.

[34]   David Lazer and others, 'Computational Social Science' en (2009) 323(5915) Science 721; David MJ Lazer and others, 'Computational Social Science: Obstacles and Opportunities' (2020) 369(6507) Science 1060.

[35]   A recent book-length effort to establish it as such is Justin Grimmer, Margaret E Roberts, and Brandon M Stewart, *Text as Data: A new Framework for Machine Learning and the*

Coming back to terms tailored to the legal domain, 'legal informatics' can be said to be concerned with research and development of tools that enable 'computational legal studies', 'legal analytics' or 'legal data science' (these three are taken to be near-synonyms). 'Computational law' is slightly ambiguous because 'law' can refer to the academic discipline, in which case the term is synonymous to computational legal studies, or to the law itself, in which case it can mean an effort to formalise or infer legal knowledge, automate legal reasoning and potentially replace lawyers and judges. AI & Law is similarly broad, or even broader. In this context, 'formalising' the law or legal knowledge means something like translating it into a logical or programming language that, in the extreme, makes law computable in what used to be called legal expert systems in the 1980s/1990s,[36] but has long been a dream of logicians. However, despite decades of work in logic-based as well as statistical/ML approaches to legal knowledge representation,[37] little if any of it has permeated into international practice. While the formalisation of national law has certainly seen more interest from the AI & Law community, it is not as though international law has been completely off the radar. Erich Schweighofer, one of the early scholars in this field, wrote his habilitation thesis on knowledge representation and automatic text analysis of public international and European law in the 1990s, using both formal logic and neural networks.[38] One of his case studies included 100 multilateral agreements with a focus on environmental protection.[39] However, neither the source code of the program[40] nor the data generated by the analysis seem to be published and thus comparability, verifiability and reusability are limited. Also, his work, like many others in this field, seems to focus more on methodology than on results, and thus opts for more complex and less transparent methods and techniques. This is perhaps to be expected in a methodologically orientated subfield.

Overall, the dearth of open source programs and data in international legal informatics, and lack of uptake of existing approaches and prototypes by major international law databases like UNTS led me to seek inspiration in other fields, in particular bioinformatics. By way of example, the Gene Ontology provides a

---

Social Sciences (Princeton University Press 2022).

[36] For a retrospective by one of its most vocal critics, see Philip Leith, 'The Rise and Fall of the Legal Expert System' (2016) 30(3) International Review of Law, Computers and Technology 94.

[37] See e.g. Trevor Bench-Capon and others, 'A History of AI and Law in 50 Papers: 25 years of the International Conference on AI and Law' (2012) 20 Artificial Intelligence and Law 215.

[38] Erich Schweighofer, *Legal Knowledge Representation: Automatic Text Analysis in Public International and European Law* (Kluwer Law International 1999).

[39] Ibid , p.190.

[40] The analysis uses the expert system KONTERM described in Erich Schweighofer and Werner Winiwarter, 'Legal Expert System KONTERM - Automatic Representation of Document Structure and Contents' (DDEXA '93, Springer 1993).

structured, controlled vocabulary of over 40'000 'terms' or 'classes' describing the functions of genes in order to support the computational representation of biological systems. These ontology terms have been used to annotate over 1.5 million gene products from over 5000 species described in over 170'000 scientific publications,[41] and all that in only 24 years of scientific collaboration with a growing community of developers and users.[42] The Protein Data Bank was established in the 1970s and has accumulated experimentally determined 3D structure data on over 100'000 proteins and other macromolecules in less than 50 years.[43] Biomedical ontologies proved so successful that it soon became necessary to establish a coordinating body to ensure semantic interoperability and quality control across dozens of specific domain ontologies.[44]

While the initial tools and standards used for the development of these ontologies were specifically geared towards bioinformatics, the recommended tools and standards today are more general and co-developed by a much larger community, namely that working towards the ideal of the 'semantic web' and 'linked open data'.[45] Both also heavily borrow tools, practices and ideals of free and open source software development, which underpins and enables the information age. These communities' success in building software, standards and databases together that are widely adopted in academia and beyond suggests that the epistemic challenge in international law might also be best tackled in a long-term collaborative effort.

An alternative framing of the same vision is articulated in the Findable, Accessible, Interoperable and Reusable (FAIR) Data Principles, whose authors argue that all research objects should be FAIR "both for machines and for people".[46] Interoperability is defined in this context as "the ability of data or tools from non-cooperating resources to integrate or work together with minimal effort".[47] Thus, this vision supports more individualistic and ad hoc research projects, which nevertheless produce

---

[41]   Seth Carbon and Chris Mungall *Gene Ontology Data Archive* version 2022-10-07 (Zenodo July 2018, with statistics presented at http://geneontology.org/stats.html (accessed 15 Oct 2022, consulting release version 2022-10).

[42]   Michael Ashburner and others, 'Gene Ontology: Tool for the Unification of Biology' (2000) 25(1) Nature Genetics 25; The Gene Ontology Consortium, 'The Gene Ontology Resource: 20 years and still GOing strong' (2018) 47(D1) Nucleic Acids Research D330.

[43]   wwPDB consortium, 'Protein Data Bank: the Single Global Archive for 3D Macromolecular Structure Data' (2018) 47(D1) Nucleic Acids Research D520.

[44]   Barry Smith and others, 'The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration' en (2007) 25(11) Nature Biotechnology 1251.

[45]   For an articulation of the original vision by the inventor of the World Wide Web, see Tim Berners-Lee, James Hendler, Ora Lassila, and others, 'The Semantic Web' (2001) 284(5) Scientific American 28, and Tim Berners-Lee and James Hendler, 'Publishing on the Semantic Web' (2001) 410(6832) Nature 1023.

[46]   Mark D Wilkinson and others, 'The FAIR Guiding Principles for Scientific Data Management and Stewardship' (2016) 3(1) Scientific Data 1, p.3.

[47]   Ibid , p.2.

data that can be integrated into a coherent whole. The 2021 UNESCO Recommendation on Open Science also refers to FAIR data principles and stresses that "convergence between the various semantic artefacts (particularly vocabularies, taxonomies, ontologies and metadata schema) is essential for the interoperability and reuse of data for interdisciplinary research."[48] The end result is the same, but the ad hoc approach may require more effort to achieve convergence.

What becomes glaringly obvious when looking at genomics, proteomics, astronomy, particle physics and similar fields is that international law data is minuscule compared to these big data domains. Our datasets are measured in megabytes or at most gigabytes, not terabytes or petabytes. This means that the computational challenges are not comparable, with implications for storage, processing, funding, analysis and accessibility.[49] International law data can be stored and processed on a consumer laptop, no need for high-performance computing, expensive data centres and novel algorithms for distributed computing. Presenting international law data as 'big data' could be unnecessarily intimidating to newcomers, induce hype, or lead new researchers to disregard the literature and computational tools of the last century because they are convinced by the argument that the epistemic challenge is new and requires entirely novel approaches.[50] In the extreme, text-as-data converts may abandon traditional legal research questions and give in to a fully 'data-driven' approach because of a few thousand texts, while molecular biologists are undeterred by their petabytes of data and proceed to collaboratively define hundreds of thousands of biological processes and structures. Machine learning is by no means shunned in bioinformatics,[51] and indeed leverages expert-curated databases for prediction.[52]

The present thesis consequently employs best practices of the bioinformatics and open science communities to build a draft ontology of public international law based on consensus terminology such as that defined in the Vienna Convention on

---

[48] UNESCO Recommendation on Open Science, SC-PCB-SPP/2021/OS/UROS, para. 18(f).

[49] For real big data challenges see e.g. Vivien Marx, 'The Big Challenges of Big Data' (2013) 498(7453) Nature 255.

[50] E.g. Alschner and Skougarevskiy (n 33); Wolfgang Alschner, Joost Pauwelyn, and Sergio Puig, 'The Data-Driven Future of International Economic Law' (2017) 20(2) Journal of International Economic Law 217, p.220-221 (on the 'new frontier of data-driven research' being new or different because "data is becoming 'big'—and not just in a quantitative sense. It becomes big because [...] it requires us to make use of big data tools developed in other disciplines" among other reasons); and Deeks (n 7), p.597 ('Each of these collections contains thousands of documents, which [...] provides the type of "big data" that makes text-as-data tools effective and efficient').

[51] For a review see Maxwell W Libbrecht and William Stafford Noble, 'Machine Learning Applications in Genetics and Genomics' (2015) 16(6) Nature Reviews Genetics 321.

[52] E.g. John Jumper and others, 'Highly Accurate Protein Structure Prediction with AlphaFold' (2021) 596(7873) Nature 583, on an AI application that was trained on data from the Protein Data Bank.

the Law of Treaties (VCLT).[53] An overview and source code excerpt is included in Appendix E and the full version is available for download from the thesis code archive. This is of course only a proof of concept, as collaboration and iterative improvement will be necessary to avoid idiosyncrasies, assess needs, and make it maximally useful to international lawyers and beyond. The ontology was drafted in the Web Ontology Language (OWL) 2[54] using the Protégé ontology development software[55] with the Basic Formal Ontology (BFO) as top-level ontology.[56] In this conception, all ontologies "consist of (1) a central backbone taxonomy, in which all the nodes of the ontology are linked together via is_a relations, together with (2) further relations defined between the nodes of the ontology. In addition, each node consists of (3) a term along with, when necessary, (4) synonyms for the term, and crucially (5) a definition of the term that makes use of the Aristotelian genus and differentia structure."[57] The taxonomy should follow the principle of asserted single inheritance,[58] which means that each term has at most one parent. Aristotelian definitions are very simple. By way of illustration, to define *bilateral treaty* as a subclass of *treaty* one would simply repeat the parent term and then add the defining element, e.g. "*bilateral treaty* is_a *treaty* which has exactly two parties."

There is an existing example of a collaborative international law-related ontology using semantic web standards, namely the United Nations System Document Ontology (UNDO),[59] but it does not follow best practices of scientific ontology development. For instance, the terms 'Treaty', 'Convention' and 'Protocol' are all categorised as separate subclasses of 'NormativeDocument' which also includes non-binding documents. The respective definitions in UNDO are as follows:

Treaty An international agreement concluded between States, and bilateral or multilateral agreements governed by international law to which international organizations are parties.

---

[53] VCLT (n 18).

[54] Pascal Hitzler and others (eds), OWL 2 Web Ontology Language: Primer (Second Edition) (W3C Recommendation, 2012) ⟨https://www.w3.org/TR/2012/REC-owl2-primer-20121211/⟩ accessed 20 October 2016.

[55] Mark A Musen, 'The Protégé Project: A Look Back and a Look Forward' (2015) 1(4) AI Matters 4.

[56] Robert Arp, Barry Smith, and Andrew D Spear, *Building Ontologies with Basic Formal Ontology* (MIT Press August 2015).

[57] Ibid , p.36.

[58] Ibid , p.79.

[59] Silvio Peroni, Monica Palmirani, and Fabio Vitali, 'UNDO: The United Nations System Document Ontology' (Claudia d'Amato and others eds, Springer 2017), ontology available at https://w3id.org/un/ontology/undo (accessed 15 Dec 2022).

Convention Mandatory regulations subject to ratification. Instruments usually negotiated under the auspices of an international organization.

Protocol International agreement that supplements a previous treaty or international agreement.

Instead of the three being sibling classes, *treaty* should be a parent class, and 'supplementary agreement' could be entered as a subclass of treaty, defined e.g. as 'a treaty which supplements an existing treaty', if there is a need for such a category. The use-case and conceptualisation of the 'Convention' class is unclear, especially how it could be construed as a distinct category, mutually exclusive with regard to 'Treaty'. The latter's definition is closer to the customary understanding of treaties in international law, albeit missing the written form that the VCLT uses in its definition of the concept.[60] In short, UNDO may be sufficient for internal document classification at some UN system entities, but as a component of an international law database it would not be suitable, as it does not conform to existing conceptualisations and definitions of such foundational terms as 'treaty'.

Defining and implementing a formal ontology in software is only the first step. The second is populating the ontology with instances to produce a database. While instances of the classes defined by the ontology could in principle be entered and edited manually with Protégé, a more expedient and scalable approach is to process texts automatically using natural language processing (NLP) software and populate the ontology with instances found by the algorithm. The General Architecture for Text Engineering (GATE) software toolkit[61] conveniently provides ontology population and visualisation features.[62] The texts processed by the algorithms developed for this project consist of treaty texts and the content of UNTS treaty pages. Following reproducible data science principles, the texts were gathered programmatically with custom-built web scrapers from online databases, preprocessed and annotated entirely by automated means. A range of NLP approaches and techniques were trialled, including unsupervised topic modelling, but a more 'hard-coded' approach using iterative regular expressions pattern matching proved better suited for identifying legally relevant information in treaty texts.

Note that while the language and format chosen allows for some automated reasoning and logical inference (Protégé itself has various automated reasoning plug-

---

60   VCLT (n 18), art 2(a).

61   https://gate.ac.uk described in Hamish Cunningham and others, 'Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics' (2013) 9(2) PLOS Computational Biology 1.

62   Kalina Bontcheva and others, 'Evolving GATE to Meet New Challenges in Language Engineering' (2004) 10(3/4) Natural Language Engineering 349.

ins), the aim here is efficient and accurate information representation, not automating legal reasoning and professional legal services. A large, well-built database can of course facilitate legal research and access to legal information, but most likely will never exhaustively cover international law, as it is in constant evolution and may well be fundamentally indeterminate.[63] 'Ontology' is thus perhaps not an ideal term from certain scholarly perspectives, but 'vocabulary' or 'domain model' is not necessarily much clearer. A related objection to this exercise could be that it is about power and politics, trying to impose one particular interpretation of international law over another. I have tried to avoid this by using consensus vocabulary and (mostly) uncontroversial definitions, and by being transparent about difficulties encountered with fuzzy or shifting boundaries of concepts, especially more recent ones like 'non-compliance procedure'. Scientific ontology development sees updating of terms and definitions as a natural part of the process, not as a sign of failure.

Apart from building a database containing information about the content of multilateral agreements, this research project originally also aimed to collect data on political success, compliance and effectiveness in order to assess what kind of treaty design achieves the best outcomes in the domain of environmental problem-solving. However, compliance and effectiveness data were found to be insufficiently available, complete, and fine-grained even in the largest databases. As for political success, only treaty adoption and general entry into force dates are reliable and complete in UNTS, thus the only outcome analysis conducted was the prediction of treaties' speed of entry into force from their text features.[64] Several types of linear regression and machine learning models were compared and the best performing model type was a deep neural network, predicting out-of-sample treaties' entry into force with a precision of a bit over 2.5 years $\pm$ 0.5 years, which is not terrible considering that the treaties in the dataset took between 0 and 17 years to enter into force.

Finally, this thesis is divided into four chapters. The first chapter presents the approach taken in the sourcing and selection of treaty texts, as well as giving a broad overview of text analysis methods and tools developed. Chapter 3 describes the process and findings of legal information extraction (the types of provisions included in the ontology), while Chapter 4 focuses on the subject matter of the selected treaties. Chapter 5 then uses both sets of variables to try to predict entry into force of treaties. The thesis concludes with an outlook to the future of international legal informatics and data science. The appendices contain excerpts of most of the

---

[63]   Martti Koskenniemi, *From Apology to Utopia: the Structure of International Legal Argument* (Cambridge University Press 2005).

[64]   More precisely, the prediction target was the time elapsed between the date of adoption and the date of entry into force minus the waiting period stipulated by the treaty, if any.

programming scripts developed for this dissertation, and the full source code and data are available for download from the thesis code repository at https://gitlab.com/martinakunz/phd (archived at https://doi.org/10.5281/zenodo.10078710).

2 Selection of treaty texts and methods

2.1 Treaty text sources and quality

Any text analysis first needs to concern itself with obtaining sufficiently high quality texts. Collecting, processing and analysing texts by automated means additionally requires sources providing machine-readable texts[65] in a consistent format.

The most authoritative source of treaty texts would be the respective treaty depositary, but each depositary has a different treaty text access point (if any), file format, text layout and so forth. The second best source is the United Nations Treaty Series (UNTS), because pursuant to article 102 I of the UN Charter,[66] all international agreements entered into by any member of the United Nations have to be registered with the UN Secretariat and be published by it. UNTS publishes these agreements in their authentic language(s), along with translations into English and French where appropriate. The series currently consists of over 70'000 agreements[67] and their related acts in over 3'000 volumes.[68] Unfortunately, even recent electronic UNTS volumes contain only image files of authentic treaty texts, which means that they are not readily machine-readable. Extracting text from such files is an error-prone process and requires in-depth verification and correction.

Ecolex,[69] an environmental law database maintained by UNEP, FAO and IUCN, is the third best source of environmental agreements, but unfortunately the format of the texts is inconsistent, i.e. sometimes containing lengthy introductions, footnotes or endnotes that are not part of the treaty text as such, presumably because the documents were obtained from treaty secretariats.

The fourth official source considered is InforMEA,[70] a more recent and steadily growing UN information portal on MEAs which provides easy access to consistently

---

[65] 'Machine-readable' can have different meanings, here I only mean files from which a computer program can extract an accurate and complete string representation of the text, not that semantic content is encoded in structured data.

[66] Charter of the United Nations (adopted 26 June 1945, entered into force 24 October 1945).

[67] This is an estimate based on the results of the UNTS-searchbot in Appendix B.1.

[68] https://treaties.un.org/Pages/LatestVolumes.aspx (accessed 1 Feb 2023).

[69] https://ecolex.org (accessed 10 Sept 2022).

[70] https://informea.org (accessed 10 Sept 2022).

formatted machine-readable treaty texts. During the initial treaty text collection phase of this thesis (2015-2016) InforMEA contained only 43 treaties in total and was thus not suitable. The current version contains 34 global and 461 regional agreements, but important global treaties, e.g. those concluded under the auspices of the IMO, are still missing.

The fifth official source of treaty texts are governmental treaty series, especially those countries which tend to participate in many multilateral environmental agreements. The official treaty series published by the UK, Ireland, Canada, Australia and the US were consulted, and of these only Australia publishes its treaties in HTML format, through a collaboration with the Australasian Legal Information Institute (AustLII). Coverage is extensive, going back to 1900, and the treaty library includes not only treaties that Australia signed or ratified, but also those it merely considered participating in. However, relying exclusively on the Australian treaty collection would still introduce a political bias, e.g. against ILO treaties, which seem less popular in Australia than elsewhere. Also, some of the more recent treaties, such as the Paris Agreement, are were only published in PDF format, not HTML, at the time of AustLII data collection (July 2020). The relevant part of the web scraper is included in Appendix D.3.

Due to these limitations of more authoritative data sources, academic databases were considered as well. There are two large databases of international environmental agreements, one is the abovementioned IEAdb by Ronald Mitchell[71] and the other is hosted by the Center for International Earth Science Information Network (CIESIN) at Columbia University. CIESIN's treaty text collection, the "Environmental Treaties and Resource Indicators Treaty Texts v1 (1940-2000)"[72], has not been updated since 2000 and texts are published with line breaks stemming from the PDF conversion process, which makes it much harder to automatically detect where sentences begin and end. Mitchell seems to have built on the CIESIN collection (among others) and mostly removed unnecessary line breaks.

However, in terms of accuracy and completeness of treaty texts, AustLII's HTML treaty collection is far superior to Mitchell's. Apart from missing or extraneous line breaks and whitespace disrupting text flow, IEAdb texts also contain typographical errors likely resulting from inaccurate optical character recognition, footers from the

---

[71] Mitchell, *International Environmental Agreements Database Project* (n 22).

[72] CIESIN/Columbia University, UNEP, Fletcher School of Law and Diplomacy/Tufts University, British Columbia Ministry of Environment/Lands/Parks, Antarctic Cooperative Research Centre, and American Society of International Law, 2002, Socioeconomic Data and Applications Center (SEDAC) Collection of Treaty Texts. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4251G48 (accessed 10 Sept 2019).

original documents which ended up in the midst of a treaty provision due to lack of clean-up after conversion from PDF, and in some cases final clauses or appendices are entirely omitted, even when present in the source. In fact, treaty annexes were often entered as separate instruments in the database, complicating reconstruction of the full original treaty text and accurate agreement counts. The case of the treaty establishing the Center for International Forestry Research (CIFOR)[73] may serve as an example to illustrate this issue. The source used for the text of the agreement is the online version of the Australian Treaty Series,[74] i.e. sufficiently authoritative and of high quality. However, instead of keeping the Center Constitution appended to the Establishment Agreement as it was in the source, the IEAdb version 2014.3 lists them as two separate MEAs (with the same adoption and entry into force date), and version 2018.1 contains the Establishment Agreement as a MEA, and the Constitution as a multilateral environmental "non-agreement", with the "Agreement Type (level 2)" set to "Statute-nonbinding". This is a rather odd choice from a legal perspective, given that Article 1 of the Establishment Agreement states that CIFOR "shall operate in accordance with the Constitution appended hereto and forming an integral part of this Agreement." As for agreement counts, the most egregious example is perhaps the original MARPOL agreement,[75] which was split up into eight different documents (the main agreement, its five annexes and two protocols), each entered separately into the IEAdb.[76] The counts and summary statistics published by Mitchell et al should thus be interpreted in light of this kind of inflation.

Due to these and other quality concerns, in all instances where AustLII provided an HTML version of the treaty text, its version was chosen for the present research project, using IEAdb texts for the remainder. It is hoped that authoritative data providers will increasingly adopt machine-readable publication methods, which would reduce barriers to large-scale text analytics. This is where one of the prime advantages of scripted data processing becomes apparent. Whereas manual text analysis and information extraction would require another full pass through the treaty texts by the number of annotators chosen, programmatic extraction only needs a few tweaks before running the algorithm on a higher quality text collection if and when it becomes available. In fact the present project initially used only IEAdb texts, switching to a mixed approach once the difference in quality became too striking to ignore. The IEAdb web scraper is included in Appendix D.2 and the

---

[73]   Establishment Agreement for the Center for International Forestry Research (adopted 5 March 1993, entered into force 5 March 1993) 1736 UNTS 170.

[74]   http://austlii.edu.au/au/other/dfat/treaties/ATS/1993/13.html (accessed 10 Sept 2022).

[75]   International Convention for the Prevention of Pollution from Ships, 1973 (adopted 2 November 1973, entered into force 2 October 1983) 12 ILM 1319.

[76]   Appendix D.4 shows how these parts were concatenated for the present project.

text cleaning script for both AustLII and IEAdb texts can be found in Appendix D.4.

A closely related issue is the accuracy of treaty metadata in the database, such as title, adoption date and entry into force date. For instance, when analysing regime evolution it is of paramount importance that the adoption date be correct, and yet consolidated versions of treaties are sometimes published under an earlier adoption date in the IEAdb.[77] Additionally, when combining treaty-related data from multiple databases, a common identifier is needed for the purpose of matching records. The treaty title and adoption date is an obvious candidate,[78] the UNTS volume and page number, or the UNTS treaty page URL would work as well, but unfortunately none of these are used consistently across databases and UNTS volume publication tends to have a multi-year delay. The need for manual verification and matching in turn reduces the number of treaties that can realistically be analysed within a given project timeframe.

2.2 Treaty selection

The question of treaty selection, while constrained by the availability and quality of texts, is also subject to its own set of considerations, mostly determined by the research question and methodology. As the present research project was concerned with analysing as large a number and as wide a range of treaties as possible, both for mapping the legal toolkit and for drawing inferences from observed outcomes, it made sense to start with the largest machine-readable English language MEA collection available, i.e. Mitchell's IEAdb. In addition to MEAs, the IEAdb also contains agreements classified as 'non-environmental' (MNAs) and 'uncoded' (MuAs), non-multilateral agreements, and texts classified as 'non-agreement'. Initially only those categorised as MEAs were used, but once a number of inadequacies of these classifications came to the fore, such as the one with separately listed annexes mentioned in the previous section, a more thorough selection process starting from the full IEAdb collection version 2018.1[79] was undertaken. The process described hereafter is implemented in the Python script contained in Appendix C for reproducibility

---

[77] E.g. the Convention on Wetlands of International Importance Especially as Waterfowl Habitat (adopted 2 February 1971, entered into force 21 December 1975) 996 UNTS 245, with IEAdb ID 2793; see archived text at https://web.archive.org/web/20200925202550/https://iea.uoregon.edu/treaty-text/1971-wetlandsentxt (accessed 1 Dec 2022). It includes Article 10bis on amendments, which was not in the original version of the agreement.

[78] The title alone is not enough, e.g. the title "Agreement concerning financial co-operation." occurs over 500 times in UNTS.

[79] Available for download in csv format at https://iea.uoregon.edu/base-agreement-list/csv-all?attach=page (accessed 13 Sept 2019).

and illustration purposes. The final script is the result of iterative improvement and each line of code was tested in an interactive interpreter session to ensure it works as expected. Filtering out agreements for which there was no text available in the database reduced the set drastically, from 5665 to 1965 entries. Of these, the 300 agreements classed as bilateral and 'other' non-multilateral (e.g. national legislation), as well as 185 non-English texts were excluded, leaving 1490 agreements.

2.2.1   Successive versions and parts of treaties

Next, 555 obvious amendments were removed, such as amending agreements where the title starts with the word "Amendment". While it would be feasible to automatically reconstruct the full version of a treaty for each time period and party concerned, this would require considerable effort given that amendment insertions and deletions are not always defined consistently within regimes, let alone across regimes.

In cases where the treaty only entered into force after an amending agreement had been adopted, such as UNCLOS-82 and MARPOL-73, this fact was added as an indicator variable to be taken into account in the analysis of political success.

As for annexes and protocols, if they were listed separately in the IEAdb but were adopted at the same time and in the same act as the treaty, then their text was appended to the treaty for joint analysis. If their adoption and entry into force was different from the main treaty's, more like supplementary protocols to a framework convention, then they were included as a treaty in their own right so long as their text was sufficiently self-contained.

Regarding treaties which have been terminated or replaced by a new treaty ratified by all or a large majority of parties to the earlier treaty, such as a number of law of the sea conventions and several successive versions of the International Tropical Timber Agreement, only the most recent treaty available in the Australian Treaty Library or IEAdb was selected. As detailed information representation on regime evolution by way of amendment was deemed too time-consuming for the purposes of the present research project, it seemed unjustified to include all successive versions of regimes which evolve by replacing rather than amending treaties. This exclusion criterion cannot be gauged from treaty metadata alone and was thus determined during the case-by-case selection process after the initial automated filtering. For some areas, such as maritime labour law and health & sanitation law, the IEAdb contained dozens of superseded treaties but not the actual superseding treaty, i.e. the 2006 Maritime Labour Convention and the International Health Regulations (any version).

2.2.2 Open and global treaties

Additionally, closed multilateral agreements, treaties with a restricted geographic scope, and treaties concluded under the auspices of a regional/restricted organisation were excluded. Whereas off-the-shelf NLP tools can be used on any number of agreements, customised legal information extraction requires extensive tweaking and testing to ensure accuracy. Hence, some restriction was necessary, and choosing open multilateral treaties with global application seemed most appropriate for the present research project. Indeed, many of today's pressing environmental problems require global coordination. While regional treaty systems such as that on Long-Range Transboundary Air Pollution[80] provide a fertile ground for experimentation with legal tools and techniques, ultimately what counts is scalability to the global level, and what works at a regional level might not work at a global level. Moreover, even when some states are not directly affected by a given environmental issue, such as desertification or sea level rise, global regimes such as those set up by the UNCCD[81] and UNFCCC[82] can facilitate and coordinate international assistance to tackle such unequally distributed problems.

However, compared to random sampling, this selection strategy is decidedly more challenging. One approach would be to distinguish based on the number of parties, e.g. 150 or more could count as global, but this would introduce a selection bias because some treaties with global aspirations have not yet reached their potential. Including the (as yet) failures as well as the successes is crucial for making progress towards answering the research question. Another option would be to rely on an existing list of global or open multilateral treaties. Several partially overlapping lists were examined but found inconsistent. Discrepancies with regard to Ecolex are outlined in Table 2.1 below, after introducing the procedure chosen.

Ultimately treaty provisions on participation and geographic scope are decisive, but for the sake of expediency, as many selection decisions as possible were made based on the treaty title alone. From the set of 935 treaties obtained by the filtering process described above, 142 agreements can be subtracted because the treaty title names its restricted set of parties. Another 63 contain the name of an inland watercourse, lake or mountain range, 87 refer to regional seas, straits and islands, and an impressive 117 mention a specific ocean or ocean current. 505 agreements remain

---

[80] Convention on Long-Range Transboundary Air Pollution (adopted 13 November 1979, entered into force 16 March 1983) 1302 UNTS 217, and its protocols.

[81] United Nations Convention to Combat Desertification in those Countries Experiencing Serious Drought and/or Desertification, Particularly in Africa (adopted 14 October 1994, entered into force 26 December 1996) 1954 UNTS 3.

[82] United Nations Framework Convention on Climate Change (adopted 9 May 1992, entered into force 21 March 1994) 1771 UNTS 107.

after these exclusions, from which 110 can be removed based on a reference to a
continent (except for "Africa" in the UNCCD's treaty title) and a further 51 due
to the inclusion of the name or acronym of a regional organisation such as ASEAN
or some other indication of restricted membership, such as "Nordic", "Arab", or
"Tripartite". Finally, 31 soft law instruments were excluded based on fairly uncon-
troversial keywords in the instrument title such as "Council Decision", "Code of
Conduct" or "Plan of Action", before proceeding to a case-by-case manual selection
among the remaining 300+ agreements.

To mitigate against any potential biases in the IEAdb and compare different ap-
proaches, data from Barrett's table of treaties[83] and Ecolex's list of global treaties[84]
were manually added to the IEAdb spreadsheet.

While openness of treaty participation is a spectrum, the openness threshold for
filtering purposes was set to treaties which allow at least all parties to a global treaty
to accede upon approval by existing parties. The most open treaties allow all states
to sign, ratify and accede to them on an equal basis, without special invitation or
approval.

Recognising that 'global'[85] treaty is not well-defined, here it is understood as
a treaty whose parties represent at least 75% of the total number of states in the
world and all continents except Antarctica. Accordingly, intergovernmental organ-
isations based on a global treaty are global organisations. Most treaties and orga-
nizations which are global today were not global in their early days, including the
UN and its specialised agencies. Thus, treaties making accession conditional upon
UN membership would not have reached this openness threshold during the UN's
initial phase, but subsequently they did, and that is what matters for this filtering
exercise. By contrast, treaties open only to members of the Organisation for Eco-
nomic Co-operation and Development (OECD) fail to reach the threshold, because
the OECD is not a global organisation, quite far from it.[86] In fact, the existence
of ideologically restricted organisations like the OECD motivated this condition of
global membership. While the OECD grew out of a regional organisation, now its
constituent convention is nominally open to "any Government prepared to assume

---

83   Barrett, *Environment and Statecraft* (n 12), p.165-194 ("Appendix 6.1. Multilateral Environ-
      mental Agreements").
84   https://www.ecolex.org/result/?q=&type=treaty&xdate_min=&xdate_max=&tr_field_
      of_application=Global yields 170 agreements as at Dec 2021.
85   'Universal' would have been more aligned with established practice (as in "universal ratifica-
      tion", "universal jurisdiction", or as used in the 1975 Vienna Convention on the Representation
      of States in their Relations with International Organizations of a Universal Character), but
      representing the whole Universe seems a tad overambitious at our current state of science and
      technology, without knowing whether there are other civilisations outside our solar system.
86   38 states parties as at 25 May 2021.

the obligations of membership" upon invitation from the OECD Council.[87] This accession provision is not unlike that of the UN Charter and of constitutions of its specialised agencies, but in practice organisations within the UN system have about five times more members than the OECD and almost any state may join. Hence, treaties restricted to UN members are much more open than treaties restricted to OECD members. Openness is thus understood as a measure of barriers to entry and is inherited by downstream treaties. For instance, if a state not presently a member of the ILO were interested in acceding to an ILO convention, it would most likely be able to become a member of the organisation fairly swiftly and thereafter accede to any treaty adopted under its auspices.

This conceptualisation of openness contrasts with Barrett's more restrictive binary classification, qualifying as 'open' only those treaties which are open to all states, and all others as 'restricted'. Thus, an agreement open to all members of a global organisation receives the same label as a trilateral treaty. A protocol to a convention open only to parties of the convention, even if the convention itself is open to all states, is likewise classified as restricted. Therefore, it is unsurprising that out of Barrett's list of 298 agreements only 52 are categorised as open. This is an example of a classification scheme which seems insufficiently holistic in its narrow, atomistic focus on a single treaty at a time, without considering the legal context.

The UNTS index gathered by the UNTS searchbot (see Appendix B.1) has a similar column called 'Treaty Type' in the source which categorises agreements into 'Bilateral', 'Open Multilateral' and 'Closed Multilateral'.[88] For instance, the above-mentioned OECD Convention is classified as a closed multilateral agreement. Supplementary protocols and amending agreements to open multilateral treaties inherit the openness classification, in line with a holistic perspective.

As for restrictions of geographic scope in treaty provisions, in some cases the convention area is defined directly in terms of coordinates or agreed geographic denominations, and in others it is specified indirectly as the range of a species of interest. Most geographically restricted agreements are also closed, but not all of them, hence the need for this additional selection criterion.

The Ecolex database has a filtering criterion called "Field of Application" which can be set to "global" and returns about 170 agreements including amendments. While largely overlapping with the approach taken here of considering aspects of

---

[87] Convention on the Organisation for Economic Co-operation and Development (adopted 14 December 1960, entered into force 30 September 1961) 888 UNTS 179, art 16.

[88] The individual treaty pages also have an 'Agreement type' field, but it only distinguishes between bilateral and multilateral treaties.

both treaty participation and geographic scope, there are nonetheless some treaties included in the Ecolex global application list that would not fit the criteria used here. These exceptions are listed in Table 2.1.

Table 2.1: Agreements misclassified as global in Ecolex

Year	Agreement Title	Ecolex ID
1920	Treaty regulating the Status of Spitsbergen and conferring the Sovereignty on Norway	TRE-000156
1964	Convention for the International Council for the Exploration of the Sea	TRE-000502
1991	European Energy Charter	TRE-160046
1993	Memorandum of Understanding Concerning Conservation Measures for the Siberian Crane	TRE-001295
1993	Convenio Regional sobre cambios climáticos	TRE-146580
2014	Agreement between the International Atomic Energy Agency, the Government of the Republic of Ghana and the Government of the People's Republic of China for Assistance in Securing Low Enriched Uranium for a Research Reactor	TRE-160037

This presentation of contrasting classifications only scratches the surface of participation and geographic scope restrictions. Still, despite the complexity, it seems possible to agree on well-defined criteria for classifying treaties along these lines, at least among international lawyers. As there is no consensus yet on what counts as an open and/or global treaty, these terms were not added to the PILO ontology (in Appendix E), only the uncontroversial terms 'bilateral treaty' and 'multilateral treaty' (as subtypes of treaty).

The use cases of such classifications beyond obtaining a treaty set with similarly wide geographic scope and participation potential are manifold. Research and advocacy regarding treaties that states could in principle accede to but do not, and the patterns in such abstentions over time, can be interesting and fruitful. In many cases it may well be an oversight due to the sheer amount and pace of international legal and political developments, rather than a deliberate act not to join, especially with treaties that do not have a secretariat or organisation actively promoting participation. Additionally, global treaties are more convincing evidence for (global) *opinio juris* than regional/restricted agreements, and may be more likely to reflect customary international law. Hence, a treaty database with these filtering options could be more useful for litigation, arbitration, and academic research on custom than one that merely distinguishes between bilateral and multilateral treaties.

2.2.3 Environmental focus and relevance

Finally, the thorny question of selecting 'environmental' treaties remains to be addressed. As mentioned in the introductory chapter, "natural environment" and "nature" are fuzzy and socially constructed concepts, and it may be scientifically and logically more consistent to forego the distinction between natural and built environment, and between natural and artificial substances. But the fact of the matter is that these distinctions are still entrenched in law and literature. Thus, 'environmental' law and policy is typically concerned with the natural environment or some of its components or features. Moreover, in many scholarly accounts "environmental" actually means "environmentalist" in the sense that the law or policy in question is required to be explicitly and primarily aimed at environmental protection. Yet another step in this direction would be to include only those that attribute intrinsic value to environmental protection, but that is less common. Still, this concern with and reification of environmental protection *per se* risks drawing unhelpful boundaries. If two treaties have exactly the same content except for one provision laying out the objective of the agreement, wherein one mentions environmental protection whereas the other does not, then the former would be included in the environmentalist scholar's analysis whereas the latter would not. Similar variations would make the treaty pertain to the body of "sustainable development law" or "natural resource law".

Mitchell's classification rests upon a purpose-based distinction, defining an IEA as "an intergovernmental document intended as legally binding with a primary stated purpose of preventing or managing human impacts on natural resources."[89] This classification is said to be undertaken by searching for relevant terms in agreement titles, preambles, or articles setting out agreement goals.[90] In practice, however, according to information contained in the database itself (v.2018.1),[91] the vast majority of treaties were classified based on the title (96%), a further 2% based on the preamble, and only 27 (0.5%) out of 5809 records with environmental classification source information mention a treaty provision. In the much smaller set of treaties selected for the present study, 10% of treaties received their IEAdb label based on their text (preamble only), the rest were judged by the title. Needless to say, in some cases the primary purpose can be gauged from the treaty title, but in many cases it cannot. What counts as a 'primary' and 'secondary' purpose is also tenuous at

---

[89] Ronald B Mitchell, 'International Environmental Agreements: A Survey of Their Features, Formation, and Effects' (2003) 28(1) Annual Review of Environment and Resources 429, p.432.

[90] ibid, p.433.

[91] The IEAdb spreadsheet conveniently includes a column called "Source for E (environmental) code".

best, especially in the many cases where treaties do not explicitly state or rank their objectives in this way. For instance, UNCLOS[92] and the Marrakesh Agreement[93] are both classified as environmental treaties in the IEAdb because they mention environmental protection in their respective preamble. While many scholars would probably consider these agreements highly relevant for environmental protection, few would see it as their primary purpose or objective.

By contrast, the present study used two criteria for deciding borderline cases: environmental focus (envFocus) and environmental relevance (envRelevance), both as a score of 0-5. envFocus was based on the number of environmental provisions or mentions in the treaty text, and envRelevance on how relevant the treaty could be for prevention or management of environmental problems if it were to achieve its potential. Neither depends on the primary purpose or stated aims of the treaty, which may be centered on worker protection, public health, economic efficiency, or whatever else happened to be most important to the treaty negotiators at the time.

Rather than dwelling on definitions, it may be more illustrative to compare the outcome of different conceptualisations in a few borderline domains. Discrepancies are particularly pronounced in the set of treaties adopted under the auspices of the International Labour Organisation (ILO). Table 2.2 lists those included in the present study as well as their categorisations by Mitchell, Barrett and Ecolex.

Table 2.2: Inclusion of ILO treaties

Year	Short Title	IEAdb	IEAdb Basis	Barrett	Ecolex
1960	Radiation Protection Convention	MNA	Title	No	No
1971	Benzene Protection Conv.	MNA	Title	Yes	No
1974	Occupational Cancer Convention	MNA	Title	Yes	No
1977	Working Environment Conv.	MEA	Title	Yes	Yes
1981	Occupational Safety & Health Conv.	MNA	Title	Yes	No
1985	Occupational Health Services Conv.	MNA	Title	Yes	No
1986	Asbestos Safety Conv.	MNA	Title	Yes	No
1988	Safety & Health in Construction Conv.	MNA	Title	Yes	No
1989	Indigenous & Tribal Peoples Conv.	MuA	Title	No	Yes
1990	Chemicals Safety Conv.	MNA	Title	Yes	Yes
1993	Major Industrial Accidents Conv.	MEA	Title	Yes	No
1995	Safety & Health in Mines Conv.	MEA	Preamble	Yes	No
2001	Safety & Health in Agriculture Conv.	MNA	Title	No	No

As this table shows, most inclusion decisions for the IEAdb were made based on the title alone. Of the three compilations, Barrett's is the most inclusive with regard

---

[92]   United Nations Convention on the Law of the Sea (adopted 10 December 1982, entered into force 16 November 1994) 1833 UNTS 3.

[93]   Marrakesh Agreement Establishing the World Trade Organization (adopted 15 April 1994, entered into force 1 January 1995) ATS 1995 8.

to treaties concluded under the auspices of the ILO. The only agreement classified as environmental in all three of them is the Working Environment Convention, perhaps because it contains the term "air pollution" in its full title.[94] However, the primary purpose of this convention is clearly the protection of workers, not of the environment per se, thus its inclusion as a MEA in the IEAdb seems unjustified given the latter's selection criteria.

Nevertheless, international labour law is precisely one of the domains that should not be overlooked by scholars interested in environmental protection and sustainable development. After all, workers are at the frontline of industrial pollution and often the first to suffer its consequences. Educating and empowering workers to report and correct unsafe practices may be one of the most effective ways to prevent and mitigate pollution. Indeed, the Chemicals Safety Convention explicitly notes in its preamble that "the protection of workers from the harmful effects of chemicals also enhances the protection of the general public and the environment",[95] and yet Mitchell somehow considered the treaty title sufficient to classify it as non-environmental.

For the purposes of the present research project, all of the agreements listed in Table 2.2 were included, and selection decisions were made based on the operative part of the agreement, not merely the title or preamble. The full IEAdb contains nine additional ILO treaties, one of which is classified as environmental, but it was excluded from the present list on account of its pre-1945 adoption date. The same 1921 treaty (on white lead in paint) also featured in Barrett's list, other than that neither Barrett nor Ecolex included further labour conventions. For the sake of completeness, the ILO's list of up-to-date conventions[96] was browsed for additional candidates and five more treaty texts were examined in-depth, but found insufficient in environmental focus and relevance compared to other treaties.

In the case of treaties focused on health and safety, if they were purely concerned with measures such as personal protective equipment for workers without attempting to reduce the danger at the source or to prevent leakage into the environment, then they were not deemed to be sufficiently environmental. For instance, the Convention (No. 152) concerning Occupational Safety and Health in Dock Work[97] did not

---

[94] Convention (No 148) concerning the Protection of Workers against Occupational Hazards in the Working Environment Due to Air Pollution, Noise and Vibration (adopted 20 June 1977, entered into force 11 July 1979) 1141 UNTS 106.

[95] Convention (No 170) concerning Safety in the Use of Chemicals at Work (adopted 25 June 1990, entered into force 4 November 1993) 1753 UNTS 190, fourth preambular paragraph.

[96] "Up-to-date Conventions and Recommendations", available at www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12020:0::NO:::.

[97] Convention (No 152) concerning Occupational Safety and Health in Dock Work (adopted 25 June 1979, entered into force 5 December 1981) 1260 UNTS 3.

meet the requirements because its main environmental aspects are provisions relating to fire fighting and handling of dangerous cargo, only referring to international regulations applying to the transport of dangerous goods rather than adding to them. While this reference does provide port authorities with an additional ground for refusing dangerous goods which are inadequately packed, marked or labelled, SOLAS[98] and other international agreements governing the transport of dangerous goods already have their own compliance control and enforcement mechanisms. Convention No. 152 also lacks provisions regarding recycling and waste disposal, which are not uncommon in other ILO treaties. To give but one example, the Safety and Health in Agriculture Convention requires states parties to ensure that there is "a suitable system for the safe collection, recycling and disposal of chemical waste, obsolete chemicals and empty containers of chemicals so as to avoid their use for other purposes and to eliminate or minimize the risks to safety and health and to the environment".[99] Apart from waste treatment, its provisions on the sound management of chemicals also mandate preventive and protective measures for the preparation, handling, application, dispersion, storage and transportation of chemicals. The environmental impact of large-scale agricultural use of certain types of pesticides and fertilisers on biodiversity, groundwater quality, eutrophication and acidification is well established. Thus, ignoring this treaty merely because its title or preamble does not mention environmental protection seems rather short-sighted.[100]

Similar observations could be made about other areas of international law that have been neglected by environmentalist lawyers and are not featured in large environmental treaty databases like Ecolex and IEAdb even though they arguably should. Granted, there are borderline cases and my own classification may not be internally consistent. This part of the selection process was very tedious and I only undertook it because I noticed some omissions in Ecolex and Mitchell's MEA list and was curious to investigate and see if I could do better. In my view, databases on a subject with fuzzy boundaries should be overinclusive rather than underinclusive, so as to accommodate different interpretations and conceptions of the domain. Moreover, an expansive approach is also more likely to bring agreements to the fore that have unrecognised impacts or underutilised potential for environmental problem-solving. Much of transport and energy law has or could have environmental relevance. Including agreements from some of these areas in the study set

---

[98]   International Convention for the Safety of Life at Sea, 1974 (adopted 1 November 1974, entered into force 25 May 1980) 1184 UNTS 278.

[99]   Convention (No 184) concerning Safety and Health in Agriculture (adopted 21 June 2001, entered into force 20 September 2003) 2227 UNTS 241, art 12(c).

[100]   In the case of Barrett's list the absence of this treaty is likely due to his cut-off date, which is some time in 2001.

also provides a more diverse picture of treaty design, with more opportunities for comparative analysis and cross-fertilisation.

As for the larger epistemic challenge, a legal data science approach to subject matter analysis will be discussed in Chapter 4, and if a collaborative long-term version of it succeeds, the kind of manual selection undertaken here will not be necessary anymore. What counts as 'environmental' will likely remain contested, however, and for future treaty analytics projects I would recommend choosing a more easily delimitated domain. Still, the finding that current environmental law databases are underinclusive in some regards will hopefully be of use to researchers and database maintainers. Ultimately, improving UNTS online subject matter indexing and search may be more impactful than enhancing domain-specific databases, but more on this in Chapter 4.

2.3 Text analysis methods and applications

Having covered text collection and selection, the choice of text analysis methods and tools remains to be explained in more detail than was possible in the introduction.

The search and selection of methods and tools was guided by the goal of obtaining comprehensive, intelligible, reproducible, accurate, verifiable, structured data for each treaty which is useful for comparative analysis of treaty design and for prediction of outcome variables in the context of this research project but that can also be re-used and built upon by other researchers and made publicly available in the form of an online database (or added to existing UN databases).

Comprehensiveness of treaty features modelling was aspirational and subordinated to other desired attributes. Capturing every detail of the agreements would have entailed not being able to cover a wide range of design features because each additional piece of information takes time to accurately extract and verify. An overly fine-grained typology would likewise have led to largely idiosyncratic information about each treaty which would have precluded quantitative analysis at worst and not contributed to it at best. For instance, if each aspect of amendment provisions were captured and combined into a subtype of amendment provision, then many treaties would be the sole member of their class. Instead, only one subtype of amendment provision was constructed with the label of 'simplified amendment provision', as discussed in Chapter 3.1.3. The goal of comprehensiveness thus favoured breadth over depth.

By intelligibility I mean that the variables (and variable names) express information that is understandable to a domain expert and is consistent with domain

knowledge. This is closely related to utility for the legal research community (e.g.
for testing existing hypotheses on this treaty set), legal education, treaty negoti-
ations, litigation, activism, journalism, and for access to legal information more
generally. For example, Mitchell's coded agreements dataset[101] includes a variable
called 'RESV' which stands for "Reservations, terminations, withdrawals, opt out
clauses, escape clauses, and suspension provisions"[102] – a hodgepodge of provisions
with little logic behind it from a lawyer's perspective, and a variable name that
could at most hint at reservations, but not at the rest of the mixed bag. The group-
ing is explained in the coding manual as being based on the possibility of limiting
the application of the treaty by or for a given party,[103] but they are such different
legal scenarios that the practical and analytical utility is limited. Moreover, treaty
termination, as this term is understood in international law, is not actually included
in this variable, but is instead classified under a variable called 'EIF' - "Entry into
Force",[104] yet another example of inaccurate variable names and imprecise delimita-
tions. Clustering of variables into higher-level categories can of course be useful for
analysis, presentation and communication, but in this case the lowest level variables
(RESV and EIF) are arguably not informative or precise enough.

The desire for meaningful, useful information outputs also precludes the use
of standard bag-of-words, latent semantic analysis and unsupervised learning ap-
proaches more broadly. By definition, unsupervised methods do not leverage domain
knowledge, and thus cannot solve the epistemic challenge faced by international
lawyers. This kind of NLP can elucidate some aspects or answer some questions,
but typically not the questions legal scholars are most interested in.

Instead, in this thesis, variables were modelled as closely as possible after gener-
ally accepted domain knowledge in international law and terminology used in inter-
national codification efforts such as the Vienna Convention on the Law of Treaties[105]
and other work of the International Law Commission (ILC).[106]

Accuracy refers to the absence of false positives (overinclusion) and false nega-
tives (underinclusion). The feasibility and effort required for accuracy depends on
the size of the treaty set, the number of variables, how clearly defined these variables
are (in literature/domain knowledge), and how recognisable they are in the text. For
instance, entry into force, amendment, denunciation, termination and dispute settle-

---

[101]   See https://iea.uoregon.edu/coded-agreements for an introduction, the coding manual, and
       other useful resources.
[102]   Ronald B Mitchell and Steven Rothman, 'International Agreement Coding Manual' version
       1.0 (2007) p.21.
[103]   Ibid.
[104]   Ibid, p.21-22.
[105]   VCLT (n 18).
[106]   https://legal.un.org/ilc/ (accessed 2 Feb 2022).

ment provisions are very well understood and use similar terminology across treaties, whereas national reporting, non-compliance procedures and sanctions provisions are not as homogenous across the full set of treaties (only within subsets) and are thus more prone to borderline cases and potential misclassifications. Such difficulties are described in the relevant subsections of Chapters 3 and 4.

Reproducibility (in the context of data science) means that anyone with the same computer programs installed could run my scripts and obtain the exact same results, from treaty text collection,[107] pre-processing, information extraction, database construction, analytics, to figures and tables etc. In political science, the term is sometimes used to refer only to the tail end of the research process, i.e. producing figures from data tables, very far from the meaning it takes in hard sciences. The 'coding manual' (instructions for manual data entry) is often incomplete and uses different variable names than in the paper, appendix and data tables shared (if any). Many of the research papers based on Mitchell's IEAdb by authors other than himself fail to provide a unique, unambiguous identifier (such as the integer ID given by Mitchell, the url of the text, or the treaty title exactly as included in the IEAdb), which would facilitate reproducibility and comparison of results. That is if they even provide a table with the values assigned to each treaty rather than merely a table of summary results, which is of little use for assessing accuracy of the data. As mentioned in the section on data collection, automated reproducibility also facilitates updating of the analysis when higher quality texts become available or when expanding the set of treaties covered. Tweaking a web scraping or information extraction script is child's play compared to writing it from scratch.

Verifiability is related to reproducibility but is important at all stages of a project, from initial variable selection, information extraction, analysis, through to presentation, dissemination and usability of results. In this context, it required extracting information underlying each text-based variable, even when the final analysis only uses a boolean (true/false) transformation. For instance, the variable `provisAppProv` expresses whether a given treaty contains a provisional application provision or not, and the variable `provisAppProvStr` contains the exact string used to construct the binary variable, and thus presents the evidence for its accuracy. Sharing this textual evidence is still rare unfortunately, Mitchell being an exception with his coded agreements dataset. In today's fast-paced and high-volume information society, being able to quickly judge whether or not a paper, database or other information output is based on accurate data can make all the difference to successful uptake and impact of academic work. Web publications have the advantage of more com-

---

[107] Assuming remote sources stay the same, but if not the run could start from my own copy of the texts.

pact and yet well-grounded, extensible information representation, obviating the need for a reader to go back and forth between indices, glossaries, figures, tables, text and appendices of a lengthy tome.

More broadly, the kind of outputs best suited to illustrate the benefits of structured treaty data without reinventing the wheel are arguably (i) a treaty profile page with structured summary information about relevant aspects of the treaty with, in its online version, embedded excerpts of text that allow one to immediately verify the accuracy of the displayed variables (e.g. the string of the national reporting provision),[108] (ii) a treaty text page with the full text of the agreement and annotations showing what parts of it were extracted and how they were transformed into structured information,[109] (iii) overall summary statistics of the treaty corpus with the possibility to narrow down by criteria such as topics, types of provisions, parties, time periods etc.,[110] (iv) an ontology page with class definitions and instance lists,[111] and (v) a search interface that allows queries for a single or combination of variables and returns lists of treaties and provisions/excerpts matching the search criteria (e.g. all transport-related treaties with compulsory dispute settlement provisions). The intended audience is as much practitioners as it is academics. Treaty negotiators may wish to have a wider range of sample provisions at hand to draw inspiration from, or a more precise set to support their proposed draft. Activists and legal counsels may benefit from the ability to narrow down the search space of applicable law by legally relevant criteria, especially under-resourced NGOs and law firms in developing countries. Journalists and pundits would probably appreciate a resource to double-check their assumptions without having to read through lengthy texts.

All these outputs were produced in a proof of concept version for the present thesis. To make them usable for a broader audience would likely require extensive consultations, surveys, and user experience testing in an iterative R&D process. The widest possible reach would probably only be achieved if some or all of an improved version could be integrated into UN and other public sector websites and databases. The data, models, and tools used and developed in the course of this project are all geared towards such an ambitious uptake. The environmental focus of some of the subject matter variables, while important for this thesis, may or may not be worth including in the first collaborative version. The addition of other focus areas, such as human rights or finance, may also help attract more interest.

---

[108]   See Appendix G for an example.
[109]   See the GATE screenshots in Figures 2.1 and 4.3 below.
[110]   Chapters 3 and 4 illustrate this with summary tables and figures.
[111]   Appendix E provides a snapshot of PILO, but the interactivity of the desktop or web version of Protégé cannot be conveyed adequately in a static medium.

Additional outputs include instructions, files and code to automate treaty references, such that referencing a provision of e.g. the Minamata Convention on Mercury in a suitable editor[112] is as simple as `[[cite:MercuryConv][art 27(4)]]` which would automatically be converted to a proper OSCOLA style treaty citation and also add an entry to a treaty index, if present.

Specifically, the tool best suited for customised information extraction meeting the selection criteria outlined above was found to be regular expression (regex) pattern matching, widely used in software applications dealing with text. Regex modules define a specialised syntax allowing the specification of a text search pattern in a sophisticated and succinct way. Matches can be counted or extracted and stored in a variable. For fast trial & error the regular expression module of the Python programming language was used, but for the collaborative version the Java regular expression plugin within the General Architecture for Text Engineering (GATE)[113] software may be better. GATE facilitates collaboration between programmers and non-programmers, among others by providing clear and intuitive data visualisation and search functions in the graphical user interface. It also provides the possibility to populate an ontology and export the resulting knowledge base in standardised formats. Figure 2.1 shows a screenshot of sample treaty provisions annotations visualised in the GATE Developer desktop software.[114] Annotation sets and types can be selected in the right-hand side bar, and an annotation list can be displayed at the bottom, which displays all the annotation features, including those used to populate the ontology. Appendix F.2.1 contains code excerpts for generating these annotations in a syntax that will be explained in section 3.1.1 below.

To start with a simple example, the shortest regular expression used in the present thesis is `"(?i)ozone.(?:deplet|layer)\w*"`. This pattern first turns on case-insensitive matching (with `(?i)`) and then instructs the engine to find any occurrence of 'ozone' followed by any character other than a newline (the dot has a special meaning in regex patterns),[115] and ending with either 'deplet' or 'layer' (vertical bars divide between alternatives). `\w*` completes the word with additional alphanumeric characters, if present. The search returns 50 results across the full treaty set, 45 of which are references to the ozone layer in lower, upper or title case, and the

---

[112] GNU Emacs Orgmode with the org-ref package and a Texlive installation with the oscola package was used here, though it could be adapted to other Biblatex-capable editors.

[113] Described in Cunningham and others (n 61).

[114] The excerpt stems from the International Convention on Civil Liability for Bunker Oil Pollution Damage (adopted 23 March 2001, entered into force 21 November 2008) ATS 2009 14, arts 14-15.

[115] The dot can be used to match any character, even a newline, with the dotAll flag `(?s)`. A single dot matches one character, `.?` matches zero or one, `.*` matches zero or more, and `.+` matches one or more.
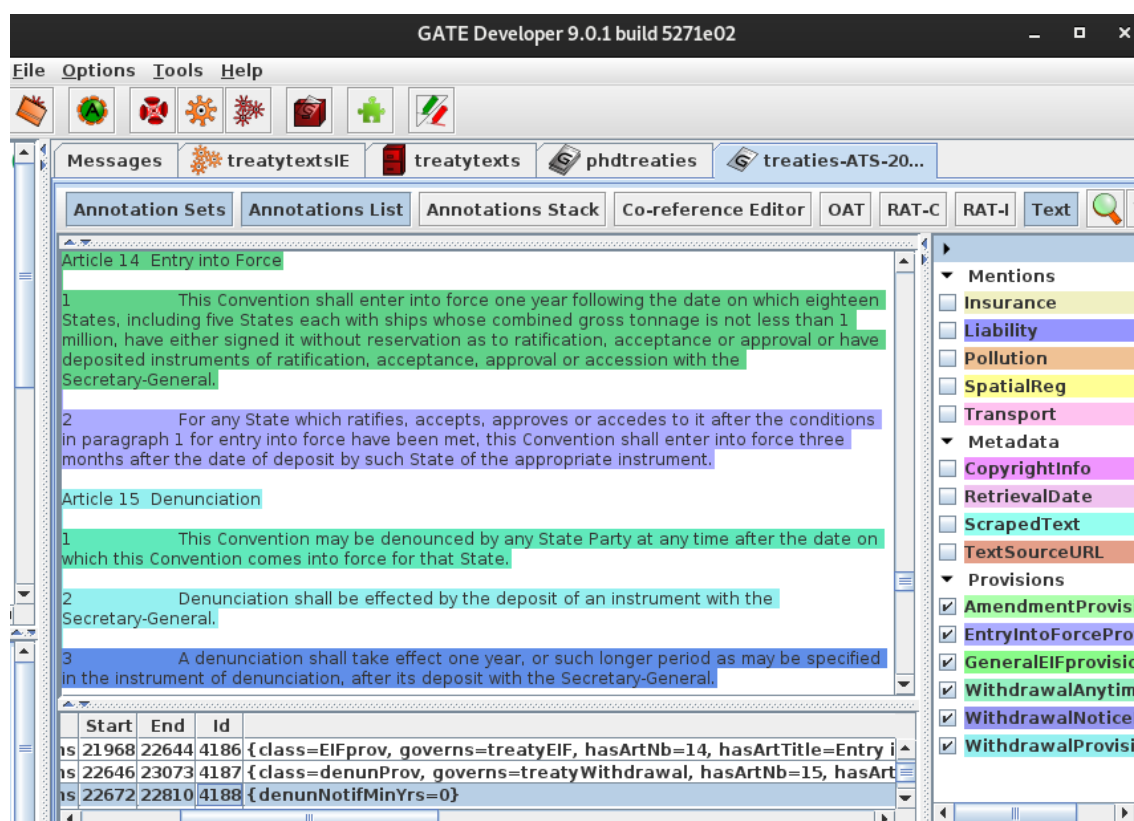
Figure 2.1: Treaty provisions annotations in GATE

rest occur in expressions such as 'ozone depleting potentials' or 'ozone-depleteing[sic] substances'. All other occurrences of 'ozone' were extracted with a suitable context window (a few words before and after the term)[116] to ensure no relevant mentions are missed.[117] This sample pattern does not even come close to showing the power of regular expressions, but rather than expounding on it here, more examples and explanations will be given in the results sections of Chapters 3 and 4. The point of this example is mostly to illustrate the simplicity and readability of regular expressions, even for less tech-savvy readers. In fact, I only learned about regular expressions during my doctoral studies and now believe they should be taught to any undergraduates having to ingest large amounts of texts, including law students. Thus, collaborators willing to contribute to regex crafting would gain a transferable skill that can be used in many other text analysis projects.

Finding the right text excerpts is of course only the first step in the analysis pipeline. What happens next depends on the target variables. For legal information extraction (Chapter 3), text excerpts were converted into boolean, numeric, or cate-

---

[116] E.g. `(?i)(.{,30}ozone.{,30})` extracts up to 30 non-newline characters on either side (if present). This kind of false negative search can be constrained to treaty texts with no or few mentions of the target term, making verification more expedient.

[117] See Chapter 4.1.2 for more information.

gorical variables. For treaty subject matter analysis (Chapter 4), occurrence counts were divided by respective text lengths for comparison of feature importance. Both types of text-derived variables warrant their own discussion and analysis, to which we now turn.

3 Treaty provisions analysis

This chapter is concerned with some of the main recurring components of multilateral treaty design, grouped into three clusters: treaty life cycle (3.1), treaty organs (3.2), and compliance and effectiveness (3.3). The first cluster covers provisions on provisional application, general entry into force, amendments, supplementary agreements, withdrawal and termination. It illustrates how a wide range of temporal and other stipulations can be identified and transformed into variables suitable for statistical analysis.

The second cluster deals with information regarding organisational auspices, provisions on legal personality, on the secretariat or similar administrative organ, and on the supreme plenary decision-making body often called Conference of the Parties (COP). This set of variables attempts to take the institutional configuration and legal context of treaties seriously. The categorical variable of organisational auspices, which is not a type of provision, but still legal and political information, provides a means for feature inheritance, such that e.g. ILO conventions have the `secretariat` variable set to `True` even though the agreement texts contain no provision on the secretariat. The ILO Constitution, which would provide the legal basis for administering the conventions adopted by the organisation, is not in the present dataset. This workaround of capturing evidence of organisational auspices and then adjusting relevant variables accordingly, attains the goal of accurately representing the legal and organisational reality. This is an example of the required holistic perspective mentioned in the introductory chapter. Another auxiliary variable used in the same way is `supplAgTo` which links supplementary agreements with their supplemented treaties for lookup and copy of variables in case of an incorporation clause.

As for the third cluster, it encompasses provisions broadly aiming at compliance, implementation and effectiveness. Specifically, the six types of provisions covered govern periodic national reporting of treaty implementation, international verification of compliance, general treaty reviews, non-compliance procedures, sanctions or other consequences for non-compliance, and judicial or arbitral dispute settlement. With the exception of dispute settlement provisions, the main challenge in this cluster is consistent identification and delimitation of the types of provisions in-

stantiated in treaties governing very different problems and activities. Perhaps some of these are not proper 'types' of provisions, but simply new labels for variations of long-standing practices previously known under a different name.

In practical terms, treaties that have named articles facilitate the task of information extraction with regular expressions, because they mark the core provision to capture, no matter how many times relevant keywords like 'amend', 'report' or 'dispute' occur elsewhere in the agreement. Likewise, provisions that tend to be expressed in self-standing articles, like those on entry into force and dispute settlement, are easier targets than ones which are often entangled with others and have to be interpreted in the context of a long and convoluted provision on something else. Which excerpt to select as textual evidence in cases where the target is only part of a subparagraph in an lengthy enumeration of, say, the duties of the COP is debatable. The guiding principle used is that the excerpt should be just enough to allow for verification of the constructed variable.

Regarding data analysis, since the aim is to make it accessible to lawyers without statistical training, the analysis is only slightly more quantified than what one would find in normal doctrinal analysis. Rather than vaguely describing the occurrence of a type of provision as "infrequent" or "not uncommon", I put numbers to this, both absolute and relative (as a percentage of the total). Figure 3.1 displays a high-level overview of the main variables, their type (boolean, numerical, or categorical), the number of treaties for which the variable is constructed, as well as their prevalence (as a percentage) or their mean value in the treaty set at hand. More details of the calculations and how they can be used for descriptive analytics are given in section 3.1.1 below.

As for associations between variables, a straightforward and intuitive way to explore these is to look at subset prevalence (or conditional probabilities) as compared to full set prevalence (or independent probabilities), when focusing on a particular type of provision encountered in treaties. This will also be explained in section 3.1.1, analysing the subset of treaties which include provisional application provisions. A related measure introduced in section 3.3 is the Jaccard similarity coefficient, calculated by dividing the intersection by the union of two sets. Numerical variables constructed based on treaty life cycle provisions lend themselves to additional statistical descriptions of mean, range, and correlations, with suitable caveats. Another dimension of interest to many legal scholars is the evolution over time, i.e. associations with the `treatyAdoptionYear` variable. This is described in the text and visualised for some of the variables of the third cluster by showing how many treaties were adopted with vs. without a given provision per decade, starting

Cluster	Variable	Type	N. Treaties
		Type	**N. Treaties**
Start of Life	Provisional Application Provision	bool	14 (12%)
	Entry into Force Provision	bool	119 (100%)
	—Min. Nb. of Ratifs for EIF	num	112 (19.8)
	—Additional EIF Conditions	bool	47 (39%)
	—EIF Waiting Period (Years)	num	110 (0.4)
Change	Amendment Provision	bool	115 (97%)
	—Simplified Amendment	bool	60 (50%)
	Adoption of Annexes or Protocols	bool	26 (22%)
End of Life	Denunciation Provision	bool	113 (95%)
	—Denunciation Waiting Period (Years)	num	97 (2.7)
	—Other Denunciation Conditions	bool	33 (28%)
	—Denunciation Effect Delay (Years)	num	107 (0.9)
	—Other Denunciation Effect	bool	7 (6%)
	Termination Provision	bool	15 (13%)
	—Auto-Termination Memb. Threshold	num	4 (13.5)
	—Auto-Expiry	bool	6 (5%)
Treaty Organs	Legal Personality	bool	25 (21%)
	Organisational Auspices	cat	108 (91%)
	Secretariat	bool	111 (93%)
	Conference of Parties	bool	85 (71%)
Compliance & Effectiveness	National Reporting	bool	89 (75%)
	International Verification	bool	63 (53%)
	Progress Review	bool	73 (61%)
	Non-Compliance Procedures	bool	66 (55%)
	Sanctions	bool	58 (49%)
	Dispute Settlement	bool	74 (62%)
	—Compulsory Dispute Settlement	cat	41 (34%)

Figure 3.1: Treaty provisions summary results

with the 1940s. Which combinations of treaty design features are common, rare, or inexistent is also something that legal scholarship tends to be interested in and that can easily be quantified. A fine-grained analytical tool for this are multivariate contingency tables (also called frequency tables or truth tables). Illustrations are given in section 3.2 on institutional variables (with source code) as well as in the introduction to the third cluster (section 3.3). This is not to say that everything can or should be quantified, nor that too much weight should be attached to these numbers. Changing the composition of the treaty set analysed would likely change all the numbers in the present thesis. As mentioned in the introduction, the point is not so much to have the last word on treaty design, but to show ways to support and complement doctrinal legal analysis.

3.1   Treaty life cycle

The 'life cycle' of an international agreement could be said to start with a proposal and end with termination or obsolescence. As the focus of the present study is treaty texts, pre-adoption phases of negotiation and drafting are absent from the discussion, as are any subsequent developments not foreseen in the text. The question of interest

is how the treaty life cycle is envisaged and shaped by its designers, even though reality may unfold differently. After all, contracting parties are free to amend the agreement in accordance with its provisions and the Law of Treaties. For instance, the Ramsar Wetlands Convention[118] did not contain a provision on amendment in its original version, but was amended to include one.[119] One of the reasons to comprehensively regulate the treaty life cycle from the outset is that non-uniform entry into force of amendments to such provisions can create fragmented regimes and unnecessary complications, especially as membership grows. In the following, the main types of provisions governing treaty life cycle will be outlined along with the results of the analysis for the present treaty set.

3.1.1 Provisional application

Provisional application provisions govern the application of all or part of a treaty pending its entry into force for prospective parties.[120] A number of ubiquitous treaty provisions apply *ipso facto* from the moment of adoption, such as those on the depositary and entry into force.[121] This is not provisional application *stricto sensu* and thus not covered here, except inasmuch as provisional application provisions themselves fall into this category of treaty provisions applying from the time of treaty adoption.[122] It is worth noting that a more elaborate version of such provisions which apply necessarily before entry into force, in particular those setting up interim and preparatory organs and procedures, could be interesting to analyse from a treaty life cycle perspective. For the present study, however, the scope was restricted to provisions governing provisional application of either the entirety of the treaty, or of a part which continues to apply after entry into force.

Provisional application, its characteristics and difficulties have not ceased to intrigue scholars.[123] The International Law Commission (ILC) spent nearly a decade

---

[118] Ramsar Wetlands Convention (n 77).

[119] Protocol to amend the Convention of 2 February 1971 on Wetlands of International Importance Especially as Waterfowl Habitat (adopted 3 December 1982, entered into force 1 October 1986) 1437 UNTS 344, art 1, adding art 10bis to the Convention.

[120] VCLT (n 18), art 25(1)(a) ("A treaty or a part of a treaty is applied provisionally pending its entry into force if the treaty itself so provides").

[121] Ibid , art 24(4).

[122] Ibid, *in fine* ("...and other matters arising necessarily before the entry into force of the treaty apply from the time of the adoption of its text").

[123] See e.g. Daniel-Henri Vignes, 'Une notion ambiguë: l'application à titre provisoire des traités' (1972) 18(1) Annuaire Français de Droit International 181; Martin A Rogoff and Barbara E Gauditz, 'The Provisional Application of International Agreements' (1987) 39 Maine Law Review 29; even monographs like Albane Geslin, *La mise en application provisoire des traités* (Edition A Pedone 2005), and Anneliese Quast Mertsch, *Provisionally Applied Treaties: Their Binding Force and Legal Nature* (Brill 2012); and for a contextualisation and critique of recent

elaborating a Guide to Provisional Application of Treaties, including twelve draft
guidelines and a draft annex containing examples of provisional application provi-
sions in bilateral and multilateral treaties.[124] Whether as a prelude or justification
for a descriptive or prescriptive argument, the emergence,[125] variation, prevalence
and/or evolution of provisional application clauses are commonly described.[126] Such
descriptions, much like this assertion, tend to be unquantified but quantifiable and
verifiable with a suitable dataset and imposing certain constraints for practical rea-
sons (e.g. time period, languages and sources covered). Their association with other
types of provisions, for example on trade[127] or institutional arrangements, are also
of interest in order to describe and explain the observed occurrence patterns. Espe-
cially when treaty practice and its interpretation are varied and in flux, statements
as to the frequency or correlates of (sub)types of provisions can impact law-making,
application and litigation. Therefore, addressing the epistemic challenge involved
in observing and accurately describing legal developments becomes all the more
important.

In terms of modelling provisional application in PILO, including provisional ap-
plication provisions as a distinct subtype of treaty provision is probably uncontro-
versial. Even if the provision occurs in a paragraph of an article ostensibly about
entry into force,[128] it is still an instance of a provisional application provision. The
provision governs provisional application, which is a treaty process. 'Provisional
entry into force' could be defined as the process boundary at the start of the provi-
sional application process, and 'definitive' or regular entry into force would be the
instant which ends provisional application and starts regular application for con-
tracting parties. However, some scholars see a difference between 'provisional entry
into force' and 'provisional application' (beyond the instant vs. period distinction),

---

codification efforts see Gerhard Hafner, 'Provisional Application of Treaties' (2021) 24(1) Austrian Review of International and European Law Online 67.

[124] ILC, 'Draft guidelines and draft annex constituting the Guide to Provisional Application of Treaties, with commentaries thereto', *Yearbook of the International Law Commission*, 2021, vol II, Part Two, para 52, UN Doc A/76/10 ('ILC Guide to Provisional Application'), with a bibliography at pp. 87-94.

[125] See e.g. Hafner (n 123), at p.68 ("it is said that the treaties of Münster and Osnabrück of 1648 were the first instance of provisional treaty application").

[126] See e.g. ibid, at p.68-69 ("In recent decades, the number of treaties provisionally applied has increased, as the cases of the Agreement relating to the implementation of Part XI of the UNCLOS or the Energy Charter Treaty indicate", footnotes omitted).

[127] ILC Guide to Provisional Application, p.72 ("parts of treaties containing trade provisions are frequently subject to provisional application").

[128] E.g. Statutes of the International Centre for Genetic Engineering and Biotechnology (adopted 13 September 1983, entered into force 3 February 1994) 1763 UNTS 91, art 21 entitled 'Entry into force', para. 3 ("Until the entering into force of the Statutes in accordance with paragraph 1 hereinbefore, they shall be applied provisionally upon signature within the limits allowed by national legislation").

with the former being more like a subtype of entry into force and the latter constituting a class of its own.[129] Be that as it may, the ILC Guide to Provisional Application covers both and the latter term was favoured over the former,[130] thus the same approach was followed here. As the sample size is relatively small, no further subdivisions or feature extraction were undertaken.

True and false positives

In the treaty set at hand, 14 agreements contain provisional application provisions (12%), 9 of which govern commodities from the agriculture, forestry and fisheries sectors. Identifying relevant provisions with regular expressions is relatively straightforward, as all but three are in articles entitled 'Provisional application' or 'Notification of provisional application'. However, this search pattern also matches a false positive, namely Article 31 of the WTO Agreement on Subsidies and Countervailing Measures, which is about expiration of the application of certain provisions *after* entry into force for any party:[131]

> Article 31 Provisional Application
>
> The provisions of paragraph 1 of Article 6 and the provisions of Article 8 and Article 9 shall apply for a period of five years, beginning with the date of entry into force of the WTO Agreement. Not later than 180 days before the end of this period, the Committee shall review the operation of those provisions, with a view to determining whether to extend their application, either as presently drafted or in a modified form, for a further period.

The Marrakesh Agreement is a bit of a special case in that it had to manage a complex, partially overlapping transition between the old and new trade regime,[132] with the core component of the old trade regime having been applied provisionally for nearly 50 years.[133] After becoming an infamous example of stretching the definition of "provisional" application, the new version of the General Agreement on

---

[129] See ILC Secretariat Memorandum on the negotiating history of VCLT Article 25, UN Doc A/CN.4/658, pp 15-19 (on the shift in terminology).

[130] See ILC Guide to Provisional Application, Commentary to Guideline 1, para 2.

[131] WTO Agreement (n 93), Agreement on Subsidies and Countervailing Measures, art 31. The provisions lapsed but were still taken into account for interpretation of the agreement during arbitration, see Wolfgang Müller, *WTO Agreement on Subsidies and Countervailing Measures: A Commentary* (Cambridge University Press 2017), p.591.

[132] Patrick M Moore, 'The Decisions Bridging the GATT 1947 and the WTO Agreement' (1996) 90(2) American Journal of International Law 317.

[133] Protocol of Provisional Application of the General Agreement on Tariffs and Trade (adopted 30 October 1947, entered into force 1 January 1948), 55 UNTS 308.

Tariffs and Trade (GATT 1994) explicitly excludes the Protocol of Provisional Application and any provisions on provisional application in protocols of accession.[134] Hence, regular expressions need to be tweaked so as not to match these mentions of provisional application. This can be done either by finding a formulation that only true provisional application provisions share, or by adding a negative condition that excludes the WTO Agreement. For named articles, an exclusion criterion is the easiest way to handle it, because the article heading is otherwise sufficient for a match. Ideally, this exclusion criterion would be self-explanatory and also prevent future false negatives when the dataset is expanded. In other words, it should clearly and robustly exclude false matches. For instance, the word "years" only occurs in the wrongly matched Article 31, and thus could be used to eliminate it from the set of provisional application clauses in this treaty set. However, someone running the source code without reading the explanation may be puzzled by this exclusion criterion and not trust the results, or, if the treaty set is different, may miss true provisional application clauses that happen to contain the word. Hence, a better choice for the negative condition is "beginning with the date of entry into force". It goes to the essence of why Article 31 is not a provisional application provision in the sense of the law of treaties, and it is a phrase that is unlikely to occur in other treaties.[135]

With this modification, the regular expression for capturing named articles on provisional application with its corresponding code and results becomes:[136]

Extracting named articles on provisional application

```
1  eos = r'(?=\n\n(?:Article|Regulation|Chapter|Part|In Witness Whereof|Annex|Appendix|[A-Z]+ ?\d? ?[.-]
       ↪ [A-Z]{3,}))'
2  target = 'provisAppProv'
3  provdf.loc[target,'namedArt'] = (r'(?ism)^(?P<provisAppProvStr>(?:Article \w+\s+)?' +
4    r'[^\n]*Provisional Application\n' +
5    r'(?!\n*[^\n]+beginning with the date of entry into force).+?)')
6  new = textseries.str.extract(provdf.loc[target,'namedArt'] + eos).dropna()
7  new[target] = True
8  new['treatyLabel'] = textdf.treatyLabel[new.index]
9  print(len(new), 'treaties')
10 print(new.set_index('treatyLabel'))
```

---

[134]   WTO Agreement (n 93), Annex 1A, GATT 1994, para. 1(a) and 1(b)(ii).
[135]   The WTO Agreement is the only treaty in the present set of 119 agreements that contains this exact phrase (in any provision).
[136]   As with all code result blocks in this thesis, the results are not copy-pasted into this document but produced directly by running the code in the Emacs Orgmode source document, which is exported to PDF via LATEX. The file treatytextIE.org in the supplementary materials contains the full code and results for reproduction and reuse.

```
Results
```

```
11 treaties
                                          provisAppProvStr  provisAppProv
treatyLabel
OliveOilAg79          NOTIFICATION OF PROVISIONAL APPLICATION\n\nArticle 45\n\...          True
CocoaAg93             Article 55 Notification of provisional application\n\n1...           True
NuclearAccNotifConv   Article 13 Provisional application\n\nA State may, upon...          True
NuclearAccAssistConv  Article 15 Provisional application\n\nA State may, upon...          True
SugarAg92             Article 39 Notification of provisional application\n\n1...           True
GrainsTradeConv95     Article 26 Provisional application\n\nAny signatory Gov...           True
AntiPersMineBanConv   Article 18 Provisional application\n\nAny State may at ...           True
FishStocksAgmt        Article 41 Provisional application\n\n1. This Agreement...           True
TropicalTimberAg06    Article 38 NOTIFICATION OF PROVISIONAL APPLICATION\n\nA ...          True
FoodAssistConv        Article 14 Notification of Provisional Application\n\nAn...          True
IUUfishingPortStateAg Article 32 Provisional application\n\n1.\tThis Agreement...          True
```

The first line defines the end-of-section marker `eos`, a positive lookahead to the next section (article, chapter, testimonium etc.). In the full script, this variable is assigned at the beginning and then used for all relevant provision patterns, but it is included here for completeness of the explanation. The second line specifies the name of the target variable, a contraction of 'provisional application provision'. The main regular expression starts on line 3 and is separated into three parts for clarity (Python uses the + operator for string concatenation, not only for numerical summation). The regular expression is directly saved to the provisions dataframe (`provdf`) in the `provisAppProv` row under the `namedArt` column, so that the dataframe always contains the latest version.

The main regex starts with turning on case-insensitive (`?i`), 'dot-all' (`?s`),[137] and multi-line matching (`?m`), combined to (`?ism`) and valid for the full string. Case-insensitive matching means that the regex pattern and the processed text can have any combination of upper- and lowercase characters. Dot-all mode makes the dot special character match newlines as well, which is convenient when a provision may be composed of multiple paragraphs. Multi-line mode enables the use of the start-of-string anchor `^` and the end-of-string anchor `$` for each line in the document, not only at the very start and end of the full text. The text of the provision is extracted by means of a named capture group (`?P<provisAppProvStr>...`). Naming the group has the advantage of generating the column name automatically when the regex is passed to the `str.extract()` function on line 6 and the resulting matches are assigned to a new data table. The `.+?` at the end of the capture group, when combined with the `eos`, signals that the match should extend up until the first occurrence of the `eos` pattern (i.e. right before the next article). Without the 'lazy' (also called 'reluctant') repetition operator `?` the match would continue until the last occurrence

---

[137] The 's' stands for 'single-line' matching but that can be a confusing name because it may seem mutually exclusive with multi-line matching, so it is often called dotall mode.

of the `eos` pattern at the end of the treaty text because by default `.+` calls for 'greedy' matching, capturing everything when combined with the dot-all instruction. The first part of the regex (on line 3) looks for an optional 'Article' at the beginning of a new line, followed by a space and one or more alphanumeric characters `\w+` (not only digits `\d+` because Roman numerals are still used for numbering articles in treaties).[138] This is enclosed in a non-capturing group which can occur zero or one time `(?:)?`. Setting it to be optional means that the first result, which puts the article heading before the article number,[139] can also be matched by this regex pattern. In fact the results would be the same if this non-capturing group were removed altogether, because `[^\n]*` matches any non-newline characters occurring before the article heading (on line 4). It is worth including it regardless since some treaty segmentation formats put a line break between the article number and its heading, which would be matched by `\s+` (one or more whitespace characters) in `(?:Article \w+\s+)?`. Indeed, these are the three types of article title arrangements in the treaty set at hand (for those that have named articles): (1) before the 'Article' + number, (2) on the same line (with a range of separators from spaces to dashes or colons), or (3) on a subsequent line (with one or more line breaks in-between). Hence, including this optional non-capturing group makes the regex pattern more robust to different formatting styles. Line 4 sets out the key non-optional part of the pattern, namely `Provisional Application\n`. The newline at the end is necessary[140] because it makes the pattern match article headings rather than an expression in the middle of a sentence. "Notification of ", occurring in 5 of the 11 results produced by this pattern, is covered by `[^\n]*`. Finally, line 5 specifies the above-mentioned negative lookahead, checking whether the article heading is followed (after zero or more additional newlines) by a paragraph containing the phrase "beginning with the date of entry into force". It only tests its occurrence in the first textual paragraph following the article heading because there is no need to look further in this case. If we had used `.+` instead of `[^\n]+`, the regex engine would have searched for the expression until the very end of the treaty text, which would be computationally inefficient and inappropriate, because its occurrence in another provision would not be relevant for the present inclusion decision. This begs the question of how to test the presence or absence of an expression within a given treaty provision which spans multiple paragraphs. We will see an example of this in the section on entry into force provisions (3.1.2).

---

[138]   31 out of 119 treaties (26%) in the given treaty set.

[139]   International Olive Oil Agreement, 1979 (adopted 30 March 1979, entered into force 1 January 1981) 1219 UNTS 135, art 45.

[140]   The line-end anchor `$` would serve the same purpose.

To conclude the explanation of this code block, after extracting the relevant strings from the treaty texts, dropping empty rows with `.dropna()` and assigning the result to a new data table (line 6), the table receives an additional column called 'provisAppProv', a binary variable with its value set to `True` for the treaties returning regex matches (line 7). This additional variable may seem redundant since it could also be derived from the extracted strings (the `provisAppProvStr` column) at the moment of analysis. This is technically true here, but in the case of inherited provisions there is not always a corresponding string to extract, and thus it proved more convenient to construct the relevant variables in this incremental fashion with immediate verification. Line 8 retrieves the treaty label for presentation purposes from the `textdf` based on the common index, line 9 outputs the number of treaties with positive matches from this regex pattern, and line 10 produces an overview table of the results with treaty labels as index.

The corresponding code block in the `treatytextIE.org` file also updates `textdf` with the new results in a single line (`textdf.update(new, overwrite=True)`) and has a second code block immediately following this one, printing out[141] the full text of the extracted provisions for verification purposes, amounting to approximately 1000 words in this case. This kind of length is perfectly suitable for overinclusion verification by reading through them, possibly with the full treaty text at hand in case anything is unclear or needs to be interpreted in the context of other provisions.[142] The regex pattern could of course be modified to extract a given context string around the target excerpt, as was done for the shorter expressions in Chapter 4, but for provisions it was usually faster to open the full text in a separate window and scroll manually, because the relevant context could be anywhere. Once the main inclusion decision is made, the coding focuses on the precise extraction of the right excerpt and the readability and computational performance of the regex pattern. In an interactive regex editing session it would be wise to automatically scan for differences in output before updating the dataframe,[143] because sometimes even a purportedly cosmetic fix can negatively impact the results.

The second `provisAppProv` regex pattern captures the provisional application provisions of another three treaties in a similar fashion, but in this case the provisions are extracted only from those texts that have not yielded a match with the first pat-

---

[141] As in using Python's `print()` function to write to a text file or standard output stream, not printing to paper.

[142] This is one of the reasons for using the treaty text filename as index for `textdf` and `textseries`, to be able to access it quickly.

[143] For instance by running `print(len([s for s in new.provisAppProvStr if s not in textdf.provisAppProvStr.values]), 'different from current version')` and if there is a non-zero difference, outputting diverging strings.

tern. This residual matching strategy is implemented by searching over a subset of
texts with `textseries[textdf[target].isna()]`[144] on line 5 of the following code block:

Other provisAppProv excerpts

```
1  provdf.loc[target,"otherExtract"] = (r'(?im)^(?P<provisAppProvStr>(?:Article.+\n+)?.*' +
2    r'(?:enter into force provisionally' +
3    r'|Pending entry into force of this' +
4    r'|shall be applied provisionally upon signature).+)')
5  new = textseries[textdf[target].isna()].str.extract(provdf.loc[target,"otherExtract"]).dropna()
6  new[target] = True
7  new['treatyLabel'] = textdf.treatyLabel[new.index]
8  print(len(new), 'treaties')
9  print(new.set_index('treatyLabel'))
```

Results

```
3 treaties
                                             provisAppProvStr  provisAppProv
treatyLabel
ICGEBstatutes  3.\tUntil the entering into force of the Statutes in accordanc...        True
CoffeeAg83     (2)\tThis Agreement may enter into force provisionally on 1 Oc...        True
UNCCD          Article 7  Timetable for preparation of action programmes\n\nP...        True
```

Without this restriction, four of the named provisions from the first extraction
pattern would be overwritten with subsequent paragraphs (of entry into force provi-
sions) in the same treaties. In the Java-based GATE pipeline, multiple annotation
and extraction is the desired result, but for the Python pipeline it is preferable to
have one key excerpt per provision, ideally the one which provides the best evidence
for the constructed variables. The regular expression is saved to a different column
in `provdf` named 'otherExtract' which is used for all residual miscellaneous patterns.
In other regards this code block is like the previous one, except that it uses three
key phrases in the body of the article instead of headings to find relevant provisions
(lines 2-4). As with the negative lookahead above, the phrases are selected to be as
self-explanatory, transparent and generalisable as possible, but without guarantee
that they will work beyond the given treaty set. Another difference is that only the
paragraph and, optionally, an immediately preceding article heading are captured,
because some provisions have dozens of paragraphs and extracting the full article in
each case would make verification more burdensome and the evidence less targeted.
Hence no end-of-section lookahead (`eos`) is added to the main regex before extrac-
tion on line 5, dot-all mode and lazy quantifiers are not required. In fact, applying
a reluctant repetition operator to the `.+` at the end of the `provisAppProvStr` capture
group (line 4 *in fine*) in the absence of a lookahead would make it (`.+?`) match only
a single character, since that is the minimum needed to satisfy the condition.

---

[144] The `.isna()` function returns `True` for those rows which have a 'not available' NA value in the
       target column.

Together these two code blocks capture the 14 provisional application provisions found in the treaty set. The crucial component are the two regular expressions, which can be used with other, similarly capable regex matching engines, such as in the R, Java or Perl programming languages, possibly with minor tweaks. To illustrate such adaptations, the following code block contains the relevant excerpt of the GATE pipeline:[145]

GATE Jape provisAppProv regexps

```
1  // end of section macro
2  eos=\n\n(?:Article|Regulation|Chapter|Part|In Witness Whereof|Annex|Appendix|[A-Z]{2,} ?\d? ?[.-] [A-
       ↪ Z]{3,})
3  // named arts
4  |(?ism)^Article (\w+)\s+\b([^\n]*Provisional Application)\n(?!\n*[^\n]+beginning with the date of
       ↪ entry into force).+?(?=<<eos>>)
5  |(?ism)^Article (\w+)\s+\b(Timetable[^\n;.]*)\n+Pending entry into force.+?(?=<<eos>>)
6  0 => ProvisionalApplication class="provisAppProv",governs="provisionalApplication",hasArtNb=$1,
       ↪ hasArtTitle=$2
7  // named arts (title before art)
8  |(?ism)^((?:Notif\w+ of )?Provisional Application)\n+Article (\w+)\n+.+?(?=<<eos>>)
9  0 => ProvisionalApplication class="provisAppProv",governs="provisionalApplication",hasArtNb=$2,
       ↪ hasArtTitle=$1
10 // art paras
11 |(?i)\n(.*(?:enter into force provisionally|shall be applied provisionally upon signature).+)
12 1 => ProvisionalApplication class="provisAppProv",governs="provisionalApplication"
```

It is more compact because these are only the regex patterns and annotation instructions which are passed to the Java Regexp Annotator plugin[146] to annotate the documents in the corpus as part of a pipeline. The plugin makes recourse to the Java regex engine, but has some limitations, such as only extracting full matches or numbered capture groups, not named groups as used in the Python examples above. Still, the syntax is simple, readable and beginner-friendly: Lines starting with double slashes are comments (to structure and explain the code) and are ignored, just like empty lines. Macro lines are of the form 'name=pattern' (see line 2) and any subsequently occurring `<<name>>` is substituted with the pattern (e.g. lines 4, 5 and 8). Pattern lines start with a vertical bar followed by a Java regex, and each line in a multi-line pattern (e.g. lines 4-5) is interpreted as an alternative. Finally, action lines specify the name of the annotation, in this case 'ProvisionalApplication', and what part of the match it is to cover, if any. For instance, `0~=>` indicates that the annotation should extend over the full match, whereas `1~=>` means it will correspond to the string captured by the first capturing group of the pattern definition (simple parentheses are capturing groups). The annotation name is followed by an optional

---

[145] See Appendix F.2.1 for more.
[146] See the documentation at https://gatenlp.github.io/gateplugin-StringAnnotation/ JavaRegexpAnnotator (accessed 19 Jan 2022), archived at https://web.archive.org/web/ 20220119234759/gatenlp.github.io/gateplugin-StringAnnotation/JavaRegexpAnnotator.

comma-separated feature assignment list, with "class" referring to the PILO class the annotation instance falls under (a special feature when used with the ontology-aware transducer). The dollar sign with number `$n` is substituted with the string of the respective capture group. For example, in the first set of named article patterns (lines 4-5), the first plain parentheses of the pattern definition enclose the article number and the second surround the article heading. Accordingly, the corresponding action line (6) specifies that the annotation should have a feature called `hasArtNb` with the value extracted by the first capture group, and another feature called `hasArtTitle` with the string extracted by the second capture group. Since the order of the groups in the regex pattern matter, the pattern and action rules for the named article with the heading occurring before the article number need to be defined separately (lines 8-9). Another difference with regard to the Python regexps is that the pattern for Article 7 UNCCD is added to the first set of named article patterns, because this allows for simultaneous extraction of article number and title. These features are used to generate data properties at the ontology population stage. In the Python pipeline it did not seem worth extracting these features as it would have added two additional columns per provision without providing more information than could be gleaned from the article string itself. The rationale for capturing them in the GATE pipeline is that it can be useful to have these data items in a database for selective display, and that properly referencing provisions in nested paragraphs will benefit from having as much structured data as possible for each instance. Indeed, constructing the proper reference for individual (sub)paragraphs automatically is less straightforward due to the range of numbering and structuring systems used. This work is more focused on content and thus only captures a minimal amount of structural/numbering information where readily available.

While readability may be hampered somewhat by the limited page width and long lines of code,[147] these examples are hopefully beginning to show the versatility, corrigibility and transparency of regular expressions for reproducible text annotation and extraction to support legal research beyond the present work. For instance, some scholars consider 'provisional entry into force' a distinct category not to be confused with 'provisional application', while others have a preference for the latter term but see the former as an instance of the same class. For present purposes the two formulations are captured under the same umbrella, as evidenced e.g. by line 11 of the preceding code block, but distinguishing between the two and counting them separately to observe trends would constitute a relatively minor adaptation (depending on the size of the treaty set), and the resulting source code could be

---

[147]   Rounded hook right arrows indicate line continuation.

shared and reused in turn. A researcher manually copy-pasting provisions into a separate Word file who expands their classification scheme mid-way would likely need to re-read already processed texts to see if there are any additional provisions that match the new scheme, which could take a significant amount of time. A machine learning classifier would be even less transparent and corrigible. For example, addressing the initial false positive of the WTO Agreement may require training the classifier on a large, artificially constructed dataset specifically geared towards preventing this kind of classification error. As we have seen, statistically, the occurrence of 'provisional application' in an article heading is a good indicator for provisional application provisions. But in law the devil is often in the details, and regular expressions excel at handling details.

True and false negatives

Assessing and tackling underinclusion (false negatives) is in some ways a rather different task from the one discussed in the previous section. It only matters if the aim is to exhaustively capture all relevant matches for a given dataset. If, instead, the goal is merely to output a number of correct examples to analyse and cite in a publication, then the elimination of false positives is all that counts. For the present chapter, all efforts were made to avoid underinclusion. Searching and capturing false negatives can be a time-consuming task, especially in a diverse set of treaties. While ultimately only careful reading through the entire agreement texts can provide reassurance that no relevant provision was missed, regex search can still help eliminate false negatives and estimate the magnitude of the verification task. Publishing the source code of the search (where available) may also aid other researchers verify, adapt or expand the code and dataset.

To this end, the following code block illustrates a regex-based false negatives search for provisional application provisions:

False negative search

```
1  provdf.loc[target,'searchPattern'] = (
2    r'(?i)(?:provisional\w* (?:appl\w+|\w+ into force)' +
3    r'|(?:appl\w+|\w+ into force) (?:\w+ \w+ )?provisionally' +
4    r'|pending entry into force)')
5  unmatched = textdf[textdf[target].isna() & textseries.str.contains(provdf.searchPattern[target])]
6  print(len(unmatched), 'unmatched treaties:')
7  print(unmatched.shortTitle)
8  print(textseries[unmatched.index].str.count(provdf.searchPattern[target]).sum(), 'matches to verify')
```

Results

```
2 unmatched treaties:
```

```
fn
treaties-ATS-1994-31.txt              UNCLOS
treaties-ATS-1995-8.txt      WTO Agreement
Name: shortTitle, dtype: object
7 matches to verify
```

As before, lines 1-4 define the regex search pattern and save it to a column of the
`provdf` dataframe for reuse and incremental improvement. The search is undertaken
on line 5, where the treaties which do not already have a `provisAppProv` value and
yet return `True` when searching for the regex with the `str.contains()` function, are
stored in a new dataframe called `unmatched`. Lines 6-8 produce the output seen in the
results block.[148] The total count of matches to verify can also be a useful indicator
to estimate the effort needed for a particular regex fine-tuning task. For example,
searching for "(?i)provisional" in the same treaty set would yield 122 matches, a
much larger verification task. When limiting it to instances not followed by the
word 'measure' (with a negative lookahead `'(?i)provisional(?!  measure)'`), only 69
matches remain. In some cases, perusing an overview of the matches with a suitable
context window may already be sufficient if one is familiar with the treaties in
question. Naturally, there are convenience functions for this purpose:

Overview of matches to verify

```
1  window = '30'
2  patternplus = r'(.{,'+ window +'}'+ provdf.searchPattern[target] +'.{,'+ window +'})'
3  print(textseries[textdf[target].isna()].str.extractall(patternplus).dropna())
```

Results

```
                                                                            0
                          match
treaties-ATS-1994-31.txt 0      (ii)\tadopt and apply provisionally, pending approval by the Asse
                         1       Preparatory Commission shall apply provisionally pending thei...
treaties-ATS-1995-8.txt  0      nt (excluding the Protocol of Provisional Application), as rec...
                         1      the provisions (a) concerning provisional application and with...
                         2      Part II of GATT 1947 shall be applied provisionally to the ful...
                         3                                   Article 31  Provisional application
```

`window` on line 1 denotes the maximum number of characters the regex engine
should extract before and after the search pattern. Since the dot is used here in
normal mode (not matching newlines), the last row of the results is much shorter
than the others, delimited by newlines on either side. The matches from the WTO
Agreement were explained earlier, and as for UNCLOS, once can already guess from
this excerpt that provisional application in this case takes place pending approval by

---

[148]   In a typical interactive Python session, the code on lines 1-7 would be expressed in a single
line, without an intermediate storage object like `unmatched`, but long lines of code do not look
good on A4 paper.

the Assembly, and concerns the rules, regulations and procedures of the International Seabed Authority. The file `treatytextIE.org` contains the full paragraphs of these matches, as well as code and results for a wider search pattern.

Ideally, all keywords and expressions searched should be included to bolster the case for the present selection being adequate and for facilitating research on a other treaties, but this being a wide-ranging exploratory research project, it was not practical to do so.

Finally, when combining traditional reading with automated checks, it can be helpful to mark a given treaty as 'done' with respect to a particular provision, so as to exclude it from further false negatives search. This is another reason for having a separate binary column for each text variable. For instance, removing UNCLOS from further false negatives search can be done by setting its `provisAppProv` value to `False` with:

```
textdf.loc[textdf.treatyLabel=="UNCLOS", "provisAppProv"] = False
```

As a result, the `textdf[target].isna()` filter used in the preceding code blocks would skip UNCLOS. Memorable treaty labels and variable names expedite these kinds of incremental checks and edits.

Description of identified provisions

Even in such a small sample, provisions are more varied than one might expect. A first distinguishing element is the time span during which provisional application is encouraged or allowed. The longest possible time span would be between treaty adoption and entry into force for a given party. The second longest and more common one is from the moment of simple signature to the entry into effect of the agreement for the signatory.[149] At the other end of the spectrum are treaties which allow provisional application only *after* general entry into force, pending individual entry into force. Agreements governing trade in agricultural commodities tend to follow this pattern.[150] A second differentiating factor is concerned with consent to

---

[149] See e.g. Convention on Early Notification of a Nuclear Accident (adopted 26 September 1986, entered into force 27 October 1986) 1439 UNTS 275, art 13 ("A State may, upon signature or at any later date before this Convention enters into force for it, declare that it will apply this Convention provisionally"); and identically, the Convention on Assistance in the Case of a Nuclear Accident or Radiological Emergency (adopted 26 September 1986, entered into force 26 February 1987) 1457 UNTS 133, art 15.

[150] See e.g. International Olive Oil Agreement (n 139) art 45; International Sugar Agreement, 1992 (adopted 20 March 1992, entered into force 10 December 1996) 1703 UNTS 203, art 39; International Cocoa Agreement, 1993 (adopted 16 July 1993, entered into force 22 February 1994) 1766 UNTS 3, art 55; and International Tropical Timber Agreement, 2006 (adopted 27 January 2006, entered into force 7 December 2011) 2797 UNTS 75, art 38; but not Grains

provisionally apply a treaty—a matter of particular importance for open multilateral agreements adopted under the auspices of a global organisation, given that some negotiating states may never join the agreement and conversely, some states which did not participate in the conclusion of the treaty may be interested in becoming a party at a later stage. While notification of consent to apply the convention provisionally is not always required,[151] it is nevertheless the design option used by almost all treaties in the study set which provide for provisional application. Agreements with an opt-in rule for provisional application can be further subdivided based on the level of flexibility afforded to willing states, the most flexible allowing any unilateral declaration, expressed at any time, and effective at the time desired by the declaring state (before definitive entry into force for the party concerned). A third high-level delineation can be drawn between treaties that are to be provisionally applied in their entirety and treaties that limit the scope of provisional application to a part, which could be as small as a single paragraph.[152]

One of the political reasons for including such provisions could be to seize the momentum gained by treaty adoption, encouraging willing states to take steps towards implementation right away, before waiting for dozens of other states to complete potentially lengthy ratification procedures. The Convention to Combat Desertification and the Ottawa Anti-Personnel Mine Ban Convention, with their minimum number of instruments of consent of 50 and 40 respectively,[153] could be a case in point. However, the Nuclear Accident Notification and Assistance conventions have a very low entry into force threshold (only three ratifications and a waiting period of one month),[154] yet still contain a provisional application clause.[155] Perhaps this inclusion is due to an increased sense of urgency at the time, the agreements having been adopted in the wake of the Chernobyl disaster. The importance of reciprocity

---

Trade Convention, 1995 (adopted 7 December 1994, entered into force 1 July 1995) 1882 UNTS 195, art 26.

[151] See Convention to Combat Desertification (n 81), Annex I art 7 ("Pending entry into force of this Convention, the African country Parties, in cooperation with other members of the international community, as appropriate, shall, to the extent possible, provisionally apply those provisions of the Convention relating to the preparation of national, subregional and regional action programmes.").

[152] See e.g. Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction (adopted 18 September 1997, entered into force 1 March 1999) 2056 UNTS 211 art 18 ("Any State may at the time of its ratification, acceptance, approval or accession, declare that it will apply provisionally paragraph 1 of Article 1 of this Convention pending its entry into force").

[153] Convention to Combat Desertification (n 81), art 36(1); Ottawa Anti-Personnel Mine Ban Convention (n 152), art 17(1).

[154] Nuclear Accident Notification Convention (n 149), art 12(3); Nuclear Accident Assistance Convention (n 149), art 14(3).

[155] Nuclear Accident Notification Convention (n 149) art 13; Nuclear Accident Assistance Convention (n 149) art 15.

and scale for the functioning of the regime may also be an influential design consideration, such as for the above-mentioned commodity agreements which employ various techniques to achieve simultaneous entry into force for all contracting parties. Provisional application after general entry into force can thus be a means to synchronise entry into force of the agreement.

As for associations with other design features, there are many ways in which these could be analysed. One way that may be particularly intuitive to a lawyer (without any university-level training in statistics) involves comparing conditional with independent probabilities. In other words, comparing the frequency of design features occurring in the subset of treaties covered in this section with those of the full treaty set. Independent probability is the probability of a given type of provision occurring in the full set, e.g. $\frac{14}{119} = 12\%$ for provisional application provisions.[156] This independent prevalence is likewise calculated for all other variables. Conditional probability of some design feature *given* the presence of a provisional application provision represents the prevalence of the design feature in this smaller subset of treaties. For instance, provisions about legal personality of a treaty organ (or organisation) occur in 21% (25) of the agreements constituting the full set, whereas in the subset under consideration they are present in 50% (7) of the agreements.[157] The difference between conditional and independent probability can be assessed as a ratio (dividing the former by the latter) or as a percentage point difference (subtracting the latter from the former). Both have their merits and demerits. In the example given, legal personality provisions are 2.4 times more likely to occur in treaties with provisional application provisions (and vice versa), or have a 29 percentage points higher prevalence. The ratio is symmetric[158] unlike the respective conditional probabilities and percentage point differences. That is, provisional application provisions are also 2.4 times more likely to occur in treaties with legal personality provisions than in the full set, but the percentage point increase is only 16, from 12% to 28%. Figure 3.2 shows the absolute and relative prevalence, ratio ('Div' column) and percentage point difference ('Diff' column)[159] for the other variables covered in this chapter with regard to the `provisAppProv` subset.

---

[156] By convention, probabilities are expressed as a number between 0 and 1, i.e. 0.12 in this case, but expressing this number as a percentage is equivalent and, I believe, easier to grasp and remember for less quantitatively minded readers.

[157] $P(legalPers|provisAppProv) = \frac{7}{14} = 0.5$.

[158] This follows from a transformation of Bayes' theorem, $\frac{P(A|B)}{P(A)} = \frac{P(B|A)}{P(B)}$, but a detailed explanation is beyond the scope of this work.

[159] A cool-warm gradient background colour scheme is applied to the 'Diff' column to highlight positive and negative extremes.

Cluster	Variable	Type	Full Set (n=119)	Subset (n=14)	Div	Diff
Start of Life	Entry into Force Provision	bool	119 (100%)	14 (100%)	1.0	0.0
	—Min. Nb. of Ratifs for EIF	num	112 (19.8)	11 (21.6)	1.1	1.8
	—Additional EIF Conditions	bool	47 (39%)	8 (57%)	1.4	17.6
	—EIF Waiting Period (Years)	num	110 (0.4)	7 (0.1)	0.4	-0.3
Change	Amendment Provision	bool	115 (97%)	14 (100%)	1.0	3.4
	—Simplified Amendment	bool	60 (50%)	6 (43%)	0.8	-7.6
	Adoption of Annexes or Protocols	bool	26 (22%)	1 (7%)	0.3	-14.7
End of Life	Denunciation Provision	bool	113 (95%)	14 (100%)	1.1	5.0
	—Denunciation Waiting Period (Years)	num	97 (2.7)	11 (0.8)	0.3	-1.8
	—Other Denunciation Conditions	bool	33 (28%)	3 (21%)	0.8	-6.3
	—Denunciation Effect Delay (Years)	num	107 (0.9)	13 (0.6)	0.7	-0.3
	—Other Denunciation Effect	bool	7 (6%)	1 (7%)	1.2	1.3
	Termination Provision	bool	15 (13%)	8 (57%)	4.5	44.5
	—Auto-Termination Memb. Threshold	num	4 (13.5)	0 (nan)	nan	nan
	—Auto-Expiry	bool	6 (5%)	6 (43%)	8.5	37.8
Treaty Organs	Legal Personality	bool	25 (21%)	7 (50%)	2.4	29.0
	Organisational Auspices	cat	108 (91%)	14 (100%)	1.1	9.2
	Secretariat	bool	111 (93%)	14 (100%)	1.1	6.7
	Conference of Parties	bool	85 (71%)	12 (86%)	1.2	14.3
Compliance & Effectiveness	National Reporting	bool	89 (75%)	11 (79%)	1.1	3.8
	International Verification	bool	63 (53%)	9 (64%)	1.2	11.3
	Progress Review	bool	73 (61%)	11 (79%)	1.3	17.2
	Non-Compliance Procedures	bool	66 (55%)	9 (64%)	1.2	8.8
	Sanctions	bool	58 (49%)	7 (50%)	1.0	1.3
	Dispute Settlement	bool	74 (62%)	8 (57%)	0.9	-5.0
	—Compulsory Dispute Settlement	cat	41 (34%)	5 (36%)	1.0	1.3

Figure 3.2: Design features of treaties with provisional application provisions

The numbers in parentheses of the 'Full Set' and 'Subset' columns represent the set prevalence as described for boolean variables like legal personality provisions. For numeric variables such as the minimum number of ratifications required for entry into force (`EIFnConsent`), the numbers in parentheses express the set average. In this case, the two are nearly identical. The average `EIFnConsent` for the 11 agreements in the subset which stipulate this condition is 22 such instruments, whereas for the full set it is 20 instruments. Accordingly, 'Div' shows the ratio between the subset and the full set average[160] and 'Diff' the result of the subtraction of the full set average from the subset average. Numeric variables are marked with 'num' in the 'Type' column and the measuring unit is expressed in the variable name. Categorical variables ('cat') display a binarised version of the set prevalence, i.e. whether some category of the relevant variables occurs in the (sub)set in question, regardless which category it is.

The table may be somewhat dense, but the granular representation aids in the prevention of misinterpretation. For instance, auto-expiration provisions, a subtype of termination provisions, are 8.5 times more likely to occur in treaties with pro-

---

[160] $\frac{21.6}{19.8} = 1.1$.

visional application provisions when compared with the full set. However, in total there are only six agreements with such provisions, not exactly a large sample, and their auto-expiry and prolongation pattern could be more of a quirk of these (commodity) agreements than anything particularly interesting. It does stand to reason that agreements which are designed to synchronise entry into force would also have termination provisions to assist with an orderly ending of the regime and its organs. Thus, the higher prevalence of termination and legal personality provisions certainly make sense. Conversely, the nearly 15 percentage point decrease of the prevalence of provisions on adoption of annexes or protocols (a third as likely as in the full set) does not have an obvious explanation. In any case, this kind of descriptive analytics is not intended for causal inference, but for exploratory analysis and hypothesis generation.

3.1.2 Entry into force

Entry into force (EIF) provisions are typically among the final clauses of a treaty and are one of the main indicators for distinguishing treaties from non-binding agreements in borderline cases, because these provisions specify the point in time when a treaty becomes binding for its contracting parties. Open multilateral treaties usually contain two types of EIF provisions: initial or general entry into force of the treaty per se (genEIF), and subsequent or individual entry into force (indivEIF) for states expressing their consent to be bound after the conditions for initial entry into force are met. EIF provisions as discussed here are strictly about entry into force for validly expressed consent by eligible parties, whereas the conditions for participation are typically stipulated in one or more separate provisions on signature, ratification and accession.

One of the most common design patterns of genEIF provisions in large-scale multilateral treaties sets out the minimum number of expressions of consent to be bound required (`EIFnConsent`) and a time period to elapse once this condition is fulfilled (`EIFdelay`). In some cases there is no delay and the treaty enters into force on the day the requisite number of ratifications is reached,[161] and in other cases

---

[161] E.g. International Convention for the Regulation of Whaling (adopted 2 December 1946, entered into force 10 November 1948) 161 UNTS 72, art X(4); Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies (adopted 27 January 1967, entered into force 10 October 1967) 610 UNTS 205, art XIV(3); Treaty on the Prohibition of the Emplacement of Nuclear Weapons and Other Weapons of Mass Destruction on the Sea-Bed and the Ocean Floor and in the Subsoil Thereof (adopted 11 February 1971, entered into force 18 May 1972) 955 UNTS 115, art X(3); Agreement to Promote Compliance with International Conservation and Management Measures by Fishing Vessels on the High Seas (adopted 24 November 1993, entered into force

a target date of entry into force is included in EIF provisions rather than, or in addition to, a time span.[162] There are many types of conditions beyond the simple number of ratifications, reflecting political, technological or economic considerations, i.e. material inequality despite formal equality of states. For the purpose of this analysis, the occurrence of such additional conditions is coded as a binary variable `EIFaddlCond` which is `True` when such conditions exist and `False` otherwise. IndivEIF provisions are simpler in that they only specify an EIF delay, if anything. This delay is often the same as that of genEIF, and if not, it is typically shorter rather than longer. For the sake of simplicity, only genEIF provisions were processed for the final analysis of the treaty set, and therefore the remainder of this section focuses on genEIF.

The considerations taken into account when negotiating the content of EIF provisions may well differ significantly from one regime to the next. Problem characteristics, expected cost of problem-solving, perceived urgency, need for global cooperation, risk of regulatory evasion, and organisational auspices may all influence the decision. EIFdelay may additionally depend on the time required for national implementation, and whether parties commit to bringing their domestic laws and regulations in compliance with the treaty *before* it becomes effective. Some agreements erect minimal barriers for genEIF, e.g. ILO conventions typically only require two ratifications to enter into force.[163] In other cases EIF conditions are such that all the biggest polluters or producers need to be on board before a treaty can take effect.

Before discussing the results of the analysis, it may be worth giving an example of how the relevant variables are constructed. Merely extracting the text of the provision and adding a binary variable like for provisional application provisions would not be sufficiently informative in this case, because all treaties have EIF provisions. Instead, `EIFdelay` and `EIFnConsent` are captured and transformed into comparable numeric values, and `EIFaddlCond` is set to `True` when the former two are not sufficient for calculating the date of entry into force. About a dozen regular expression patterns are used for this task, each with many alternatives, some of which matching over 40 treaties and others barely a handful. The following code

---

[footnotes]

24 April 2003) 2221 UNTS 91, art XI(1); and International Plant Protection Convention, 1997 (adopted 17 November 1997, entered into force 2 October 2005) 2367 UNTS 223, art XXII.

[162] E.g. Food Assistance Convention (adopted 25 April 2012, entered into force 1 January 2013) ATS 2014 12, art 15(1) ("This Convention shall enter into force on 1 January 2013 if by 30 November 2012 five Signatories have deposited instruments of ratification, acceptance, or approval.").

[163] But the low EIFnConsent threshold is compensated by a long EIFdelay of twelve months, see e.g. Safety and Health in Agriculture Convention (n 99), art 23(2), and likewise all the other ILO conventions in the treaty set.

block displays an information extraction and transformation process covering 23 treaties, with results for only three of them due to space constraints:

EIF provision IE sample code

```
1  pattern = r"(?i)\n\n(?P<EIFprovStr>Article[^.]*.?\s?Th\w+(?: present)? (?!amendment)(?:\w+ )?\w+
      ↪ shall \w+ into force (?:on the )?(?P<EIFdelay>[^.]+? (?:day|month|year)[s]?) (?:after|
      ↪ following) (?:the \w+ \w+ which (?:the )?(?:Governments of )?)?(?:not less than|at least|the)
      ↪ ? ?(?P<EIFnConsent>[^.]{0,20}?) (?:States|Governments|Parties)[^.]*? have (?:become parties|
      ↪ deposited \w* ?instrument|[^.]*?(?:ratif|consent to be bound))[^.]+[.])"
2  new = textseries.str.extract(pattern).dropna()
3  new.EIFdelay = new.EIFdelay.transform(lambda x: word2yrs(x) if x is not np.nan else x)
4  new.EIFnConsent = new.EIFnConsent.transform(lambda x: number[x] if x is not np.nan and re.search('[A-
      ↪ z]', x) is not None else x).astype(int)
5  new.loc[new.EIFprovStr.str.contains('(?:constitut\w+ not less than|account\w+ in total for|provided
      ↪ that|nuclear capacity|gross ton\w+)'), 'EIFaddlCond'] = True
6  new['treatyLabel'] = textdf.treatyLabel[new.index]
7  print(new.EIFprovStr.count(), 'treaties')
8  print(new[new.treatyLabel.str.contains('(?:Wetlands|OPRC|Paris)')].set_index('treatyLabel'))
9  print('\n', new.EIFprovStr[new.treatyLabel.str.contains('(?:Wetlands|OPRC|Paris)')].values)
```

Results

```
23 treaties

                             EIFprovStr  EIFdelay  EIFnConsent EIFaddlCond
treatyLabel
ParisAgreement  Article 21\n\n1. This Agreement shall enter in...      0.08           55        True
WetlandsConv    Article 10\n\n1.\tThis Convention shall enter ...      0.33            7         NaN
OPRC            Article 16  Entry into force\n\n(1)\tThis Conv...       1.00           15         NaN


 ['Article 21\n\n1. This Agreement shall enter into force on the thirtieth day after the date on
      which at least 55 Parties to the Convention accounting in total for at least an estimated 55
      per cent of the total global greenhouse gas emissions have deposited their instruments of
      ratification, acceptance, approval or accession.'
 'Article 10\n\n1.\tThis Convention shall enter into force four months after seven States have become
      Parties to this Convention in accordance with paragraph 2 of Article 9.'
 'Article 16  Entry into force\n\n(1)\tThis Convention shall enter into force twelve months after the
      date on which not less than fifteen States have either signed it without reservation as to
      ratification, acceptance or approval or have deposited the requisite instruments of
      ratification, acceptance, approval or accession in accordance with article 15.']
```

The regex pattern is defined in the first line of code, with nested named capturing groups extracting the relevant strings to the respective variables (`EIFprovStr`, `EIFdelay`, and `EIFnConsent`). The negative lookahead '(?!amendment)' is required to ensure amendment EIF provisions are discarded. Matching strings are extracted on line 2 and stored in a new data table for further processing. `EIFdelay` is extracted with the respective unit (days, months or years), and then converted to years with the custom function `word2yrs()` on line 3. `EIFnConsent` is mostly specified in number words, and undergoes custom transformation on line 4 to a numeric variable. `EIFaddlCond` is appended to the data table on line 5, based on regular expression matching performed on the extracted EIF provisions. Of the three results shown,

only the Paris Agreement specifies an additional EIF condition.[164] The remaining lines of code pick the three examples from the 23 rows, format and print them out for the purpose of this illustration. In the regex crafting process, the full set of extracted provisions and transformed variables is printed out in this way for immediate verification. No additional false negatives search is necessary given that in this case the provision is extracted for all treaties in the set. Appendix F.2.1 contains this and other genEIF regex pattern rules used in the Java-based GATE app for comparison. The main difference is that capture groups are unnamed, since the Java Regexp Annotator plugin employs numbered instead of named capture groups, as mentioned in the previous section.

As this excerpt illustrates, EIF provisions are similar enough to allow accurate information extraction and transformation into a structured representation suitable for quantitative analysis, but they are also complex enough to require legal domain knowledge and careful tweaking to construct regular expression patterns yielding correct results. With a large enough training set and a sophisticated enough neural network architecture[165] it may be possible to obtain these variables from treaty texts without handcrafted regular expressions. However, the result would almost certainly not be error-free, and would be much harder to correct than tweaking a regular expression. Moreover, unlike other types of data, especially data produced by internet users, the treaty universe is relatively small. If the goal is to most efficiently obtain a maximally accurate reproducible database containing these variables, then the regex approach is probably the best option.

`EIFnConsent` is extracted from 112 out of 119 treaties, the remaining ones specifying their EIF conditions in other terms, such as all states included in some list,[166] and/or states jointly reaching a given percentage of trade volume[167] or votes defined

---

[164] Paris Agreement (adopted 12 December 2015, entered into force 4 November 2016) ATS 2016 24, art 21(1) ("accounting in total for at least an estimated 55 per cent of the total global greenhouse gas emissions").

[165] E.g. Andrew Trask and others, 'Neural Arithmetic Logic Units' (S Bengio and others eds, NIPS, 2018) vol 31 use a combination of LSTMs and their proposed Neural Arithmetic Logic Units to learn to convert number words to scalar representations with relatively high accuracy (p.6).

[166] E.g. Treaty Banning Nuclear Weapon Tests in the Atmosphere, in Outer Space and Under Water (adopted 5 August 1963, entered into force 10 October 1963) 480 UNTS 43, art III(3); and Comprehensive Nuclear-Test-Ban Treaty (adopted 10 September 1996) 35 ILM 1444, art XIV(1).

[167] See e.g. the International Cocoa Agreement (n 150) art 56(1) ("This Agreement shall enter into force definitively on 1 October 1993 or any time thereafter if by such date Governments representing at least five exporting countries accounting for at least 80 per cent of the total exports of countries listed in annex A and Governments representing importing countries having at least 60 per cent of total imports as set out in annex B have deposited their instruments of ratification, acceptance, approval or accession with the depositary."). Note that the threshold for importing countries is only specified in percentage terms and the two

in an annex to the treaty.[168] In the set of 112, `EIFnConsent` ranges from 2 to 65, with three quarters of the agreements requiring up to 25 expressions of consent to be bound, as shown in Table 3.1.

Table 3.1: Min. number of ratifications required

EIFnConsent	N. Treaties
2–5	31
6–17	24
18–25	28
26–65	29

Unsurprisingly, apart from 16 treaties opting for a minimal `EIFnConsent` of 2, there is a clear preference for numbers divisible by 5, the most frequent ones being 5, 15 and 20.[169] This aesthetic factor serves as a reminder that treaty negotiators are human beings, not the perfectly rational agents sometimes postulated by game theorists.

As for additional or alternative conditions for entry into force apart from the number of instruments of consent to be bound, 47 treaties, i.e. about 40% of the treaties in the study set have such conditions, 41 of which also specify a minimum number of ratifications, with a similar distribution as the full treaty set.

Regarding `EIFdelay`, this variable has a numeric value for 110 treaties, the remaining 9 being indeterminate in that there is a waiting period of sorts, but it cannot be ascertained definitively based on the text alone. In 7 of the 9 this is due to a target EIF date, which can only ever be aspirational given the need for state consent. Among the 20 agreements for which `EIFdelay` is found to be 0, some EIF provisions could be interpreted to leave the question of a delay open, because of expressions using "after" rather than "upon" receipt of the final instrument of ratification required for entry into force. Thus, in practice these treaties may not have entered into force immediately upon fulfilment of EIF conditions, but as a design pattern, these provisions seemed sufficiently similar to group them together in the no-delay category.

Like `EIFnConsent`, `EIFdelay` also skews towards the lower end of the spectrum and exhibits numerical aesthetics (or habits), with 1, 3 and 12 months being the most

---

conditions are cumulative, so five ratifications are not sufficient.

[168] International Sugar Agreement (n 150) art 40(1).

[169] 60% of EIFnConsent numbers in this set are divisible by 5, which is about triple the amount that occurs with randomly drawn integers from the same range. Incidentally, even numbers also make up 60%, but this is closer to their normal probability of 50%.

frequent waiting periods stipulated,[170] as shown in Table 3.2. In cases where entry into force falls on the first day of the month following the one during which conditions are met, half a month is added to the time span to account for the additional delay. In reality this delay will range from 1 to 31 days, but assuming each day of any given month is equally likely to be the one on which conditions for entry into force are fulfilled, half a month is the average additional delay for all treaties that have this feature.[171]

Fractions of years may not be the most intuitive unit for `EIFdelay` because the duration of the latter tends to be short and expressed in days or months, but this conversion facilitates comparison with multi-year temporal variables, such as pre-agreed waiting periods for treaty denunciation. The conversion from smaller units of time spans to years is also what causes some of the minor differences in numeric values, e.g. between 0.49 years (180 days) and 0.5 years (six months).[172]

Table 3.2: Pre-agreed EIF waiting period

EIFdelay (Years)	N. Treaties
0.0	20
0.08	16
0.16	1
0.21	2
0.25	33
0.33	1
0.46	1
0.49	2
0.5	4
1.0	27
1.04	1
1.5	1
2.0	1

As for correlations between treaty features, the strongest pairwise correlation for EIF-related variables is exhibited by `EIFdelay` and `denunNotifMinYrs` (Pearson coeff. 0.51), and the second strongest is between adoption year and `EIFnConsent` (0.32).

---

[170]   This is consistent with Anthony Aust's finding that "the normal range is from thirty days to twelve months", see Anthony Aust, *Modern Treaty Law and Practice* (3rd ed., Cambridge University Press 2013), p.151.

[171]   Technically the average number of working days per month divided by two may have been more accurate, because receipt by the depositary is what counts and most treaty depositaries presumably do not receive instruments of consent on weekends and public holidays, but *in casu* waiting periods for the few treaties concerned were expressed in months and thus adding half a month was the simplest solution.

[172]   The year was taken to have 365.25 days on average.

Indeed, only since the 1980s has `EIFnConsent` started to reach 50 or more instruments of consent, probably at least partially due to the increasing number of states in the international community following decolonisation.

Future work could expand on additional conditions for entry into force beyond the binary variable `EIFaddlCond`, or capture information on entry into force of amendments, annexes and supplementary protocols (only partially addressed in Section 3.1.3 below) and compare it with initial EIF conditions.

3.1.3 Change

The ability to adapt to change occurring in the natural, scientific, technological, economic, social, political or legal domains is one of the most important characteristics of treaty systems governing environmental issues. The need for responsiveness and agility contrasts with the need for legal stability and predictability, as well as with governments' need for control. In the extreme, the latter can lead to a fragmented treaty regime, where no two parties are bound by the same rules at the same time. However, a number of legal innovations have been made in international law over the past few decades to strike a balance between these concerns.

The most obvious way for regimes to change their normative framework is through an amendment, but even this traditional mechanism has become more flexible and multifaceted. Another tried and tested approach is to provide for subsequent adoption of annexes or protocols as a way to extend the treaty system. Adopting a new annex is of course itself a treaty amendment so long as the annex is defined as an integral part of the agreement. Still, treaties with multiple appendices (including annexed protocols and regulations) often frame these as discrete units, similar to treaties encouraging the adoption of separate protocols to achieve treaty objectives. To minimise the overlap between these two change management techniques in the information extraction algorithm, provisions merely governing amendment of initial annexes without adoption of new ones are categorised as amendment provisions, whereas those mentioning the possibility of adoption of further annexes are additionally classified as evidence of the second technique. Of course, nothing stops parties from subsequently amending a treaty to include additional annexes, but the variable of interest here is extensibility by design. Provisions governing the adoption of supplementary protocols *not* annexed to the main treaty are conceptually distinct from amendment provisions and thus not subject to overlaps.

Techniques not covered here are treaty interpretation,[173] delegated law-making,

---

[173] Except in the context of dispute settlement, see section 3.3.6.

and soft law. While important in practice,[174] information extraction from treaty texts alone seems insufficient to yield interesting results regarding these features. Even if the scope of the analysis were to include additional documents such as decisions by treaty organs, programmatically assessing the extent to which these techniques are used by parties and treaty bodies to evolve and adapt to change, rather than 'only' for clarifying the existing normative framework, would be a considerable challenge. Traditional doctrinal research may be better suited for drawing the fine and contested line between these different use cases and examining implications for treaty design and effectiveness. Automatic expiry with planned extension or replacement of the treaty can also be a change management technique, but will be discussed under the heading of termination in section 3.1.5.

As for treaty amendment, provisions governing this process are ubiquitous in multilateral environmental agreements. In the set at hand, all but four treaties[175] are subject to amendment provisions, making this the second most common type of provision after entry into force. Due to the wide variety of rules governing everything from initial proposal to entry into effect of amendments, only two variables are constructed for each treaty: (i) whether it contains an amendment provision (`amendProv`), and (ii) whether there are any simplified amendment provisions (`simplAmendProv`). Both are binary variables and the corresponding text is extracted for verification and further analysis. Logically, `simplAmendProv` is as a subtype of `amendProv` even when its scope of application is narrower or its denomination does not include the term 'amendment'.[176] This begs the question of what counts as simplified, which in turn depends on the baseline. For present purposes, the fallback rules of arts 39-40 VCLT constitute a useful baseline, according to which the rules on conclusion and entry into force of regular treaties also apply to amending agreements,[177] and all contracting parties to a multilateral treaty have the right to be notified and

---

[174]   Jutta Brunnée, 'COPing with Consent: Law-Making Under Multilateral Environmental Agreements' (2002) 15(01) Leiden Journal of International Law 1.

[175]   The exceptions are International Convention for the Protection of Birds (adopted 18 October 1950, entered into force 17 January 1963) 638 UNTS 185; Ramsar Wetlands Convention (n 77) (subsequently amended to include them, as mentioned in the introduction to section 3.1); Convention on the Law of the Non-Navigational Uses of International Watercourses (adopted 21 May 1997, entered into force 17 August 2014) 2999 UNTS 120; and International Convention for the Suppression of Terrorist Bombings (adopted 15 December 1997, entered into force 23 May 2001) 2149 UNTS 256.

[176]   See e.g. Montreal Protocol on Substances that Deplete the Ozone Layer (adopted 16 September 1987, entered into force 1 January 1989) 1522 UNTS 3, art 2(9) on 'adjustments' to Annex A and to the phase-out schedule in art 2 of the Protocol; or International Convention for the Protection of New Varieties of Plants (adopted 19 March 1991, entered into force 24 April 1998) ATS 2000 6, art 38 on 'revision' of the Convention. Revision is sometimes construed to be more comprehensive or consequential than amendment, but is treated as the same category here.

[177]   VCLT (n 18) , art 39.

take part in the decision on whether to act on a proposal to amend the treaty, its negotiation and conclusion.[178] Also, the "amending agreement does not bind any State already a party to the treaty which does not become a party to the amending agreement".[179] These rules apply except insofar as the treaty otherwise provides, which is exactly what most treaties do. To give a sense of the range, the shortest single article on amendment comprises 53 words,[180] and the longest 1101 words,[181] a twentyfold increase. Clearly, there is more to this difference than mere verbosity. Some treaties also have more than one article dedicated to amendment, e.g. placing `simplAmendProv` in a separate article to emphasise the distinction.[182]

Named amendment provisions are very straightforward to identify and capture with regular expressions pattern matching. Out of the 64 treaties with named amendment provisions, all but two have the term 'amendment' in their article heading, and the remainder can be found with 'Revision of the (?:Treaty|Convention)'.[183] In the Python pipeline, a negative lookahead is added to prevent matching of articles which are only about amendments to treaty appendices, whereas the GATE pipeline annotates annotates them all. Thus, the Java Regexp Annotator pattern for the GATE app is as follows:

GATE Jape amendProv named article regex

```
1  |(?ism)^(Article (\w+)\s*.?\s*\b([^.;\n]*(?:Amendment|Revision of the)[^.;\n]*)\n+.+?)(?=<<eos>>)
2  1 => AmendmentProvision class="amendProv",governs="treatyAmendmentProcess",hasArtNb=$2,hasArtTitle=$3
```

The `[^.;\n]*` character class before and after the amendment/revision keyword matches matches anything that is not a period, semicolon or newline. Generally, article headings do not contain these characters, whereas the first paragraph of an unnamed article does. The word boundary metacharacter `\b` right before the article title capture group ensures that no unnecessary whitespaces or other structural elements (like a dash or colon matched by the optional dot `.?` are captured as part of the article heading. The whitespace metacharacters `\s*` in turn allows for zero or more spaces, tabs or newlines between the article number and the title.

---

[178] Ibid , art 40(2).
[179] Ibid , art 40(4).
[180] Bunker Oil Pollution Damage Liability Convention (n 114), art 16.
[181] MARPOL73 (n 75), art 16.
[182] See e.g. Convention on Supplementary Compensation for Nuclear Damage (adopted 12 September 1997, entered into force 15 April 2015) 3038 UNTS 94, art XXV ('Amendment by simplified procedure'); or Basel Protocol on Liability and Compensation for Damage Resulting from Transboundary Movements of Hazardous Wastes and their Disposal (adopted 10 December 1999), art 23 ('Amendment of Annex B').
[183] 'Revision' alone leads to false positives, such as art 84 entitled 'Revision of conviction or sentence' of the Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 3. 'Revision of the' is sufficient and may be more generalisable to agreements with other denominations.

Unnamed articles are slightly more cumbersome to extract because amendment or amended versions of a treaty may be *mentioned* by a number of provisions which do not *govern* the amendment process in any meaningful way. Still, only seven key clauses are needed to capture unnamed amendment provisions of 44 treaties. The following code block shows the relevant part of the Python pipeline as well as results for four treaties:

Unnamed amendProv regex

```
1  provdf.loc[target,"unnamedArt"] = (r'(?im)^(?P<amendProvStr>Article \w+\n+.*' +
2    r'(?:amendments to this \w+ may' +
3    r'|(?:propose|recommend).* amendment' +
4    r'|for revision or amendment' +
5    r'|for the purpose of revising' +
6    r'|adopt a new \w+ revising this' +
7    r'|consider the revision' +
8    r'|this \w+ may be revised)' +
9    r'.+(?:\n+.+(?:amend|revising).+)*)')
10 new = textseries.str.extract(provdf.loc[target,'unnamedArt']).dropna()
11 new[target] = True
12 new['treatyLabel'] = textdf.treatyLabel[new.index]
13 print(len(new), 'treaties')
14 print(new.tail(4).set_index('treatyLabel'))
```

Results

```
44 treaties
                                                              amendProvStr  amendProv
treatyLabel
OilPollFundConv71      Article 45\n\n1.\tA Conference for the purpose of revising or ...    True
OccSafetyHealthConv    Article 29\n\n1.\tShould the Conference adopt a new Convention...    True
KyotoProt              Article 20\n\n1. Any Party may propose amendments to this Prot...    True
NuclearTerrorSupprConv Article 26\n\n1.\tA State Party may propose an amendment to th...    True
```

This extraction pattern is similar to the `namedArt` pattern in that it requires the string to start with 'Article' and article number, but then there must be one or more newlines, and the first paragraph of the article must contain one of the subpatterns on lines 2-8. Like the 'Other provisAppProv excerpts' pattern seen in section 3.1.1, this unnamed amendProv pattern also foregoes the dot-all mode (`?s`) and only matches the paragraphs that contain the target expressions. Line 9 captures any subsequent paragraphs that contain 'amend' or 'revising'. There may be other subsequent paragraphs that would also be relevant, but perfect boundaries of strings (capturing all and only the relevant paragraphs) was not within the scope of this exploratory work. The main point is that it is possible to craft regular expressions which match precisely what we want them to.

Simplified amendment proved much more challenging because it is not a well-defined, clear-cut, unambiguously specified mechanism. Treaty amendment can be

facilitated in a number of ways. The techniques considered here are primarily concerned with adoption and/or entry into force of amendments. First, the adoption decision-making process may be rendered less costly for parties through voting by correspondance instead of convening a conference.[184] When amendment conferences involve dispatching delegates to an expensive location for repeated multi-day sessions, the cost of the process may discourage financially disadvantaged state parties from participating. The carbon footprint of such in-person meetings may also be considerable for global treaties. Thus, treaties with a working system for adopting amendments by correspondance could serve as an example for other aspiring global environmental agreements going forward. The process can be expedited by omitting a vote altogether, merely circulating amendment proposals and considering them accepted if no objection is received within a certain time period.[185] Indeed, this is the default mechanism for amending CITES Appendices I and II between meetings. A postal vote is only held if at least one party objects to the proposed amendment.[186] CITES Appendix III is perhaps the easiest part of any multilateral treaty to amend.[187] Amending it is essentially a unilateral decision which enters into force 90 days after its communication, with other parties only able to enter a reservation.[188] However, this is an exception more akin to national communications than treaty amendment.

Second, regarding entry into force of the amendment, an important design consideration is whether simultaneous entry into force of the amendment for all parties is necessary to ensure consistency and effectiveness of the treaty regime. If it is, then the principal options are (i) requiring ratification by all parties before the amendment becomes effective,[189] (ii) entry into force without ratification,[190] and (iii)

---

[184] See e.g. Convention on International Trade in Endangered Species of Wild Fauna and Flora (adopted 3 March 1973, entered into force 1 July 1975) 993 UNTS 243, art XV(2)(a) ("Any Party may propose an amendment to Appendix I or II for consideration between meetings by the postal procedures set forth in this paragraph."); UNCLOS (n 92), art 313(1) ("A State Party may... propose an amendment to this Convention, other than an amendment relating to activities in the Area, to be adopted by the simplified procedure set forth in this article without convening a conference").

[185] 12 months in the case of the simplified amendment procedure under UNCLOS art 313, see para 3 ("If, 12 months from the date of the circulation of the communication, no State Party has objected to the proposed amendment or to the proposal for its adoption by the simplified procedure, the proposed amendment shall be considered adopted"). There is no voting in this procedure.

[186] CITES (n 184) , art XV(2)(f) and (g).

[187] Pursuant to ibid , art XVI.

[188] Ibid , art XVI(2).

[189] E.g. Statute of the International Renewable Energy Agency (adopted 26 January 2009, entered into force 8 July 2010) 2700 UNTS 48, art XV(B)(2) ("after all the Members have consented to be bound by the amendment in accordance with their respective constitutional processes").

[190] E.g. Convention on the Conservation of Migratory Species of Wild Animals (adopted 23 June

entry into force for all parties upon acceptance by a subset of parties.[191] In some cases, entry into force conditions are made contingent on the adoption decision, with amendments adopted by consensus entering into force without ratification.[192] Simultaneous entry into force simplifies the amendment process in that it avoids fragmentation of the treaty regime, even when it means waiting for full ratification by all parties.

If a simultaneous and uniformly binding amendment is not required, then there are two further options. One is a differentiated opt-in approach, i.e. general entry into force of the amending agreement upon submission of their respective instruments of consent by a subset of parties, whilst the other parties are still bound by the previous version of the rules until they ratify the amendment. This is not simplified amendment and only mentioned for the sake of completeness.

The second approach is a differentiated opt-out system, whereby the amendment enters into force for all parties which have not objected to it within a given time period.[193] An additional twist of this approach is exemplified by the Minamata Convention, which provides that parties can opt out of this simplified (annex) amendment system through a declaration in their (initial) instruments of consent to be bound, in which case amendments enter into force on the 90th day after submission of their ratification thereof (differentiated opt-in system).[194]

Apart from treaty amendment, some agreements also foresee the subsequent

---

1979, entered into force 1 November 1983) ATS 1991 32 art XI(5) ("An amendment to the Appendices shall enter into force for all Parties ninety days after the meeting of the Conference of the Parties at which it was adopted, except for those Parties which make a reservation in accordance with paragraph 6 of this Article.").

[191] E.g. Partial Test Ban Treaty (n 166), art II(2) ("The amendment shall enter into force for all Parties upon the deposit of instruments of ratification by a majority of all the Parties, including the instruments of ratification of all of the Original Parties"); for a higher threshold see e.g. Convention on the International Hydrographic Organisation (adopted 3 May 1967, entered into force 22 September 1970) 751 UNTS 41, art XXI(3) ("The amendment shall enter into force for all Contracting Parties three months after notifications of approval by two thirds of the Contracting Parties have been received").

[192] See e.g. Agreement for the Implementation of the Provisions of the United Nations Convention on the Law of the Sea of 10 December 1982 relating to the Conservation and Management of Straddling Fish Stocks and Highly Migratory Fish Stocks (adopted 4 August 1995, entered into force 11 December 2001) 2167 UNTS 3, art 48(2) ("if a revision to an Annex is adopted by consensus at a meeting of States Parties, it shall be incorporated in this Agreement and shall take effect from the date of its adoption or from such other date as may be specified in the revision. If a revision to an Annex is not adopted by consensus at such a meeting, the amendment procedures set out in Article 45 shall apply").

[193] E.g. International Convention for the Safe and Environmentally Sound Recycling of Ships (adopted 15 May 2009) IMO Doc SR/CONF/45, art 18 para 2.6.2 ("An amendment to the Annex shall enter into force with respect to all Parties six months after the date on which it is deemed to have been accepted, except for any Party that has... notified its objection to the amendment").

[194] Minamata Convention on Mercury (adopted 10 October 2013, entered into force 16 August 2017) 55 ILM 586, , art 27(4).

adoption of annexes or protocols. Such additions are typically conceived as an optional or necessary future extension of the treaty system, without specifying their content in the agreement at hand because scientific, technological or economic knowledge is not mature enough, or because there is no political agreement yet on some anticipated component. For instance, the Basel Hazardous Wastes Convention stipulates:[195]

> The Parties shall co-operate with a view to adopting, as soon as practicable, a protocol setting out appropriate rules and procedures in the field of liability and compensation for damage resulting from the transboundary movement and disposal of hazardous wastes and other wastes.

Depending on the importance of the extension to the overall treaty system, and on the amount or specificity of normative content in the initial agreement, this approach has been dubbed the 'framework convention – protocol' approach.

The 1985 Vienna Convention for the Protection of the Ozone Layer[196] with its more well-known Montreal Protocol[197] exemplify this approach. However, the Basel Convention cited above is itself rather extensive and specific already, with highly technical annexes. In some cases, in particular regimes governing international maritime, air, and road transport, the extensions are framed as annexes rather than additional protocols, but the effect is similar. The advantage of annexes is that they can more readily be adopted through a simplified procedure without ratification.

For present purposes, anticipating and preparing for future extensions to the new legal architecture arguably makes it easier to update it in due course, not least because there is likely to be a treaty secretariat that can support negotiations of any new protocols or annexes to the initial agreement.

Information on such extensibility is encoded in a binary variable called `annexProtProv` [198] with its accompanying string variable containing the text of the provision. Only 26 treaties meet the inclusion criteria, even though 80 agreements mention the words 'annex', 'appendix' or 'protocol'. Simplified amendment provisions are more prevalent at 60 out of 115 treaties with amendment provisions (52%). Unsurprisingly, the two are closely related. Indeed, 22 treaties contain both types of provisions.[199] For

---

[195] Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and their Disposal (adopted 22 March 1989, entered into force 5 May 1992) 1673 UNTS 57, , art 12.

[196] Vienna Convention for the Protection of the Ozone Layer (adopted 22 March 1985, entered into force 22 September 1988) 1513 UNTS 293.

[197] Protocol on Substances that Deplete the Ozone Layer (n 176).

[198] Short for 'annex/protocol provision'.

[199] The three treaties with an `annexProtProv` but no `simplAmendProv` are the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects (adopted 10 October 1980, entered

`annexProtProv`, this constitutes 88% or a 38 percentage point increase compared to the full set, whereas for the `simplAmendProv` subset the intersection represents 37%, a 16 percentage point increase with regard to the full set. In other words, `simplAmendProv` nearly subsumes `annexProtProv`, but the reverse is far from true.

Notable differences in their respective associations with other provisions include provisional application clauses which see their prevalence drop to a third of the full set figure in the `annexProtProv` subset, while remaining nearly unchanged from the base rate in `simplAmendProv` treaties. Both subsets have a higher `EIFnConsent` threshold than average (33 and 26 instruments respectively), treaties in the `annexProtProv` subset are about half as likely (-16 p.p.) to stipulate additional entry into force conditions, whereas this likelihood is 12 percentage point higher than average in `simplAmendProv` treaties. By contrast, both are about half as likely to set out special denunciation conditions. This is perfectly reasonable given that a treaty regime may develop in a direction that some states are not fully on board with, and this fear could lead them to argue against any burdensome withdrawal conditions at the initial treaty negotiation stage. Curiously, none of the `annexProtProv` treaties have termination provisions, whereas `simplAmendProv` treaties do, at the normal prevalence of ~12%. This may be because many of the `annexProtProv` treaties are designed to be ever-lasting umbrella treaties for their problem domain, not even envisaging termination as a possibility. Legal personality provisions also differ considerably in their subset prevalence, falling from 21% to 8% (0.4) in `annexProtProv` treaties, while rising to 32% (1.5) in `simplAmendProv` agreements. National reporting provisions are present in all `annexProtProv` treaties but at 83% only slightly more likely to occur in the `simplAmendProv` subset. Progress review and dispute settlement provisions are also about 1.4 times more common in `annexProtProv` agreements while staying at or near the full set prevalence in `simplAmendProv` treaties. Thus, despite the large overlap between the two subsets, they do differ along a number of normative dimensions and are worth including as separate variables.

3.1.4   Withdrawal

How a party can withdraw from a treaty may be just as important as how it can join. Many treaties specify both an initial period during which notice of withdrawal cannot validly be given (`denunNotifMinYrs`), ranging from one year to ten years in the study sample, and a delayed effect of a valid notification (`denunEffectYrs`). Some

---

into force 2 December 1983) 1342 UNTS 137; WHO Framework Convention on Tobacco Control (adopted 21 May 2003, entered into force 27 February 2005) 2302 UNTS 166; and Treaty on the Prohibition of Nuclear Weapons (adopted 7 July 2017, entered into force 22 January 2021) UN Doc A/CONF 229/2017/8.

agreements allow withdrawal anytime after acceptance or entry into force for a party with only a few months' notice. At the other end of the spectrum are treaties which even after the initial waiting period can only be denounced at periodically recurring intervals. For instance, the ILO Convention concerning the Prevention of Major Industrial Accidents provides:[200]

1. A Member which has ratified this Convention may denounce it after the expiration of ten years from the date on which the Convention first comes into force, by an act communicated to the Director-General of the International Labour Office for registration. Such denunciation shall not take effect until one year after the date on which it is registered.

2. Each Member which has ratified this Convention and which does not, within the year following the expiration of the period of ten years mentioned in the preceding paragraph, exercise the right of denunciation provided for in this Article, will be bound for another period of ten years and, thereafter, may denounce this Convention at the expiration of each period of ten years under the terms provided for in this Article.

Stable commitment periods and pre-agreed delays are useful for at least four reasons. First, if the delay between the formation of an intention to withdraw and its effect is long enough, it could prevent denunciations (or threats thereof) from being misused as tools in near-term domestic or international politics. Longer delays also provide more time to change intentions to withdraw. Second, other parties may wish to reconsider their own participation in the treaty given the denunciation, either by withdrawing themselves or by strengthening their support of the treaty such as through additional funding to compensate for the loss of a contributing party. Third, any institutions set up by the denunciated treaty likely need time to adapt to the change in membership.[201] Fourth, governmental and non-governmental actors implementing the treaty in both remaining and withdrawing parties will require sufficient advance notice to accommodate the changes and update their practices within and across borders.

These reasons might lead one to hypothesise that the more effort and time it takes to implement a treaty or adjust to a change in membership and budget, the longer the stable commitment period and pre-agreed delay between intention and effect of withdrawal. However, in practice it appears that the organisational auspices

---

[200] Convention (No 174) concerning the Prevention of Major Industrial Accidents (adopted 22 June 1993, entered into force 3 January 1997) 1967 UNTS 232, art 25.

[201] E.g. Wilfred Jenks argued for a minimum of two years' notice for withdrawal from permanent organisations, see CWilfred Jenks, 'Some Constitutional Problems of International Organizations' (1945) 22 British Yearbook of International Law 11, p.23.

under which a treaty was negotiated is much more influential. For instance, all but one of the ILO conventions in the set at hand stipulate a 10-year stable commitment period,[202] and none of them set up a new entity with legal personality. The second longest period is five years, found mostly in IMO conventions. Conversely, all six nuclear safety conventions adopted under the auspices of the IAEA stipulate no initial waiting period (and only 6-12 months' notice). Table 3.3 displays the range of values `denunNotifMinYrs` takes and the corresponding number of treaties.

Table 3.3: Initial waiting period before withdrawal

Years	N. Treaties
0	39
1	9
2	11
3	10
4	3
5	13
10	12

Zero is by far the largest category. This is typically formulated as parties being able to withdraw from the treaty "at any time" after its entry into force.[203] Denunciation provisions not specifying any initial waiting period or absence thereof are nevertheless captured in the 'anytime' pattern.[204] This is consistent with VCLT articles 54 and 56 on treaty denunciation, which do not stipulate such a condition. 58 treaties do specify a non-zero initial commitment period, i.e. just over half (51%) of the agreements with withdrawal provisions (n=113).

As for `denunEffectYrs`, both the average and range of values is smaller. Table 3.4 summarises the results.

---

[202] The exception being the Convention (No 115) concerning the Protection of Workers against Ionising Radiations (adopted 22 June 1960, entered into force 17 June 1962) 431 UNTS 41, which prescribes a five-year initial waiting period at art 18(1).

[203] See e.g. Convention relating to Civil Liability in the Field of Maritime Carriage of Nuclear Material (adopted 17 December 1971, entered into force 15 July 1975) 974 UNTS 255, art 7(1) ("The present Convention may be denounced by any Contracting Party to it at any time after the date on which the Convention comes into force for that State").

[204] See e.g. New Varieties of Plants Protection Convention (n 176), art 39(1) ("Any Contracting Party may denounce this Convention by notification addressed to the Secretary-General").

Table 3.4: Withdrawal notice period treaty counts

Years	N. Treaties
0.08	1
0.25	11
0.33	1
0.49	1
0.5	8
1.0	84
2.0	1

Twelve months is by far the most frequently chosen minimum notice period (84 treaties) and coincides with the fallback rule provided by Article 56(2) of the Vienna Convention on the Law of Treaties.[205] Three months is the second most common number chosen (11 treaties), followed by six months (9 treaties). In total, `denunEffectYrs` was extracted from 107 out of 113 denunciation provisions.

Similar to entry into force provisions, the most prevalent variables do not tell the full story. Both validity and effects of denunciation may be subject to additional rules. These are captured as binary variables `denunCondOther` for other denunciation conditions affecting validity (n=33), and `denunEffectOther` for provisions governing when a valid withdrawal notification shall take effect in cases not covered by `denunEffectYrs` alone (n=7). The latter typically involves withdrawal taking effect at the end of the year in which (or following that in which) notice of denunciation is received.[206] A fixed notice period and a month- or year-end alignment can of course co-exist,[207] like in the case of entry into force conditions.

Conditions affecting validity are more diverse and consequential. Perhaps it would be best to subdivide them into more meaningful categories than this catch-all. The largest category is timing-related (n=18). The recurring 5 or 10-year waiting periods of ILO conventions are the most extreme example. Less onerous conditions have an annual date by which the treaty can be denounced with appropriate notice.[208] Other agreements allow withdrawal only after the first revision

---

[205] Provided the conditions of art 56(1) for denunciation *per se* are met, a withdrawing party must give at least twelve months' notice.

[206] E.g. Fishing Vessels Compliance Promotion Agreement (n 161), art XIV *in fine* ("Withdrawal shall become effective at the end of the calendar year following that in which the notice of withdrawal has been received by the Director-General").

[207] E.g. IHO Convention (n 191), art XXII(1), which stipulates a five-year initial waiting period, one year's notice, and a delayed effect on 1 January of the year following expiration of the notice period.

[208] E.g. at the end of the fiscal year in Grains Trade Convention (n 150), art 29; or on June 30th of any year in Whaling Convention (n 161), art XI.

conference.[209] The second largest subcategory of `denunCondOther` is formed by treaties which stipulate a substantive condition. A standard formulation used in the seven WMD treaties of the sample set provides that each state party "shall in exercising its national sovereignty have the right to withdraw from the Convention if it decides that extraordinary events, related to the subject matter of the Convention, have jeopardized the supreme interests of its country."[210] Humanitarian law treaties in turn disallow withdrawal by parties engaged in an armed conflict for the duration of such conflict.[211] In this case it is not so much the organisational auspices but the subject matter of the treaty which calls for a higher denunciation threshold.

Overall, the treaty set studied provides a rich sample of withdrawal provisions and the wording is similar enough to make information extraction with regular expressions relatively straightforward.

3.1.5   Termination

Differing usage in certain treaties notwithstanding,[212] 'termination' of a multilateral agreement denotes the end of the lifespan of a treaty per se, for all parties,[213] and as such is the counterpart of its general entry into force as opposed to individual entry into force. Provisions governing consequences of termination continue to apply to the extent necessary, just as provisions on its entry into force apply from adoption of the agreement.

Unlike amendment and withdrawal, termination provisions are much rarer in the sort of open multilateral treaties under examination here. Even in the case of replacement through a subsequent treaty, the pre-existing agreement typically remains in force for as long as at least some parties have not yet expressed their consent to

---

[209] Convention on the Liability of Operators of Nuclear Ships (adopted 25 May 1962) 2 ILM 727, art XXVII(1); and Vienna Convention on Civil Liability for Nuclear Damage (adopted 21 May 1963, entered into force 12 November 1977) 1063 UNTS 265, art XXVI(B).

[210] Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction (adopted 10 April 1972, entered into force 26 March 1975) 1015 UNTS 163, art XIII(2), and nearly identically in the other withdrawal provisions of WMD treaties.

[211] Technically, they only stipulate that withdrawals "shall not take effect before the end of the armed conflict", see e.g. Certain Conventional Weapons Convention (n 199), art 99(1), thus perhaps nominally more about effect than validity of a withdrawal notice.

[212] E.g. Agreement for the Establishment of the Global Crop Diversity Trust (adopted 4 October 2003, entered into force 21 October 2004) 2366 UNTS 205, art 6 is entitled 'Termination' but is in fact about denunciation (as the wording of the provision makes clear), and the 1963 version of the Convention on Civil Liability for Nuclear Damage (n 209) likewise uses 'termination' to refer to withdrawal of a party, e.g. in arts XX and XXV, although it also contains a separate provision on 'denunciation' (art XXVI). This distinction was dropped in the amending protocol of 1997 in favour of a more standard use of terms.

[213] See VCLT (n 18) arts 54-56.

be bound by the new treaty. To be clear, provisions governing termination and re-placement of another (prior) treaty are not captured here,[214] only those concerning termination of the agreement they are a part of. After all, the present focus lies on the design of the life cycle of a given treaty and in particular whether its ending is envisaged or not. Provisions deferring the duration/extension decision to a review conference are included so long as termination is at least implicitly acknowledged as a possibility.[215] Articles merely echoing the customary rule of unlimited duration[216] are not counted as (non-)termination provisions, because the information gain from including such rules is insufficient for present purposes.

The first question that treaty drafters face is whether to govern termination at all. In the overwhelming majority[217] of agreements covered here, the answer was no. In the minority that do contain termination provisions (n=15), the extent to which termination and its effects are governed differs considerably. At one end of the spectrum are conventions which only go as far as putting the topic of duration on the agenda of a review conference[218] or allowing a party to propose termination at any time after its entry into force,[219] but leaving the decision to terminate and any questions regarding its consequences entirely up to future decision-makers. At the other end are treaties which unambiguously specify a condition which, when met, automatically triggers termination without any scope for reconsideration.[220] In-between are provisions with varying degrees of openness to ad hoc considerations

---

[214] E.g. Convention on Road Traffic, 1968 (adopted 8 November 1968, entered into force 21 May 1977) 1042 UNTS 17, art 48 ("Upon its entry into force, this Convention shall terminate and replace, in relations between the Contracting Parties, the International Convention relative to...").

[215] E.g. Treaty on the Non-Proliferation of Nuclear Weapons (adopted 1 July 1968, entered into force 5 March 1970) 729 UNTS 161, art X(2) ("Twenty-five years after the entry into force of the Treaty, a conference shall be convened to decide whether the Treaty shall continue in force indefinitely, or shall be extended for an additional fixed period or periods").

[216] E.g. Ramsar Wetlands Convention (n 77), art 11(1) ("This Convention shall continue in force for an indefinite period"); similarly Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction (adopted 3 September 1992, entered into force 29 April 1997) 1974 UNTS 45, art XVI(1); and Treaty on the Prohibition of Nuclear Weapons (n 199), art 17(1).

[217] More precisely, 87% (104 out of 119) treaties.

[218] See Nuclear Non-Proliferation Treaty (n 215), art X(2) op. cit.

[219] Food Assistance Convention (n 162), art 17(2), also specifying that the proposal must be communicated in writing to the Secretariat which shall circulate it to all parties at least six months in advance of its consideration, much like amendments governed by art 16(1).

[220] Vienna Road Traffic Convention (n 214), art 51; International Convention on the Establishment of an International Fund for Compensation for Oil Pollution Damage, 1971 (adopted 18 December 1971, entered into force 16 October 1978) 1110 UNTS 57, art 43(1); and International Convention on Liability and Compensation for Damage in Connection with the Carriage of Hazardous and Noxious Substances by Sea (adopted 3 May 1996) IMO Doc LEG/CONF 10/8/2, art 51(1)(a). Of course, treaty amendment remains an option, but this can be a lengthy and onerous process with a complicated outcome if not accepted by all parties.

if and when the matter arises. Along this axis, a line can be drawn to separate auto-terminating treaties from those that only cease to be in force upon further action by parties.

There are two main design patterns for auto-termination provisions, namely membership thresholds and expiration. The former provides that the treaty shall cease to be in force when the number of parties drops below a certain threshold. Four treaties in the study set contain this type of provision, with a threshold of three,[221] five,[222] six,[223] and forty[224] members respectively, corresponding to about a third of the number of instruments of consent needed for entry into force in the case of the first two, half the entry into force threshold in the third case, and the full number of ratifications in the fourth. The question of legal relevance of the entry into force threshold for continuance in force of the treaty was acknowledged but dismissed by the Vienna Convention on the Law of Treaties.[225]

Only one of the four agreements provides an override option, stipulating that the treaty shall be automatically terminated once the condition obtains, "unless the remaining Contracting Parties unanimously decide otherwise".[226] This contrasts with the second common design pattern of specifying an expiration date or period, which, at least in this set of treaties, always contains a way to override the default by agreeing on an extension. Six agreements display this pattern,[227] all of which are agricultural commodity agreements.

As for associations with other variables, Figure 3.3 displays an overview of pairwise comparisons. The most striking surges in prevalence are provisional application provisions (4.5 times more common in treaties with termination provisions), legal personality provisions (nearly three times more frequent), and additional EIF conditions (2.2 times more likely). Conversely, the inclusion of dispute settlement provisions drops to half the normal rate and provisions on adoption of annexes or supplementary agreements do not occur at all in this subset.

---

[221] Oil Pollution Damage Compensation Fund Convention (n 220), art 43(1).

[222] Vienna Road Traffic Convention (n 214), art 51.

[223] HNS Sea Carriage Liability & Compensation Convention (n 220), art 51(1)(a).

[224] International Treaty on Plant Genetic Resources for Food and Agriculture (adopted 3 November 2001, entered into force 29 June 2004) 2400 UNTS 303, art 33(1).

[225] VCLT (n 18), art 55.

[226] Plant Genetic Resources Treaty (n 224), art 33(1) *in fine*.

[227] International Olive Oil Agreement (n 139), art 49; International Coffee Agreement, 1983 (adopted 16 September 1982, entered into force 11 September 1985) 1333 UNTS 119, art 68; International Sugar Agreement (n 150), art 45; International Cocoa Agreement (n 150), art 61; Grains Trade Convention (n 150), art 33; and International Tropical Timber Agreement (n 150), art 44.

Cluster	Variable	Type	Full Set (n=119)	Subset (n=15)	Div	Diff
Start of Life	Provisional Application Provision	bool	14 (12%)	8 (53%)	4.5	41.6
	Entry into Force Provision	bool	119 (100%)	15 (100%)	1.0	0.0
	—Min. Nb. of Ratifs for EIF	num	112 (19.8)	12 (17.5)	0.9	-2.3
	—Additional EIF Conditions	bool	47 (39%)	13 (87%)	2.2	47.2
	—EIF Waiting Period (Years)	num	110 (0.4)	8 (0.4)	1.0	0.0
Change	Amendment Provision	bool	115 (97%)	15 (100%)	1.0	3.4
	—Simplified Amendment	bool	60 (50%)	7 (47%)	0.9	-3.8
	Adoption of Annexes or Protocols	bool	26 (22%)	0 (0%)	0.0	-21.8
End of Life	Denunciation Provision	bool	113 (95%)	15 (100%)	1.1	5.0
	—Denunciation Waiting Period (Years)	num	97 (2.7)	11 (0.6)	0.2	-2.0
	—Other Denunciation Conditions	bool	33 (28%)	4 (27%)	1.0	-1.1
	—Denunciation Effect Delay (Years)	num	107 (0.9)	14 (0.6)	0.7	-0.3
	—Other Denunciation Effect	bool	7 (6%)	1 (7%)	1.1	0.8
Treaty Organs	Legal Personality	bool	25 (21%)	9 (60%)	2.9	39.0
	Organisational Auspices	cat	108 (91%)	14 (93%)	1.0	2.6
	Secretariat	bool	111 (93%)	14 (93%)	1.0	0.1
	Conference of Parties	bool	85 (71%)	13 (87%)	1.2	15.2
Compliance & Effectiveness	National Reporting	bool	89 (75%)	10 (67%)	0.9	-8.1
	International Verification	bool	63 (53%)	9 (60%)	1.1	7.1
	Progress Review	bool	73 (61%)	11 (73%)	1.2	12.0
	Non-Compliance Procedures	bool	66 (55%)	8 (53%)	1.0	-2.1
	Sanctions	bool	58 (49%)	9 (60%)	1.2	11.3
	Dispute Settlement	bool	74 (62%)	5 (33%)	0.5	-28.9
	—Compulsory Dispute Settlement	cat	41 (34%)	4 (27%)	0.8	-7.8

Figure 3.3: Normative profile of treaties with termination provisions

Commodity agreements certainly leave their mark on this subset's normative profile, but civil liability treaties and agreements setting up smaller, potentially non-permanent institutions beyond those managing agricultural commodities also benefit from clarifying their ending. As with other types of provisions, factors influencing this profile probably include a mix of organisational habits and design needs.

3.2 Treaty organs

Treaty organs, as this term is used here, refers to permanent or temporary roles and bodies established or mandated by a treaty to perform certain functions, such as depositary, secretary-general, governing council, general assembly, expert committees, trust funds etc. An organ set up by one treaty can perform the same or a different role for another. For instance, the UN Secretary-General is a treaty organ created by the UN Charter, but assumes the role of depositary for a whole host of multilateral treaties. The degree of institutionalisation expresses the extent to which normative or other powers are conferred to treaty organs. At one end of the spectrum are agreements without any provisions on treaty organs other than the

depositary. At the other end are treaties establishing a full-blown international or-
ganisation with legislative, executive and adjudicative functions. There is a plethora
of configurations in-between these two extremes and an exhaustive typology would
be beyond the scope of this project, even for the limited number of treaties con-
sidered. Instead, only four institutional aspects were examined. First, whether the
treaty creates an entity with legal personality, such as an intergovernmental organ-
isation or tribunal. Second, whether the treaty was adopted under the auspices of
an existing international organisation, and if yes, which one. Third, whether the
agreement is administered by a secretariat, either newly established or pre-existing.
Fourth, whether the treaty institutionalises a Conference of Parties (COP). Infor-
mation on the depositary was not deemed worth extracting from treaty texts given
that it is included in the UNTS database and thus already in a structured format.

In terms of treaty design, these four aspects may be less interesting than others,
because, with the exception of legal personality, they may be determined more by
the institutional context than anything else. Moreover, the secretariat and COP
are such common organs that it is their absence rather than presence in a treaty
that may be illuminating. As for legal personality, this seems mostly given by the
treaty objective and the main functions its organs are intended perform. Neverthe-
less, these four aspects are basic parameters characterising a treaty and thus worth
capturing. Having extracted the relevant treaty provisions, future analyses of spe-
cific functions, meeting frequency, locations, budget, decision-making rules etc. may
enable a more fine-grained picture of these institutional aspects.

In the treaty set at hand, 25 agreements explicitly establish an entity with legal
personality.[228] Of these, only six specify it to be *international* legal personality.[229]
Eight employ a standard formulation on legal capacity to the same effect as legal
personality, which appears to be limited to capacity under domestic law of states
parties.[230] The others make no mention of the jurisdiction where this legal[231] or

---

[228] On this concept, see Jan Klabbers, 'The Concept of Legal Personality' (2005) 11 Ius Gentium
        35.
[229] E.g. IRENA Statute (n 189), art XIII(A) ("The Agency shall have international legal person-
        ality").
[230] E.g. Constitution of the World Health Organization (adopted 22 July 1946, entered into force
        7 April 1948) 14 UNTS 185, art 66 ("The Organization shall enjoy in the territory of each
        Member such legal capacity as may be necessary for the fulfilment of its objective and for the
        exercise of its functions"). This formulation stems from Article 104 of the UN Charter, which
        is incorporated by reference in Constitution of the United Nations Educational, Scientific and
        Cultural Organization (adopted 16 November 1945, entered into force 4 November 1946) 4
        UNTS 275, art XII. The drafting committee of the UN Charter thought it unnecessary to
        explicitly mention international legal personality alongside domestic legal personality. For an
        overview of these concepts and their uses, see Jan Klabbers, *An Introduction to International
        Institutional Law* (Cambridge University Press 2002), pp. 49-57.
[231] E.g. International Tropical Timber Agreement (n 150), art 17(1) ("The Organization shall

'juridical'[232] personality ought to exist. As for what kind of entity is endowed with such personality, about a third of them are specialised agencies and related organisations of the UN.

Organisational auspices differs from other variables in that there is typically no treaty provision stipulating the organisational auspices the treaty is concluded under. It is part of the pre-existing legal context that the treaty is embedded in. Organisations can encourage or facilitate the conclusion of treaties without subsequently playing a role in its implementation, especially when the treaty establishes its own organs and processes. The determination is not always clear-cut, because the concept itself is not formally defined. It is understood in a broad sense here to encompass not only situations where a treaty organ of an international organisation adopts a convention based on a constitutional mandate, but also looser configuration, such as UNEP facilitating the adoption of and performing administrative functions for environmental agreements. Technically, UNEP is not an international organisation but a programme established by the UN General Assembly. Still, it is worth including it as one of the nine `orgAuspices` categories given the focus on environmental agreements.[233] The variable is constructed for 108 (91%) of the 119 multilateral agreements and nine organisations (IAEA, ILO, IMO, FAO, UN, UNEP, UNESCO, WHO and WIPO). Constitutions of organisations are not categorised as adopted under its own auspices, since the organisation did not exist at the moment of adoption of its constituent instrument. Relevant information is expressed in the title,[234] preamble,[235] in the provisions on the secretariat[236] and depositary,[237] among others. Regex pattern matching on the full text is perfectly suitable for extracting

---

have legal personality").

[232] IHO Convention (n 191), art XIII; and ICGEB Statutes (n 128), art 13(1).

[233] Other entities established by the UNGA and playing a role in the adoption and administration of the treaties examined are not included due to the small number of treaties.

[234] E.g. the full titles of UNCLOS (n 92), Framework Convention on Climate Change (n 82), and Convention to Combat Desertification (n 81).

[235] E.g. ILO conventions begin the preamble with "The General Conference of the International Labour Organization," and end it with "adopts this [date] the following Convention[...]:".

[236] E.g. Convention on Nuclear Safety (adopted 20 September 1994, entered into force 24 October 1996) ATS 1997 5, art 28(1) ("[The IAEA] shall provide the secretariat for the meetings of the Contracting Parties"); and Stockholm Convention on Persistent Organic Pollutants (adopted 22 May 2001, entered into force 17 May 2004) 2256 UNTS 119, art 20(3) ("The secretariat functions for this Convention shall be performed by the Executive Director of the United Nations Environment Programme, unless the Conference of the Parties decides, by a three-fourths majority of the Parties present and voting, to entrust the secretariat functions to one or more other international organizations").

[237] E.g. Nuclear Damage Supplementary Compensation Convention (n 182), art XVIII(2) ("The instruments of ratification, acceptance or approval shall be deposited with the Director General of the International Atomic Energy Agency who shall act as the Depositary of this Convention").

any kind of textual evidence wherever it occurs, but careful verification is key. For instance, UNESCO is depositary of both the World Heritage Convention[238] and the Wetlands Convention,[239] but the former was adopted under its auspices whereas the latter was not.

The need for this variable stems from the fact that in some agreements certain provisions are omitted entirely if they are already exhaustively covered in the constitutive treaty, such as on the secretariat and dispute settlement. Additionally, as we have seeen in previous sections, agreements adopted under the auspices of the same organisation often have very similar final clauses. Structured data on organisational auspices facilitates the generation and testing of hypotheses on organisational habits as a factor influencing treaty design.

A related piece of information is whether a given treaty is a supplementary agreement to another treaty in the data set. This is only necessary for agreements which state that provisions of the supplemented treaty apply mutatis mutandis. Such clauses trigger a lookup and copy operation in the information extraction algorithm, such that legal features from the convention are copied to the protocol they apply to.

Which features are 'inherited' or incorporated explicitly or implicitly from an organisation or treaty needs to be determined on a case-by-case basis and encoded accordingly. The focus here was not on mapping the network of legal and organisational relations but on eliminating false negatives in the set of target variables constructed for each treaty. The most striking example are ILO conventions. Perhaps due to the prolific treaty-making activity of this organisation, the conventions adopted under its auspices are very streamlined and largely omit unnecessary provisions. Specifically, the binary variables on the secretariat, national reporting, international verification, treaty progress review, non-compliance procedure, sanctions, and dispute settlement were set to `True` based on provisions in the ILO Constitution. Of these, only review provisions[240] occur in all conventions of the set at hand, and national reporting provisions occur in some of them. Not making an adjustment based on the legal context would make ILO conventions seem underdeveloped. As for treaties adopted under the auspices of UNESCO, IMO, IAEA, WHO and UNEP,

---

[238]   Convention for the Protection of the World Cultural and Natural Heritage (adopted 16 November 1972, entered into force 17 December 1975) 1037 UNTS 151.

[239]   Ramsar Wetlands Convention (n 77).

[240]   E.g. Convention (No 176) concerning Safety and Health in Mines (adopted 22 June 1995, entered into force 5 June 1998) 2029 UNTS 207, art 22 ("At such times as it may consider necessary, the Governing Body of the International Labour Office shall present to the General Conference a report on the working of this Convention and shall examine the desirability of placing on the agenda of the Conference the question of its revision in whole or in part.") and nearly identical in other ILO conventions.

only the secretariat variable is set to `True`, because the other variables can be derived from the text of the convention itself. Agreements adopted under the auspices of the UN more broadly have no feature inheritance, since this category includes a wide range of institutional contexts with different features.

A more atomistic approach to treaty analysis is exemplified by Scott Barrett's variable on whether treaties "create a supporting secretariat or administrative organization".[241] The operationalisation of this appears to mostly exclude treaties without a need to have a provision on the secretariat such as treaties adopted under the auspices of a treaty-administrating organisation like the ILO, IMO and IAEA, as well as treaties entrusting UNEP with administrative functions.[242] In other words, perhaps the focus is on institutional proliferation (creating a new secretariat for every treaty) rather than whether administration is part of treaty design (in its broader context). However, there are many exceptions that belie this interpretation of the variable.[243]

For present purposes, the `secretariat` variable is more concerned with whether a treaty is designed to have an international administrator than where or how provisions to this effect are expressed. Administration is taken to consist in more than basic depositary functions,[244] but of course the depositary can be asked to fulfill such additional roles as convening review conferences. The only cases where this kind of role assignment is not deemed sufficient for the `secretariat` variable are treaties with a governmental rather than international depositary. There are four treaties in the dataset that have a provision on review conferences of the parties but not on the secretariat, because they task their depositary governments with the organisation of said conferences.[245] In total, a secretariat was found to be provided for in 111 (93%) of the 119 agreements.

The plenary organ is rather diverse in terms of naming, including not only Conference or Meeting of the Parties and Assembly or Congress in its name, but also

---

[241] Barrett, *Environment and Statecraft* (n 12) , p.163.

[242] See the table in ibid, p.165-194 ("Appendix 6.1. Multilateral Environmental Agreements").

[243] One ILO convention (Working Environment Convention (n 94)) and two IAEA conventions (Nuclear Accident Notification Convention (n 149) and Nuclear Accident Assistance Convention (n 149)), all without secretariat provisions, are classified as creating an administrative organisation in Barrett's conceptualisation. UNEP agreements are split about 50/50 (five 'No' and four 'Yes'), and IMO treaties are categorised as 'Yes' in six treaties and 'No' in three, without an immediately apparent logic.

[244] Listed in VCLT (n 18), art 77(1)(a) to (g).

[245] Nuclear Non-Proliferation Treaty (n 215), art VIII(3); WMD Seabed Emplacement Prohibition Treaty (n 161), art VII; Biological Weapons Convention (n 210), art XII; and Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I) (adopted 8 June 1977, entered into force 7 December 1978) 1125 UNTS 3, art 7.

Council, Commission or Committee in some instances.[246] The decisive factor for
the `copProv` variable is whether the body in question is open to representatives of
all parties or not. In the case of treaties having more than one plenary organ, the
highest-level one[247] was chosen for information extraction. Both agreements with
entire and sometimes lengthy articles on this treaty organ[248] and treaties only includ-
ing a minor provision on the COP in the context of an article focused on amendment,
revision or review were considered. To avoid making this variable co-extensive with
the one on amendment provisions (given that regular amendment necessitates a
meeting of the parties), mere mention of the COP in the context of amendment was
deemed insufficient, there has to be some additional function or task envisaged like
treaty progress review or compliance verification. Such provisions were found in 85
(71%) of the 119 agreements. Unlike the approach chosen with the secretariat, for
`copProv` only treaties with a provision on the plenary body are counted, whatever
the organisational auspices. This is reasonable because the baseline for measuring
institutionalisation is different. Every treaty can have a meeting of the parties,
there is no need for a provision on it. By default, a meeting of all the parties is
the supreme decision-making body. By contrast, a secretariat or director-general
cannot materialise out of nowhere, without appropriate legal basis.

Regarding false negative search and ensuring consistency, examining conspicuous
combinations proved useful, such as the above-mentioned cases with COP provisions
but no secretariat. The Python pipeline makes this exceedingly easy. For instance,
counting the number of treaties in each of the occurring combinations of the main
four variables in this section (using a binarised version of `orgAuspices`)[249] takes one
line of code:

Contingency table of institutional variables

```
1  print(textdf[['legalPers','orgAuspicesB','secretariat','copProv']].value_counts())
```

---

[246] E.g. Whaling Convention (n 161), art III(1) ("The Contracting Governments agree to estab-
     lish an International Whaling Commission, hereinafter referred to as the Commission, to be
     composed of one member from each Contracting Government."); or Food Assistance Conven-
     tion (n 162), art 7(1) ("A Food Assistance Committee (the "Committee"), consisting of all of
     the Parties to this Convention, is hereby established").

[247] E.g. the Ministerial Conference in case of the WTO.

[248] E.g. CITES (n 184), art XI; Migratory Species Conservation Convention (n 190), art VII;
     Convention on Biological Diversity (adopted 5 June 1992, entered into force 29 December 1993)
     1760 UNTS 79, art 23; Framework Convention on Climate Change (n 82), art 7; Convention to
     Combat Desertification (n 81), art 22; and Persistent Organic Pollutants Convention (n 236),
     art 19.

[249] This means that all treaties that have some `orgAuspices` (whatever they may be) are assigned
     `True` and the others count as `False`.

```
Results
legalPers  orgAuspicesB  secretariat  copProv
False      True          True         True     55
                                      False    29
True       True          True         True     17
           False         True         True      7
False      True          False        True      4
           False         False        False     2
                         True         True      2
           True          False        False     2
True       True          True         False     1
dtype: int64
```

Instead of the absolute number of treaties, the proportion could be displayed (by passing the option `normalize=True` to the `.value_counts()` function), but that may be more of a distraction here. For instance, one treaty corresponds to 0.84% and two treaties make up 1.68% of the total, which seems more complicated than simple treaty counts.[250] Which combinations are surprising relies on background knowledge and understanding of the treaty set and variables at hand. In this case, the occurrence pattern seems mostly reasonable. The sole instance without a COP despite having `True` in all other institutional variables is the Agreement for the Establishment of the Global Crop Diversity Trust,[251] which only sets up non-plenary organs.

As for associations with other variables, these are discussed in previous and subsequent sections and are thus not repeated here.

3.3 Compliance and effectiveness

Due to the wide range of treaty objectives and problem characteristics, compliance and effectiveness mechanisms take on many shapes and forms. Some of the types of provisions discussed in this section partially overlap, but every effort was made to limit the scope of the variables sufficiently so as to capture the essence of each. Apart from dispute settlement, variables are often named differently in the agreements covered, e.g. instead of "sanctions" euphemisms such as 'consequences' or 'suspension of benefits' are used. Without aiming at exhaustiveness, in the following subsections, provisions on national reporting (3.3.1), international verification of compliance (3.3.2), treaty review (3.3.3), non-compliance procedures (3.3.4), sanctions (3.3.5), and dispute settlement (3.3.6) will be briefly introduced and relevant results discussed.

---

[250] Preferences may vary and both absolute and relative numbers are given in other tables. Here the aim is partly to show how much work a single line of code can do.

[251] Global Crop Diversity Trust Agreement (n 212).

When looking at all the possible combinations of these six binary variables, the scenario where they are all present[252] is the single most common at 24 (20%) out of 119 agreements, followed by all present except for sanctions provisions in 9 (8%) treaties, and all absent in 9 (8%) other treaties. That still leaves 77 (65%) treaties with less common combinations. In total, 36 out of 64 possible configurations have at least one member. 110 (92%) treaties contain at least one of the provisions in this cluster, and the median (50% mark) lies at 4 provisions out of 6.

One way to quantify the overlap between binary variables is to calculate the Jaccard similarity of each pair of variables.[253] This measure simply divides the intersection of the two variables, i.e. the number of treaties where both are present, by the number of all treaties where at least one of them is present (the union of the two sets). The result is a number between 0 and 1 where 0 means no overlap and 1 means identical. Figure 3.4 shows the result for each pair of the six variables, with the diagonal corresponding to the comparison of each variable with itself, and the triangles above and below the diagonal mirroring each other.

	natRepProv	verifProv	reviewProv	ncpProv	sanctionsProv	disputeProv
natRepProv	1.000	0.567	0.653	0.598	0.441	0.567
verifProv	0.567	1.000	0.528	0.697	0.635	0.473
reviewProv	0.653	0.528	1.000	0.527	0.365	0.531
ncpProv	0.598	0.697	0.527	1.000	0.550	0.538
sanctionsProv	0.441	0.635	0.365	0.550	1.000	0.435
disputeProv	0.567	0.473	0.531	0.538	0.435	1.000

Figure 3.4: Compliance provisions' Jaccard similarity

Based on this metric, the two most closely connected (similarly distributed) variables are verification and NCP provisions (70% overlap), followed by national reporting and treaty review provisions (65%). The most distant (or dissimilar) pair is formed by review and sanctions (37%), which makes sense. Overall, 11 out of 15 pairs exhibit over 50% overlap.

3.3.1   National reporting

Periodic reporting on implementation, effectiveness, or other variables of interest to a treaty organ is an obligation found in a wide range of treaties and the basis for

---

[252]   This set intersection corresponds to the joint probability of all six types of provisions occurring in a treaty.

[253]   For an overview of this and other binary similarity measures, see Seung-Seok Choi, Sung-Hyuk Cha, Charles C Tappert, and others, 'A Survey of Binary Similarity and Distance Measures' (2010) 8(1) Journal of Systemics, Cybernetics and Informatics 43.

several other mechanisms in this cluster. Every treaty examined had some information exchange aspect. To narrow down the scope of this variable while remaining sufficiently flexible, four factors are considered: originator, addressee, frequency, and content. Whether or not it is called 'reporting' is irrelevant.

First, the originator has to be a state or regional integration organisation, at least at the moment of submission. The report may include data from the private sector or any other actors on a voluntary or mandatory basis, but it is to be submitted on behalf of the contracting party. Also, all parties ought to have some reporting obligation, even if the content requirements differ depending on party characteristics.

Second, the addressee must be a treaty organ, any interested contracting party, or the general public.[254] Bilateral communications between parties do not qualify, but confidential reports to a treaty organ do.[255] Treaty organs such as the secretariat or COP are the most common addressees, which may or may not be required to publish the content or a summary thereof.[256]

Third, while the frequency is often stipulated to be a regular interval, regularity is not included as a requirement for the purpose of this analysis. An initial report and subsequent updates as the situation changes is sufficient.[257] The main question here is whether the treaty sets up a system which allows for an assessment of the evolution over time for as long as the treaty remains in force. Merely ad hoc communications about some emergency situation, such as the reporting provisions of the Nuclear Accident Notification Convention,[258] do not qualify.

Fourth, the content is deliberately left as open as possible, as long as it has some substance beyond mere contact information,[259] and provided it is not purely event-driven, such as disaster communication.[260] While most obligations consist in self-reporting, i.e. information about the party's own actions to implement the treaty, reports on other states' or private sector actions and data gathered from areas beyond national jurisdiction are also covered and sufficient.

89 (75%) of treaties are found to have a national reporting mechanism meeting these four criteria.

---

[254] E.g. Outer Space Treaty (n 161), art XI, and Agreement Governing the Activities of States on the Moon and Other Celestial Bodies (adopted 5 December 1979, entered into force 11 July 1984) 1363 UNTS 3, art 5(1), mention the public and the international scientific community among the addressees.

[255] E.g. Framework Convention on Climate Change (n 82), art 12(9).

[256] E.g. ibid, art 12(10).

[257] E.g. Ramsar Wetlands Convention (n 77), arts 2(1), 2(5) and 3(2).

[258] Nuclear Accident Notification Convention (n 149), arts 2 and 5.

[259] Unlike e.g. ibid, art 7.

[260] E.g. Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched into Outer Space (adopted 22 April 1968, entered into force 3 December 1968) 672 UNTS 119, art 1.

As for system design considerations, there is clearly no one-size-fits-all solution, and details may be best fleshed out by COP decisions, but it is still worth giving adequate thought to the reporting system at the treaty negotiation stage. In particular, if the treaty's *raison d'être* is to achieve progress towards some measurable goal, then it stands to reason that relevant outcome data should be collected. In some treaties this is made explicit by requiring parties to not just report on measures taken to implement the treaty, but also on effectiveness thereof.[261] However, for the present analysis no such distinction is made in the information extracted from treaty provisions, among other reasons because treaty implementation may well include setting up a monitoring system at the national level to evaluate the effect of any measures taken, in which case reporting on implementation of this system would be sufficient. Another aspect to consider is whether the scope of reporting should extend to actions by, or relations with, other parties and non-parties, to enable cross-checking of reports and detect potential loopholes. Similarly, in the case of preventive treaties and agreements with a mutual inspection regime, a high degree of transparency may facilitate the operation of any verification mechanism. Despite this link between reporting and verification, it seems wise to frame reporting obligations as a standalone provision. After all, the information gathered and potentially published may serve other purposes beyond compliance and effectiveness assessments, such as education, capacity-building, training, research, harmonisation, coordination and cooperation, which for some states may enhance incentives to submit accurate, complete and timely reports. A more practical issue is the growing reporting burden for states that are parties to dozens or even hundreds of multilateral agreements with periodic reporting requirements, affecting compliance with such obligations and potentially willingness to expand treaty commitments. While streamlining of national reporting has mostly been addressed at the implementation stage, in more recent treaties it is already covered by the treaty text itself. For instance, the 2013 Minamata Convention on Mercury provides that the COP, when deciding on the timing and format of the reports, shall take into account "the desirability of coordinating reporting with other relevant chemicals and wastes conventions".[262]

Regarding associations with other provisions, non-compliance procedures, international verification, and treaty review provisions are most closely associated, as expected. Also, all but one treaty with national reporting obligations have a secre-

---

[261] See e.g. Convention on Biological Diversity (n 248), art 26 (". . . reports on measures which it has taken for the implementation of the provisions of this Convention and their effectiveness in meeting the objectives of this Convention"), and very similarly, Persistent Organic Pollutants Convention (n 236), art 15(1), and Minamata Convention on Mercury (n 194), art 21(1).

[262] Ibid , art 21(3).

tariat.[263]

3.3.2   International verification

Some treaty regimes have a mechanism for international verification of compliance as a substitute for, or complement to, national reporting. Multiple subtypes can be distinguished based on the authority, location, timing, communication, and method of verification.

Perhaps the most important distinction to draw is between remote and on-site verification, on-site being any location, facility, vessel, spacecraft or installation under the jurisdiction or control of a state party, or any area beyond national jurisdiction. Oon-site inspection is not inherently superior to remote verification. For some problem domains high quality data can reliably be gathered remotely, without requiring access to the territory or property of parties to be verified. Still, there are matters where only on-site verification is able to generate sufficient trust and cooperation. Treaties governing weapons of mass destruction and their precursors fall into this category. While the specifics are unique to each regime, important factors to consider when designing such systems include (i) inspection authority, typically either a treaty body or other states parties, (ii) announced vs unannounced visits, and (iii) verification triggered by an alleged breach on the one hand, and periodic or continuous monitoring on the other.

As for remote verification, there are all kinds of configurations, but the choice between ex ante and ex post verification may be the most influential in practice. The two are not mutually exclusive, a continuous or periodic monitoring system can and should be complemented by a complaint-triggered verification mechanism, but only a proactive approach can be preventive. Of course, in an ex post verification regime, states parties may set up their own monitoring system, so as to trigger the treaty-based procedure before any serious damage occurs. This kind of system may be aided by not requiring conclusive evidence that a breach has in fact occurred, exemplified by the more cautious formulation of Article V(3) of the ENMOD Convention[264] as compared to the nearly identical but more stringent Article VI(1) of the Biological Weapons Convention[265] adopted five years earlier.[266] In this kind of

---

[263]   The exception is Geneva Conventions Additional Protocol I (n 245), art 84. The depositary government receives and distributes national reports, which does not count as an international secretariat, because it is a government.

[264]   Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques (adopted 10 December 1976, entered into force 5 October 1978) 1108 UNTS 151, art V(3).

[265]   Biological Weapons Convention (n 210), art VI(1).

[266]   Specifically, a party which "has reason to believe" (art V(3) ENMOD) vs "finds" (art VI(1)

scenario establish a national monitoring system anyway, the difference with regard to international monitoring and (preventive) verification is mainly one of institutionalisation and burden-sharing. Still, coverage, fairness, transparency and quality of international monitoring may be more effective than a patchwork of national monitoring with potentially politicised triggering of international investigations. The simplest version of international verification may be to require sufficiently detailed national reports combined with a mandate for a treaty organ to routinely cross-check the information contained in these reports, e.g. comparing trade records of exporters with those of corresponding importers. Given that reporting compliance can be patchy, such a mandate ought to extend to consultation of external data sources, such as other international agencies, commercial or academic databases.

International verification regimes are often the first step of a non-compliance procedure or a pre-condition for imposing sanctions, but the focus of verification is the assessment of compliance, whatever the consequences of a finding of non-compliance. In fact, some NCP provisions eschew all mention of verification due to its inquisitive and potentially adversarial aspect, and instead frame the mechanism entirely in terms of promoting compliance and providing assistance.[267] For instance, treaties containing an NCP provision aiming to promote compliance and "address cases of non-compliance" without otherwise mentioning anything resembling international verification, such as the protocols to the CBD,[268] are not deemed to have an implied verification provision. Similarly, agreements suspending parties' voting rights and benefits in case of delays in payment of membership fees or submission of national reports do not *ipso facto* establish a verification system, even though technically international verification of a treaty obligation (to pay fees or submit reports) is implied. Also, judicial dispute settlement typically entails the establishment of facts, but even compulsory adjudication provisions do not count as `verifProv` in the context of this analysis unless they explicitly mention international fact-finding such as in-situ investigations.

Even with these restrictions of the scope of the `verifProv` variable, there is still a sizeable number of treaty provisions meeting the criteria, namely 63, representing a little over half of the agreements in the dataset. Unsurprisingly, the closest

---

Biological Weapons Convention) that another party is acting in breach of treaty obligations may lodge a complaint with the UNSC to request an international investigation.

[267] E.g. Paris Agreement (n 164), art 15 ("A mechanism to facilitate implementation of and promote compliance with the provisions of this Agreement... that shall be expert-based and facilitative in nature and function in a manner that is transparent, non-adversarial and non-punitive"). This is not classified as a verification provision.

[268] Cartagena Protocol on Biosafety to the Convention on Biological Diversity (adopted 29 January 2000, entered into force 11 September 2003) 2226 UNTS 208, art 34, and CBD-ABSprot, art 24.

associations of this variable are with NCP and sanctions provisions, as shown in
Figure 3.4 above. There are no noteworthy differences in prevalence of treaty life
cycle and organs provisions, except for a 1.4 times longer withdrawal waiting period
and a 1.3 times higher prevalence of special denunciation conditions (9 percentage
points increase), as well as special termination provisions being more common in
this subset (all automatically expiring treaties have verification provisions). Two
variables discussed in the next chapter really stand out in terms of their associa-
tion with verification provisions, and for good reason. All but one of the 13 treaties
mentioning weapons of mass destruction have an international verification provision,
representing a 1.7 times higher prevalence compared to the average. The second one
captures mentions of 'inspection' and its synonyms in agreement texts, and only
seven of those treaties do not have a verification provision as defined in this section
(because some treaties provide for domestic inspections without any international
verification). Conversely, 21 treaties have `verifProv` set to `True` without mention of
inspections. 42 treaties have both, corresponding to a Jaccard similarity coefficient
of 60%.[269]

3.3.3 Treaty review

Treaty review provisions (`reviewProv`) as understood here are concerned with imple-
mentation and effectiveness of the treaty as a whole, not as applied by a particular
state party,[270] unlike the other types of provisions in this section. Even the actions
of non-parties may fall under the purview of such a review, given that membership
growth is of paramount importance to some agreements. Typically, treaty review
is undertaken by the COP based on a report by the secretariat or subsidiary or-
gans, sometimes with the input of scientific advisory bodies. For present purposes,
the nature of the body conducting the review is irrelevant. Also, many agreements
do not contain a dedicated article focused solely on periodic review, but instead
include it in an article on the COP,[271] or on amendment/revision.[272] Even those

---

[269] $\frac{42}{42+21+7} = 0.6$.

[270] Otherwise the variable would be too similar to national reporting and non-compliance proce-
dures.

[271] E.g. Ozone Layer Protection Convention (n 196), art 6(4) ("The Conference of the Par-
ties shall keep under continuous review the implementation of this Convention"). A very
similar formulation is used in Basel Hazardous Wastes Convention (n 195), art 15(5), and
in Rotterdam Convention on the Prior Informed Consent Procedure for Certain Hazardous
Chemicals and Pesticides in International Trade (adopted 10 September 1998, entered into
force 24 February 2004) 2244 UNTS 337, art 18(5).

[272] E.g. ILO conventions all contain a provision like art 42 of the Convention (No 169) concerning
Indigenous and Tribal Peoples in Independent Countries (adopted 27 June 1989, entered into
force 5 September 1991) 1650 UNTS 384 ("At such times as it may consider necessary the

that do have dedicated articles are harder to capture accurately, when compared to provisions with more standardised language like entry into force or dispute settlement. 'Review', 'evaluation' and 'assessment' occur frequently in the heading and body of treaty review articles, but also in provisions on risk or safety assessment. Positive and negative lookarounds proved essential for eliminating false positives in the pattern matching process. In terms of fomulations, the net is deliberately cast wide, on the assumption that tracking grogress and evaluating effectiveness may be a more recent focus. What matters for the `reviewProv` variable is the meaning behind the words. Hence, a review of the 'working'[273] or 'operation'[274] of a treaty is deemed to be equivalent to a review of implementation or effectiveness. A proper 'review' or 'evaluation' need not be stipulated either, calling on the COP to "discuss the implementation" of the convention[275] or convening meetings to "consider general problems concerning the application" of the treaty system[276] is sufficient. This is where it is sometimes hard to draw a line between `reviewProv` and provisions on national reports or non-compliance procedures. For example, the Nuclear Safety Convention contains a provision on 'Review meetings',[277] which are held explicitly "for the purpose of reviewing [national] reports" (para 1),[278] "reviewing specific subjects contained in the reports" (para 2), "to discuss the reports submitted by other Contracting Parties and to seek clarification of such reports" (para 3). This is a borderline case, but is not classified as a `reviewProv` for present purposes,[279] because it is focused on the application of the treaty by each party rather than taking a macro perspective and evaluating the treaty as a whole. Perhaps there should be a superclass on review of treaty implementation or effectiveness, under which `reviewProv` would fall (e.g. changing the label to 'general review provision'), with another class comprising specific/national review provisions.

73 (61%) of the agreements in the study set have one or more provisions meeting

---

Governing Body of the International Labour Office shall present to the General Conference a report on the working of this Convention and shall examine the desirability of placing on the agenda of the Conference the question of its revision in whole or in part"). A "report on the working of this Convention" is taken to be the output of a review process.

[273] Ibid.

[274] E.g. Nuclear Non-Proliferation Treaty (n 215), art VIII(3) ("Five years after the entry into force of this Treaty, a conference of Parties to the Treaty shall be held in Geneva, Switzerland, in order to review the operation of this Treaty with a view to assuring that the purposes of the Preamble and the provisions of the Treaty are being realised…"), and nearly identically, WMD Seabed Emplacement Prohibition Treaty (n 161), art VII.

[275] Ramsar Wetlands Convention (n 77), art 6(2)(a).

[276] Geneva Conventions Additional Protocol I (n 245), art 7.

[277] Nuclear Safety Convention (n 236), art 20.

[278] These are reports by each party "on the measures it has taken to implement" the convention, pursuant to its art 5.

[279] Nor is it classified as a `verifProv` or `ncpProv`, as it does not quite meet the conditions for those variables either.

`reviewProv` conditions. Additional factors like the periodicity of reviews are not extracted, only the text of the provision. Unsurprisingly, there is an over 50% overlap with all the provisions of this cluster except for sanctions, as mentioned in the introduction (see Figure 3.4). COP provisions also occur in all but four of the treaties having a review provision (95%), representing a 23 percentage point increase with regard to the full set average. In two of these four exceptions, a non-plenary organ conducts the review, but the other two could also potentially be classified as COP provisions.[280]

To follow up on the impression that review provisions may be a more recent phenomenon, Figure 3.5 displays the number of treaties adopted with review provisions vs. the number adopted without review provisions per year, grouped into decades. It shows that since the 1970s, treaties with review provisions consistently outnumber those without such provisions by a large margin.
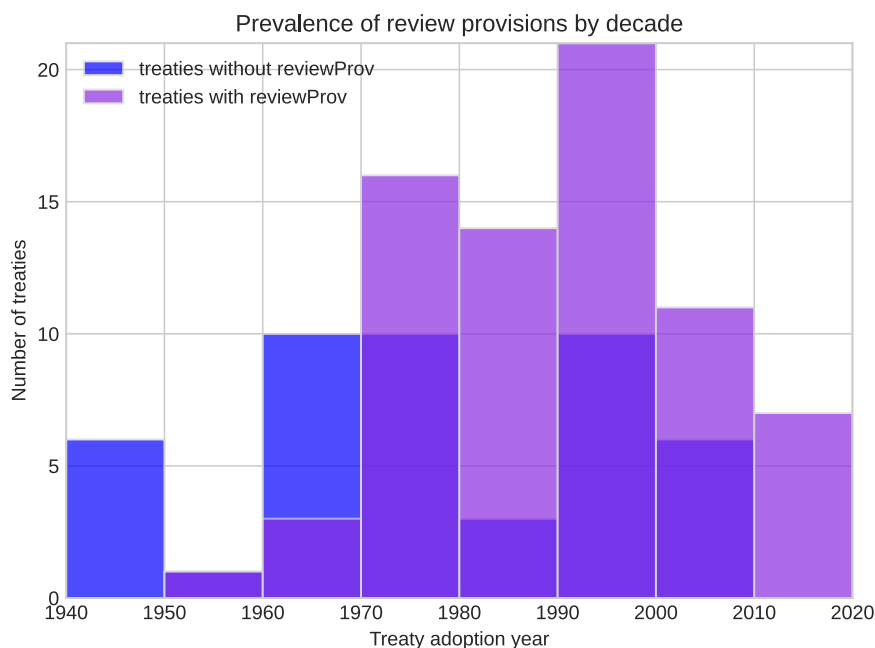
Figure 3.5: Inclusion of review provisions over time

Organisational auspices also add a potential explanatory factor. For instance, all the agreements adopted under the auspices of UNEP include a rather detailed

---

[280] International Convention on Oil Pollution Preparedness, Response and Cooperation (adopted 30 November 1990, entered into force 13 May 1995) ATS 1995 12, art 13 entitled 'Evaluation of the Convention' ("Parties shall evaluate within the Organization the effectiveness of the Convention in the light of its objectives, particularly with respect to the principles underlying co-operation and assistance"), and Protocol on Preparedness, Response and Cooperation to Pollution Incidents by Hazardous and Noxious Substances (adopted 15 March 2000, entered into force 14 June 2007) ATS 2007 41, art 11 likewise on 'Evaluation of the Protocol'. These were not classified as COP provisions since it could be a subsidiary body that is tasked with the review within the IMO.

implementation and effectiveness review provision. If these treaty drafting trends continue, review provisions will remain important for multilateral environmental agreements and may well be further fleshed out to establish a clear legal basis. Future work could extract information on frequency, delegation, participation, transparency, inputs, scope and purpose of review, in order to analyse patterns and provide treaty negotiators more samples to draw from based on lessons learned.

3.3.4   Non-compliance procedures

Non-compliance procedures (NCPs) are designed to deal with implementation and compliance difficulties in a constructive, non-litigious way, focusing on how to resolve such difficulties and bring parties back into compliance as swiftly as possible. This may involve the delivery of assistance of some sort, in particular in the case of disadvantaged states. The agreement may itself contain rules on dealing with non-compliance, or it may contain a mandate for a treaty organ to adopt them subsequently. NCPs can be triggered by a number of actors, depending on the treaty at hand, including the non-compliant party itself (e.g. asking for assistance), other states parties, treaty bodies, and less commonly, civil society actors such as non-governmental organisations (NGOs).[281] The Cartagena Protocol on Biosafety may serve to illustrate this facilitative approach:[282]

> Article 34 Compliance
>
> The Conference of the Parties serving as the meeting of the Parties to this Protocol shall, at its first meeting, consider and approve cooperative procedures and institutional mechanisms to promote compliance with the provisions of this Protocol and to address cases of non-compliance. These procedures and mechanisms shall include provisions to offer advice or assistance, where appropriate. They shall be separate from, and without prejudice to, the dispute settlement procedures and mechanisms established by Article 27 of the Convention.

International environmental lawyers typically consider the Montreal Protocol to be the first MEA to include a provision on NCPs,[283] and see these mechanisms as

---

[281]   See F. R. Jacur, 'Triggering Non-Compliance Procedures', in Non-Compliance Procedures and Mechanisms and the Effectiveness of International Environmental Agreements, pp. 373–87.

[282]   CBD Biosafety Protocol (n 268), art 34.

[283]   Protocol on Substances that Deplete the Ozone Layer (n 176), art 8 ("The Parties, at their first meeting, shall consider and approve procedures and institutional mechanisms for determining non-compliance with the provisions of this Protocol and for treatment of Parties found to be in non-compliance").

a major development of and for this area of international law.[284] The concept of compliance itself has also been developed in various directions, e.g. as a continuum or process. For present purposes, however, a more traditional view will be taken, according to which compliance describes a state of conformity with a treaty, and non-compliance is equivalent to a breach of treaty. It is understandable that slightly less strong language than 'breach' or 'violation' may be used when a state party e.g. fails to submit an annual report. Non-compliance with a reporting obligation would almost certainly not amount to a 'material breach' as defined in Article 60(3) VCLT, except if timely reporting is "essential to the accomplishment of the object or purpose of the treaty".[285] However, it is still a breach of an international obligation. Consequently, for the `ncpProv` variable, any denomination of non-compliance is acceptable, and there are indeed many ways to refer to such problems.

In the study set, 66 (55%) of agreements match the `ncpProv` pattern, 24 of which were adopted before the Montreal Protocol. This discrepancy with the perceived novelty of the mechanism may be due to the broader conception of agreements governing environmental problems explained in Chapter 2.2, or with the broader conception of NCPs for the purpose of the present analysis. The two are not unrelated. A fair comparison across a more diverse set of treaties calls for consideration of different types of obligations, contexts, constraints and objectives. In the following, I explain this conception and some of its difficulties.

Whether the procedure is triggered ex ante (in anticipation of a violation) or ex post (as a complaint or report about an alleged breach) is irrelevant for present purposes. Which treaty organ is involved is not taken into account for classification, as long as there is one. That is, there has to be a multilateral aspect to count as an NCP, purely bilateral procedures are insufficient. Essentially, the question is whether the collective interest in full implementation and compliance is reflected in an institutional mechanism that can be seized to deal with challenges thereto. Logically, it would seem reasonable to start with international verification of compliance (`verifProv`), followed by a cooperative, facilitative attempt to remedy any situation of non-compliance, and if this fails, or if the breach is grave, to provide for sanctions. However, some treaties only have an NCP provision like Article 34 cited above, without proper verification and sanctions provisions. One could argue that a light version of verification and sanctions is implied in a procedure that aims to

---

[284] E.g. Philippe Sands, *Principles of International Environmental Law* (Second Edition, Cambridge University Press 2003), p.205f ("One of the most significant developments in the field of international environmental law has been the emergence of non-compliance procedures under various multilateral environmental agreements... The first was the non-compliance procedure established under the 1987 Montreal Protocol").

[285] VCLT (n 18), art 60(3)(b).

"address cases of non-compliance", but for present purposes this kind of provision only counts as an `ncpProv`, not containing a latent `verifProv` and `sanctionsProv`.[286] 'Advice or assistance' is hardly a negative consequence.

The multilateral process can be voluntary, based on ad hoc consent. In these cases, it is sometimes difficult to distinguish between provisions on diplomatic means of (bilateral) dispute resolution and provisions on a voluntary (institutional) NCP. For example, in the Biological Weapons Convention states parties "undertake to consult one another and to cooperate in solving any problems which may arise in relation to the objective of, or in the application of the provisions of, the Convention. Consultation and cooperation pursuant to this Article may also be undertaken through appropriate international procedures within the framework of the United Nations and in accordance with its Charter".[287] This provision has the cooperative and constructive side of NCPs, but could be seen as mostly bilateral, because UN procedures are optional in this case ('may also be undertaken...'). The provision that follows is about complaints to the UN Security Council about breaches of the Convention, and thus clearly institutionalised, but also more confrontational.[288] Still, taken together and considering the mandate of the UNSC under the Charter, it does seem appropriate to classify this Convention as having an `ncpProv` (as well as a `verifProv` and `sanctionsProv`). By contrast, the treaty prohibiting the emplacement of weapons of mass destruction on the seabed and ocean floor has a provision on consultations and cooperation in case of doubts concerning the fulfilment of obligations, but it is clearly about international verification of compliance.[289] There is no sense that a nuclear weapon might be misplaced on the international seabed and some friendly advice and assistance could bring a state party back into compliance. There are bilateral and multilateral verification procedures (`verifProv`), followed by referral to the UN Security Council, "which may take action in accordance with the Charter"[290] (`sanctionsProv`), without an NCP in-between (and for good reason).

Note that in this conception of NCPs, the body conducting or facilitating such procedures could be established by a different treaty, e.g. by the constituent instrument of the international organisation under whose auspices a given agreement is adopted. For instance, the ILO Constitution sets out a procedure under which an industrial association of employers or of workers can make a 'representation' to the

---

286   Such provisions are not found elsewhere in the protocols to the CBD either, e.g. CBD Biosafety Protocol (n 268), art 35 on 'Assessment and review' is a typical `reviewProv` calling for an 'evaluation of the effectiveness of the Protocol' every five years.

287   Biological Weapons Convention (n 210), art V.

288   ibid, art VI.

289   WMD Seabed Emplacement Prohibition Treaty (n 161), art III.

290   ibid, art III(4).

ILO "that any of the Members has failed to secure in any respect the effective observance within its jurisdiction of any Convention to which it is a party" (art 24(1)), and a member can likewise file a 'complaint' to the ILO "if it is not satisfied that any other Member is securing the effective observance of any Convention which both have ratified" (art 26(1)). Complaints can be referred to a Commission of Inquiry which establishes the facts and makes recommendations. This is usually called a supervisory procedure but bears enough similarity to non-compliance procedures to be in the same category. As with other variables, the fact that these NCP provisions are in the Constitution, without repetition in individual Conventions, does not matter. The 13 ILO Conventions in the study set are thus classified as having an `ncpProv` without a corresponding excerpt from the text of the agreement. The Supplementary Protocol on Liability and Redress to the above-mentioned Biosafety Protocol is similarly classified as having an `ncpProv` without actually mentioning (non-)compliance, based on a generic incorporation clause of the provisions of the Convention and Protocol.[291]

As for the kind of multilateral involvement required, at the very least treaty organs should be apprised of any situation of potential non-compliance so that they may consider taking action. For example, under Article IX(4) of the 1948 Whaling Convention, each state party is obliged to "transmit to the [International Whaling] Commission full details of each infraction of the provisions of this Convention by persons or vessels under the jurisdiction of that Government as reported by its inspectors. This information shall include a statement of measures taken for dealing with the infraction and of penalties imposed". The Commission is also authorised to "make recommendations to any or all Contracting Governments on any matters which relate to whales or whaling and to the objectives and purposes of [the] Convention".[292] This advisory role of the COP is common in NCPs. While in this case (and others) the two provisions (on communication of violations and on recommendations to parties) are not in the same Article, nor labelled as a '(non-)compliance procedure', the core elements are still there. It makes sense for parties to exchange information about difficulties they face with national implementation and enforcement, and share advice on how to overcome them. In maritime law, both flag and port state enforcement actions and their outcomes typically need to be reported to the IMO. Some provisions on casualty investigations specify that while findings are

---

[291] Nagoya - Kuala Lumpur Supplementary Protocol on Liability and Redress to the Cartagena Protocol on Biosafety (adopted 15 October 2010, entered into force 5 March 2018) 50 ILM 108, art 16(3) ("Except as otherwise provided in this Supplementary Protocol, the provisions of the Convention and the Protocol shall apply, mutatis mutandis, to this Supplementary Protocol").

[292] Whaling Convention (n 161), art VI.

circulated to all parties, "[no] reports or recommendations of the Organization based upon such information shall disclose the identity or nationality of the vessels concerned, or in any manner fix or imply responsibility upon any vessel or person".[293] This is NCP style avoidance of entering the terrain of legal responsibility for a breach, presumably for the purpose of encouraging accurate information-sharing among parties. Only issuing 'reports or recommendations' in response to findings of violations is also typical of NCPs, although many treaties do provide for more coercive measures if the intially lenient approach fails. Specifically, 67% of treaties with NCPs also have at least one `sanctionsProv`, and 74% have dispute settlement provisions, but the largest overlap is between NCP and verification provisions (70% Jaccard similarity).

With regard to the popularity of NCP provisions over time, Figure 3.6 shows a similar chart to the one in the previous section. The trend is much like that of review provisions, with the noteable exception of the 1980s where review provisions occur in 82% of newly concluded agreements, whereas NCP provisions are only included in 47% of them (n=17). Apart from that decade, NCP provisions occur in well over 50% of new treaties adopted per decade since the 1970s, and in the 2010s even reaching 100%.
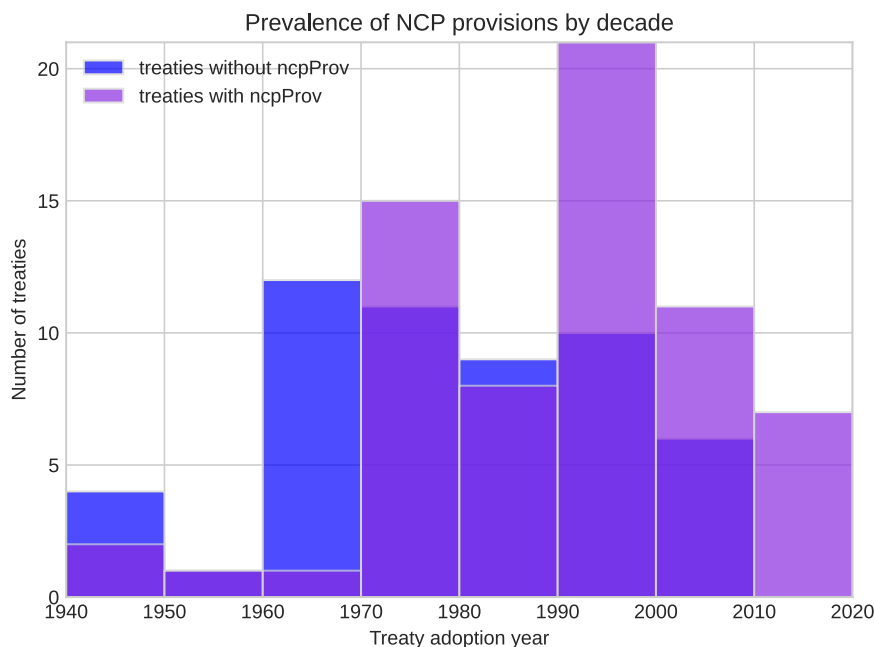
Figure 3.6: Inclusion of NCP provisions over time

---

[293]   E.g. International Convention for the Safety of Fishing Vessels, 1977 (adopted 2 April 1977), art 7(2), on investigations of casualties occurring on fishing vessels under a party's jurisdiction.

3.3.5 Sanctions

Sanctions are measures which penalise non-compliance with treaty obligations. They range from temporary suspension of voting rights in a treaty body to the use of force authorised by the UN Security Council under Chapter VII of the UN Charter. As mentioned in the introduction to this section, the term 'sanctions' may not appear in the treaty, as long as there is a provision on negative consequences that may or must imposed upon non-performance. Mere publication of a breach ('naming and shaming') does not count for present purposes.[294] The entity or person subject to sanctions does not matter, so long as there is an international element. Thus, provisions on sanctions to be imposed by states parties on their own nationals, such as provisions on flag state enforcement measures,[295] are disregarded, while provisions on port state enforcement measures[296] count as `sanctionsProv`. The aim may not necessarily be punitive, it could simply be to remove any danger to persons, property or the environment. Indeed, international maritime safety control measures are ostensibly about removing threats,[297] and under some treaties, ships are entitled to compensation for any loss or damage resulting from undue detention or delays.[298] Provisions on such control measures still count as `sanctionsProv`, because detention and delays due to non-compliance certainly are negative consequences.

Sanctions provisions in treaties may serve as a deterrent, fairness tool, and as a way to bring an infringing party back to compliance. Treaty negotiators are not oblivious to the fact that imposing sanctions can be counterproductive, leading a party to withdraw from the regime altogether. A number of sanctions provisions allow for an exception to the rule if non-compliance is due to circumstances beyond

---

[294] E.g. Brown Weiss classifies publication of infractions as coercive measures, see Weiss (n 10), p.1584f.

[295] E.g. Fishing Vessels Compliance Promotion Agreement (n 161), art III(8) ("Each Party shall take enforcement measures in respect of fishing vessels entitled to fly its flag which act in contravention of the provisions of this Agreement, including, where appropriate, making the contravention of such provisions an offence under national legislation. Sanctions applicable in respect of such contraventions shall be of sufficient gravity as to be effective in securing compliance with the requirements of this Agreement and to deprive offenders of the benefits accruing from their illegal activities. Such sanctions shall, for serious offences, include refusal, suspension or withdrawal of the authorization to fish on the high seas").

[296] E.g. Agreement on Port State Measures to Prevent, Deter and Eliminate Illegal, Unreported and Unregulated Fishing (adopted 22 November 2009, entered into force 5 June 2016) ATS 2016 21, art 9.

[297] E.g. MARPOL73 (n 75), art 5(2) ("...the Party carrying out the inspection shall take such steps as will ensure that the ship shall not sail until it can proceed to sea without presenting an unreasonable threat of harm to the marine environment").

[298] E.g. International Convention on Standards of Training, Certification and Watchkeeping for Fishing Vessel Personnel (adopted 7 July 1995, entered into force 29 September 2012) 3357 UNTS, art 8(3)(1). This is also a `sanctionsProv`, because "all possible efforts shall be made to avoid a vessel being unduly detained or delayed".

the party's control, as determined by a treaty organ such as the COP.[299] In some agreements there is also a clear aim to make sanctions proportional to the breach, taking into account its severity and frequency or duration in the case at hand.[300] The details of the sanctions regime can be left for the COP to flesh out, but mere mention that procedures will be developed for 'treatment of Parties found to be in non-compliance' are deemed insufficient given that NCP provisions can be purely facilitative.[301]

What kind of obligation gives rise to sanctions when breached does not matter for the purpose of the `sanctionsProv` variable. It could 'only' be the payment of financial contributions by a party or person under an obligation to pay. This could seem to make this category too broad, but considering that some of the agreements in the set are funds or trusts whose resources are the primary impact vector, this breadth is arguably justified. It may also be that the label 'sanctions' is not entirely adequate, given the wide range of negative consequences captured by this variable. The underlying question that motivated the classification is whether these treaties have 'teeth' of some kind at the international level, and naturally what 'teeth' are useful for an agreement depends on its objective, subject matter and legal context. Nearly half of the agreements in the study set have teeth in this sense (n=58, i.e. 49%). The variable is closely related with `verifProv`, `ncpProv` and `legalPers`. Out of the subject matter variables covered in Chapter 4, `sanctionsProv` is most strongly associated with mention of inspections, not civil liability or criminal responsibility.

Regarding temporal trends, `sanctionsProv` is unique in that the last decade covered has no agreements with this type of provision (out of a total of seven treaties adopted in this period and included in the present dataset). Figure 3.7 shows the respective numbers over time. Based on this chart, it would seem that the 1970s were the heyday of sanctions provisions, with 65% of agreements adopted in that decade including such provisions. However, this is could just be a coincidence due to the treaties adopted in that decade. Overall, there does not seem to be an obvious trend in popularity over time, as compared to the increased prevalence of treaty review and NCP provisions since the ~1970s.

---

[299] E.g. IRENA Statute (n 189), art XVII(A).

[300] E.g. Kyoto Protocol to the United Nations Framework Convention on Climate Change (adopted 11 December 1997, entered into force 16 February 2005) 2303 UNTS 162, art 18 ("...development of an indicative list of consequences, taking into account the cause, type, degree and frequency of non-compliance").

[301] This formulation is used in Protocol on Substances that Deplete the Ozone Layer (n 176), art 8; PIC Convention (n 271), art 17; and Persistent Organic Pollutants Convention (n 236), art 17. All of these are classified only as `ncpProv`, not `sanctionsProv` as well.
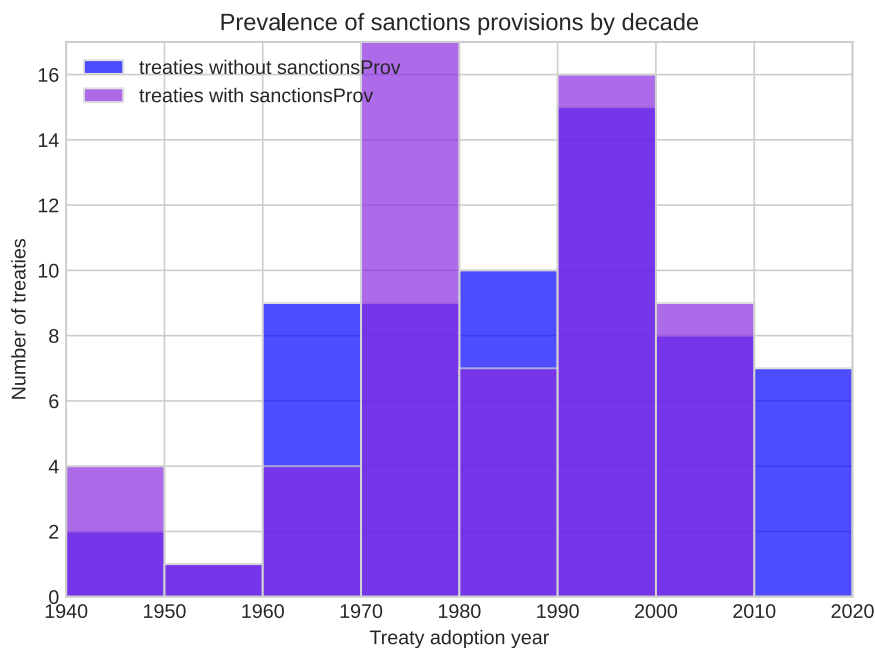
Figure 3.7: Inclusion of sanctions provisions over time

3.3.6 Dispute settlement

Dispute settlement provisions are very common in multilateral agreements and increase resilience of the system in that they clarify in advance the procedure to follow in case of a dispute over interpretation or application of the treaty rather than falling back on general international law. Similar to NCPs, the objective is often first and foremost to prevent litigation by encouraging or mandating recourse to amicable mechanisms such as negotiation, mediation or conciliation. More elaborate versions of this technique stipulate how long such mechanisms should be pursued before initiating judicial or arbitral dispute settlement, or who shall be the provider of such services. Due to the mostly voluntary and non-binding nature of these amicable mechanisms, no information is extracted on them. Only provisions which provide for international arbitration or judicial dispute settlement count as `disputeProv` for the purpose of the present analysis. 74 (62%) of the agreements in the dataset are found to have such provisions (within the agreement itself or by incorporation). Provisions which do not themselves provide for arbitration or adjudication, but task a future COP to "consider procedures for the settlement of disputes concerning the interpretation and application of [the] Convention"[302] or suchlike are not captured as `disputeProv`. Procedural details can be decided later, but the principle itself must be clear.

---

[302]   Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter, 1972 (adopted 29 December 1972, entered into force 30 August 1975) 1046 UNTS 120, art XI.

Perhaps the most important distinguishing factor of dispute settlement provisions occurring in international treaties is whether they provide for some form of compulsory adjudication or arbitration of any dispute arising out of the treaty, i.e. embodying or encouraging ex ante consent to a binding dispute settlement mechanism. There are three variations of this consent that are extracted from dispute settlement provisions found in the study set, namely opt-in, opt-out, and compulsory consent. Opt-in consent is exemplified by Article 27(3) CBD:[303]

> Article 27 Settlement of disputes
>
> **3.** When ratifying, accepting, approving or acceding to this Convention, or at any time thereafter, a State or regional economic integration organization may declare in writing to the Depositary that for a dispute not resolved in accordance with paragraph 1 or paragraph 2 above, it accepts one or both of the following means of dispute settlement as compulsory:
>
> > **(a)** Arbitration in accordance with the procedure laid down in Part 1 of Annex II;
> >
> > **(b)** Submission of the dispute to the International Court of Justice.
>
> **4.** If the parties to the dispute have not, in accordance with paragraph 3 above, accepted the same or any procedure, the dispute shall be submitted to conciliation in accordance with Part 2 of Annex II unless the parties otherwise agree.

The same solution was chosen for UNFCCC,[304] adopted at the same summit. Its supplementary agreements, the Kyoto Protocol and the Paris Agreement, incorporate the Convention's provision *mutatis mutandis*.[305] The Vienna Ozone Layer Protection Convention, after which these provisions are modelled, includes a similar formulation.[306] The Hazardous Wastes Convention,[307] UN Watercourses Convention,[308] and the Plant Genetic Resources Treaty[309] also provide for this kind of opt-in mechanism. More detailed versions specify conditions of revocation and expiry of such declarations. A total of 19 treaties are found to have this subtype of dispute settlement provision.

A *prima facie* similar design pattern, but with the default switched, is opt-out consent, found in five of the eight treaties adopted under the auspices of the IAEA,

---

[303]   Convention on Biological Diversity (n 248), art 27(3).

[304]   Framework Convention on Climate Change (n 82), art 14.

[305]   Kyoto Protocol (n 300), art 19; Paris Agreement (n 164), art 24.

[306]   Ozone Layer Protection Convention (n 196), art 11.

[307]   Basel Hazardous Wastes Convention (n 195) , art 20.

[308]   Watercourses Convention (n 175), art 33.

[309]   Plant Genetic Resources Treaty (n 224), art 22.

and three treaties adopted under UN auspices.[310] The Nuclear Accident Notification Convention illustrates the principle:[311]

> Article 11 Settlement of disputes
>
> **2.** If a dispute of this character between States Parties cannot be settled within one year from the request for consultation pursuant to paragraph 1, it shall, at the request of any party to such dispute, be submitted to arbitration or referred to the International Court of Justice for decision. Where a dispute is submitted to arbitration, if, within six months from the date of the request, the parties to the dispute are unable to agree on the organization of the arbitration, a party may request the President of the International Court of Justice or the Secretary-General of the United Nations to appoint one or more arbitrators. In cases of conflicting requests by the parties to the dispute, the request to the Secretary-General of the United Nations shall have priority.
>
> **3.** When signing, ratifying, accepting, approving or acceding to this Convention, a State may declare that it does not consider itself bound by either or both of the dispute settlement procedures provided for in paragraph 2. The other States Parties shall not be bound by a dispute settlement procedure provided for in paragraph 2 with respect to a State Party for which such a declaration is in force.
>
> **4.** A State Party which has made a declaration in accordance with paragraph 3 may at any time withdraw it by notification to the depositary.

A variation of this would be to omit any mention of opting out in the dispute settlement provision, but to allow reservations to the treaty (without excluding dispute settlement provisions from the scope of admissible reservations). Apart from the default being reversed, which likely makes a significant difference in practice, another crucial distinction is that in this case the decision to opt out of compulsory dispute settlement can only be made upon signature, ratification or accession. A declaration (or other reservation) under this kind of provision can be *withdrawn* at a later stage, but cannot be *lodged* subsequently.

Finally, compulsory dispute settlement *stricto sensu* does not allow opting out in any way other than treaty denunciation. UNCLOS and the WTO are well-known examples of this approach, but a dozen other treaties also seem to fit this pattern.

---

[310] Vienna Road Traffic Convention (n 214), art 52; Terrorist Bombings Convention (n 175), art 20; and International Convention for the Suppression of Acts of Nuclear Terrorism (adopted 13 April 2005, entered into force 7 July 2007) 2445 UNTS 89, art 23.

[311] Nuclear Accident Notification Convention (n 149), art 11.

Regarding associations with other variables, there is no sense that NCP provisions displace dispute settlement provisions. On the contrary, they co-occur more frequently than e.g. sanctions and dispute settlement provisions. As for prevalence over time, dispute settlement provisions are and have remained a relatively constant feature of international agreements, similar to national reporting provisions. The decades with the highest relative prevalence of `disputeProv` are the 1940s and the 2000s, occurring in over 80% of agreements adopted during that time period. The average per decade is 64% with a mean absolute deviation of 16%, i.e. a relatively narrow band.

4   Treaty subject matter analysis

This chapter covers what are typically considered the 'substantive' provisions and the scope *ratione materiae*. Due to the number and diversity of treaties in the study set, text analysis for this chapter is limited to mere mentions of terms and expressions, instead of extracting legal information such as whether a given activity is prohibited or not. While legal information extraction from substantive provisions was initially planned and trialled, it proved too time-consuming for a single research project. Even for a larger collaborative project it is questionable whether the effort would be worth it. Accurately modelling specific treaty obligations incumbent on parties would require not only collecting information from treaty texts, but also reservations lodged, objected to, and withdrawn by parties, as well as decisions by treaty organs on interpretation and application of exceptions and exemptions. Incorporating better semantic search functionalities into public databases would likely benefit a lot more researchers and practitioners than necessarily incomplete attempts at fully formalising individual agreements.

The aim of the present chapter is more modest. In essence, it is to provide an overview of the subject matter covered by the treaty corpus, especially aspects relevant to environmental problem-solving. The constructed variables should lend themselves to analysis of treaty design, including associations between treaty features, predictive analysis of outcome variables, and should be suitable for a wide range of structured database queries.

Apart from this specific aim and overarching project goals and criteria outlined in Chapter 2.3, three additional observations motivate the methodology underlying this chapter. First, subject matter classifications published by organisations and academics tend to be relatively high-level, often not going beyond what could be gleaned from the treaty title. Furthermore, these abstract labels may be limited by a narrow conception of topics and branches of international law rather than taking treaty texts as evidence (see Chapter 2.2.3 on environmental focus and relevance). Second, labels or tags tend to be binary variables, not admitting of degrees, with the result that the relative importance of a given issue to a treaty remains unexplored. Moreover, some classification systems allow only one tag per treaty, which is un-

fortunate when agreements straddle domains or cover important side issues. Third, some of the subject terms provided in existing databases appear rather ad hoc and fail to include important topics such as climate change (see section 4.1.1 below).

Hence, the first set of features constructed for this chapter is concerned with global environmental problems and risks (4.1). A cluster mapping environmental spheres (atmosphere, hydrosphere, lithosphere and biosphere) was also trialled but eventually discarded due to verification difficulties and lower priority compared to other variables. Another group of features representing environmental objectives, including prevention, sustainable use, mitigation, and restoration, proved too onerous to capture with sufficient accuracy. Similarly, aspects of environmental governance framing, such as anthropocentrism and longtermism, and references to general principles, declarations, customary law and other treaties did not meet the selection criteria and constraints of the present thesis. However, they should be considered for inclusion in future work. Additional clusters that do satisfy the requirements are human contributions to environmental problems (4.2), including activities such as agriculture and trade, as well as means of intervention (4.3). These two are discussed at a more aggregate level whereas the first one is presented in some detail.

Variables are constructed with a broad meaning in mind, e.g. mentions of the term 'greenhouse gas' in agreement texts are counted as references to climate change, even though technically the former are gases and the latter is a process. Future extensions could subdivide some of the variables, such as pollution into air, water and soil pollution, or distinguishing between different types of greenhouse gases. For the present project it seemed best to tackle higher-level classes in the first instance. Also, the objective was to at least extract *some* of the mentions of a given subject in a given treaty, so that the agreement would appear in a binary search (whether or not, rather than to what extent a topic is covered by a treaty). Thus, the higher end of the term counts should not be seen as definitive, especially in treaties with more than 10'000 words.[312] In contrast with the variables covered in Chapter 3, no exhaustive false negative search was conducted for the variables in this chapter. Steps taken to minimise false negatives and false positives are outlined in subsequent sections where relevant, but generally involve looking at potentially ambiguous matches with a context window (e.g. 'climate change' is not ambiguous, but 'climate' is), and examining excerpts that contain a given keyword but are not matched by the regular expression (such as all instances of 'climate' not already captured). Having crafted a regular expression for each term or topic of interest, the number of occurrences of the regular expression in a given agreement is then

---

[312] The average word count in the treaty set is about 9600, including treaty annexes. The median is about 5100 words.

divided by the number of words in the treaty. This term frequency (tf) is more useful than absolute counts because it facilitates comparison of term/topic importance across treaties, otherwise longer agreements would falsely appear to attribute more importance to a given term. Another commonly used transformation in natural language processing is term frequency - inverse document frequency (tf-idf), which takes into account how many documents a given term occurs in and weighs infrequent terms more highly. However, this would lead to variables changing slightly whenever treaties are added to or removed from the corpus, which is not ideal for an incrementally expanding project and online database. Finally, a binary variable is constructed based on whether a given regex occurs in the treaty or not. Each regular expression thus yields a list of matches, an absolute count, a relative/proportional count, and a boolean transformation. When term frequencies are included in tables hereafter, they are multiplied by 10'000 to facilitate visual comparison of magnitudes. Grouping these term frequencies into three categories representing zero, few and many matches was also attempted, with the aim to construct an ordinal variable that is more expressive than a binary version but also simpler to grasp and remember than a continuous variable. Ideally it would correspond to informal, qualitative classifications of core/primary vs. ancillary/secondary treaties for a given topic, issue, activity or intervention. However, there is no suitable cut-off that can be applied to all treaties and variables in the same way.

Figure 4.1 provides an overview of the number and percentage of treaties containing the relevant search patterns (binary variables), and Figure 4.2 presents the term frequency results for five global environmental treaties, with darker colors indicating higher relative term frequency. At the individual treaty level, there are no major surprises, but at the aggregate level the number of treaties found is larger than I expected in many cases. I would venture a guess that only a few individuals with exceptional memory could spontaneously recite the list of treaties identified through the regex pattern matching process for those terms found in over twenty agreements. That is, someone familiar with the area could probably recite from memory the five treaties mentioning stratospheric ozone depletion, but not the 44 mentioning liability. Indeed, I believe working towards a shared database covering a good range of not only core but also ancillary topics or issues mentioned by treaties could prove useful for scholars and practitioners alike. It may also help to dispel myths about supposedly clear and rigid boundaries between branches or regimes of international law.

To facilitate verification, use and contributions by non-programmers, subject matter pattern matching is also implemented in a GATE application. Figure 4.3

Cluster	Variable	N. Treaties
	Climate Change	9 (8%)
	Ozone Depletion	5 (4%)
	Pollution	70 (59%)
	Waste	34 (29%)
Problems & Risks	Land Degradation	12 (10%)
	—Deforestation	6 (5%)
	—Desertification	6 (5%)
	Endangered Species	12 (10%)
	Pests & Weeds	15 (13%)
	Animal Welfare	6 (5%)
	Fishing & Hunting	25 (21%)
	Agriculture	36 (30%)
	Production	76 (64%)
Human Stressors	Storage & Stockpiling	62 (52%)
	Transport	76 (64%)
	Trade	58 (49%)
	Weapons of Mass Destr.	13 (11%)
	Quant. Restrictions	31 (26%)
	Spatial Regulation	46 (39%)
	Risk/Impact Assessment	47 (39%)
	Inspections	49 (41%)
	Data & Research	99 (83%)
Means of Intervention	Education	48 (40%)
	Training/Capacity-build.	63 (53%)
	Tech Transfer	37 (31%)
	Insurance	27 (23%)
	Liability	44 (37%)
	Criminal Responsibility	16 (13%)

Figure 4.1: Treaty subject matter summary results

shows a screenshot of the graphical user interface centred on a section of the UNCCD (arts 10-12). The key parts of the program are included in Appendix F.2.2 and the full version is available for download from the code archive of this project.

Regarding data analysis, the same analytical tools as in Chapter 3 are applied again, including independent and conditional probabilities, contingency tables, Jaccard similarity, summary statistics and time-indexed charts. An additional measure and test of independence (χ^2) is introduced in section 4.2.2 on stressor analytics to facilitate aggregate analysis of many pairs of variables at once.

4.1 Environmental problems

This cluster is concerned with identifying the environmental harms and risks tackled, or at least recognised, by a treaty. Needless to say, the framing and linking of environmental issues is an important part of treaty negotiation and regime design.

Cluster	Variable	WHC	CITES	CBD	UNFCCC	POPconv
Problems & Risks	Climate Change	0.0	0.0	0.0	102.5	0.0
	Ozone Depletion	0.0	0.0	0.0	2.4	0.0
	Pollution	0.0	0.0	0.0	2.4	53.7
	Waste	0.0	0.0	0.0	1.2	26.1
	Land Degradation	1.9	0.0	2.2	8.3	0.0
	—Deforestation	0.0	0.0	0.0	1.2	0.0
	—Desertification	0.0	0.0	1.1	7.1	0.0
	Endangered Species	1.9	10.8	9.7	0.0	0.0
	Pests & Weeds	0.0	0.0	1.1	0.0	0.0
	Animal Welfare	0.0	11.9	0.0	0.0	0.0
Human Stressors	Fishing & Hunting	0.0	5.4	0.0	0.0	0.0
	Agriculture	0.0	3.2	1.1	3.6	3.0
	Production	0.0	0.0	0.0	4.8	43.2
	Storage & Stockpiling	0.0	0.0	0.0	1.2	7.5
	Transport	0.0	8.6	1.1	1.2	1.5
	Trade	0.0	136.9	0.0	4.8	13.4
	Weapons of Mass Destr.	0.0	0.0	0.0	0.0	0.0
Means of Intervention	Quant. Restrictions	0.0	1.1	0.0	0.0	0.7
	Spatial Regulation	77.8	0.0	5.4	1.2	6.0
	Risk/Impact Assessment	0.0	0.0	3.2	6.0	10.4
	Inspections	0.0	1.1	0.0	0.0	0.7
	Data & Research	17.5	4.3	23.8	27.4	23.8
	Education	3.9	0.0	7.6	10.7	7.5
	Training/Capacity-build.	5.8	0.0	6.5	7.1	3.7
	Tech Transfer	1.9	0.0	10.8	11.9	3.0
	Insurance	0.0	0.0	0.0	1.2	0.0
	Liability	0.0	0.0	2.2	0.0	0.0
	Criminal Responsibility	0.0	0.0	0.0	0.0	0.0

Figure 4.2: Subject matter results for major MEAs

Especially the complex problems of the 21st century require due consideration of the various interplays. Thus, it is not surprising that UNFCCC[313] leads the pack with seven out of ten environmental problems mentioned, closely followed by UNCCD[314] referring to six of them. Figure 4.4 shows the Jaccard similarity between the different (binary) variables, in the same vein as Figure 3.4 in the introduction to section 3.3 above. Unlike those provisions, here many pairs have no or very little overlap.

This section presents the specifics and results of the information extraction process with respect to these ten issue areas. Other environmental problems and risks such as acidification, eutrophication and natural disasters are not sufficiently prominent in the set of treaties under consideration to warrant extraction by regex pattern matching.

---

[313] Framework Convention on Climate Change (n 82).
[314] Convention to Combat Desertification (n 81).
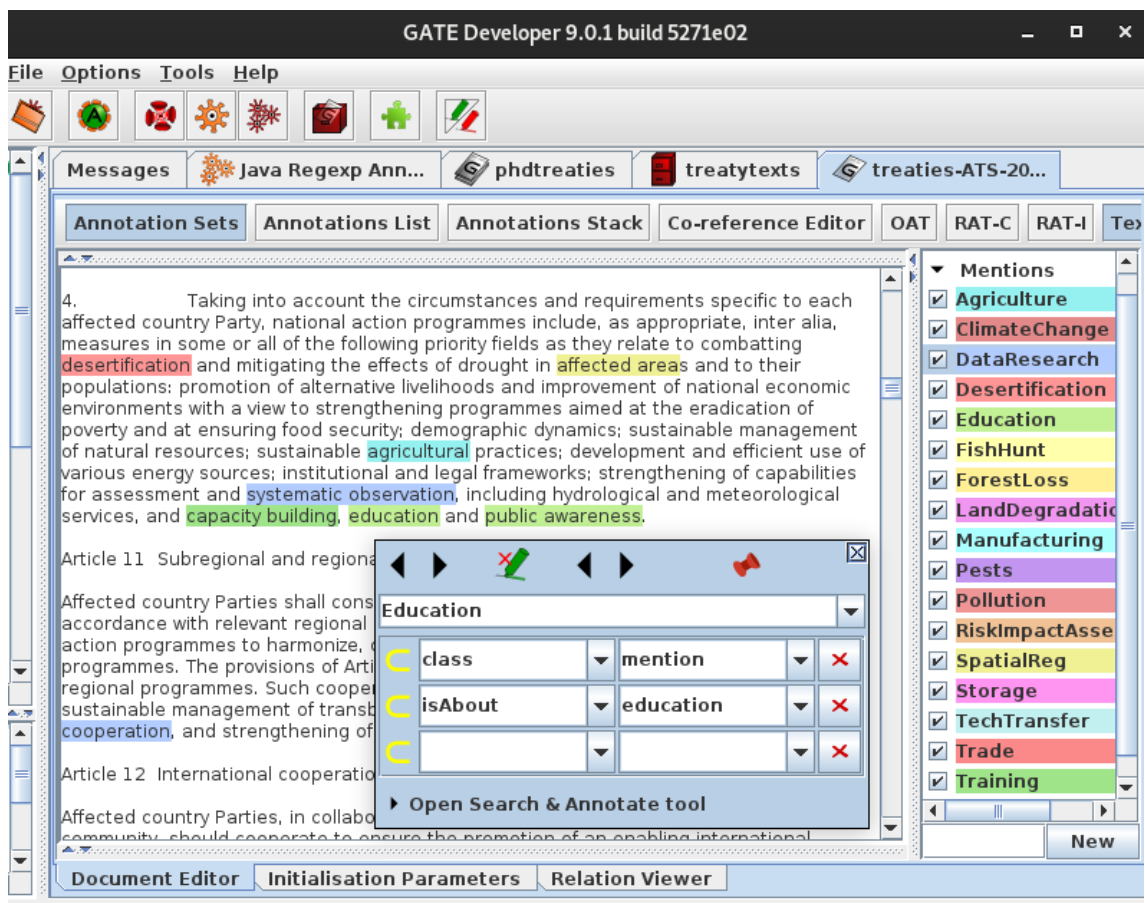
Figure 4.3: Treaty mentions annotations in GATE

	climateChange	ozoneDeplet	pollution	waste	landDegrad	deforest	desertif	endangSpec	pests	animalWelfare
climateChange	1.000	0.273	0.039	0.049	0.400	0.667	0.364	0.000	0.043	0.000
ozoneDeplet	0.273	1.000	0.042	0.114	0.133	0.222	0.100	0.000	0.000	0.000
pollution	0.039	0.042	1.000	0.444	0.051	0.027	0.041	0.065	0.133	0.027
waste	0.049	0.114	0.444	1.000	0.070	0.053	0.026	0.095	0.167	0.026
landDegrad	0.400	0.133	0.051	0.070	1.000	0.500	0.500	0.143	0.174	0.000
deforest	0.667	0.222	0.027	0.053	0.500	1.000	0.500	0.000	0.050	0.000
desertif	0.364	0.100	0.041	0.026	0.500	0.500	1.000	0.059	0.167	0.000
endangSpec	0.000	0.000	0.065	0.095	0.143	0.000	0.059	1.000	0.350	0.286
pests	0.043	0.000	0.133	0.167	0.174	0.050	0.167	0.350	1.000	0.235
animalWelfare	0.000	0.000	0.027	0.026	0.000	0.000	0.000	0.286	0.235	1.000

Figure 4.4: Problems' Jaccard similarity

4.1.1 Climate change

The regular expression tailored to capture references to climate change is as follows:

Table 4.1: Climate change regex pattern

Variable	Regular expression pattern
climateChange	`(?i)(?:climate (?:change|protection)|global warming|greenhouse (?:gas|effect)|(?:impact on|change[s]?  in(?:  the Earth's)?|effects on .{0,30}|stabili\w+ of the|protect the) climate)`

All remaining mentions of 'climate', 'greenhouse', and of specific greenhouse gases like CO2 and methane were analysed to make sure no reference is missed. Out of a total of 158 mentions of 'climate', 116 (73%) occur within the expressions of 'climate change' or 'climate protection'. This means the verification task is greatly sped up thanks to the use of regular expressions as compared to manual annotation. Global warming is denoted three times, the greenhouse effect twice, while 'greenhouse gas' (including its plural) has 69 occurrences.[315] The remaining pattern consists in five alternatives (and their variants) all ending in 'climate'. This subpattern was constructed based on excerpts of the 42 occurrences of 'climate' that were not already captured. The pattern could be made more generally applicable to include variations such as 'protection of the climate' which would not be matched by the extant version (only 'protect the climate' gets matched), but for the current treaty set it is sufficient. The eight excerpts (with context) matched by this subpattern are as follows:

Table 4.2: Climate change subpattern matches

Treaty	Excerpts
IRENA Statute	"the stabilisation of the climate system"
Ozone Layer Convention	"including changes in climate"
"	"the impact on climate parameters"
"	"effects on human health, environment and climate"
UNFCCC	"change in the Earth's climate"
"	"to protect the climate system" (twice)
"	"should protect the climate system"

Note that in the case of the Ozone Layer Convention,[316] these are the only three mentions of climate change or climate impacts, meaning that the treaty does

---

[315] There is also one reference to greenhouses in the original (gardening) sense, hence the need for conditioning the word on being followed by 'gas' or 'effect'.

[316] Ozone Layer Protection Convention (n 196).

not appear in a simple keyword search for 'climate change'. Conversely, a broader full-text search for 'climate' would have yielded false negatives such as the Geneva Conventions' Additional Protocol II[317] and the PIC Convention.[318]

Expressions which have only recently gained popularity such as 'global heating', 'climate emergency' and 'climate crisis' do not occur in the treaty set and are thus not added to the regex pattern. References to the Climate Change Convention (as opposed to climate change per se) are not excluded from the search pattern as doing so would arguably lead to an underinclusion of relevant treaties. For instance, UNCCD calls for coordination between the three Rio Conventions[319] and emphasises synergies between combatting desertification and achieving the objectives of the UNFCCC.[320] These are the only two (indirect) references to climate change by UNCCD. A total of nine treaties match the climate change regex pattern, as displayed in Table 4.3.

Table 4.3: Climate change treaties regex counts

Year	Name	Count	Fraction
1985	Ozone Layer Protection Convention	3	5.1
1992	Framework Convention on Climate Change	86	102.5
1994	Convention to Combat Desertification	2	1.2
1997	Kyoto Protocol	49	56.6
2006	International Tropical Timber Agreement	1	1.0
2009	IRENA Statute	3	6.7
2010	CBD Access and Benefit-Sharing Protocol	2	2.7
2012	Global Green Growth Institute Agreement	2	5.0
2015	Paris Agreement	50	68.4

The three primary climate change treaties clearly stand out, both in terms of absolute and relative counts of pattern matches. However, it is the secondary treaties which may be more interesting or surprising to less seasoned international environmental lawyers or non-lawyers. Even domain-specific treaty search engines fail to list some of them when performing a climate change keyword search. For instance, Ecolex omits three of the nine treaties listed in Table 4.3, namely the Ozone Layer Protection Convention, the Convention to Combat Desertification, and the IRENA

---

[317] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II) (adopted 8 June 1977, entered into force 7 December 1978) 1125 UNTS 609 ("be afforded safeguards as regards health and hygiene and protection against the rigours of the climate and the dangers of the armed conflict").

[318] PIC Convention (n 271) ("The relevance of such incidents to other States with similar climate, conditions and patterns of use of the formulation").

[319] Convention to Combat Desertification (n 81), art 8(1).

[320] ibid, preamble.

Statute,[321] while the IEA Database Project does not even include climate change as a subject to search agreements by.[322] As for UNTS online, it similarly fails to provide 'climate change' (or even 'climate') as a subject term, despite having no qualms including all sorts of specific subjects like 'Dairy farming', 'Desalination' and 'Dredging'.[323] Especially a monumental challenge like climate change governance requiring broad-based societal, political and economic support on a global level should be made as easily accessible and understandable as possible. If relevant multilateral agreements cannot be found effortlessly through open access databases by interested students, scholars, activists, philanthropists, investors and citizens, then clearly legal knowledge management and access to legal information have a long way to go.

Turning now to associations between climate change and other topics, the most closely related environmental problems in the given treaty set are deforestation, land degradation, desertification, and ozone depletion. The absolute or relative counts are not correlated because each treaty has its own core focus, but analysing the binary variables shows that the sets overlap to a large extent. Figure 4.5 displays the results of the pairwise analysis between `climateChangeB` and each of the other binary subject matter variables.

This representation is similar to Figure 3.2 in Chapter 3.1.1, except that all variables are boolean here. A probabilistic interpretation is again quite straightforward, except that it is about mentions instead of provisions this time. If the variables were pairwise independent, then the percentage of treaties mentioning a given subject in the full treaty set would correspond to the percentage of treaties exhibiting this feature in the subset at hand. In other words, the conditional probability of encountering a mention of some topic in a climate change treaty would be the same as the probability of that mention occurring in any other treaty. For instance, when drawing a random treaty text from the full set, the likelihood of it mentioning ozone depletion is only 4%, whereas in the subset of climate change treaties this probability is 33%, almost 8 times higher. As before, the 'Div' column contains the multiplication factor between the two percentages,[324] and the 'Diff' column shows

---

[321] See https://www.ecolex.org/result/?q=&type=treaty&xkeywords=climate+change&xdate_ min=&xdate_max=&tr_type_of_document=Multilateral&tr_field_of_application= Global (accessed 1 Dec 2021).

[322] See https://iea.uoregon.edu/environmental-agreements-subject (accessed 1 Dec 2021)). It does have climate change as a lineage, which includes the UNFCCC and all multilateral and bilateral agreements stemming from it, but this is a different type of classification.

[323] See https://treaties.un.org/Pages/AdvanceSearch.aspx?tab=UNTS&clang=_en (accessed 1 Dec 2021) – Filter by: Subject term.

[324] The percentage of the subset divided by that of the full set (using float numbers, not the rounded integers displayed in the table).

Cluster	Variable	Full Set (n=119)	Subset (n=9)	Div	Diff
Problems & Risks	Ozone Depletion	5 (4%)	3 (33%)	7.9	29.1
	Pollution	70 (59%)	3 (33%)	0.6	-25.5
	Waste	34 (29%)	2 (22%)	0.8	-6.3
	Land Degradation	12 (10%)	6 (67%)	6.6	56.6
	—Deforestation	6 (5%)	6 (67%)	13.2	61.6
	—Desertification	6 (5%)	4 (44%)	8.8	39.4
	Endangered Species	12 (10%)	0 (0%)	0.0	-10.1
	Pests & Weeds	15 (13%)	1 (11%)	0.9	-1.5
	Animal Welfare	6 (5%)	0 (0%)	0.0	-5.0
Human Stressors	Fishing & Hunting	25 (21%)	2 (22%)	1.1	1.2
	Agriculture	36 (30%)	6 (67%)	2.2	36.4
	Production	76 (64%)	8 (89%)	1.4	25.0
	Storage & Stockpiling	62 (52%)	3 (33%)	0.6	-18.8
	Transport	76 (64%)	2 (22%)	0.3	-41.6
	Trade	58 (49%)	5 (56%)	1.1	6.8
	Weapons of Mass Destr.	13 (11%)	0 (0%)	0.0	-10.9
Means of Intervention	Quant. Restrictions	31 (26%)	2 (22%)	0.9	-3.8
	Spatial Regulation	46 (39%)	3 (33%)	0.9	-5.3
	Risk/Impact Assessment	47 (39%)	5 (56%)	1.4	16.1
	Inspections	49 (41%)	0 (0%)	0.0	-41.2
	Data & Research	99 (83%)	9 (100%)	1.2	16.8
	Education	48 (40%)	8 (89%)	2.2	48.6
	Training/Capacity-build.	63 (53%)	9 (100%)	1.9	47.1
	Tech Transfer	37 (31%)	8 (89%)	2.9	57.8
	Insurance	27 (23%)	3 (33%)	1.5	10.6
	Liability	44 (37%)	2 (22%)	0.6	-14.8
	Criminal Responsibility	16 (13%)	0 (0%)	0.0	-13.4

Figure 4.5: Climate change treaties summary results

the percentage point difference. Thus, it can be seen for instance that endangered species and animal welfare are not mentioned at all in climate change treaties. It is important to bear in mind, however, that in some cases this may simply be due to the time at which the relevant treaties were adopted and the evolving popularity of terms and framings (see section 4.1.5).

As for contributing activities, agriculture (including forestry) is more than twice as likely to be mentioned in climate change treaties, while transport is only about a third as likely to appear in such texts and weapons of mass destruction are not mentioned at all. None of the differences in likelihood are as striking as the ones seen with the most closely associated environmental problems. This is unsurprising given that climate change treaties are not typically focused on any one activity or sector (except for the IRENA Statute and the Tropical Timber Agreement). Transport is of course a crucial but also contentious sector to decarbonise, and is mentioned in two of the three core climate change treaties.

Regarding means of intervention, the strongest pairwise associations are with education and technology transfer, both more than twice as likely to be mentioned

in climate change treaties. Capacity-building is not far behind, being mentioned in all of the climate change agreements. Data & research is equally ubiquitous, but full sample prevalence is also much higher and therefore its presence here is less surprising. The fact that inspections, liability and criminal responsibility are much less likely to be mentioned in a climate change treaty than in any random treaty drawn from the full sample is perhaps an unfortunate but accurate reflection of political reality.

Figure 4.6 presents the results of the legal information extraction process from the subset of climate change treaties as compared to the full set, generated with the same source code as Figure 3.2 in Chapter 3.1.1 on provisional application.

Cluster	Variable	Type	Full Set (n=119)	Subset (n=9)	Div	Diff
Start of Life	Provisional Application Provision	bool	14 (12%)	2 (22%)	1.9	10.5
	Entry into Force Provision	bool	119 (100%)	9 (100%)	1.0	0.0
	—Min. Nb. of Ratifs for EIF	num	112 (19.8)	9 (36.7)	1.9	16.9
	—Additional EIF Conditions	bool	47 (39%)	3 (33%)	0.8	-6.2
	—EIF Waiting Period (Years)	num	110 (0.4)	8 (0.2)	0.5	-0.2
Change	Amendment Provision	bool	115 (97%)	9 (100%)	1.0	3.4
	—Simplified Amendment	bool	60 (50%)	6 (67%)	1.3	16.2
	Adoption of Annexes or Protocols	bool	26 (22%)	6 (67%)	3.1	44.8
End of Life	Denunciation Provision	bool	113 (95%)	9 (100%)	1.1	5.0
	—Denunciation Waiting Period (Years)	num	97 (2.7)	9 (2.6)	1.0	-0.1
	—Other Denunciation Conditions	bool	33 (28%)	0 (0%)	0.0	-27.7
	—Denunciation Effect Delay (Years)	num	107 (0.9)	8 (0.8)	1.0	-0.0
	—Other Denunciation Effect	bool	7 (6%)	1 (11%)	1.9	5.2
	Termination Provision	bool	15 (13%)	2 (22%)	1.8	9.6
	—Auto-Termination Memb. Threshold	num	4 (13.5)	0 (nan)	nan	nan
	—Auto-Expiry	bool	6 (5%)	1 (11%)	2.2	6.1
Treaty Organs	Legal Personality	bool	25 (21%)	3 (33%)	1.6	12.3
	Organisational Auspices	cat	108 (91%)	7 (78%)	0.9	-13.0
	Secretariat	bool	111 (93%)	9 (100%)	1.1	6.7
	Conference of Parties	bool	85 (71%)	9 (100%)	1.4	28.6
Compliance & Effectiveness	National Reporting	bool	89 (75%)	7 (78%)	1.0	3.0
	International Verification	bool	63 (53%)	2 (22%)	0.4	-30.7
	Progress Review	bool	73 (61%)	8 (89%)	1.4	27.5
	Non-Compliance Procedures	bool	66 (55%)	5 (56%)	1.0	0.1
	Sanctions	bool	58 (49%)	3 (33%)	0.7	-15.4
	Dispute Settlement	bool	74 (62%)	7 (78%)	1.3	15.6
	—Compulsory Dispute Settlement	cat	41 (34%)	6 (67%)	1.9	32.2

Figure 4.6: Design features of climate change treaties

A notable difference in numeric variables is that the average `EIFnConsent` is nearly two times higher in the subset of climate change treaties (37) than in the full set (20). Incidentally, in the three core climate change treaties it is even higher, at 53 instruments of consent, but otherwise the quantitative variables in the two subsets are very similar. As compared to the full set, the only other continuous variable that stands out is the waiting period between the conditions for entry into force being

met and EIF itself, which is about half as long in the subset than it is in the full set (a bit over two months vs. almost five months).

Pairwise associations between boolean variables diverge a bit more between the full set, the subset of 9 and the subset of 3. Still, they're not particularly surprising to anyone familiar with the three primary climate change treaties. The biggest jump in probability is that of finding a provision on the adoption of annexes or protocols to a treaty, which is over three times as high in the subset of 9 and 4.76 times as high in the core climate change treaties, where it is 100%. Simplified amendment provisions also occur in all three core treaties (twice as much as in the full set), as do national reporting provisions, review provisions and dispute settlement provisions (in some cases applicable *mutatis mutandis*), though none of them amounting to a doubling of incidence. Conversely, some binary variables occur less frequently in the core climate change treaties than in the larger subset or the full set. In particular, none of the core treaties contains provisional application clauses, provisions on other conditions for, or effects of, denunciation, treaty termination provisions, or provisions establishing an entity with legal personality. It will come as no surprise that provisions on international verification of compliance and on sanctions are much less common in climate change treaties than in the full set, a difference of -31 and -15 p.p. respectively.

Last but not least, the two categorical variables are not fully represented in the table. The core climate change treaties were all adopted under the auspices of the UN. The ancillary climate change treaties were also concluded under UN or UNEP's auspices, except for the IRENA Statute and the Global Green Growth Institute Agreement which were adopted independently of the United Nations. This entails a slightly lower set average prevalence of the `orgAuspices` variable. As for dispute settlement, two agreements have no provision on judicial or arbitral dispute settlement[325] and one has a dispute settlement provision with ad hoc consent by the parties to the dispute,[326] all counting as no compulsory dispute settlement mechanism. The remaining six agreements, including the three primary climate change treaties, all provide for the possibility to opt into compulsory dispute settlement by means of a declaration submitted upon ratification or at any time thereafter.[327] This counts as a basis for (opt-in) compulsory dispute settlement and represents a steep increase in prevalence as compared to the full set (nearly twice as common, or +32 p.p.).

---

[325] International Tropical Timber Agreement (n 150); and Agreement on the Establishment of the Global Green Growth Institute (adopted 20 June 2012, entered into force 18 October 2012) ATS 2014 34.

[326] IRENA Statute (n 189), art XVI.

[327] See e.g. Framework Convention on Climate Change (n 82), art 14(2).

4.1.2 Stratospheric ozone depletion

Unlike 'climate', 'ozone' is not polysemous. The only issue in identifying references to the problem of stratospheric ozone depletion is that some mentions of ozone in treaties denote ground-level ozone pollution, which is a separate health and environmental concern. Consequently, the optimal regex pattern for the treaties at hand is as simple as:

```
(?i)ozone.(?:deplet|layer)\w*
```

All mentions of ozone and the stratosphere were analysed to avoid underinclusion. Five treaties contain the pattern, and as expected, two of them stand out in terms of absolute and relative counts. Table 4.4 lists the details. As with climate change conventions, references to the core ozone treaties (both containing 'ozone layer' in their title)[328] are not excluded from the match.

Table 4.4: Ozone layer treaties regex counts

Year	Name	Count	Fraction
1985	Ozone Layer Protection Convention	32	54.8
1987	Prot. on Substances that Deplete the Ozone Layer	14	31.7
1992	Framework Convention on Climate Change	2	2.4
1997	Kyoto Protocol	1	1.2
2009	Ship Recycling Convention	1	0.7

The reason why the Montreal Protocol[329] has comparatively few mentions is that it uses 'controlled substances' as the main term to refer to ozone-depleting substances. As it nonetheless contains enough matches to distinguish itself from the secondary ozone treaties, there is no need to address this further. However, researchers devising a count-based classification scheme may choose to substitute non-matching terms with definitions in the treaty. Given the low number of treaties in this subset, the full subject matter results for the five treaties can be displayed (in Figure 4.7). It is worth noting that in both the ozone and climate change regimes the protocols extend and implement rather than replace the conventions, and thus the lack of mention of climate change in the Ozone Layer Protocol does not imply a change of mandate. The protocols simply focus on specific measures rather than on general obligations and contextualisation, as evidenced by the frequent mentions of quantitative restrictions in the protocols, which are absent from the conventions.

---

[328] Ozone Layer Protection Convention (n 196); and Protocol on Substances that Deplete the Ozone Layer (n 176).

[329] ibid.

Cluster	Variable	OzoneLayer Conv	OzoneLayer Prot	UNFCCC	Kyoto Prot	ShipRecyc. Conv
Problems & Risks	Climate Change	5.1	0.0	102.5	56.6	0.0
	Ozone Depletion	54.8	31.7	2.4	1.2	0.7
	Pollution	3.4	0.0	2.4	0.0	5.9
	Waste	0.0	4.5	1.2	6.9	197.0
	Land Degradation	0.0	0.0	8.3	1.2	0.0
	—Deforestation	0.0	0.0	1.2	1.2	0.0
	—Desertification	0.0	0.0	7.1	0.0	0.0
	Endangered Species	0.0	0.0	0.0	0.0	0.0
	Pests & Weeds	0.0	0.0	0.0	0.0	0.0
	Animal Welfare	0.0	0.0	0.0	0.0	0.0
Human Stressors	Fishing & Hunting	1.7	0.0	0.0	0.0	0.0
	Agriculture	3.4	0.0	3.6	12.7	0.0
	Production	3.4	83.7	4.8	6.9	0.7
	Storage & Stockpiling	0.0	4.5	1.2	0.0	4.4
	Transport	0.0	2.3	1.2	4.6	3.7
	Trade	3.4	52.0	4.8	4.6	0.0
	Weapons of Mass Destr.	0.0	0.0	0.0	0.0	0.0
Means of Intervention	Quant. Restrictions	0.0	29.4	0.0	25.4	0.0
	Spatial Regulation	0.0	0.0	1.2	1.2	0.7
	Risk/Impact Assessment	3.4	0.0	6.0	2.3	4.4
	Inspections	0.0	0.0	0.0	0.0	8.1
	Data & Research	70.3	43.0	27.4	12.7	3.0
	Education	0.0	4.5	10.7	2.3	0.0
	Training/Capacity-build.	3.4	0.0	7.1	4.6	8.9
	Tech Transfer	3.4	2.3	11.9	5.8	0.7
	Insurance	0.0	4.5	1.2	1.2	0.0
	Liability	0.0	0.0	0.0	0.0	0.0
	Criminal Responsibility	0.0	0.0	0.0	0.0	0.0

Figure 4.7: Subject matter results for ozone treaties

To put this subset into a broader context, Figure 4.8 shows the conditional probabilities of each of the other subject matter variables occurring in one of these five agreements. This being the smallest subset means that a single treaty's mentions can have an outsized impact on percentages, as in the case of desertification, where the conditional probability quadruples compared to the baseline. An interesting difference when comparing ozone with climate change treaties is the high prevalence of references to waste. In the present subset waste is mentioned in 4 out of 5 treaties, nearly three times as much as in the full set average, whereas in the climate change subset the prevalence of such mentions is below average (-6 percentage points), despite the large overlap between the two subsets. Technology transfer is again the most prominent intervention, understandably so.

As for legal information extraction, the results are similar to the previous section in that provisions on the adoption of annexes and protocols, on simplified amendment procedures, and on compulsory dispute settlement are all much more

Cluster	Variable	Full Set (n=119)	Subset (n=5)	Div	Diff
Problems & Risks	Climate Change	9 (8%)	3 (60%)	7.9	52.4
	Pollution	70 (59%)	3 (60%)	1.0	1.2
	Waste	34 (29%)	4 (80%)	2.8	51.4
	Land Degradation	12 (10%)	2 (40%)	4.0	29.9
	—Deforestation	6 (5%)	2 (40%)	7.9	35.0
	—Desertification	6 (5%)	1 (20%)	4.0	15.0
	Endangered Species	12 (10%)	0 (0%)	0.0	-10.1
	Pests & Weeds	15 (13%)	0 (0%)	0.0	-12.6
	Animal Welfare	6 (5%)	0 (0%)	0.0	-5.0
Human Stressors	Fishing & Hunting	25 (21%)	1 (20%)	1.0	-1.0
	Agriculture	36 (30%)	3 (60%)	2.0	29.7
	Production	76 (64%)	5 (100%)	1.6	36.1
	Storage & Stockpiling	62 (52%)	3 (60%)	1.2	7.9
	Transport	76 (64%)	4 (80%)	1.3	16.1
	Trade	58 (49%)	4 (80%)	1.6	31.3
	Weapons of Mass Destr.	13 (11%)	0 (0%)	0.0	-10.9
Means of Intervention	Quant. Restrictions	31 (26%)	2 (40%)	1.5	13.9
	Spatial Regulation	46 (39%)	3 (60%)	1.6	21.3
	Risk/Impact Assessment	47 (39%)	4 (80%)	2.0	40.5
	Inspections	49 (41%)	1 (20%)	0.5	-21.2
	Data & Research	99 (83%)	5 (100%)	1.2	16.8
	Education	48 (40%)	3 (60%)	1.5	19.7
	Training/Capacity-build.	63 (53%)	4 (80%)	1.5	27.1
	Tech Transfer	37 (31%)	5 (100%)	3.2	68.9
	Insurance	27 (23%)	3 (60%)	2.6	37.3
	Liability	44 (37%)	0 (0%)	0.0	-37.0
	Criminal Responsibility	16 (13%)	0 (0%)	0.0	-13.4

Figure 4.8: Subject matter of ozone depletion treaties

prominent in ozone layer treaties than in the full set.

4.1.3 Pollution and waste

Both pollution and waste are construed broadly and required a considerable amount of tweaking as evidenced by the long regex patterns in Table 4.5. For instance, a negative lookahead is needed to prevent 'waste' from being matched in expressions like 'waste of economic and human resources'.[330]

The pollution pattern matches as many as 70 (59%) treaties, while references to waste are present in 34 (29%) of the agreements. The overlap between the two sets is nearly complete, only two waste treaties do not match the pollution pattern,[331] pollution seemingly subsuming waste concerns. However, when taking into account the different degrees of focus modelled here, the subsumption hypothesis becomes

---

[330] This expression occurs in the preamble of the Treaty on the Prohibition of Nuclear Weapons (n 199), but 'waste' is rarely used in this sense in the study set.

[331] The Protocol on Substances that Deplete the Ozone Layer (n 176); and the Kyoto Protocol (n 300).

Table 4.5: Pollution and waste regex patterns

Variable	Regular expression pattern
pollution	`(?i)(?:pollut|contamin|toxic|poison|hazard\w+ substance|deleterious|carcinogenic (?:subst|agent)|chemical[^.]+?(?:hazard|bioaccum)|chemical.+risk.+ to the environment|radiation hazard|hazard\w* [^\n]* (?:radiation|radioact)|nuclear (?:incident|accident|war))`
waste	`(?i)(?:(?<!fraud, )waste(?! of (?:human|economic|financial))|recycl|garbage|landfill|sewage)\w*`

less tenable and waste does emerge as its own concern. Indeed, almost half of all waste treaties are at least equally if not more focused on waste than on pollution, and six are clearly core waste treaties. Likewise, the agreements with the highest term frequency for pollution are nearly exclusively focused on pollution. With 72 (60.5%) of agreements in the study set containing references to either one or both of these environmental problems, the full list of treaties would be too long to reproduce here. Hence, Table 4.6 only includes agreements with at least one of the two variables having a scaled term frequency of 60 or more. The absence of MARPOL[332] may seem surprising, but it is only because MARPOL is such a lengthy agreement (including annexes) and the table is based on relative rather than absolute counts. Ordering treaties by the latter would put UNCLOS at the top and MARPOL in second position. Indeed, marine pollution is probably the type of pollution with the most agreements focused on it.

While the pollution regex pattern combines all kinds of pollution into a single expression, other variables such as organisational auspices can aid in distinguishing between different sub-focuses of treaties. For instance, 20 out of 21 agreements adopted under the auspices of the IMO are partly about pollution, 9 of which also mention waste. Similarly, 12 out of 13 ILO treaties in the treaty set match the pollution pattern, but only 5 of them also appear in the subset on waste. By contrast, the waste focus is much more prevalent in treaties adopted under the auspices of the IAEA, where it appears in 6 out of 8 treaties, only one less than pollution, which is mentioned in 7 out of 8.

As for associations with other environmental problems, the strongest one is between waste and ozone depletion, with a nearly three times higher prevalence than in the full set, as already seen in the preceding section. Pests & weeds are also more likely to be mentioned in both waste (1.6 times, 8 percentage point increase) and pollution treaties (only slight increase).

---

[332] MARPOL73 (n 75).

Table 4.6: Pollution and waste regex counts

Year	Treaty	pollution	waste
1962	Nuclear Ships Operators Liability Convention	61.0	21.1
1969	Oil Pollution Damage Liability Convention	69.3	0.0
1972	London Dumping Convention	37.5	68.0
1974	Occupational Cancer Convention	82.7	0.0
1977	Working Environment Conv.	80.0	0.0
1986	Nuclear Accident Assistance Conv.	61.6	0.0
1989	Basel Hazardous Wastes Convention	11.2	281.6
1990	Chemicals Safety Conv.	65.6	3.3
1990	Oil Pollution Preparedness, Resp. & Coop. Conv.	138.5	0.0
1996	London Dumping Protocol	36.2	95.0
1997	Spent Fuel and Radioactive Waste Safety Conv.	9.7	120.2
1999	Hazardous Wastes Liability & Compensation Prot.	1.8	122.0
2000	HNS Pollution Preparedness, Resp. & Coop. Prot.	114.9	0.0
2001	Bunker Oil Pollution Damage Liability Conv.	67.2	0.0
2009	Ship Recycling Convention	5.9	197.0

With regard to contributing activities, production, storage and transport are all somewhat more commonly mentioned in pollution and waste treaties. For example, storage & stockpiling rise 33 percentage points to 85% probability of occurring in waste agreements, a 1.6-fold increase, similar to mentions of pollution in the waste subset. Weapons of mass destruction are one anthropogenic stressor that differs in the direction its conditional probability takes, in that it is more likely to occur in pollution treaties than in the full set (1.3 times), but slightly less common in waste treaties (0.8 times).

The means of intervention associated with pollution and waste treaties are largely as expected. Quantitative and spatial restrictions, risk/impact assessments and inspections are all more frequently mentioned in these two subsets, and waste treaties additionally tend to stress insurance, technology transfer and training/capacity-building. Overall, because waste is a smaller subset, it exhibits starker differences with regard to the full set and may thus be more interesting to include here (Figure 4.9).

As for treaty provision variables, one that stands out in both the pollution and waste subsets concerns the establishment of entities with legal personality. It is much less prevalent in these subsets than in the full set (only 5/70 for pollution treaties, a third of the independent probability, and not much more in waste treaties). Termination provisions are also less likely to occur in pollution and waste agreements, the probability dropping from 13% in the full set to 6% in pollution and 3% in waste treaties. Interestingly, while the pollution subset has very similar incidence

Cluster	Variable	Full Set (n=119)	Subset (n=34)	Div	Diff
Problems & Risks	Climate Change	9 (8%)	2 (6%)	0.8	-1.7
	Ozone Depletion	5 (4%)	4 (12%)	2.8	7.6
	Pollution	70 (59%)	32 (94%)	1.6	35.3
	Land Degradation	12 (10%)	3 (9%)	0.9	-1.3
	—Deforestation	6 (5%)	2 (6%)	1.2	0.8
	—Desertification	6 (5%)	1 (3%)	0.6	-2.1
	Endangered Species	12 (10%)	4 (12%)	1.2	1.7
	Pests & Weeds	15 (13%)	7 (21%)	1.6	8.0
	Animal Welfare	6 (5%)	1 (3%)	0.6	-2.1
Human Stressors	Fishing & Hunting	25 (21%)	9 (26%)	1.3	5.5
	Agriculture	36 (30%)	10 (29%)	1.0	-0.8
	Production	76 (64%)	29 (85%)	1.3	21.4
	Storage & Stockpiling	62 (52%)	29 (85%)	1.6	33.2
	Transport	76 (64%)	31 (91%)	1.4	27.3
	Trade	58 (49%)	17 (50%)	1.0	1.3
	Weapons of Mass Destr.	13 (11%)	3 (9%)	0.8	-2.1
Means of Intervention	Quant. Restrictions	31 (26%)	20 (59%)	2.3	32.8
	Spatial Regulation	46 (39%)	20 (59%)	1.5	20.2
	Risk/Impact Assessment	47 (39%)	24 (71%)	1.8	31.1
	Inspections	49 (41%)	19 (56%)	1.4	14.7
	Data & Research	99 (83%)	31 (91%)	1.1	8.0
	Education	48 (40%)	14 (41%)	1.0	0.8
	Training/Capacity-build.	63 (53%)	23 (68%)	1.3	14.7
	Tech Transfer	37 (31%)	15 (44%)	1.4	13.0
	Insurance	27 (23%)	12 (35%)	1.6	12.6
	Liability	44 (37%)	12 (35%)	1.0	-1.7
	Criminal Responsibility	16 (13%)	4 (12%)	0.9	-1.7

Figure 4.9: Subject matter of treaties on waste

rates for variables of the compliance & effectiveness cluster (with the exception of a 5.6 percentage point increase in provisions on international verification of compliance), the waste subset has five double figure leaps in percentage points when compared with the full set (see Figure 4.10). Perhaps the pollution subset is a bit too heterogeneous, yielding values close to the general average.

4.1.4   Deforestation, desertification and land degradation

As these three problems are related, both in nature and in treaties, they are discussed jointly in this section. The search patterns used for the three variables are included in Table 4.7.

For deforestation, all references to forests and woodlands were inspected, but only those explicitly mentioning loss or degradation of forests were retained. The wider keyword search for desertification likewise included any reference to deserts, drylands, droughts and suchlike. Land degradation was a bit harder to capture, as

Cluster	Variable	Type	Full Set (n=119)	Subset (n=34)	Div	Diff
Start of Life	Provisional Application Provision	bool	14 (12%)	2 (6%)	0.5	-5.9
	Entry into Force Provision	bool	119 (100%)	34 (100%)	1.0	0.0
	—Min. Nb. of Ratifs for EIF	num	112 (19.8)	32 (22.7)	1.1	2.9
	—Additional EIF Conditions	bool	47 (39%)	15 (44%)	1.1	4.6
	—EIF Waiting Period (Years)	num	110 (0.4)	32 (0.5)	1.3	0.1
Change	Amendment Provision	bool	115 (97%)	33 (97%)	1.0	0.4
	—Simplified Amendment	bool	60 (50%)	21 (62%)	1.2	11.3
	Adoption of Annexes or Protocols	bool	26 (22%)	11 (32%)	1.5	10.5
End of Life	Denunciation Provision	bool	113 (95%)	34 (100%)	1.1	5.0
	—Denunciation Waiting Period (Years)	num	97 (2.7)	31 (3.2)	1.2	0.5
	—Other Denunciation Conditions	bool	33 (28%)	10 (29%)	1.1	1.7
	—Denunciation Effect Delay (Years)	num	107 (0.9)	34 (0.9)	1.1	0.1
	—Other Denunciation Effect	bool	7 (6%)	0 (0%)	0.0	-5.9
	Termination Provision	bool	15 (13%)	1 (3%)	0.2	-9.7
	—Auto-Termination Memb. Threshold	num	4 (13.5)	0 (nan)	nan	nan
	—Auto-Expiry	bool	6 (5%)	0 (0%)	0.0	-5.0
Treaty Organs	Legal Personality	bool	25 (21%)	3 (9%)	0.4	-12.2
	Organisational Auspices	cat	108 (91%)	33 (97%)	1.1	6.3
	Secretariat	bool	111 (93%)	34 (100%)	1.1	6.7
	Conference of Parties	bool	85 (71%)	22 (65%)	0.9	-6.7
Compliance & Effectiveness	National Reporting	bool	89 (75%)	31 (91%)	1.2	16.4
	International Verification	bool	63 (53%)	23 (68%)	1.3	14.7
	Progress Review	bool	73 (61%)	21 (62%)	1.0	0.4
	Non-Compliance Procedures	bool	66 (55%)	25 (74%)	1.3	18.1
	Sanctions	bool	58 (49%)	17 (50%)	1.0	1.3
	Dispute Settlement	bool	74 (62%)	28 (82%)	1.3	20.2
	—Compulsory Dispute Settlement	cat	41 (34%)	18 (53%)	1.5	18.5

Figure 4.10: Design features of waste treaties

Table 4.7: Deforestation, desertification and land degradation regex patterns

Variable	Regular expression pattern
deforest	`(?i)(?:deforest|degraded forest|forest coverage loss|forest decay|forest fire[s]?)`
desertif	`(?i)(?:desertif|\Warid (?:zone|area|region))`
landDegrad	`(?i)(?:(?:soil|land) (?:degradation|erosion|deterior\w+)|(?<!photo)(?:degrad|deterior)\w+\b(?! treatment)[^.]*?  (?:land|soil|area(?![s]?  beyond the limits of national))|maintain\w* soil fertility|conserv\w* \w* ?soil)`

the longer regex pattern shows, and has more potential overlaps with other environmental problems. To narrow the scope of this environmental concern, references to land contamination, waste disposal on land, and land as a carbon sink are only extracted under the headings of pollution, waste and climate change respectively. All mentions of soil and land in agreement texts not returning a match were examined as part of the false negatives search.

The land degradation search pattern itself matches seven treaties, while deforestation and desertification occur in six agreements each, but as the latter two are construed as subcategories of land degradation, any match in the two is propagated upwards, yielding a total of twelve matches for land degradation. Table 4.8 summarises the results for all three variables. Aside from the obvious exception of desertification in the UNCCD,[333] the three environmental problems are only a minor topic in the treaties that mention them.

Table 4.8: Deforestation, desertification and land degradation regex counts

Year	Treaty	deforest	desertif	landDegrad
1972	World Heritage Convention	0.0	0.0	1.9
1979	International Olive Oil Agreement	0.0	0.0	1.0
1989	Basel Hazardous Wastes Convention	0.0	0.0	0.9
1992	Convention on Biological Diversity	0.0	1.1	2.2
1992	Framework Convention on Climate Change	1.2	7.1	8.3
1994	Convention to Combat Desertification	1.2	70.0	78.1
1997	Non-Navigational Uses of Intl Watercourses Conv.	0.0	1.8	1.8
1997	Kyoto Protocol	1.2	0.0	1.2
2001	Plant Genetic Resources Treaty	0.0	0.0	1.1
2006	International Tropical Timber Agreement	1.0	1.0	3.0
2009	IRENA Statute	4.4	2.2	6.7
2015	Paris Agreement	1.4	0.0	1.4

One might have expected there to be a considerable overlap between treaties concerned with deforestation and treaties focusing on endangered species, biodiversity and habitat loss. However, there is very little intersection to be gauged from treaty texts. None of the treaties on endangered species and only two treaties on biodiversity loss mention deforestation. CBD is not one of them, in fact it lacks any explicit mention of forests, but uses broader concepts such as 'habitat' and 'ecosystem' instead. Also, animal welfare is not mentioned in any of the three subsets, nor are WMDs. Fishing & hunting and criminal responsibility are both mentioned in only one of the agreements. Apart from the above-mentioned association with climate change, agriculture is also strongly correlated with all three of these land variables (more than doubling its prevalence). As for the means of tackling land

---

[333] Convention to Combat Desertification (n 81).

degradation problems, only data & research is mentioned in all treaties, but education, training/capacity-building and technology transfer are a close second, being highlighted in all but one of them. Spatial regulation and risk/impact assessment, despite both only being mentioned in seven out of the twelve agreements, still reach a nearly 20 p.p. increase (and 27 p.p. in the two smaller subsets). Inspections and liability exhibit a noteable drop, especially for treaties mentioning deforestation. The long overdue treaty on forest conservation could close some of these gaps.

4.1.5 Endangered species and related problems

Information regarding biodiversity, habitat, or ecosystem loss was harder to extract than anticipated. On the one hand, treaties covering particular branches of the tree of life often use branch-specific terminology. For example, the Birds Protection Convention lacks any mention of the term 'habitat' but aims to tackle the "rapid disappearance of suitable breeding grounds for birds as a result of human intervention".[334] The London Dumping Convention contains no instance of the terms 'ecosystem', 'habitat', or 'species', but calls for the prevention of harm or damage to 'marine life'[335] and mentions spawning and nursery areas as types of locations to avoid when choosing a dumping site.[336] On the other hand, the popularity of concepts has changed over time. For instance, the Ramsar Wetlands Convention does not mention the terms 'ecosystem' or 'ecosystem services', but does highlight the "fundamental ecological functions of wetlands as regulators of water regimes".[337] Figure 4.11 shows the timeline of four interrelated terms with their respective occurrence counts summed across all treaties adopted in a given year.[338] Note that the y axis is not the same, 'species' has by far the highest count, with CITES and CMS driving the peaks in 1973 and 1979 respectively.[339]

Therefore, while habitat and biodiversity loss were originally part of the subject matter variables, only endangered species form part of the final analysis, using the search pattern in Table 4.9. To mitigate the risk of false negatives, I examined every instance of 'species' and 'extinct' in unmatched treaties, among other searches.

---

[334] Birds Protection Convention (n 175) , art 11.

[335] London Dumping Convention (n 302), arts I and V(1).

[336] Ibid , Annex III(B)(1).

[337] Ramsar Wetlands Convention (n 77), preambular para. 2.

[338] For the purpose of this illustration, 'habitat', 'ecosystem', and 'species' were searched with very simple one-word regex patterns, while biodiversity was represented with a slightly more complex regex to account for variations of the same concept: `'(?i)((?:bio\w*|ecol\w+|genetic).?(?:diversity|richness|variability))'`.

[339] The total counts are 37 for 'habitat' across 11 treaties, 57 for 'ecosystem' across 18 treaties, 165 for 'biodiversity' across 14 treaties, and a whopping 407 for 'species' across 20 treaties.
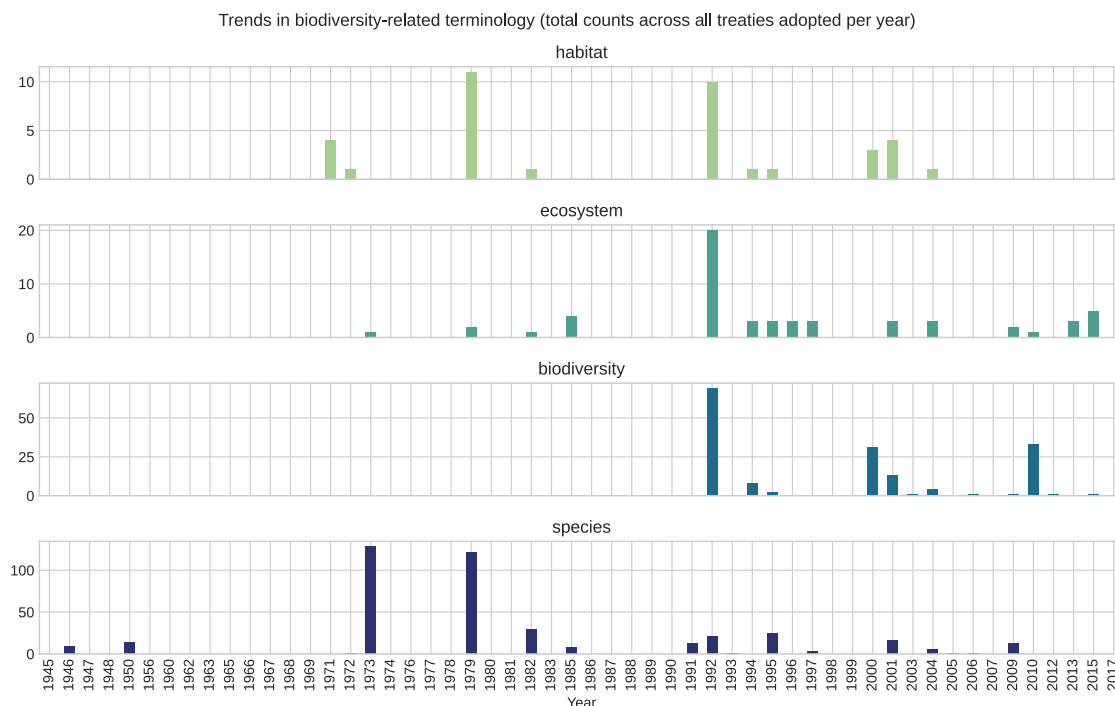
Figure 4.11: Biodiversity-related terminology over time

Table 4.9: Endangered species regex pattern

Variable	Regular expression pattern
endangSpec	`(?i)(?:(?:threaten|endanger|extermin|destruction)[^.]{,70}`
	`species|species[^.]{,70} (?:threat|endanger|in danger)[^.]{,20}`
	`(?:survival|extinction)|(?:survival|perpet\w+) of \w+ species|recovery`
	`.. \w+ (?:populations of )?species)`

References to the extermination of pests or to the eradication of invasive species[340] were considered false positives, because the target variable is construed as (natural and anthropogenic) threats to the survival of species we do *not* want to go extinct.[341] The difference between disappearance from a certain geographic area and global extinction was not made so long as the text excerpts were concerned with the survival of species in one way or another (the last remaining population of an endangered species may well be confined to a narrow geographic range). There are of course many more agreements tackling issues of overfishing or other overexploitation issues leading to economically suboptimal results without threatening the existence of a species as such. The low counts of the Whaling Convention and the Fish Stocks Agreement in Table 4.10 are a case in point.

---

[340] E.g. Convention on Biological Diversity (n 248), art 8(h) ("Prevent the introduction of, control or eradicate those alien species which threaten ecosystems, habitats or species;").

[341] There are no explicit references to extinction risks of *homo sapiens* in the treaties considered, thus no need to distinguish between these problem domains.

Table 4.10: Endangered species regex counts

Year	Name	Count	Fraction
1946	Whaling Convention	1	2.4
1950	Birds Protection Convention	5	37.6
1972	World Heritage Convention	1	1.9
1973	CITES	10	10.8
1979	Migratory Species Conservation Convention	5	9.3
1982	UNCLOS	1	0.1
1992	Convention on Biological Diversity	9	9.7
1995	Fish Stocks Agreement	1	0.8
2001	Ship Anti-Fouling Systems Convention	2	2.6
2001	Plant Genetic Resources Treaty	1	1.1
2004	Ballast Water Management Convention	1	0.8
2009	IUU Fishing Port State Measures Agmt	1	1.3

In fact, 8 of the 12 of the treaties listed in this table only have one or two matches for the endangered species regex. Still, even if they only mention the issue in passing, having it within the purview of the convention can prove invaluable for environmentalists to use as a lever for action when the time is ripe.

As for associations with other subject matter variables (shown in Figure 4.12), land degradation, pests & weeds, and animal welfare are more likely to be mentioned in treaties on endangered species, as is fishing & hunting for obvious reasons. Among the means of intervention, spatial regulation (protected areas), inspections (e.g. of fishing vessels), and technology transfer are more prevalent in this subset than in the full set, while insurance is only mentioned in one of them.

With regard to legal information extraction, the treaty design patterns in this subset are very similar to those of the full set, except that withdrawal conditions are looser. The average denunciation waiting period, at a bit under a year, is about a third as long as that of the full sample average, and the incidence rate of other denunciation conditions is also about a third as high as that of the full set (only one treaty out of the twelve has any).

4.1.6 Pests and weeds

The target variable of this section is also focused on the biosphere, but less on its loss than on its unwanted growth and spread. It is construed to comprise any growth or movement of nonhuman organisms that humans see as a problem for whatever reason, including damage to agricultural productivity, harm to local wildlife, biodiversity loss, spread of diseases, etc. In some cases pests or their hosts are themselves migratory and in other cases humans are responsible for transport, knowingly or not.

Cluster	Variable	Full Set (n=119)	Subset (n=12)	Div	Diff
Problems & Risks	Climate Change	9 (8%)	0 (0%)	0.0	-7.6
	Ozone Depletion	5 (4%)	0 (0%)	0.0	-4.2
	Pollution	70 (59%)	5 (42%)	0.7	-17.2
	Waste	34 (29%)	4 (33%)	1.2	4.8
	Land Degradation	12 (10%)	3 (25%)	2.5	14.9
	—Deforestation	6 (5%)	0 (0%)	0.0	-5.0
	—Desertification	6 (5%)	1 (8%)	1.7	3.3
	Pests & Weeds	15 (13%)	7 (58%)	4.6	45.7
	Animal Welfare	6 (5%)	4 (33%)	6.6	28.3
Human Stressors	Fishing & Hunting	25 (21%)	8 (67%)	3.2	45.7
	Agriculture	36 (30%)	4 (33%)	1.1	3.1
	Production	76 (64%)	6 (50%)	0.8	-13.9
	Storage & Stockpiling	62 (52%)	5 (42%)	0.8	-10.4
	Transport	76 (64%)	9 (75%)	1.2	11.1
	Trade	58 (49%)	5 (42%)	0.9	-7.1
	Weapons of Mass Destr.	13 (11%)	0 (0%)	0.0	-10.9
Means of Intervention	Quant. Restrictions	31 (26%)	4 (33%)	1.3	7.3
	Spatial Regulation	46 (39%)	8 (67%)	1.7	28.0
	Risk/Impact Assessment	47 (39%)	6 (50%)	1.3	10.5
	Inspections	49 (41%)	8 (67%)	1.6	25.5
	Data & Research	99 (83%)	12 (100%)	1.2	16.8
	Education	48 (40%)	6 (50%)	1.2	9.7
	Training/Capacity-build.	63 (53%)	7 (58%)	1.1	5.4
	Tech Transfer	37 (31%)	6 (50%)	1.6	18.9
	Insurance	27 (23%)	1 (8%)	0.4	-14.4
	Liability	44 (37%)	3 (25%)	0.7	-12.0
	Criminal Responsibility	16 (13%)	2 (17%)	1.2	3.2

Figure 4.12: Subject matter of treaties on endangered species

As the regular expression pattern in Table 4.11 shows, terminology used in treaties is domain-specific in some cases, e.g. 'deratting' of ships, but there are also generic terms like 'pests', 'weeds', 'invasive species' or 'infestation' that clearly convey the perceived problem of the growth or spread of some macro- or microorganisms without specifying the species or genus concerned. This facilitates pattern matching of this class of problems across very different contexts. Capturing this variable in regional treaties would probably be more difficult if they use many specific terms for the problems they govern.

Table 4.11: Pests and weeds regex pattern

Variable	Regular expression pattern
pests	`(?i)(?:(?:invasive|alien|exotic) \w* ?\w* ?species|species,`
	`(?:invasive|alien|exotic)|\bpest[s]?\b|\w*infestation|parasite|\bweeds|damaging`
	`fields, vineyards|deratting|spread of harmful \w* ?organism)[s]?`

The results of the search are listed in Table 4.12, showing that pests are a marginal concern in all but one of the treaties. The WTO (SPS) Agreement yields a high absolute count of mentions, but a low relative count given the length of the text. Still, it is an interesting mix of agreements that are probably not often seen as having something in common. It is a slightly larger set than the previous one on endangered species, which also has one core treaty focused on the issue. However, pests are a more cross-cutting issue, exhibiting some overlap with all environmental issues covered in this section except for ozone depletion, whereas the `endangSpec` variable has three blank cells in the Jaccard similarity table (Figure 4.4 above). Still, these two are each other's most overlapping variable with a Jaccard similarity of 0.35, a relatively low number compared to some of the other pairs.

Table 4.12: Pests & weeds regex counts

Year	Name	Count	Fraction
1950	Birds Protection Convention	1	7.5
1965	Maritime Traffic Facilitation Convention	4	4.4
1972	London Dumping Convention	1	2.3
1979	Migratory Species Conservation Convention	2	3.7
1982	UNCLOS	2	0.3
1992	Convention on Biological Diversity	1	1.1
1994	WTO Agreement	27	2.3
1994	Convention to Combat Desertification	1	0.6
1997	Non-Navigational Uses of Intl Watercourses Conv.	2	3.6
1997	International Plant Protection Convention	66	114.5
1998	PIC Convention	2	2.4
2001	Ship Anti-Fouling Systems Convention	1	1.3
2001	Plant Genetic Resources Treaty	2	2.1
2004	Ballast Water Management Convention	5	4.0
2005	International Health Regulations	6	3.0

In terms of other associations, when looking at the profile of the pests subset (in Figure 4.13), waste, land degradation and animal welfare are also much more prevalent in this subset than the full set average. Stressors are largely as excpected, although transport is only mentioned in 11 out of the 15 treaties (73%), a small increase compared to the full set. As for interventions, spatial regulation and risk assessment stand out with their two-fold increase in this subset, data/research is mentioned in all of these pest control treaties, and inspections and training/capacity-building occur in 60% of them. Regarding treaty provisions, all provisions in the compliance & effectiveness cluster except sanctions are more prevalent in this subset than in the full set (over 10 percentage points higher).

Care should be taken not to overinterpret these results, especially with hindsight,

Cluster	Variable	Full Set (n=119)	Subset (n=15)	Div	Diff
Problems & Risks	Climate Change	9 (8%)	1 (7%)	0.9	-0.9
	Ozone Depletion	5 (4%)	0 (0%)	0.0	-4.2
	Pollution	70 (59%)	10 (67%)	1.1	7.8
	Waste	34 (29%)	7 (47%)	1.6	18.1
	Land Degradation	12 (10%)	4 (27%)	2.6	16.6
	—Deforestation	6 (5%)	1 (7%)	1.3	1.6
	—Desertification	6 (5%)	3 (20%)	4.0	15.0
	Endangered Species	12 (10%)	7 (47%)	4.6	36.6
	Animal Welfare	6 (5%)	4 (27%)	5.3	21.6
Human Stressors	Fishing & Hunting	25 (21%)	7 (47%)	2.2	25.7
	Agriculture	36 (30%)	7 (47%)	1.5	16.4
	Production	76 (64%)	10 (67%)	1.0	2.8
	Storage & Stockpiling	62 (52%)	9 (60%)	1.2	7.9
	Transport	76 (64%)	11 (73%)	1.1	9.5
	Trade	58 (49%)	9 (60%)	1.2	11.3
	Weapons of Mass Destr.	13 (11%)	1 (7%)	0.6	-4.3
Means of Intervention	Quant. Restrictions	31 (26%)	4 (27%)	1.0	0.6
	Spatial Regulation	46 (39%)	11 (73%)	1.9	34.7
	Risk/Impact Assessment	47 (39%)	11 (73%)	1.9	33.8
	Inspections	49 (41%)	9 (60%)	1.5	18.8
	Data & Research	99 (83%)	15 (100%)	1.2	16.8
	Education	48 (40%)	8 (53%)	1.3	13.0
	Training/Capacity-build.	63 (53%)	9 (60%)	1.1	7.1
	Tech Transfer	37 (31%)	7 (47%)	1.5	15.6
	Insurance	27 (23%)	4 (27%)	1.2	4.0
	Liability	44 (37%)	6 (40%)	1.1	3.0
	Criminal Responsibility	16 (13%)	1 (7%)	0.5	-6.8

Figure 4.13: Subject matter of treaties on pests & weeds

but the subset does seem to have a distinct profile despite the few mentions to pest control in nearly all of the treaties.

4.1.7   Animal welfare

This variable is concerned with health and wellbeing of both wild and domesticated animals from an animal welfare perspective. Countless human activities can cause animal suffering or nuisance, including farming, fishing, hunting, tourism, shipping, diving, circus operation, science, transport, construction and mining. While animal welfare concerns are not prominent in international law, six of the 119 agreements do mention them at least on the margins.[342] The regex pattern is shown in Table 4.13 and the results are listed in Table 4.14 below.

---

[342] E.g. the International Plant Protection Convention (n 161), while ostensibly about plant protection, mentions "internationally approved principles governing the protection of plant, human and animal health, and the environment" in its preamble.

Table 4.13: Animal welfare regex pattern

Variable	Regular expression pattern
animalWelfare	`(?i)(?:animal (?:welfare|health)|welfare of \w* ?(?:living` `\w+|animal|livestock|pet)[s]?|(?:animal|living` `specimen|wildlife|fish|whale|bird)[^.]{0,100} (?:cruel` `treatment|care[d]?  for)|(?<!superfluous injury or )unnecessary` `suffering|end the suffering|harassing|(?<!veterinary) treat\w+ of` `(?:fish|birds|whales|animals)|conditions in which \w+ may be` `transported or \w+ in captivity)`

This regular expression is the fourth longest in terms of character count, and by far the longest (most inefficient) when dividing this count by the number of treaties (6) or the number of mentions (20) found with it. This is partly because terms like 'welfare', 'wellbeing', '(mis)treatment', 'care' or 'suffering' are equally applied to humans in labour law, human rights, humanitarian and criminal law contexts. Therefore, additional descriptors are needed before and after these keywords to avoid overinclusion, making the expression longer. These tweaks to capture all and only animal welfare concerns also make the regex less readable and probably less generalisable. There would certainly be ways to improve it, but for present purposes it is sufficient.

Table 4.14: Animal welfare regex counts

Year	Name	Count	Fraction
1946	Whaling Convention	2	4.8
1950	Birds Protection Convention	2	15.0
1973	CITES	11	11.9
1979	Migratory Species Conservation Convention	1	1.9
1994	WTO Agreement	3	0.3
1997	International Plant Protection Convention	1	1.7

Most fisheries treaties are conspicuously absent. Large umbrella treaties like UNCLOS and CBD are also missing from this list. However, this is not entirely surprising as they are more about protection of the biosphere at an abstract level, and more about the survival of populations and ecosystems than about their wellbeing. From this perspective, CITES is an exception as it is primarily focused on threats of extinction, while still racking up the largest count of animal welfare mentions of any treaty in the study set. CMS also sneaks in animal welfare concerns by including 'harassing' as part of its definition of 'taking'.[343] With a larger treaty sample, it

---

[343] Migratory Species Conservation Convention (n 190), art I(1)(i) ("'Taking' means taking, hunting, fishing, capturing, harassing, deliberate killing, or attempting to engage in any such

would be interesting to distinguish animal suffering from physical health concerns, which may be motivated by human health or economic considerations.

As for associations with other variables, animal welfare is only significantly associated with endangered species and pest control treaties (5-6 times higher prevalence). It is the most distant subset from the other environmental problem domains as gauged by the Jaccard similarity metric (in Figure 4.4 above).

4.2   Human contribution to problems

This section is concerned with human activities contributing to environmental problems and risks. The activities and processes governed by treaties may not coincide with direct stressors, or they may not represent the most important causal factors in the problems targeted. This is because not all human behaviour is equally governable. Political, legal, social, economic, moral and practical considerations preclude or hamper (international) governance of certain activities. Less readily governable or enforceable activities and processes may still be mentioned but not constitute the main focus of a treaty. In some cases the main reason for governing an activity may be to monitor and assess the state of the environment or of more direct human stressors. Thus, one of the most important tasks of treaty negotiators is identifying activities that *can* be governed, and that are as effective and efficient as possible in addressing the environmental problem at hand, whether directly or indirectly.

As the governance of these activities merely possesses instrumental, not intrinsic, value, the variables of this cluster will not be analysed individually like in the previous section, but instead jointly as a group.

4.2.1   Selection of stressor variables

The selection of activities to include was guided by a life-cycle approach to environmental regulation, i.e. ideally covering the full spectrum of processes from extraction, to transformation, trade, consumption, disposal, and everything in-between. The general criteria outlined in the introduction of this chapter imposed some constraints, which meant excluding or narrowing the scope of some of the variables initially designed for this section. For instance, mining and other inorganic resource extraction was supposed to be part of the first set of variables on environmental extraction, together with fishing, hunting, forestry and farming. However, the ini-

---

conduct').

tial regex pattern only matched three treaties,[344] and some of the false negatives, like outer space treaties, do not use convenient terms like 'mining' but instead much broader ones that could also apply to organic resources, like 'exploitation' or 'appropriation'.[345] This kind of terminology makes regex crafting more effortful and inefficient if the result only comprises a handful of agreements.

As for the extractive activities that did make it into the selection, fishing, hunting, gathering, trapping and similar killing and catching of wildlife is combined into one variable (dominated by fisheries). Agriculture, construed broadly so as to encompass forestry, aquaculture and any other growing or breeding of natural resources in a defined territory, represents the other variable. References to the Food and Agriculture Organization (FAO) or its Constitution are excluded from the agriculture search pattern, because a number of fisheries treaties refer to the FAO in the context of treaty depositary or secretariat provisions and would thus have counted as mentioning agriculture when they really only meant the organisation. The FAO Constitution itself includes fisheries and forestry in its definition of agriculture,[346] but having a separate category for fisheries seemed more interesting for present purposes.

The second stage is the processing of raw materials into intermediate or final products. This was one of the more difficult variables to extract accurately, because 'production' and 'processing' have countless meanings, and only relying on clearer terms like 'manufacturing' or 'industrial processing' would have been insufficient. As environmental impacts of human activities is what the present thesis is interested in, natural/nonhuman production processes are excluded, as are processes with low environmental impacts, such as the production of information, including reports, documents, evidence, records etc.[347] Also, care was taken to avoid overlaps with the previous two variables as far as possible. Hence, crop production falls under agriculture, whereas clothes production would fall under this one. Energy production is included in this variable as well, but should probably be captured separately and

---

[344] UNCLOS (n 92), Safety and Health in Mines Convention (n 240), and Minamata Convention on Mercury (n 194).

[345] E.g. Moon Agreement (n 254), art 11(5) ("States Parties to this Agreement hereby undertake to establish an international regime, including appropriate procedures, to govern the exploitation of the natural resources of the moon as such exploitation is about to become feasible").

[346] Constitution of the Food and Agriculture Organization of the United Nations (adopted 16 October 1945, entered into force 16 October 1945) ATS 1945 9, art XVI ('In this Constitution the term "agriculture" and its derivatives include fisheries, marine products, forestry, and primary forestry products').

[347] To be fair, information production in a climate-controlled office probably has a higher environmental impact than traditional subsistence farming, but subsistence agriculture/fisheries is typically excluded from the scope of treaties and information production is not regulated as an activity per se, only with regard to certain types of rights and obligations.

further subdivided in future work.

The third phase of the life-cycle of a product is storage & stockpiling, be it at rest or on the move. It may seem like a minor aspect, but is in fact mentioned by slightly more than half of all treaties in the study set, in many different contexts and for many different reasons, from health and safety, to economic stability and international security. Again, natural storage/stocks/reserves are excluded, such as carbon storage by trees or fish stocks in the high seas.

Transport and trade are the obvious next types of activities to examine, and while they are of course connected, the overlap is nonetheless not complete. Trade exclusively refers to international trade in the context of this thesis, whereas transport could be domestic. Environmental transport without human intervention (e.g. through the atmosphere) is excluded, but any other kind of intentional or accidental transport or delivery of materials, energy, organisms, freight or passengers by any means, even spaceships, is included in the transport variable. Trade in turn is more about sale of goods and services than about movement per se, and either one or both of import and export could be referred to by a treaty. References to illicit trade such as human or animal trafficking, or the unlawful acquisition of nuclear weapons by non-nuclear-weapon states are included in the trade variable to make it as broad as possible. Subdivisions will certainly be required for more in-depth analyses of different types of trade referred to and governed by international agreements.

Consumption/use would have been the seventh stressor variable introduced here, but regex trials failed to reach a satisfactory accuracy level.[348] The same holds for dumping/disposal and release/discharge/emissions, which were modelled as the intentional or unintentional (re-)introduction into the environment of waste and other products generated by humans. Perhaps other languages would be better suited for extracting information on some of these variables, as e.g. 'release', 'discharge', 'dispose of' and suchlike have many different meanings in treaty texts.

Armed conflict was also considered as a variable to construct a regex pattern for, but as it is often referred to in the context of exclusion from the scope of a treaty (or from provisions on liability), weapons of mass destruction (WMD) were chosen instead. It is of course mostly the testing and use of such weapons that has major environmental impacts, but the extraction, production, storage and transport of WMD are not without hazards, and the effects of even a single use are such that it is worth capturing every mention thereof in treaties. References to WMD and to

---

[348] Consumption alone is easy enough to extract ('consum'), but 'use' is ubiquitous and would take a lot of effort to disambiguate. They are used interchangeably for some relevant meanings in the treaty set, and thus matching only the former would not be representative of the underlying concept.

each individual type of weapon are extracted, obtaining 13 treaties. For comparison, a regex matching every mention of 'nuclear' would yield twice as many treaties, but this includes civil nuclear law. From an environmental impact and treaty design perspective, it seems best to distinguish civil from military uses of nuclear energy.

4.2.2   Descriptive statistics of stressors

109 out of 119 treaties mention at least one of the seven stressors, and two agreements cover six out of seven,[349] but the average is a bit over three. The (chronologically) first treaty in this subset is the FAO Constitution,[350] adopted in 1945, mentioning fisheries, agriculture and production.

Table 4.15 displays the number of treaties matching each pattern, the percentage of the full treaty set that this number represents, the mean, median and maximum proportional count and the maximum absolute count of mentions found in any given treaty. The two maxima do not necessarily correspond to the same treaty. For instance, the treaty with the maximum absolute count of trade mentions (or any other of the extracted regex patterns for that matter) is the WTO Agreement, while the maximum proportional trade count stems from the Grains Trade Convention.[351] The trade regex pattern also produced the highest sum of absolute and proportional counts within this cluster, each about twice as large as the next largest sum.

Table 4.15: Stressor variables summary stats

Variable	N.Treaties	Mean P	Median P	Max P	Max N
Fishing & Hunting	25 (21%)	6.4	0.0	233.9	202
Agriculture	36 (30%)	6.2	0.0	156.1	141
Production	76 (64%)	6.9	2.0	83.7	306
Storage & Stockpiling	62 (52%)	3.7	0.4	34.3	116
Transport	76 (64%)	7.4	2.3	112.2	93
Trade	58 (49%)	13.2	0.0	141.6	1450
Weapons of Mass Destr.	13 (11%)	4.0	0.0	115.2	498

Despite the wide distribution of absolute and proportional counts, the mean is very similar for four of the seven variables. The median is close to zero for most of them, as expected.

As for the distribution over time, Figure 4.14 presents the aggregated proportional counts by treaty adoption year, showing a clear peak in 1993 and 2001. The number of treaties adopted is not the primary driver of these peaks. In fact, the year

---

[349]   WTO Agreement (n 93), and Chemical Weapons Convention (n 216).
[350]   FAO Constitution (n 346).
[351]   Grains Trade Convention (n 150).

with the most new agreements is 1997 (6), followed by 2001 (5), and only then 1993 and 1994 (4 each).[352] The spike in 2001 is driven by the adoption of the two most agriculture-focused treaties of the set, the Plant Genetic Resources Treaty for Food and Agriculture[353] and the ILO Convention on Safety and Health in Agriculture,[354] with a relative count of 148.1 and 156.1 respectively. 1994 also has a disproportionately large contributor, trade in this case, explained by the adoption of two heavily trade-focused treaties in the same year, the Marrakesh Agreement and the Grains Trade Convention mentioned earlier. In a chart of absolute counts, the stacked bar of 1994, and in particular the trade component, would have dwarfed most others to the point of barely being visible. The highest bar of the present plot, the year 1993, is more multi-dimensional. It reaches such heights because it features the largest proportional count of fisheries mentions,[355] the fourth largest trade count,[356] and the second highest WMD count,[357] among other contributions.

Figure 4.14: Stressor counts per treaty adoption year

Overall, there is definitely more intense treaty activity/focus on these key economic activities and WMD in the early 1990s, after the dissolution of the Soviet Union.

---

[352] 1971, 1977 and 1989 also saw the adoption of four treaties from this cluster.
[353] Plant Genetic Resources Treaty (n 224).
[354] Safety and Health in Agriculture Convention (n 99).
[355] Fishing Vessels Compliance Promotion Agreement (n 161).
[356] International Cocoa Agreement (n 150).
[357] Chemical Weapons Convention (n 216).

Turning now to associations between the variables forming this cluster, Figure 4.15 displays the Jaccard similarity coefficient between each pair of variables.

	fishHunt	agriculture	manufact	storage	transport	trade	wmd
fishHunt	1.000	0.151	0.202	0.176	0.217	0.169	0.000
agriculture	0.151	1.000	0.333	0.256	0.273	0.382	0.021
manufact	0.202	0.333	1.000	0.586	0.520	0.473	0.085
storage	0.176	0.256	0.586	1.000	0.568	0.446	0.103
transport	0.217	0.273	0.520	0.568	1.000	0.457	0.099
trade	0.169	0.382	0.473	0.446	0.457	1.000	0.092
wmd	0.000	0.021	0.085	0.103	0.099	0.092	1.000

Figure 4.15: Stressors' Jaccard similarity

Most pairs of environmental stressors have less than 50% overlap. The least intersecting variable is WMD, having no treaty in common with the fishHunt variable and only one with agriculture,[358] constituting an over 20 percentage point drop compared to the prevalence of those two in the full set. The decrease in probability of production being mentioned in WMD treaties (-10 p.p.) may seem a bit more surprising at first glance, but less so when considering that several WMD treaties are purely focused on the prohibition of testing, emplacement or use of such weapons in common areas.[359] The low Jaccard index for WMD treaties compared to other stressors is also due to the fact that this is by far the smallest subset in this cluster, and thus the intersection divided by the union with some of the larger subsets cannot reach that of classes of similarly large size. In fact, the prevalence of storage, transport, and trade in the WMD subset is very close to the full set average at 0.54, 0.62 and 0.46 respectively.

The most overlapping pairs of variables are storage and production (59%), storage and transport (57%), and production and transport (52%). But production, transport and storage are also the three most frequently mentioned stressors overall, so their overlap is not particularly surprising. It may be worth introducing another similarity measure that takes into account the independent probability of the variables. The χ^2 statistic and its test of independence is a relatively straightforward and well-established measure for determining associations between categorical variables. To illustrate the principle, Table 4.16 contains the 2x2 contingency table for the binary agriculture and trade variables, with marginal totals.

Each of the four central cells shows the result of the absolute count minus the expected value under an assumption of independence. As we saw before, agriculture

---

[358] The Chemical Weapons Convention mentions agriculture as a peaceful purpose/use-case of chemicals.

[359] E.g. Outer Space Treaty (n 161), Moon Agreement (n 254), and WMD Seabed Emplacement Prohibition Treaty (n 161).

Table 4.16: Observed vs. expected values for trade and agriculture mentions

	Agriculture	No agriculture	Total
Trade	26 - 17.5 = 8.5	32 - 40.5 = -8.5	58 (49%)
No Trade	10 - 18.5 = -8.5	51 - 42.5 = 8.5	61 (51%)
Total	36 (30%)	83 (70%)	119 (100%)

is mentioned in 30% of the agreements and trade is mentioned in 49% of them. These are the independent probabilities of occurrence of these terms in any randomly drawn treaty from the full set. If the two variables were not related, then their prevalence in each other's subset would mirror the prevalence in the full set. This is the same reasoning as we used when looking at conditional probabilities earlier. Here, the difference is that for each of the four possible scenarios, we calculate the expected value based on the independent probability. Thus, for the top left cell representing the treaties where both terms are present, the expected value can be calculated by multiplying 36 (the total number of treaties mentioning agriculture) with 49%. Basically, trade should be mentioned in about half of the agriculture treaties, since that is its normal prevalence. In reality, however, trade is mentioned in 26 of the agriculture treaties, and only absent in 10. The difference between the observed and expected value is ± 8.5 in all four cells. For the χ^2 statistic, this difference is squared and then divided by the expected value of the respective cell. Finally, the four results are summed up to obtain an aggregate score. In short, this is the formula:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

For the trade-agriculture pair, this sum is 11.4, whereas for the production-transport pair it is only 1.9. It is 0 if and only if there is no difference between any of the observed and expected values. The related phi coefficient is used to estimate and compare effect sizes.[360] Its formula is $\phi = \sqrt{\frac{\chi^2}{N}}$ where N is the total number of observations (119). In the given examples, it is 0.31 for the former pair and 0.13 for the latter, corresponding to a moderate and weak effect respectively, when following a standard interpretation. At this compressed level, the measure is easier to compare to the the Jaccard similarity coefficient which is 0.38 for the former and 0.52 for the latter pair, i.e. the order is reversed. There are many possible binary similarity measures to choose from.[361] One criterion to consider

---

[360] It was originally introduced by Karl Pearson but is also called the Yule phi coefficient or the Matthews correlation coefficient, depending on the field. In the 2x2 case, it is equivalent to Cramér's V.

[361] Choi, Cha, Tappert, and others (n 253) present and compare 76 such measures.

is whether the bottom right cell should be taken into account in the calculation. The Jaccard coefficient does not, whereas χ^2 based measures do. In the context of this exploratory analysis, I find them to be complementary, although subset-specific tables of conditional probabilities such as the ones presented in some of the previous sections seem more easily interpretable than the chi-square score. Moreover, χ^2 discards the sign of the association by squaring the difference. An overview table of chi-square scores can still be useful to quickly spot the strongest associations between many pairs of variables, as illustrated e.g. in Figure 4.16 in the next section, but the details tend to require further investigation. As for the chi-square test of independence, it is a test of statistical independence used in hypothesis testing, not in exploratory analysis. There are ways to correct for multiple hypothesis testing (counting each pairwise comparison as a hypothesis test)[362] and doing so still yields some significant results when applied to the present dataset. The test with suitable corrections and resulting p-values can be useful in order to avoid paying attention to coincidental correlations, but should not be overinterpreted.[363]

Returning to associations between stressor variables, the two pairs with the highest Jaccard similarity in fact also have the highest χ^2 scores with the lowest p-values.[364] Trade and storage also exhibit some dependence, representing an 11 percentage point increase of the likelihood of trade being mentioned in storage treaties, and a χ^2 score of 6.2 ($p = 0.02, \phi = 0.23$) but this is not statistically significant when applying the above-mentioned corrections, whereas trade and agriculture is.[365]

4.2.3 Associations with other variables

Associations with environmental issues have been discussed in problem-specific sections (under 4.1) already. The aim here is mainly to provide an overview at the cluster level. Figure 4.16 presents the χ^2 score for each pair of variables.

---

[362] For an accessible introduction see e.g. JH McDonald, *Handbook of Biological Statistics* (3rd edition, Sparky House Publishing 2014), pp. 254-260, also available at www.biostathandbook.com/multiplecomparisons.html (accessed 1 Dec 2022).

[363] See e.g. Todd Michael Franke, Timothy Ho, and Christina A Christie, 'The Chi-Square Test: Often Used and More Often Misinterpreted' (2012) 33(3) American Journal of Evaluation 448.

[364] χ^2 of 18.97 and 15.79 with p-values of 0.000013 and 0.000071 respectively. Both are significant at the Bonferroni-corrected significance level of $\frac{0.05}{378} = 0.00013$. The phi coefficients, 0.40 and 0.36, suggest a moderate to relatively strong effect.

[365] More specifically, the association between trade and agriculture is statistically significant under the Benjamini-Hochberg procedure, but not under the more conservative Bonferroni correction, because the p value of 0.0007 is higher than the corrected significance level of 0.0001.

The two variables with the strongest pairwise associations are fishHunt and waste (with other variables), yielding five statistically significant χ^2 scores ($p < 0.005$) with moderate to relatively strong effect sizes ($0.25 < \phi < 0.45$). Agriculture also has clear links to land degradation, pollution and climate change, but while the respective phi coefficients are around 0.23 and the p values well under 0.05, only the association with land degradation is statistically significant under the applied corrections for multiple comparisons ($p = 0.0038, \phi = 0.27$).[366] Climate change and transport also just about reaches the threshold of significance under the Benjamini-Hochberg procedure ($p = 0.0068, \phi = 0.25$).[367]

	fishHunt	agriculture	manufact	storage	transport	trade	wmd
climateChange	0.009	6.119	2.642	1.374	7.316	0.181	1.194
ozoneDeplet	0.003	2.189	2.953	0.130	0.589	2.041	0.640
pollution	2.269	6.273	0.252	5.927	4.214	1.350	1.974
waste	0.856	0.016	9.471	21.015	15.385	0.030	0.216
landDegrad	1.292	8.387	0.717	1.884	2.850	0.492	1.637
deforest	0.072	3.971	3.575	0.011	2.552	0.813	0.775
desertif	0.072	3.971	0.021	0.011	2.552	0.004	0.775
endangSpec	16.765	0.060	1.112	0.582	0.717	0.267	1.637
pests	6.809	2.192	0.058	0.429	0.667	0.871	0.320
animalWelfare	14.790	3.971	0.526	0.892	1.038	0.813	0.775

Figure 4.16: Problem–Stressors' χ^2 scores

For contrast and comparison, Figure 4.17 shows the same overview table for the Jaccard similarity coefficient.

	fishHunt	agriculture	manufact	storage	transport	trade	wmd
climateChange	0.062	0.154	0.104	0.044	0.024	0.081	0.000
ozoneDeplet	0.034	0.079	0.066	0.047	0.052	0.068	0.000
pollution	0.234	0.165	0.460	0.483	0.521	0.320	0.137
waste	0.180	0.167	0.358	0.433	0.392	0.227	0.068
landDegrad	0.028	0.200	0.114	0.057	0.060	0.111	0.000
deforest	0.033	0.105	0.079	0.046	0.025	0.067	0.000
desertif	0.033	0.105	0.051	0.046	0.025	0.049	0.000
endangSpec	0.276	0.091	0.073	0.072	0.114	0.077	0.000
pests	0.212	0.159	0.123	0.132	0.137	0.141	0.037
animalWelfare	0.192	0.105	0.038	0.030	0.065	0.067	0.000

Figure 4.17: Problem–Stressors' Jaccard similarity

---

[366] This is not significant at the Bonferroni-corrected level, but is under the Benjamini-Hochberg procedure.

[367] Yoav Benjamini and Yosef Hochberg, 'Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing' (1995) 57(1) Journal of the Royal Statistical Society 289, as implemented by Python's Statsmodels library, introduced in Skipper Seabold and Josef Perktold, 'Statsmodels: Econometric and Statistical Modeling with Python' (2010).

Of note are the six darker cells in the middle of the table, due to the high overlap between the more widely distributed variables of pollution and waste on the problem side (rows) and manufacturing, storage and transport on the stressors side (columns). WMD stands out in how little overlap it has with environmental problem variables. Only three of them are mentioned in WMD treaties at all, although pollution does have a higher than normal prevalence in WMD treaties (1.3x more, or 18 percentage points higher), which does not really transpire in either of these two aggregate similarity measures.

As for associations between stressors and means of intervention, Figure 4.18 displays the chi-square results in like fashion. Here the noteworthy couplings are a bit more distributed. Nine χ^2 scores are statistically significant under the Benjamini-Hochberg procedure, ranging from WMD–criminal responsibility ($\chi^2 = 7.85, p = 0.005, \phi = 0.26$) at the lower end to transport–spatial regulation ($\chi^2 = 14.2, p = 0.00016, \phi = 0.35$) at the higher end. Without correcting for false discoveries, there would be another 18 pairs of variables reaching the significance threshold, with production, storage and WMD each having four significant associations overall.

It is worth reiterating that the χ^2 test attributes equal weight to the four cells of the 2x2 contingency table, and that it does not express the direction of the (positive or negative) association. When considering percentage point differences between subset and full set proportions, most stressors have high positive values, with the exception of the WMD variable, which is more polarised. For instance, the proportion of inspection treaties in the WMD subset is 28 percentage points higher than in the full set, while the prevalence of training/capacity-building treaties is 30 p.p. lower. Insurance is not mentioned at all in WMD treaties (-23 p.p.), and education and risk assessment are only mentioned in three of them (-17 and -16 p.p. respectively). Unsurprisingly, criminal responsibility is three times more prevalent in WMD treaties (+25 p.p.). Thus, there is certainly ample material for more in-depth investigations and analyses than is possible here.

4.3   Means of intervention

Once the causal mechanisms of the environmental problem and its anthropogenic drivers are understood and an initial set of intervention targets is identified, the question of the tools and techniques of regulatory intervention becomes most pressing. Many classifications of means or types of regulatory intervention in MEAs have been proposed in the literature, but most are very high-level. The approach taken here is more inductive. All interventions stipulated in treaties in a reasonably un-

	fishHunt	agriculture	manufact	storage	transport	trade	wmd
quantLimits	3.197	0.080	9.804	8.199	5.114	1.459	0.067
spatialReg	2.377	0.197	2.014	9.165	14.217	0.944	0.346
riskAssess	3.608	0.008	9.711	4.282	0.599	0.616	1.646
inspect	4.630	3.826	0.075	12.469	3.329	1.154	4.742
research	0.523	0.314	0.819	1.411	0.392	0.135	0.410
education	0.001	4.968	2.856	0.142	2.022	0.360	1.806
training	0.633	0.602	6.688	1.364	0.730	0.393	5.225
techTransfer	0.141	4.295	4.901	0.466	0.068	3.872	1.545
insurance	0.807	0.157	0.640	0.717	2.929	8.972	4.283
liability	1.094	2.326	0.126	3.723	12.375	1.824	0.014
criminalResp	0.057	0.460	0.191	0.128	0.993	0.012	7.848

Figure 4.18: Intervention–Stressors' χ^2 scores

ambiguous way qualify for inclusion in the treaty toolkit. The aim at this point is merely to inventorise what changes in human behaviour treaties purport to effectuate and how. The detailed formulations are often idiosyncratic and the most abstract categories (such as those of deontic logic) are of little use to evaluate and inform treaty design. The optimal degree of granularity probably lies somewhere in the middle.

Like in the previous section on stressors, the first subsection will explain aspects of the selection of variables (4.3.1), the second will introduce descriptive statistics and associations within this cluster (4.3.2), and the third will provide an overview of associations across clusters (4.3.3).

4.3.1   Selection of intervention variables

With regard to interventions, the limiting factor was not so much that there were not enough mentions to make their inclusion worthwile (as with some environmental problems like acidification or eutrophication), but that there were often too many to reach a satisfactory level of accuracy. Apart from the eleven interventions that form part of the present analysis, another nine advanced to a relatively late stage in the trial and verification process and were found in dozens of treaties. These include a financial subcluster with (i) subsidies or other financial incentives, (ii) fees, taxes or tariffs, and (iii) fines and penalties, both at the national and international level. Licences or permits for operation are mentioned in over 40 agreements, and 'registration' (other than in the sense of treaty registration) occurs in about 20, but the two overlap to some extent and may cover too much ground if joined. For instance, intellectual property registration should perhaps be separated from registration or licencing of physical property such as ships, vehicles, or objects launched into outer

space. Informed consent is also mentioned in eleven treaties, but has a similar issue of straddling different domains, in this case prior informed consent between states[368] and informed consent of individuals in criminal procedure or human rights law.[369] These meanings could easily be captured separately thanks to positive and negative lookarounds in regular expressions, but verification takes time and informed consent is mentioned in relatively few treaties compared to other interventions. At the other extreme, bans, prohibitions and moratoria are stipulated in over 60 treaties, but accurately capturing them seemed too effortful, given that legal English has developed many ways to express prohibitions. More specific concepts like 'quarantine' are easier to capture, but this one in particular occurs in only eight treaties which is below the threshold used for intervention variables. Future work certainly has a lot of potential avenues to explore that could be interesting for researchers and practitioners alike.

4.3.2 Descriptive statistics of means of intervention

All but three treaties mention at least one of the eleven means of intervention, and one agreement mentions all of them.[370] The median is four types of intervention per treaty (mean=4.26, mad=1.74),[371] but they are by no means equally distributed. Class sizes range from 16 treaties for criminal responsibility to 99 agreements mentioning data & research, as shown in Table 4.17. Data & research also has the highest mean and median proportional count of regex matches, but not the maximum. The flat distribution makes sense for this variable. After all, cooperation on data and research is important for many environment-related treaties, but is rarely the core focus of an agreement. CTBT produced the maximum proportional count for any variable[372] and the Chemical Weapons Convention the maximum absolute count,[373] both for the inspections variable and far above the respective maxima of other variables.

Some interventions probably tend to be mentioned less often even when they do form a central part of a treaty system's strategy for change. For instance, the Paris Agreement has the highest proportional count of mentions of technology transfer, with other climate and biodiversity treaties following closely, but even the maximum is only 12.3 (corresponding to 9 mentions in the text), a fraction of the sort of

---

[368] E.g. PIC Convention (n 271).
[369] E.g. Nuclear Terrorism Suppression Convention (n 310).
[370] UNCLOS (n 92).
[371] 'mad' stands for mean absolute deviation, see section 5.2.
[372] Comprehensive Nuclear Test Ban Treaty (n 166), with inspectionsP of 214.2.
[373] Chemical Weapons Convention (n 216), with inspectionsN of 967.

Table 4.17: Intervention variables summary stats

Variable	N.Treaties	Mean P	Median P	Max P	Max N
Quant. Restrictions	31 (26%)	2.2	0.0	90.2	139
Spatial Regulation	46 (39%)	2.7	0.0	77.8	141
Risk/Impact Assessment	47 (39%)	2.4	0.0	38.5	34
Inspections	49 (41%)	7.2	0.0	214.2	967
Data & Research	99 (83%)	12.7	8.9	89.7	240
Education	48 (40%)	2.9	0.0	51.3	30
Training/Capacity-build.	63 (53%)	3.5	0.8	37.8	70
Tech Transfer	37 (31%)	1.2	0.0	12.3	29
Insurance	27 (23%)	3.5	0.0	67.2	44
Liability	44 (37%)	7.0	0.0	87.1	50
Criminal Responsibility	16 (13%)	2.7	0.0	79.0	236

magnitudes reached by other variables.

Regarding trends over time, the peaks are a bit more spread out and multi-factorial than in the case of stressors, as evidenced by Figure 4.19. The 1990s were also a decade of substantial treaty-making activity for this cluster, but more in the second half than in the first. Both the peaks of 1993 and 1996 are largely driven by the two above-mentioned inspection maxima, although in the case of CTBT also reaching the maximum of the data & research variable in 1996 (thanks to its extensive monitoring system). The second highest sum of proportional intervention counts (in 1997) is made up of a larger range of treaties and variables—all but one variable (education) reaching double digits across a total of seven treaties adopted that year. 2001 was also an important year for environmental treaty-making, as already seen in the stressors section, but here almost half of the total is composed of insurance and liability references in the Bunker Oil Pollution Damage Liability Convention.[374] These two variables also feature prominently in the 1960s and 1970s, liability dominating the surge in 1971.[375] Naturally, for a more representative picture of treaty activity over time one would have to include analysis of amendments and supplementary agreements.

As for associations between intervention variables, liability is one of the rare variables having more negative than positive associations with other means of intervention. Risk assessment, data & research, education and training/capacity-building

---

[374] Bunker Oil Pollution Damage Liability Convention (n 114), with insuranceP of 67.2 and liabilityP of 69.5.

[375] Nuclear Material Maritime Carriage Liab Convention (n 203), with liabilityP of 87.1 (the max. proportional count for this variable), and Oil Pollution Damage Compensation Fund Convention (n 220), with liabilityP=57.5 and liabilityN=50 (the maximum absolute count for this variable). The other three treaties adopted in 1971 contribute a wider array of smaller values to the sum.
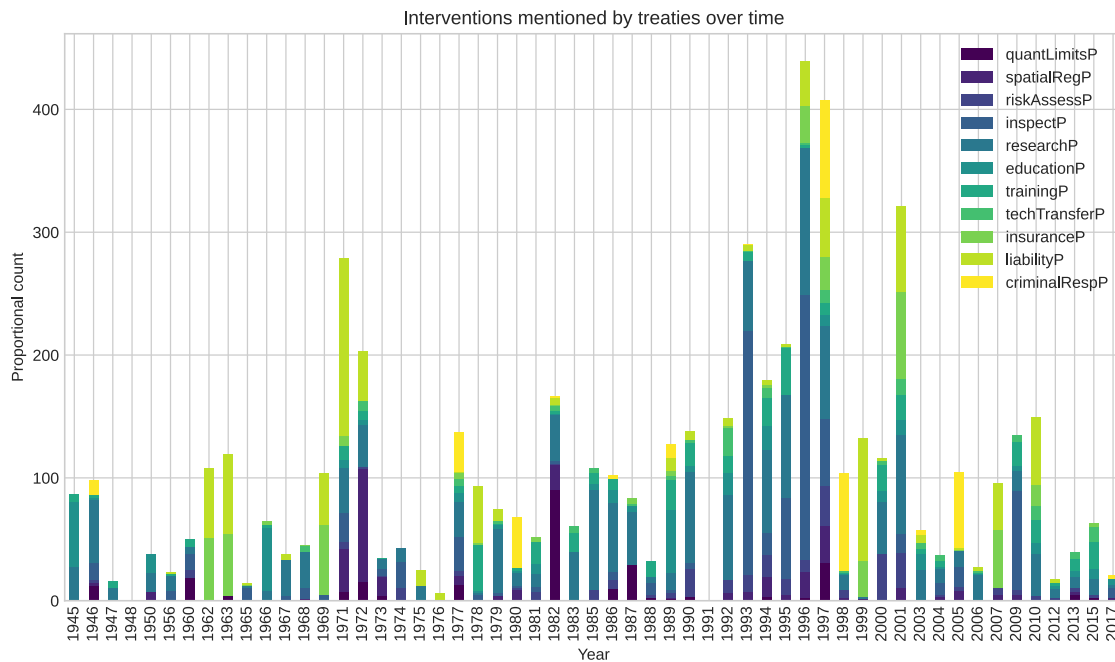
Figure 4.19: Intervention counts per treaty adoption year

are all 10 or more percentage points less prevalent in the liability subset than in the full set. Only the conditional probability of insurance is noticeably higher than average (+14 p.p.) in the liability subset. Still, agreements mentioning insurance only make up 36% of liability treaties, and conversely, the 16 treaties they have in common only make up 59% of insurance treaties (+22 p.p.).[376] This relatively low overlap between two legally linked concepts suggests that coding them separately was justified. The difference between the expected and observed values is not statistically significant under conservative continuity and false discovery rate corrections, but the variables are important for legal research and practice and merit inclusion from this standpoint alone.

Figure 4.20 shows that there are in fact a number of combinations that yield statistically significant results, most markedly with risk assessment and training/capacity-building, featuring five and four high χ^2 scores respectively.

4.3.3   Associations of interventions with other variables

Figure 4.21 presents the results of the χ^2 test of independence between environmental problem and intervention variables. The details have already been discussed in problem-specific sections, but some of the patterns may be better illustrated with this type of juxtaposition of clusters. For instance, the emphasis on 'softer' means

---

[376]   The Jaccard similarity index between the two is also rather low at 0.29.

	quantLimits	spatialReg	riskAssess	inspect	research	education	training	techTransfer	insurance	liability	criminalResp
quantLimits	nan	1.165	15.640	2.513	4.294	0.000	2.926	3.036	1.513	0.000	0.000
spatialReg	1.165	nan	7.972	4.521	0.384	0.132	2.446	1.691	0.228	0.039	1.632
riskAssess	15.640	7.972	nan	14.948	10.299	0.944	13.056	1.368	0.140	2.270	0.000
inspect	2.513	4.521	14.948	nan	3.462	0.231	2.894	0.488	1.355	1.020	1.090
research	4.294	0.384	10.299	3.462	nan	10.771	24.551	6.245	3.006	4.346	0.000
education	0.000	0.132	0.944	0.231	10.771	nan	31.907	11.986	0.000	4.127	4.912
training	2.926	2.446	13.056	2.894	24.551	31.907	nan	18.745	1.501	3.327	0.000
techTransfer	3.036	1.691	1.368	0.488	6.245	11.986	18.745	nan	0.179	0.000	0.093
insurance	1.513	0.228	0.140	1.355	3.006	0.000	1.501	0.179	nan	6.257	0.007
liability	0.000	0.039	2.270	1.020	4.346	4.127	3.327	0.000	6.257	nan	0.000
criminalResp	0.000	1.632	0.000	1.090	0.000	4.912	0.000	0.093	0.007	0.000	nan

Figure 4.20: Interventions' χ^2 scores

like education, training/capacity-building, and technology transfer among climate change, land degradation, deforestation and desertification treaties is immediately apparent from the higher χ^2 scores, in contrast to the over-representation of quantitative and spatial restrictions, risk assessments and inspections by agreements on pollution, waste and pests.

	climateChange	ozoneDeplet	pollution	waste	landDegrad	deforest	desertif	endangSpec	pests	animalWelfare
quantLimits	0.000	0.042	4.992	24.211	0.067	0.000	1.030	0.067	0.000	0.800
spatialReg	0.000	0.283	1.733	7.018	1.354	0.024	0.024	3.200	7.111	1.032
riskAssess	0.450	2.032	8.953	17.478	1.202	0.938	0.938	0.224	6.683	0.012
inspect	5.100	0.269	8.441	3.442	2.280	2.814	0.683	2.505	1.700	0.768
research	0.881	0.173	0.397	1.444	1.525	0.324	0.324	1.525	2.228	0.324
education	7.479	0.203	1.079	0.000	12.336	6.918	3.155	0.168	0.666	0.850
training	6.732	0.610	0.027	3.347	6.398	3.803	1.234	0.008	0.096	0.322
techTransfer	12.402	8.453	0.259	2.966	19.820	10.821	5.686	1.354	1.200	0.109
insurance	0.144	2.219	1.355	3.364	0.319	1.297	0.000	0.790	0.004	0.000
liability	0.353	1.630	0.000	0.001	0.000	0.389	0.000	0.349	0.000	0.000
criminalResp	0.521	0.053	0.353	0.002	0.010	0.142	0.142	0.000	0.175	0.000

Figure 4.21: Problem–Interventions' χ^2 scores

As for relationships between stressors and means, this inter-cluster overview has already been included in the relevant stressor section, and will thus not be repeated here.

5   *Forecasting entry into force*

This chapter is about prediction of entry into force based on treaty design features. As the long wait for entry into force of agreements such as the Kyoto Protocol and the Comprehensive Nuclear Test Ban Treaty has shown,[377] sufficient political support is a precondition to the achievement of any treaty objective. Two main types of indicators for political support were considered in the initial conception of this chapter. On the one hand there is treaty participation, and on the other hand there is resource allocation once the decision to participate has been taken. Such contributions may be formally committed to (e.g. the treaty organs' budget), or they may be ad hoc offers by states parties when a need arises. Given the scarcity of data on the latter type of indicator, for the present study only the former was analysed.

While states' participation in processes of negotiation, drafting and adoption of international agreements is certainly important, there is no structured data available on this for the treaties under consideration. As for individual acts of treaty participation, such as signature, ratification, and withdrawal, data published in UNTS was unfortunately found to be incomplete for a number of treaties and was therefore not included in the analysis either. This leaves treaty adoption and entry into force dates as the only reliable data points. Specifically, the length of the interval between these two dates was chosen as the prediction target. All else equal, a shorter incubation period is better than a longer one.

In brief, the first research question for this chapter is how accurately an agreement's entry into force date can be predicted given information about the treaty text and the adoption date. If the prediction is at least somewhat accurate, then the second question is whether there are any features that consistently increase or decrease treaty incubation time.

Unfortunately, unlike the chapters thus far, this chapter may not be accessible to readers without at least entry-level training in data science or machine learning.

---

[377]   A little over seven years in the case of the Kyoto Protocol (n 300), and 25 years and counting for the Comprehensive Nuclear Test Ban Treaty (n 166).

5.1   Data sourcing and preprocessing

Unlike agreement texts, treaty participation data and other metadata are published in machine-readable format in the online version of UNTS. These data were retrieved using programmatic means and merged with data harvested from the IEAdb. The latter data source proved useful for treaties not in UNTS yet for whatever reason (e.g. not in force or not registered yet). In cases of conflict between the two sources, UNTS prevailed on account of its authority.[378]

5.1.1   Preparation of prediction target

Figure 5.1 shows the frequency of the number of years[379] elapsed between adoption and entry into force for each treaty in the text set that has entered into force (n=113), with and without `EIFdelay` subtracted from the interval. The difference between the two time spans is not very large, on average less than half a year (full incubation mean=4.42, mad=3.03, shorter incubation mean=4.07, mad=2.96).[380]
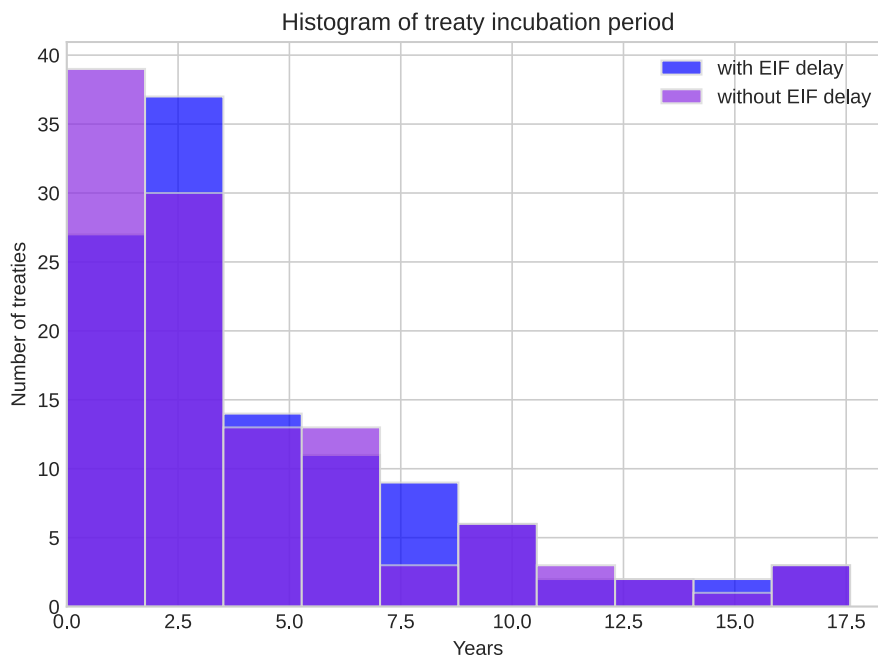
Figure 5.1: Treaty incubation years

There are several reasons to be interested in the shorter intervals. From the angle of political success, the time it takes for a treaty to meet the conditions for entry into

---

[378]   There were about a dozen discrepancies in reported treaty adoption date, and another dozen agreements studied here are not included in IEAdb at all, thus IEAdb participation data was not considered a proper alternative to UNTS for this chapter, only for filling missing values.

[379]   Technically, the number of days is measured and then converted into years for comparability with other time span variables.

[380]   'mad' stands for mean absolute deviation, see section 5.2.

force is arguably more important than the somewhat arbitrary length of the waiting period stipulated by the agreement. Especially in the lower quartiles—half of all incubation periods are below 3.08 years—a pre-agreed waiting period of one year would represent a large proportion of the time span.[381] It would in some sense be unfair to compare the entry into force speed of treaties without subtracting `EIFdelay`, because agreements with shorter or no waiting periods could seem more successful in galvanising action. This is assuming the length of the waiting period does not have a significant influence on states' decision to join. From a forecasting perspective, it also makes sense to subtract `EIFdelay` from the prediction target, because `EIFdelay` is a known quantity for treaties in the text set, i.e. a forecaster tasked with estimating the entry into force date of a treaty at the moment of adoption can already say with certainty that at a minimum, it will take the amount of time specified in `EIFdelay`. Thus, the target variable chosen for the analyses in this chapter is `incubConsent`, i.e. `incubYears` minus `EIFdelay`.

5.1.2 Selection and preprocessing of predictors

Three sets of predictors were included in the analysis, each with its own preprocessing needs and considerations. The first comprises the subject matter variables described in Chapter 4, here used in the form of binary variables coded as 0 when not found, and 1 when found in the treaty. While there are of course risks of false negatives, at least there are no missing values. The second set are the legal design variables covered in Chapter 3, which require a bit more preprocessing. All of the numerical variables have some missing values, which was perfectly fine for the more in-depth discussion in Chapter 3, but is not suitable for the present quantitative analysis. Rather than resorting to typical missing value interpolation strategies of filling with average values or suchlike, here the approach chosen is to fill the gaps with the values defined by customary international law as reflected in the VCLT. Thus, treaties which do not specify `EIFdelay` or `denunNotifMinYrs` are assigned zero for these variables, and `denunEffectYrs` is assigned 1, because 12 months is the notice period stipulated by Article 56 of the VCLT.[382]

`EIFnConsent` is a bit more complicated, because in principle the default would be two parties, but this is not appropriate for the treaty set at hand. In the information extraction algorithm for Chapter 3, `EIFnConsent` was marked as not available when-

---

[381] One year is the second most common EIFdelay chosen, see section 3.1.2.

[382] VCLT (n 18), art 56(2), assuming that the possibility to withdraw is given under paragraph 1 for those treaties without a provision on denunciation. As for the handful of agreements with a withdrawal provision but no mention of immediate or delayed effect of a notification, 12 months was assumed to be customary as well.

ever it could not be easily extracted from the text, e.g. when the number depends on some calculation of real-world economic or environmental impact of participants.[383] In some of these cases, a lower threshold for `EIFnConsent` (necessary but not sufficient) could have been extracted, but instead of going back and changing the information extraction algorithm, the six treaties in question are excluded from the present analysis. The variable `autoTermMemThresh` also has missing values but was dropped from the list anyway on account of its few instances.

The categorical variables from Chapter 3, namely `orgAuspices` and `compDispSett`, are one-hot encoded, whereby each of the values becomes its own boolean column, e.g. `orgAuspices_ILO` and `orgAuspices_UNEP`. Any of these binarised variables with less than ten instances is dropped from the feature set because it would probably add little predictive power to the model.

The third set is drawn from UNTS metadata. Specifically, adoption year is derived from the adoption date to account for possible temporal trends. Treaty adoption place is processed akin to the aforementioned categorical variables, thus reduced from 27 unique values to the four most frequent ones with 10 or more instances.[384] The number of official languages of the treaty text is also calculated from the information given in UNTS treaty pages, but it does not cover all the agreements in the already small set under consideration and was thus not used in the final analysis. The list and number of official languages could of course also be extracted from treaty texts but did not reach the threshold of expected impact and importance for the present dissertation.

Finally, variables with zero variance are removed from the set of predictors, hence excluding the binary variable `EIFprov` which is `True` for all treaties.

5.2   Predictability of entry into force

As mentioned in the introduction to this chapter, some degree of prediction accuracy is a precondition for the analysis of influential features. If a model is not a good fit for the data, then analysing its coefficients or weights is of little interest. For instance, Figure 5.2 displays the scatter plots of the two numeric EIF variables with respect to `incubConsent` with a simple linear regression line and corresponding equation.

One glance at the data suffices to see that bivariate linear models are a poor fit in this case. None of the other variables exhibit a clear linear relationship with the

---

[383]   See Section 3.1.2 on entry into force provisions.
[384]   Geneva: 25, New York: 16, London: 11, and Washington: 10.
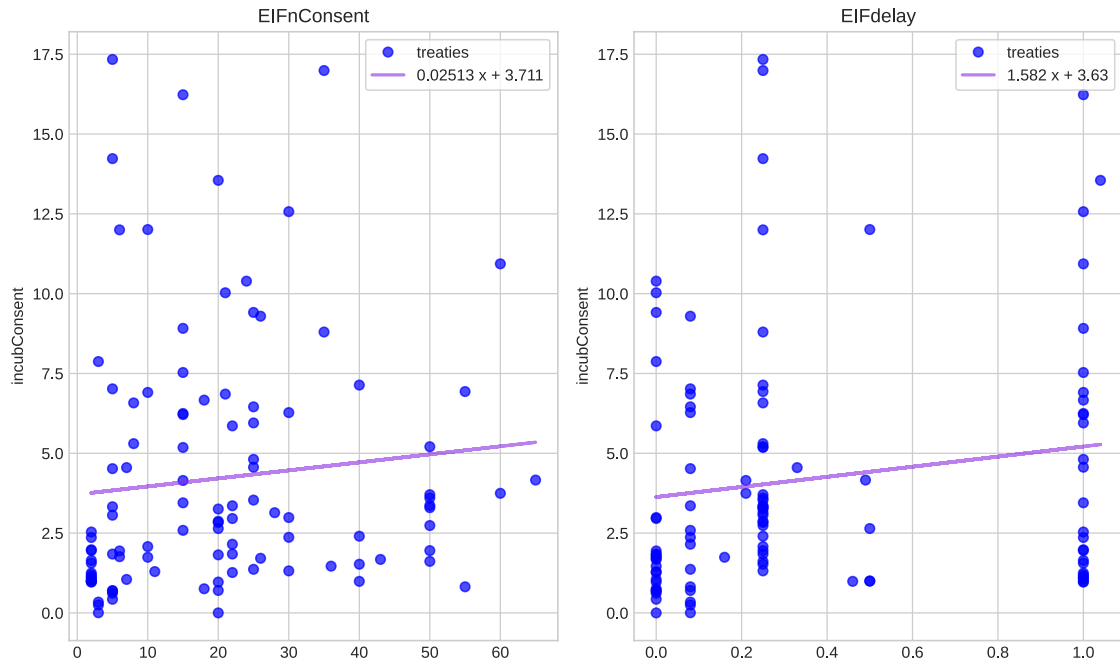
Figure 5.2: Regression of numeric EIF vars on incubConsent

target either, hence non-linear machine learning models such as random forests and neural networks were explored as well.

What counts as sufficiently accurate is an open question. In the absence of pre-existing benchmarks, the mean or median of the training set are good baselines. That is, when splitting the data into a training and a test set, as is standard practice in machine learning, the developed models should at least outperform a 'dummy' forecaster predicting the mean or median of the training set as the expected incubation time of all the treaties in the test set.

Also, note that none of the models described below are time series forecasting models. Developing proper forecasting models would have required not giving models access to any future data, e.g. by splitting the data into rolling or expanding windows of training and test data. However, as the overall dataset is rather small, this would have meant severely restricting the amount of data a model can learn from and be tested on. Future work could increase the dataset and implement a realistic time series forecasting model. This will become particularly salient once successive versions of the same treaty are included in the dataset, which was not the case here (see Chapter 2.2.1), because these successive versions likely have similar features and incubation times. For present purposes, prediction performance was merely a prerequisite for the more interesting analysis of feature importance, and it was crucial to be able to assess variability of results across many different random samples.

Another important question is what prediction error to measure and minimise within and across different models. The main choices to make are between absolute and squared errors, and between absolute and relative errors. Mean absolute error (MAE) was found to be the most intuitive and suitable for this dataset. Measures based on squared prediction errors, such as mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination R^2 attribute too much weight to outliers, which is unwarranted in this case. As can be seen in Figure 5.2, about a dozen agreements took ten years or more to enter into force, and maybe those are particularly hard to predict. There are so many extra-legal factors at play that getting a relatively accurate prediction for 75% of treaties would already be an achievement. If the predictions for the remaining quarter are far off the mark, that would be unfortunate but hardly impactful or costly in the real world. Relative error measures, such as mean absolute percentage error (MAPE), are useful when the range of values is large, but in this case the benefits probably do not outweigh the drawbacks (division by zero issues, less intuitive and communicable results etc.). Therefore, MAE is used as the main metric to select and evaluate models. As the prediction target is measured in years and not transformed or scaled in any way, MAE simply expresses the average number of years the forecast is above or below the true values. Similarly, mean absolute deviation (MAD) rather than standard deviation is used to report variability of forecast errors, because standard deviation relies on the quadratic loss.

Following best practice in machine learning for small datasets, a repeated nested cross-validation algorithm wis used to select and evaluate models. The rationale is that generalisation performance of a model can only be assessed with out-of-sample data, and if the model requires tuning of hyperparameters such as the regularisation penalty, neural network architecture, or learning rate, the same test set cannot be used for both selection of hyperparameters and evaluation of the model, because it would overfit on the test set. K-fold cross-validation helps by splitting the data into a number of non-overlapping sets, e.g. k=5 or k=10, and each subset acts as a test set once, such that in the end the average results will cover the entire dataset and account for variability in the composition of test sets. Nested or double cross-validation applies this idea recursively and hyperparameter tuning is performed in the inner, nested loop while evaluation occurs in the outer loop.

Specifically, using k=5, the full dataset is first split into five distinct and approximately equal-sized randomly sampled sets, with a test size of 20 odd treaties each. In the first iteration of the outer loop, one of these five subsets is designated as the test set and the remaining ~80 samples compose the training set. Every time

a train/test split is defined, the MAE of the mean and median predictors for the given test set is calculated as a baseline, and then the original numeric variables are scaled with z-score standardisation[385] to achieve zero mean and unit variance within the training set. The input variables of the test set are then scaled with the stored means and stds of the training set.[386] Binary variables are kept as is (0/1 values), standardising or centering them did not make a noticeable difference in performance during initial small-scale experiments.

Ordinary least squares (OLS) regression is performed directly in the outer loop as this type of model does not have hyperparameters to tune. For the other models we proceed to the inner loop, taking the training set of the given outer loop and splitting it again into 5 folds, each acting as the validation/development (dev) set exactly once,[387] and derive the baselines and z-scores again. Within this inner loop, each type of model undergoes hyperparameter tuning using exhaustive grid search to find the best hyperparameters for this particular training and dev set, and storing the results for each hyperparameter (or combination thereof for models with multiple hyperparameters). Once the inner loop is complete, the five result grids for each model type are searched to find the hyperparameters with the best average performance across all validation sets of the inner loop, and a new model with these hyperparameters is fit on the training set of the outer loop in question and evaluated on its test set. This entire procedure is repeated 10 times with different random seeds to gauge sensitivity to random factors such as random initialisation of neural network weights[388] and random data shuffling.[389] Linear models are implemented in Scikit-learn[390] and neural networks in Keras/Tensorflow.[391] The source code for this section is included in Appendix H.2 and an executable version can be downloaded from the thesis code archive.

Table 5.1 displays the results for all outer loop iterations, ordered by mean

---

[385] For each value in each variable, subtracting its mean and dividing by its standard deviation.

[386] This cannot be done across the full data set at the beginning because that way information from the test set would 'leak' into the training set.

[387] Inner loop validation sets are a bit smaller than outer test sets, containing approx. 17 treaties.

[388] He Uniform initialisation is used because it is suitable for Rectified Linear Units (ReLU), both chosen as part of the initial exploration and narrowing of the hyperparameter search space. See Kaiming He and others, 'Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification' in *International Conference on Computer Vision (ICCV)* (IEEE 2015) for initialisation details.

[389] See e.g. Christofer Fellicious, Thomas Weissgerber, and Michael Granitzer, 'Effects of Random Seeds on the Accuracy of Convolutional Neural Networks' in *International Conference on Machine Learning, Optimization, and Data Science* (Lecture Notes in Computer Science, vol 12566, Springer 2020).

[390] F Pedregosa and others, 'Scikit-learn: Machine Learning in Python' (2011) 12 Journal of Machine Learning Research 2825.

[391] The closest available DOI is that of TensorFlow v2.9.3 on Zenodo at https://doi.org/10.5281/zenodo.7604251. See the Table of Software in Appendix A for details.

test MAE. Note that multi-layer perceptrons (fully connected deep neural networks), while achieving best average performance, have slightly more variable results compared to Ridge and Lasso regression. However, the simple median predictor `medianPred` has higher test error dispersion than neural network based models, hence the distribution of outliers in training and test sets is clearly influential too.

Table 5.1: Outer loop test set MAE results

Regressor	Count	Mean	MAD	Min	25%	50%	75%	Max
MLP	50	2.60	0.51	1.61	2.10	2.55	2.94	4.37
Ridge	50	2.69	0.41	1.76	2.29	2.68	2.95	4.06
medianPred	50	2.82	0.54	1.69	2.29	2.76	3.33	4.85
Lasso	50	2.82	0.40	1.97	2.45	2.86	3.12	4.29
meanPred	50	3.07	0.36	2.15	2.69	3.09	3.32	4.62
OLS	50	6.05	0.81	3.54	5.41	5.92	6.52	9.28

These results show that the best models can predict entry into force of new treaties with a precision of about 2.6 years, whereas a dummy mean predictor would on average be off by 3 years. OLS stands out with its poor performance, whereas the other models are closer together. The best validation performance of the inner loops was slightly less than 1 year for MLP, 1.2 for Ridge and 1.3 years for Lasso regression, with the baselines nearly the same, thus cherry-picking the best model run would have produced more impressive but misleading results.

Another important factor to take into account when judging these results is that they are averages across the 21-22 treaties making up the respective test sets. Some test sets will contain more outliers than others, and thus it is perhaps more interesting to analyse prediction accuracy and its variability at the level of treaties. For the purpose of this investigation only the model type with the best average performance will be scrutinised, i.e. neural networks.

If all agreements were equally hard to predict, then the treaty-level mean absolute error and its dispersion would be the same across treaties. As expected, this is not the case here. Treaty test MAE ranges from 0.35 to 14.67 years and its variability from 0.18 to 1.37 years (mean absolute deviation). Figure 5.3 shows the two ends of the spectrum. The 10 best MAEs were all achieved with ILO conventions, and errors are consistently low across different model configurations and random initialisations for these agreements, whereas on the higher end there is more variance.

The issue of variability of model performance for the same treaty is also important for interpretability and practical utility of results. Indeed, it is probably preferable to have a few treaties with robustly high error than many treaties with a medium prediction error that is very sensitive to random factors like data shuffling
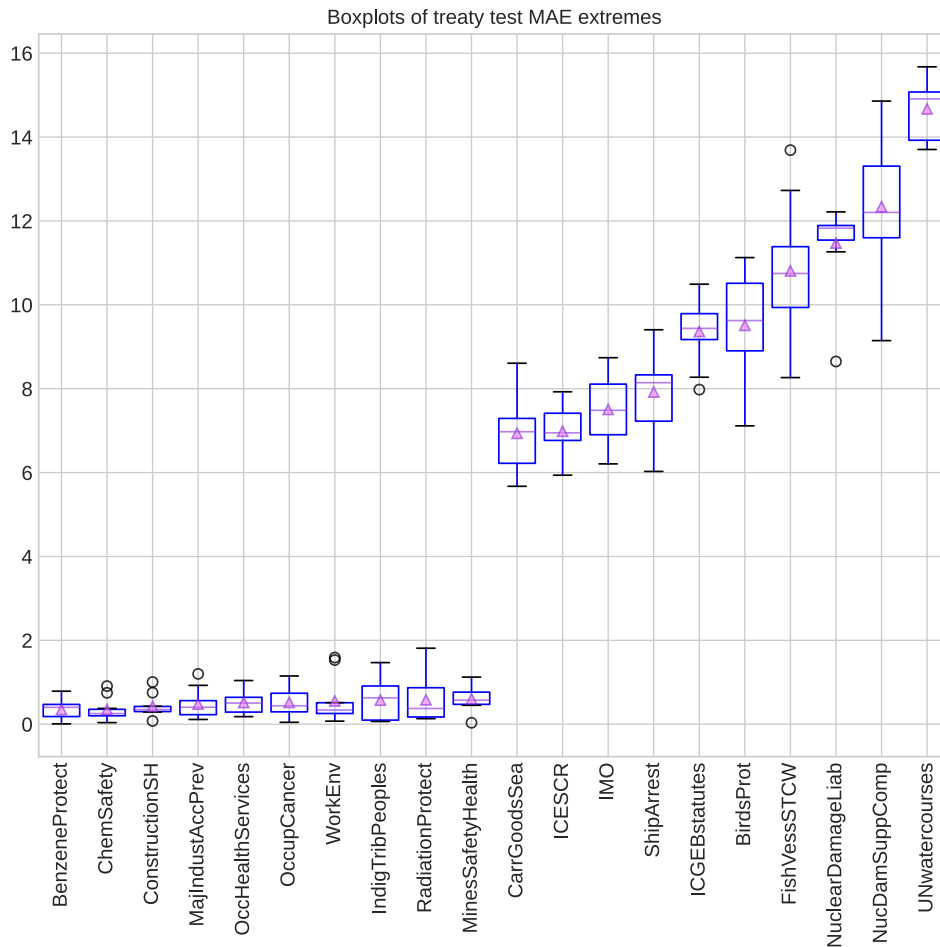
Figure 5.3: Treaty test MAE extremes

or network weights initialisation. For instance, the Vienna Convention on Civil Liability for Nuclear Damage (third from the right in Figure 5.3) took over 14 years to enter into force.[392] The predictions from the ten outer loop MLP models that had this treaty in their test set range from 2 to 5.6 years, with a mean of 2.76 and an average absolute deviation from this mean of 0.6 years. This is clearly visible in the boxplot: most of the predictions are clustered in a narrow range, with one outlier that has a lower MAE than the others, perhaps because more useful samples were included in its training set as compared to the other model runs. Still, at least it is consistently wrong and improving the prediction for this treaty may require collecting more data or choosing a different sampling scheme, such as stratified sampling, distributing relevant class members more evenly across training sets.

What may already have become apparent to readers familiar with these treaties is that `incubConsent` is highly correlated with MAE (0.84 Pearson correlation). The variability of the prediction errors is also somewhat correlated, but to a much lesser

---

[392] Convention on Civil Liability for Nuclear Damage (n 209).

extent (0.22). By way of illustration, the results for the three treaties with the highest error MAD are listed in Table 5.2, showing that `incubConsent` was actually rather low for these treaties, below average for two of the three, while the predictions are extremely variable. One of them, the London Dumping Convention,[393] is also the treaty with the largest absolute error range (6.63 years), even though its MAE is actually slightly better than average.

Table 5.2: Most variable treaty test MAE results

Treaty	Incub	MeanPred	MinPred	MaxPred	MAE	MAD	Range
DumpingConv72	2.59	5.08	2.44	9.37	2.53	1.28	6.63
AntiFoulSysConv	5.95	9.67	6.51	12.24	3.72	1.30	5.73
MicroorgDepositT.	3.06	7.11	5.07	9.86	4.05	1.37	4.79

When there is high variability across model runs, it is sometimes advantageous to create an ensemble model, e.g. by simply averaging the predictions of the individual models, and conduct the final evaluation with this result. This is indeed the case for the present dataset, decreasing the overall MAE from 2.60 to 2.49 years. However, interpretability would become harder if data from the individual models were discarded. The better approach is to examine the reasons for outliers, high variability or poor performance on certain treaties, and to improve data, training or model architecture accordingly.

One useful diagnostic is a full error chart, not only of absolute errors, because the trained models might be systematically over- or underestimating the target variable. Figure 5.4 displays all treaty test prediction errors (one boxplot per treaty, ordered by treaty adoption date). Negative errors prevail in magnitude but not in quantity. The overall mean prediction error is in fact close to zero, at -0.72.

It is also evident from this chart that there are only about a dozen influential outliers with exceedingly large absolute mean errors. Indeed, disregarding the twelve treaties with the highest MAE reduces the overall average testMAE of this class of models from 2.6 to 1.8 years, and retaining only the 75 best predicted agreements would halve it (1.3 ± 0.5 years).

Overall, predictability seems good enough to move on to the second research question for this chapter, namely whether there are any features that consistently increase or decrease treaty incubation time. This can be answered through feature permutation, an interpretability technique which involves changing (permuting) predictor values and recording how this change impacts the model's prediction.

---

[393] London Dumping Convention (n 302).
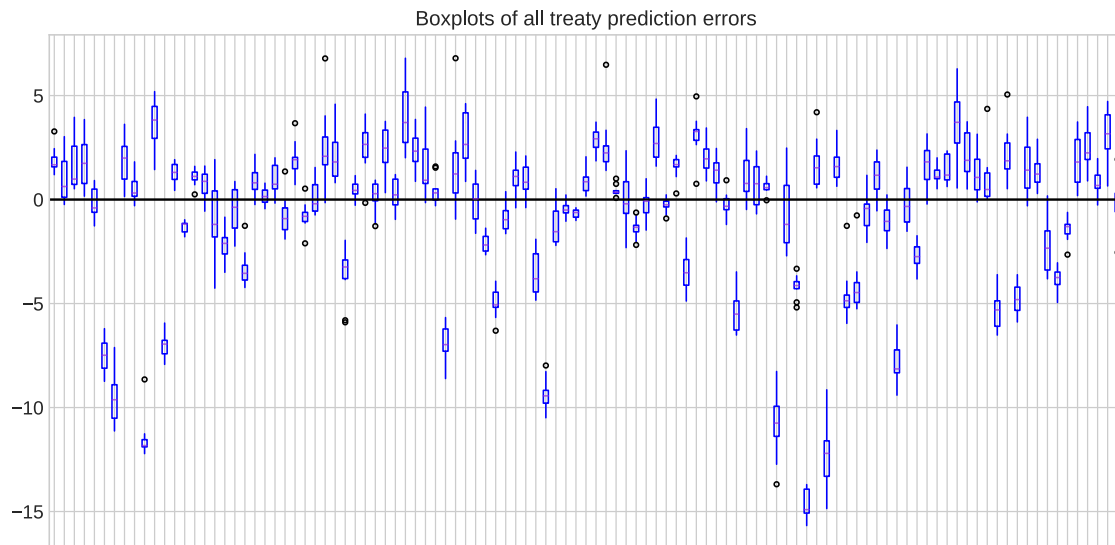
Figure 5.4: All treaty prediction errors

5.3 Insights from feature permutation

As there are over 60 variables used for prediction, not all of which treaty negotiators have meaningful control over, a customised feature permutation algorithm is used for this analysis. Metadata such as treaty adoption year and place, as well as organisational auspices, are left untouched. Regarding the subject matter variables discussed in Chapter 4, any features not presently found in the treaties, expressed by the value 0, are converted into a 1 one at a time, to see to what extent it changes the predictions. Nullifying values of existing 1s is not done, however, because that could lead to nonsensical or uninteresting test samples, such as the Paris Agreement without mention of climate change. Essentially, this approach assumes that there were political or other constraints making existing mentions indispensable, but that it would have been possible to add at least a single mention of absent topics, stressors, and means of intervention. This is probably not true for some of the variables in some of the contexts, but it serves its purpose as a simplifying assumption. As for treaty provision variables, numeric variables are not permuted because they have either too few non-zero values (`autoTermMembThresh`) or seem less likely to be amenable to change for the sake of faster entry into force. Binary variables are tested in a similar fashion to binary subject matter variables, replacing 0s with 1s where relevant and computing the difference in prediction for each permuted sample. Only binary features present in at least 20% and not more than 80% of treaties are included, because otherwise there are too few non-zero samples and/or the training data would be too imbalanced. As a result, the permutation set consists of 30 variables.

The two best performing model types, i.e. MLPs and Ridge regressors, will be

compared for this analysis. For both of them, all 50 trained outer loop models are given permuted test sets and the new predictions as well as differences to original predictions are computed. The permuted test sets are composed of the exact same 21-22 treaties that the models were originally evaluated on, except that now each of the 30 above-mentioned variables has its value flipped (if zero) one at a time, holding all other feature values fixed, to assess the impact of that single switch. This procedure yields over 17'000 permuted predictions per model type.

Figure 5.5 shows the results of the permutations for all treaties and models, grouped by model type.
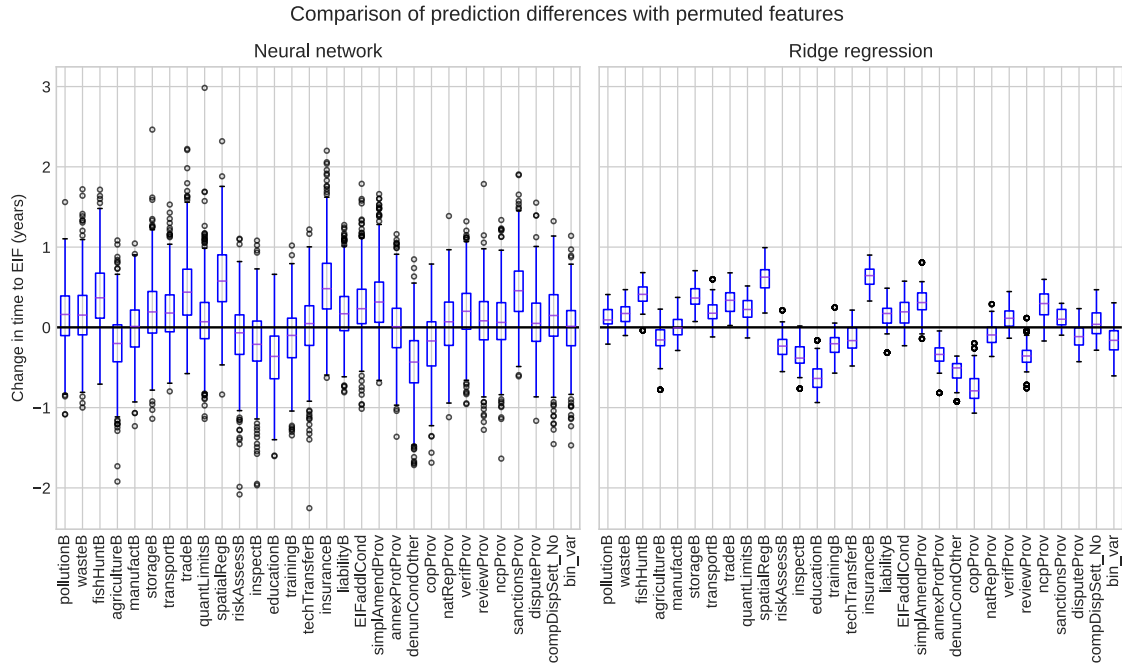
Figure 5.5: Feature permutation prediction difference for all treaties

In a nutshell, according to neural network predictions, there are no topics or provisions that, when added to a treaty, consistently speed up or slow down entry into force. Ridge regression models, by contrast, find some unambiguously positive or negative factors influencing incubation time. This is most likely because linear models are inherently more constrained. MLPs can learn very complex relationships between inputs and outputs, with many kinds of feature interactions. Thus, the answer an MLP would give to the question "which feature do we need to add to speed up entry into force of an international agreement?" would be the quintessential lawyer's answer: "it depends".

Considering that some of the treaties have such high prediction errors that it is hardly worth examining permuted predictions, and that in many cases performance is highly variable, Figure 5.6 shows the same results as Figure 5.5, but only for

the single best model for each agreement, so long as the original prediction error of these models was not more than 1 year. Filtering out poorly performing models reduces the number of outliers, without changing the big picture. Permuted MLP predictions still have a larger range, from about -2 to +2, twice that of Ridge models. The median of each variable is generally in the same vicinity for the two types of models, but there is still not a single feature that consistently increases or decreases incubation time according to neural network predictions, whereas Ridge regressors have fifteen such features.
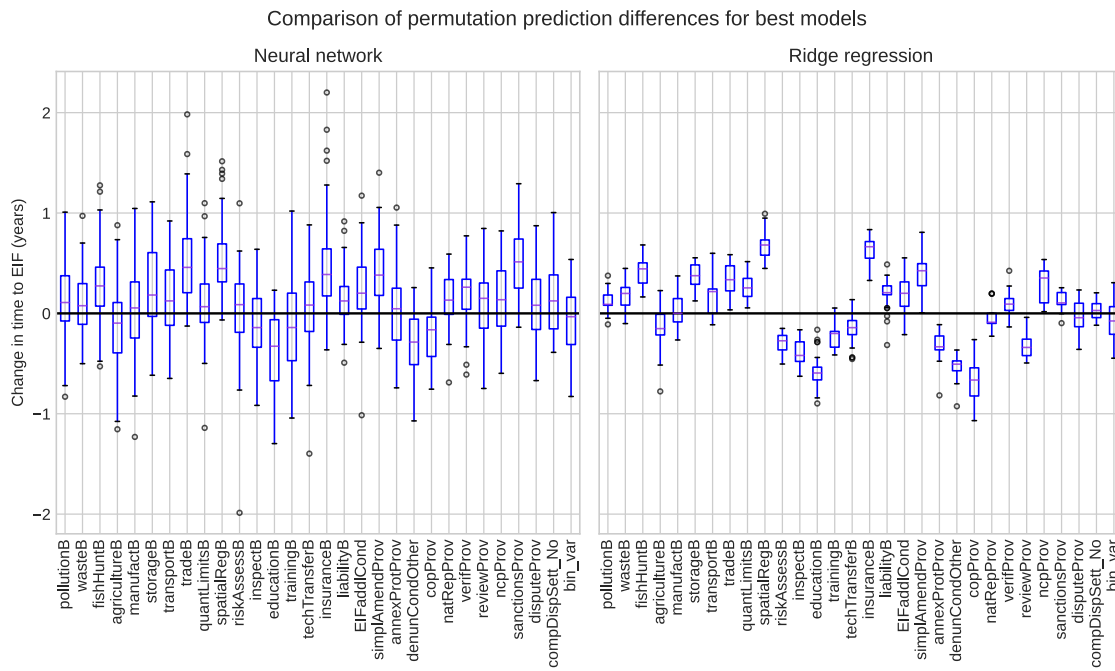
Figure 5.6: Feature permutation prediction differences for best models

Nevertheless, any given MLP test model will typically display up to a dozen or so distinctly positive or negative permutation features. For instance, Figure 5.7 shows the permutation results for a single MLP and its corresponding Ridge model with the same test data. Both performed very well, achieving the lowest test MAE for Ridge (1.76 years MAE) and second lowest test MAE for the MLP (1.64 years MAE). Permuted Ridge prediction differences have zero variance because they are the same across all treaties in the test set. These correspond to the model coefficients. In some cases the mean MLP prediction differences[394] are close to the Ridge model coefficients, but in others they have the opposite sign (see e.g. training, tech transfer, liability, annex/protocol provisions, national reporting provisions, and dispute provisions), even though these two models were trained on the exact same

---

[394] The mean is represented by a triangle, the median by a horizontal bar.

data and have similar average performance. This goes to show that caution should
be exercised when drawing lessons from these results.
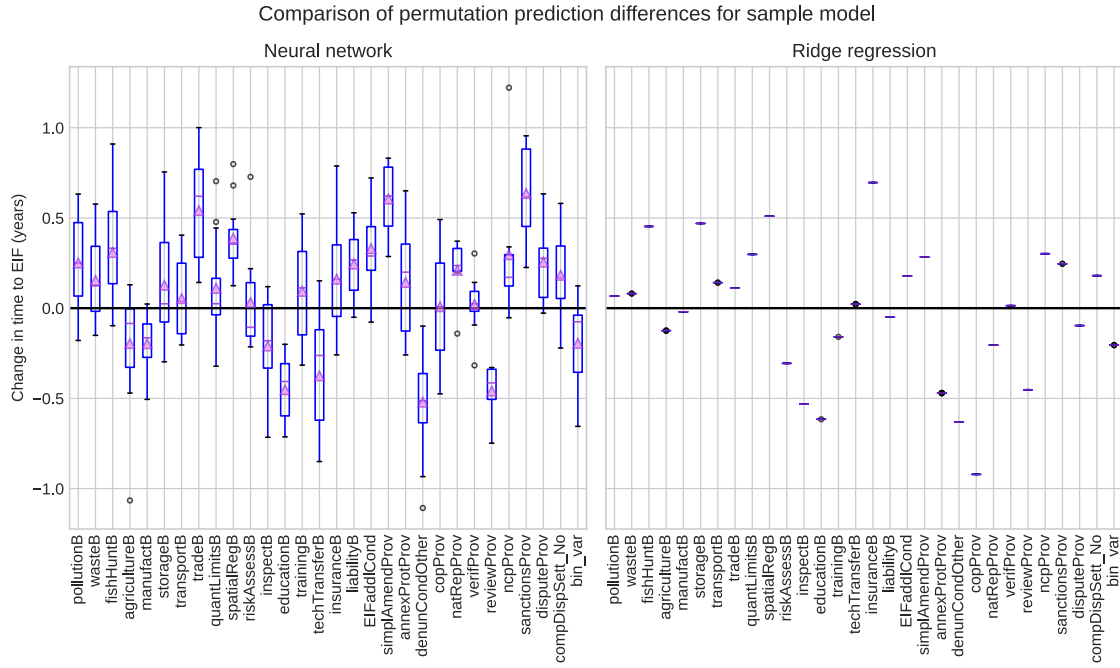
Figure 5.7: Feature permutation prediction differences single sample model

In conclusion, permutation analysis yields no straightforward results that could
be applied across the board. Which features could hasten or hinder entry into force
most likely depends on the context. Linear models do find variables that have con-
sistently positive or negative coefficients across all 50 model evaluations, but neural
networks have slightly better average performance and contradict these results. Per-
haps a larger dataset with more treaties and more explanatory variables would be
needed to achieve consistent findings. Still, these results provide a benchmark for
future work.

Conclusions and outlook

This dissertation purported to show that data science can effectively support and complement traditional legal research on environmental treaty design. Through this work I have come to believe that legal scholarship has much to gain from adopting automation, data science and AI where appropriate. Striking the right balance between human and machine contribution may be the biggest challenge, and is probably essential to achieving uptake of these technologies. Tedious data collection can be automated, whereas interpreting results remains a quintessentially human task. In this research project, I have striven to subordinate data science and automation to legal research questions and legal conceptual frameworks, even when it significantly complicated my work. At the same time, I have spared no effort to uphold the highest standards of openness, transparency, accessibility and reproducibility of research, as mentioned in the introductory chapter.

Chapter 2 thus delves into data sources, treaty selection and choice of methods. With regard to sources, I find that while a range of public, private and academic actors have amassed impressive quantities of treaty data, there is still some room for improvement before these collections can be considered 'data science ready' (more on possible improvements below). As for treaty selection, this proved to be a thornier task than I expected. I started with the longest list of potentially environmental agreements I could find (n=5665) and narrowed it down to 319 agreements in a scripted fashion, then added another dozen agreements manually due to gaps in the initial list, and finally went through the spreadsheet making case-by-case decisions based on my selection criteria, resulting in a final list of 119 open multilateral agreements with global geographic scope and some degree of environmental focus or relevance. The final section of Chapter 2 sets out the aims and criteria guiding the choice and development of text analysis methods and tools used in the core chapters of the thesis.

Chapter 3 perhaps represents the main contribution to the literature on treaty design and the law of treaties. It describes the process and results of regex-based legal information extraction on 15 commonly occurring types of treaty provisions by means of 27 variables capturing different aspects. Prevalence and variation of each

provision in the set of treaties at hand, temporal trends, as well as co-occurrence patterns with other variables are discussed, as relevant. These exploratory analyses yielded many interesting findings that merit further research, ideally with a larger treaty set. The chapter also shows that detailed, legally sound variables can be constructed in a transparent and computationally reproducible manner with evidence to support the classifications, thereby addressing the issues I identified in existing structured MEA datasets from international relations and environmental economics scholars.

Chapter 4 is concerned with treaty subject matter analysis and develops a pragmatic approach to gaining an overview of the environmental problems, anthropogenic stressors and means of intervention covered in multilateral agreements. In each of the three sections of the chapter, I describe the considerations that guided the selection of variables, including difficulties with false positives and false negatives search, before introducing descriptive statistics and associations with other variables. The ten environmental problems constituting the first section are discussed in more detail, while the seven stressors and eleven interventions are presented en bloc due to space constraints. As in Chapter 3, the variables are deliberately designed with structured database search applications in mind, in order to enhance access to international legal information by researchers and practitioners alike.

Chapter 5 employs the variables introduced in Chapters 3 and 4, as well as UNTS metadata, to try to predict entry into force of treaties in a nested five-fold crossvalidation machine learning experiment with ten repetitions. The prediction accuracy of three types of linear regression models is compared to that of multilayer perceptrons, with the latter outperforming the former by a small margin on out-of-sample data. The average prediction accuracy of MLPs on test data is a bit over 2.5 years ± 0.5 years, which is reasonably good considering that the treaties in the dataset took between 0 and 17 years to enter into force. The neural networks and Ridge regression models both outperformed the mean and median dummy predictors used as a benchmark. Outliers have a noticeable impact on prediction accuracy and its variability, hence the importance of evaluating on more than one test set. The final part of the chapter describes a customised feature permutation algorithm that examines how the predictions of the trained outer loop models change if one of 30 binary variables is present rather than absent for each treaty in the respective test set. The practical utility of this computational analysis would be the identification of treaty design features that reliably speed up entry into force. However, no such features were found to be consistent across all examined prediction models. Perhaps a larger dataset or more advanced models would yield more useful findings, but it

may also be that there are simply too many unobservable variables influencing the outcome.

Returning to the larger epistemic challenge posited at the outset of this thesis, I have argued that a collaborative, incrementally expanding database with comprehensive, intelligible, reproducible, accurate and verifiable structured data about treaties, freely accessible through suitable interfaces, would be an effective way to address it. There is nothing inherently new about the challenge, lawyers have been drawing attention to it for decades, and many formal logic-based and statistical/ML approaches have been proposed, but none have been widely adopted. The data science approach presented in this thesis is inspired by the achievements and methods of the bioinformatics, natural language processing, semantic web and open source software communities, and employs their software, open standards and best practices to build a draft ontology of public international law and to develop tools for automated treaty data collection, cleaning, merging, extraction, transformation, visualisation, descriptive and predictive analytics.

The main ideas, data, hardware, software and expertise to do this existed already in the 1990s or early 2000s at the latest (when the core semantic web standards were developed and scientific ontology development gained traction). Hence, one wonders why international legal informatics has fallen so far behind bioinformatics and fields like it. This is a bit of a conundrum, especially as treaty data are vanishingly small compared to the amounts of data collected and processed in genomics, proteomics, astronomy and high-energy physics. Treaty data consist mostly of human-readable text without the need for expensive microscopic or macroscopic technology and experiments, are published by authoritative data providers, are of lasting importance (as opposed to more ephemeral data), could in principle be analysed exhaustively, and cover topics that attract wide interest beyond the relatively small community of international lawyers, extending to other disciplines, legal practice, politics and policy-making, investors, philanthropists, journalists, educators and the general public. Instead of dwelling on past failures and missed opportunities, it may be more fruitful to ponder what a successful uptake of legal informatics and (open) data science would look like for international legal scholarship and practice.

In terms of treaty data sources, at a minimum, machine-readable texts in a non-proprietary format (e.g. Unicode TXT or HTML) would be published under an open data license by the UN Treaty Section, in all authentic treaty languages as well as all official translations available. A step up from accurate and complete string representations would be structured texts with machine-readable information about segmentation into title, preamble, articles, testimonium, optionally signature pages

and/or appendices (if any). The most elaborate standard for this task is the Akoma
Ntoso XML schema,[395] which was adapted for the mark-up of UN documents in the
AKN4UN specifications[396] and on which UNDO is based. However, given that the
treaty concept is inadequately represented in this set of standards,[397] it is probably
not suitable for treaty segmentation either. Accurate representation of amending
agreements and automated consolidation of treaty texts as they exist over time
and apply to different parties (in multilateral agreements with non-simultaneous
entry into force of amendments) would also be a task that requires international
law expertise, because the current conceptualisation in UNDO and Akoma Ntoso
more broadly is not in line with international legal practice. Apart from issues
of structure and version control, treaty texts could additionally embed machine-
readable semantic information without detriment to human readers, e.g. marking
the start and end of provisions commonly found in treaties, like those discussed in
Chapter 3. This is also part of the ambition of legal XML and the legal semantic
web. The danger is aiming too high and failing to deliver the basics. The basics
could be as simple as a plain text file available for download from UNTS treaty
pages with the language identifier added to the filename, as is already the case for
PDF text document files.[398]

Relatedly, treaty identifiers should be reconsidered and streamlined across pub-
lishers and distributors of treaty information and literature. As mentioned in Chap-
ter 2.1, data integration from multiple sources requires a common identifier, and due
to the sometimes multi-year delay in treaty registration and publication of UNTS
volumes, the reference to UNTS volume and page number is presently not sufficient,
even if it were consistently used by authors and database maintainers, which is not
the case. By contrast, scientific journals such as Nature portfolio journals require
authors to use unique resource identifiers from major scientific databases whenever
they publish papers on relevant subjects, including detailed identification of cell
lines, antibodies and model organisms. These short alphanumeric identifiers hardly
disrupt the flow of the text and can be used to automatically index, track and search
relevant literature in any language. Enforcing discipline at the source is much more

---

[395]   For an introduction and comparison with other Legal XML standards see Monica Palmirani
       and Fabio Vitali, 'Akoma-Ntoso for Legal Documents' in Giovanni Sartor and others (eds),
       *Legislative XML for the Semantic Web: Principles, Models, Standards for Document Man-
       agement* (Springer 2011).
[396]   https://www.w3id.org/un/schema/akn4un/ (accessed 12 Dec 2022).
[397]   As mentioned in the Introduction, UNDO defines 'treaty' and 'convention' as two different
       types of things, when under customary international law the denomination of a treaty is
       irrelevant, and this is just one of the problems with the standard.
[398]   Not the 'Volume in PDF' field with a link to the full UNTS volume, but the 'Text document(s)'
       field which for some treaties have the relevant excerpt of the UNTS volume in English, French
       and other languages and include identifying information in the filename.

efficient than trying to identify entities ex post with keywords, regular expressions, or ML classifiers. This is one of the advantages of having gatekeepers. While good international law journals and edited volumes typically enforce a citation style for treaties which should allow for automated identification with a suitable multilingual lookup system, national law, political science, economics, environmental sciences and many other disciplines writing about treaties seem to be unaware of treaty citation standards, sometimes only using a shorthand or acronym in the text.

As for treaty participation, reservations, declarations and other metadata, the most impactful achievement would arguably be the automation of the UNTS updating process, such that whenever treaty depositaries record new information in their own databases, this change would trigger an update of the UNTS database. Reasonably up-to-date treaty participation data is key to analysing, visualising and understanding treaty networks, growth and reach. While small-scale treaty participation analysis can resort to the latest data published by the treaty secretariat in question, larger scale or comparative analysis would really benefit from a one-stop-shop for treaty data. Unfortunately, even major multilateral agreements sometimes have grossly out-of-date participation data in UNTS (on the order of years, not months or weeks). For instance, over 12 years after its entry into force, the UNTS treaty page of the IRENA statute[399] still listed 30 ratifications as the only treaty actions data,[400] whereas the IRENA secretariat's website displayed 168 members.[401] Such discrepancies could lead to a misunderstanding of the scope and reach of an agreement by lawyers, journalists and other interested parties trusting the veracity of UN treaty databases, as well as potentially skew statistics of data analysts who fail to thoroughly examine the quality of their sources. The responsibility to register the change in membership with the UN Treaty Section lies of course with the depositary, Germany in this case. But treaties deposited with UN specialised agencies, in particular IMO and ICAO, often have incomplete and outdated UNTS treaty pages as well. Hence, a first step could be to automate updating between UN treaty databases, before implementing a system to synchronise data from depositaries outside the UN system, such as regional organisations and national governments. A notification system to remind depositaries of their registration duties could be designed by third parties in the meantime. While initial registration of agreements will probably continue to require manual processing by UN Treaty Section staff,

---

[399] IRENA Statute (n 189).
[400] See archived version at https://web.archive.org/web/20230414130313/https://treaties.un.org/Pages/showDetails.aspx?objid=08000002802aefa2&clang=_en (accessed 14 April 2023).
[401] See archived version at https://web.archive.org/web/20230414131050/https://www.irena.org/About/Membership (accessed 14 April 2023).

even if submitted online in the future,[402] subsequent updates seem more feasible and efficient to automate, especially within the UN system. A slight complication and additional data science need are consistent and unique identifiers for treaty participants. Currently, state and other participant names vary across UNTS treaty pages and between different UN treaty databases. Sometimes this is because the concerned states or other entities changed their official names over time, and what is registered is the name they had at the time the information was entered into the database, but often times it seems to be due to preferences, habits or policies of the agency publishing the information. This understandable diversity means unfortunately that analysis of networks and participation patterns requires a significant upfront effort to clean up the data and consequently limits the number of agreements that can be covered within a given research project. There is no need to change the participant names listed in UNTS treaty pages, the historical record can remain as is, simply adding a column or otherwise embedding a unique identifier that is used consistently throughout relevant databases would suffice.

Furthermore, for efficiency and reproducibility of research, it would be best to publish a version controlled corpus of all available treaty texts with relevant metadata and a persistent identifier for each release. For instance, a Gene Ontology Data Archive is published monthly to Zenodo in an automated process by the GO Consortium, each version receiving its own DOI, but there is also an overarching Concept DOI representing all versions.[403] This allows researchers to refer to the exact snapshot they used for an analysis, without having to publish a copy of the data themselves, and others can attempt to replicate the findings but also have easy access to the latest version of the data in one place, to assess whether the results still hold in light of new data. This kind of community service and commitment to reproducibility is still rare in international law, with Seán Fobbe's twin corpora of decisions of the ICJ and PCIJ being the exception to the rule.[404] With successful uptake of open, reproducible treaty data science, researchers studying a subset of agreements, such as bilateral investment treaties or preferential trade agreements, would not only publish the corpus they analysed, but also the source code of the data collection and cleaning process, or refer to existing, published versions thereof, such as the data collection tools published with this thesis.

---

[402]   UNGA Resolution A/RES/76/120, para 10, encourages the development of an online treaty registration system as an alternative to submission of electronic or paper documentation.

[403]   https://doi.org/10.5281/zenodo.1205166 is the Concept DOI in this case.

[404]   Seán Fobbe, 'Introducing Twin Corpora of Decisions for the International Court of Justice (ICJ) and the Permanent Court of International Justice (PCIJ)' (2022) 19(2) Journal of Empirical Legal Studies 491, p.499 ("It is my intention to update the CD-ICJ up to twice per year[...] Minor errors will be documented in the GitHub issue tracker and fixed with the next scheduled release."). Ideally the ICJ would provide this service with appropriate funding.

Regarding PILO, legal information extraction and subject matter indexing, there are many intermediate successes short of the UN publishing treaties in a semantic web format. A good start would be to improve access to treaties on topics of wide interest by adding relevant subject terms to UNTS and indexing agreements accordingly. This could happen independently from efforts to provide machine-readable texts and automate the updating of participation data. The question of what would be important topics to add could be tackled through surveys or by examining treaty database search history (which words occur most often in free-text/title search and what the underlying topics are), if such data were accessible to researchers. More general search engine data like Google Trends could also be analysed for treaty-related queries. The list of treaties to tag with existing and proposed subject terms could be crowd-sourced through a workshop or online questionnaire, use NLP methods like those presented in Chapter 4, ML classifiers, or a combination of methods. Perhaps more important than a one-off improvement to treaty subject matter indexing would be a well-resourced system for future improvements, because research interests and needs inevitably change, and public databases should adapt as swiftly as they can (and receive sufficient funding to do so). The UN Treaty Section is particularly well placed to make treaties discoverable on topics that traverse or encompass multiple areas of international law, do not fall under the auspices of an existing organisation, or are within the purview of an agency that has no mandate for publishing treaty information. The UNBIS Thesaurus invites suggestions from the public[405] and publishes reasonably frequent updates to its terms and definitions; perhaps the UN Treaty Section could learn from their experience. The long-term collaborative PILO database is comparatively harder to implement. It could easily become one of the little or not-so-little academic projects joining the internet graveyard of the many overambitious abandoned ideas. I have attempted to demonstrate that relatively simple NLP techniques and existing open source software tools can be used to select and process precisely what we are interested in when we read treaties and engage in traditional analysis. We can moreover store and accumulate this kind of information for subsequent retrieval, quantitative analysis, citation and presentation. It would help if more international lawyers learned some of the skills required to improve the software toolkit, but the underlying tools were chosen with mainstream international lawyers in mind. Only time will tell whether the approach I have taken in this thesis seems useful and interesting to others.

---

[405] At https://research.un.org/en/thesaurus/suggestions ("Suggestions for new concepts, changes to concepts and other matters will be reviewed by the UN Thesaurus Committee, chaired by the Dag Hammarskjöld Library." accessed 14 April 2023).

Appendices

A   Table of software

The table below lists the main software packages used to produce this thesis (see also `import` statements in subsequent source code). Prior or subsequent versions of the same packages may work as well, and their functionalities are implemented in other packages and programming languages too, but these are the ones I used. The operating system is Debian Linux, which means that some of my code may not work on Windows out of the box because of minor differences in file path specifications etc. The underlying packages are all open source and cross-platform as far as I am aware.

Short name	Version	Lang/API	License	Website	Use
Scrapy	2.4.1	Python	BSD	scrapy.org	Webscraping
Selenium	4.0.0	Python	Apache 2	selenium.dev	Webscraping
Lxml	4.6.1	Python	BSD	lxml.de	Webscraping
Requests	2.22.0	Python	Apache 2	python-requests.org	Webscraping
Protégé	5.5.0	Java	BSD 2	protege.stanford.edu	Ontology development
GATE	9.0.1	Java	LGPL 3	gate.ac.uk	Text processing
Regex	2022.3	Python	Apache 2	pypi.org/project/regex	Text processing
Apache Jena	3.4.0	Java	Apache 2	jena.apache.org	Data storage & query
Emacs Orgmode	9.5.4	Lisp	GPL 3	orgmode.org	Writing & coding
TeX Live	2020	WEB	Other free	tug.org/texlive	Typesetting
NumPy	1.22.3	Python	BSD	numpy.org	Data science
Pandas	1.4.2	Python	BSD 3	pandas.pydata.org	Data science
SciPy	1.8.1	Python	BSD	scipy.org	Data science
Statsmodels	0.13.2	Python	BSD	statsmodels.org	Data science
Scikit-learn	0.24.2	Python	BSD	scikit-learn.org	Machine learning
TensorFlow	2.9.1	Python	Apache 2	tensorflow.org	Deep learning
Keras	2.9.0	Python	Apache 2	keras.io	Deep learning
Matplotlib	3.3.4	Python	PSF	matplotlib.org	Data visualisation
Seaborn	0.11.1	Python	BSD 3	seaborn.pydata.org	Data visualisation
Imgkit	1.0.2	Python	MIT	pypi.org/project/imgkit	Data visualisation

B UNTS data collection

B.1 UNTS searchbot

This bot opens a virtual Chrome browser, visits the Advanced Search page of UNTS online, fills in the search form with predefined parameters, submits the form and then scrapes information contained in the search results pages until it reaches the last page. As the UNTS database is configured to return a maximum of 500 records per search, the bot splits the full time span into months and only requests agreements which have been adopted in a given month in any one search session, before exiting and starting again, iterating over the full range with short breaks in between to avoid overloading UNTS servers.

```python
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.support.wait import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from pyvirtualdisplay import Display
import time, os, re, math
from datetime import date, datetime
import pandas as pd


# set working directory
wd = os.path.expanduser('~/git/spiders/unts-searchbot/')


def parse_table(rows, period):
  """ Extract treaty records from UNTS search results page and return as list of dicts """
  colnb = dict(regNb=2, regDate=3, UNTStreatyType=4, adoptDate=5, UNTSvolNb=6)
  result = []
  treaty = {}
  for i in range(len(rows)):
    treaty['title'] = rows[i].find_element_by_xpath('./td[1]/a').get_attribute('title')
    treaty['UNTSurl'] = rows[i].find_element_by_xpath('./td[1]/a').get_attribute('href')
    treaty['lastRetrieved'] = str(datetime.utcnow().isoformat().split(".")[0]+'Z')
    treaty['adoptPeriod'] = str(period)
    for k,v in colnb.items():
      treaty[k] = rows[i].find_element_by_xpath('./td[{0}]'.format(v)).text
    result.append(treaty.copy())
return result


starttime = datetime.now()
# set start and end month of desired time period
prng = pd.period_range('1945-06', '2022-06', freq='M')
records = pd.Series(index=prng)
treatylist = []
idctl = '//*[@id="ctl00_ctl00_ContentPlaceHolder1_ContentPlaceHolderInnerPage_'
treatyxpath = idctl + 'drpSearchObj"]/option[@value="ts_treaty"]'
adoptxpath = idctl + 'drpAttribute"]/option[@value="conclusion_info_id"]'
addxpath = idctl + 'btnAdd"]'
submitxpath = idctl + 'btnSubmit"]'
rcxpath = '//*[@class="RecordCount"]'
```

```python
rowsxpath = idctl + 'dgTreaty"]//tr[@align="left"]'
nextxpath = '//input[@src="../Images/Paging/btn_next.jpg"]'

# loop through months (due to max results limit per search)
for period in prng:
  display = Display(visible=0, size=(1200, 1200)).start()
  s = Service('/usr/bin/chromedriver')
  driver = webdriver.Chrome(service=s)
  wait = WebDriverWait(driver, 300)
  driver.get('https://treaties.un.org/Pages/AdvanceSearch.aspx?tab=UNTS&clang=_en')
  start = period.to_timestamp('D', how='s').strftime('%d/%m/%Y')
  end =  period.to_timestamp('D', how='e').strftime('%d/%m/%Y')
  wait.until(EC.element_to_be_clickable((By.XPATH, treatyxpath))).click()
  wait.until(EC.element_to_be_clickable((By.XPATH, adoptxpath))).click()
  wait.until(EC.element_to_be_clickable((By.XPATH, idctl+'txtFrom"]'))).send_keys(start)
  wait.until(EC.element_to_be_clickable((By.XPATH, idctl+'txtTo"]'))).click()
  id = 'ctl00_ctl00_ContentPlaceHolder1_ContentPlaceHolderInnerPage_txtTo'
  driver.execute_script("document.getElementById('{id}').value='{val}'".format(id=id, val=end))
  # save screenshot of form with From-To dates filled in
  driver.save_screenshot(wd+'FormFilled.png')
  addButton = wait.until(EC.element_to_be_clickable((By.XPATH, addxpath)))
  addButton.click()
  wait.until(EC.staleness_of(addButton))
  wait.until(EC.element_to_be_clickable((By.XPATH, submitxpath)))
  # save screenshot of submitted form (for debugging)
  driver.save_screenshot(wd+'FormSubmitted.png')
  driver.find_element_by_xpath(submitxpath).click()
  recordCount = wait.until(EC.visibility_of_element_located((By.XPATH, rcxpath)))
  records[period] = int(re.search(r"Record Count : (\d+)\s*", recordCount.text).group(1))
  # save count of treaty records per month
  records.to_frame(name='recordCount').to_csv(wd+'recordcount.csv', encoding='utf-8',
                                              date_format='%Y-%m', index_label='month')
  nb = math.ceil(records[period]/10)
  for n in range(1,nb+1):
    rows =  wait.until(EC.presence_of_all_elements_located((By.XPATH, rowsxpath)))
    treatylist.extend(parse_table(rows, period))
    if n != nb:
      next = wait.until(EC.visibility_of_element_located((By.XPATH, nextxpath)))
      next.click()
      wait.until(EC.staleness_of(rows[1]))

  treatydf = pd.DataFrame(treatylist)
  treatydf.to_csv(wd+'treatydf.csv', encoding='utf-8')
  driver.quit()
  display.stop()
  time.sleep(2)

elapsed = (datetime.now() - starttime)
print('Done! Treaty records fetched: {0}. Time elapsed: {1}'.format(len(treatydf),
                                                    str(elapsed).split(".")[0]))
```

The consolidated results of the last update (run on 28 October 2022) contain:

- 66'862 unique treaty page URLs, of which 59'418 classified as bilateral treaties, 5757 as open multilateral, and 1687 as closed multilateral agreements;

- 66'784 treaty registration numbers, of which 58'269 are unique entries (the most frequent registration number is A-4789 occurring 2325 times)

When removing the first component (letter or Roman numeral) of the registration number and comparing by treaty type, the percentage of original registered and recorded agreements (Roman numerals I and II) is much higher for bilateral treaties (84%) than for multilateral treaties (37%), resulting in only about 2700 original multilateral agreements. Subsequent agreements like amendments and supplementary protocols presumably use similar language as the original agreement and thus present less of an NLP challenge.

About 20% of the treaties analysed in the substantive chapters were not found by this searchbot, even though they have an UNTS treaty page. This means that searching by treaty adoption month is not reliable. The bot itself is not the problem, the results were manually reproduced for a couple of treaties (the Minamata Convention on Mercury and the Paris Agreement). A previous run of the bot did find the Mercury Convention, which means the indexing by adoption month was subsequently removed, perhaps by accident during a cleaning/updating operation. This issue of missing records may only affect multilateral treaties because their pages are updated more regularly, in which case the total count would only be increased by about 1488 to 68'350 treaties if the 20% missing rate holds for all multilateral agreements. If this same rate also affects bilateral treaties, then the total count of treaties could be as high as 80'234.

B.2   UNTS crawler

This is the key part of a Scrapy crawler which takes a file containing treaty page URLs (gathered by the UNTS searchbot in Appendix B.1) as input, visits each page and saves the relevant data to a plain text file for downstream processing. The full crawler source code is published in the code archive of this thesis.

```python
import scrapy
from lxml import html
from datetime import datetime
import re, os

class UNTScrawler(scrapy.Spider):
  name = "unts"
  allowed_domains = ["treaties.un.org"]

  def __init__(self, urlfile=None, targetdir=None, *args, **kwargs):
    super(UNTScrawler, self).__init__(*args, **kwargs)
    if urlfile:
      try:
        with open(urlfile, "rt") as f:
          self.start_urls = [url.strip() for url in f.readlines()]
      except OSError as e:
        print(f"{type(e)}: {e}")
    else:
```

```python
      print("Please provide a file with urls, e.g. 'scrapy crawl unts -a urlfile=urls.txt'")
    if targetdir:
      if os.path.exists(targetdir):
        self.targetdir = targetdir
        print(f"Data will be saved to {targetdir}. Existing files will be overwritten.")
      else:
        raise FileNotFoundError(f"{targetdir} directory does not exist.")
    else:
      self.targetdir = "examples/"
      print(f"Data will be saved to {self.targetdir}.")

  def parse(self, response):
    # passing response body to lxml
    doc = html.fromstring(response.text)
    # pasting full urls into text fields
    els = doc.xpath('//div[@id="headerbox"]//a[contains(@href,"/")]|//div[@id="participants"]//a[
        ↪ @href]')
    domain = 'https://treaties.un.org'
    for el in els:
      el.text = el.text + ' [' + domain + el.get('href') + ']'
    # marking table headers with newlines
    for el in doc.xpath('//div[@id="headerbox"]//tr/th//*[text()]'): el.text = '\n' + el.text
    # extracting text content without whitespace
    htext = [x for x in doc.xpath('//div[@id="headerbox"]//tr//text()') if not x.isspace()]
    # concatenate into string
    hstr = '\t'.join(htext)
    # extracting treaty actions data
    ptext = doc.xpath('//div[@id="participants"]//text()')
    pstr = re.sub(r'(\t?\r\n\t*)+', r'\n', '\t'.join(ptext)).strip()
    # final join
    fullstr = '['+response.url+']\n['+str(datetime.utcnow().isoformat().split(".")[0]+'Z')+']\n\n'+
        ↪ hstr+'\n\n'+pstr
    # save as txt file
    filename = response.url.split("objid=")[-1] + '.txt'
    with open(self.targetdir + filename, 'w', encoding='utf-8') as f:
      try:
        f.write(fullstr)
      except OSError as e:
        print(f"{type(e)}: {e}")
```

C   Treaty sample selection

This is the Python script implementing the selection process described in Chapter 2.2, reducing a set of 1965 agreement texts to 319.

```python
import os
import numpy as np
import pandas as pd

ieadb = pd.read_csv(os.path.expanduser('~/git/phd/treaties_full.csv'),
                    encoding='utf-8').dropna(how='all')
# cleaning up column names
ieadb.columns = ieadb.columns.str.replace(' ','')
ieadb.rename({"IEA#(clickforadd'linfo)":'ieaid'}, axis='columns', inplace=True)
# removing entries without treaty text
ieadb = ieadb[ieadb.TreatyText.notna()]
# only multilateral agreements
ieadb = ieadb[ieadb.Inclusion.str.startswith('M')]
# only English texts (or not yet classified but likely English)
ieadb = ieadb[ieadb.Lang1.isna() | ieadb.Lang1.str.contains('EN')]
# excluding amendments
ieadb = ieadb[~ieadb.TreatyName.str.contains('(?i)^Amend|^Adjustment|Amending|^(?:\w+ ){,3}Amend')]
# restricted agreements with named parties
pattern = r'(?i)(?:Between|\bGovernment[s]?\b|Kingdom|Republic|Commonwealth|Federation|\WStates\b
    ↪ (?!.+the Moon))'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
pattern = r'(?:German|Kuwait|Norway|Niger|Chad|Congo|Senegal|Gambia|Cambodia|Nauru|Palau)'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
# named inland waters and mountains
pattern = r'(?:Rhine|Elbe|Danube|Meuse|Scheldt|Prespa|Mosel|Oder|Skagerrak|Mekong|Limpopo|Zambezi|
    ↪ Lake|River|Basin|Amazon\w*|Carpathian[s]?|Andean|Alps)\b'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
# regional seas, straits and islands
pattern = r'(?:(?:Aral|Baltic|Barents|Bering|Black|Caspian|North|Red|Wadden) Sea[s]?|Mediterranean|
    ↪ Caribbean|Gulf|Bay|Jan Mayen|Faeroe)\b'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
# oceans
pattern = r'(?:Arctic|Antarctic|Pacific|Atlantic|Indian Ocean|Benguela)\b'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
# continents
pattern = r'(?i)(?:Asia[n]?|Europe\w*|(?<!particularly in )Africa\w*|America\w*)\b'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
# regional orgs
pattern = r'(?:Arab|ASEAN|Benelux|EEC|OECD|Nordic|North|East|South|West|Tripartite|\bRegional)'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
# derivative acts
```

```
pattern = r'(?i)(?:Council Decision|Declaration|(?:Plan|Programme) of Action|Resolution(?:s| \d+)|
    ↪ Directive|Agenda|Code of Conduct|Mandate)'
ieadb = ieadb[~(ieadb.TreatyName.str.contains(pattern)|ieadb.AlternativeTreatyNames.str.contains(
    ↪ pattern))]
print(len(ieadb))
```

319

D   Treaty text harvesting & cleaning

D.1   Saving IEAdb urls for selected treaties

```python
import os
import numpy as np
import pandas as pd

tm = pd.read_csv(os.path.expanduser('~/git/phd/treaty-selection/IEA_Dataset_Agreements_332_MK.csv'),
    ↪ encoding='utf-8').dropna(how='all')
urls = tm.loc[(tm.AustLIIurl == 'x') & ((tm.selected == True) | tm.fn.str.contains('^1973.+
    ↪ PollutionFromShips')), ['fn','IEAdbUrl']]
urls.IEAdbUrl.fillna('https://iea.uoregon.edu/treaty-text/' + urls.fn.str.lower().str.replace('.','')
    ↪ , inplace=True)
urls = urls.IEAdbUrl
urls.to_csv(os.path.expanduser('~/git/phd/ieadb-scraper/ieadb_urls.txt'), encoding='utf-8', index=
    ↪ False, header=False)
print(len(urls))
```

```
24
```

D.2   Scraping selected IEAdb treaty texts

This is the Python script used for downloading IEAdb agreement texts. Additional explanations and instructions are published together with the script in the thesis code archive.

```python
import requests, re, os
from lxml import html
from datetime import datetime

with open(os.path.expanduser('~/git/phd/ieadb-scraper/ieadb_urls.txt'), 'rt') as f:
  urls = [url.strip() for url in f.readlines()]

for url in urls:
  res = requests.get(url)
  doc = html.fromstring(res.content)
  textel = doc.xpath('//div[@id="block-system-main"]')[0]
  for el in textel.xpath('//p'):
    if (el.tail is not None):
      el.tail += '\n'
  textstr = html.tostring(textel, method='text', encoding='unicode').strip()
  textstr = re.sub('\n(?:Source|SRC): ?(?:\w.+|(?=\n))', '', textstr)
  fullstr = textstr +'\n\n-----\n[Source: '+ url +' (last retrieved '+ str(datetime.utcnow().
      ↪ isoformat().split("T")[0]) +')]'
  filename = doc.xpath('//span[preceding-sibling::strong[text()="Filename: "]]/text()')[0]
  d = os.path.expanduser('~/git/phd/treatytexts/scraped/')
  if os.path.exists(d + filename):
    with open(d + filename, "r", encoding='utf-8') as f:
      oldt = f.read().split('\n\n-----\n[Source')[0].strip()
```

172

```
    if (textstr != oldt):
      print(filename, 'newt len: ' + str(len(textstr)) + ' oldt len: ' + str(len(oldt)))
      with open(d + filename, 'w', encoding='utf-8') as f:
        f.write(fullstr)
  else:
    with open(d + filename, 'w', encoding='utf-8') as f:
      f.write(fullstr)
```

D.3   Fetching treaty texts from AustLII

This is the core part of a Scrapy spider which takes a file listing treaty text URLs as input,
visits each page and extracts the relevant text without AustLII's proprietary markup and
saves it to a plain text file with provenance information for downstream processing. The
full project with further explanations is published in the thesis code archive.

Scraping AustLII treaties (as part of Scrapy spider)

```
import scrapy
from lxml import html
from datetime import datetime
import re, os

class scrapeAustLII(scrapy.Spider):
  name = "austlii"
  allowed_domains = ["austlii.edu.au"]
  with open(os.path.expanduser('~/git/phd/austlii-scraper/treatyurls.txt'), "rt") as f:
    start_urls = [url.strip() for url in f.readlines()]

  def parse(self, response):
    # passing response body to lxml
    tree = html.fromstring(response.text)
    # discarding copyrighted footnotes and footer (multiple versions)
    for el in tree.xpath('//a[@name="fn0" or @name="fn1"]/following::*'):
      el.getparent().remove(el)
    for el in tree.xpath('//a[@name="fn0" or @name="fn1"]'):
      el.getparent().remove(el)
    # drop in-text footnote anchors without removing tail
    for el in tree.xpath('//a[@name]'):
      el.drop_tree()
    # remove anything prior to treaty title
    for el in tree.xpath('//*[text()[contains(.,"Commonwealth of Australia")]] | //*[text()[contains
        ↪ (.,"Commonwealth of Australia")]]/preceding::*'):
      el.getparent().remove(el)
    # table formatting
    for el in tree.xpath('//table//br'):
      el.drop_tree()
    for el in tree.xpath('//table//*[text()]'):
      el.text = el.text.strip()
    for el in tree.xpath('//table/tr'): el.tail = '\n'
    for el in tree.xpath('//table/tr/td[position() < last()]'):
      el.tail += 'replacewithtabchar'
    # text formatting
```

```python
    for el in tree.xpath('//*[text()]'):
      el.text = el.text.replace('\n ','\n')
    for el in tree.xpath('//center/p'):
      el.text = re.sub('\n(?=[A-z])','  ', el.text)
      el.drop_tag()
    ttext = html.tostring(tree, method='text', encoding='utf-8', pretty_print=True).decode('utf-8')
    ttext = ttext.replace('replacewithtabchar', '\t')
    # adding source url
    fullstr = ttext.strip() +'\n\n-----\n[Source: '+ response.url +' (last retrieved '+ str(datetime.
        ↪ utcnow().isoformat().split("T")[0]) +')]\n[CC BY 3.0 AU Australian Government, Department
        ↪  of Foreign Affairs and Trade]'
    # save as txt file
    filename = response.url.split("dfat/")[1].replace('/','-').replace('.html','.txt')
    d = os.path.expanduser('~/git/phd/treatytexts/scraped/')
    with open(d + filename, 'w', encoding='utf-8') as f:
      f.write(fullstr)
```

D.4   Treaty text pre-processing

This Python script edits the scraped texts in order to get as close as possible to the original treaty texts, as explained in Chapter 2.1.

```python
import re, os, glob
import numpy as np
import pandas as pd
from datetime import datetime

scrapeddir = os.path.expanduser('~/git/phd/treatytexts/scraped/')
cleandir = os.path.expanduser('~/git/phd/treatytexts/cleaned/')
scrapedfns = glob.glob(scrapeddir + '*txt')
scraped = pd.DataFrame([s.split('/')[-1] for s in scrapedfns], columns=['fn'])
scraped['text'] = [open(f, 'r', encoding='utf-8').read() for f in scrapedfns]
scraped0 = scraped.copy()
scraped.text = scraped.text.str.replace('\ufeff','', regex=False)
scraped.text = scraped.text.str.replace(u'\xa0+\s+',u'\t', regex=True).str.strip()
# reconstruct MARPOL-1973
marpoldf = scraped[scraped.fn.str.contains('1973-PollutionFromShips')].copy().sort_values(by='fn')
marpoldf = marpoldf.iloc[[0]].append(marpoldf[marpoldf.fn.str.contains('Protocol')]).append(marpoldf[
      ↪ marpoldf.fn.str.contains('Annex')])
sources = marpoldf.text.str.split('-----(?=\n\[Source)').str[1]
marpoldf.text = marpoldf.text.str.split('(?=-----\n\[Source)').str[0]
marpoldf.text.iloc[0] = marpoldf.text.str.cat() + '-----' + sources.str.cat()
scraped = scraped[~scraped.fn.str.contains('1973-PollutionFromShips')].append(marpoldf.iloc[[0]],
      ↪ ignore_index=True)
scraped0 = scraped0[scraped0.fn.isin(scraped.fn)].reset_index(drop=True)
# edit start of AustLII texts
scraped.text = scraped.text.str.replace('(?is)\s+(?:Table of )?Contents.+Preamble.+(?=\n+Preamble)',
      ↪ '', regex=True)
scraped.text = scraped.text.str.replace('(?is)\s+Contents.+?(?=\n+The Contracting Governments)', '\n'
      ↪ , regex=True)
scraped.text = scraped.text.str.replace('(?is)\s+Table of Contents.+?\n(?=\n+The States Parties)', ''
      ↪ , regex=True)
# edit end of AustLII texts
```

```python
end = '(?=\n\n-----\n\[Source)'
scraped.text = scraped.text.str.replace('(?s)\nPROTOCOL CONCERNING SPAIN.+'+ end, '\n', regex=True)
scraped.text = scraped.text.str.replace('(?s)[^\n]+ORGANIZATION, as amended to 1995.+'+end, '', regex
    ↪ =True)
scraped.text = scraped.text.str.replace('(?s)\nPROTOCOL TO [^\n]+, 1969.+'+ end, '', regex=True)
scraped.text = scraped.text.str.replace('(?s)\nAMENDMENTS .+'+ end, '\n', regex=True)
scraped.text = scraped.text.str.replace('(?s)\nFOOD AID CONVENTION, 1995.+'+ end, '', regex=True)
scraped.text = scraped.text.str.replace('(?s)\sDECLARATIONS BY AUSTRALIA.+'+ end, '\n', regex=True)
scraped.text = scraped.text.str.replace('(?s)\nRESOLUTION I.+'+ end, '', regex=True)
# add 'last modified' date
today = str(datetime.utcnow().isoformat().split("T")[0])
scraped.text[scraped.text != scraped0.text] = scraped.text[scraped.text != scraped0.text].str.replace
    ↪ ('(?<=last retrieved \d{4}-\d{2}-\d{2})', ', last modified ' + today, regex=True)
# update 'cleaned' directory with new or modified texts
for i in range(len(scraped)):
  if os.path.exists(cleandir + scraped.fn[i]):
    with open(cleandir + scraped.fn[i], "r", encoding='utf-8') as f:
      oldt = f.read().split('\n\n-----\n[Source')[0]
    newt = scraped.text[i].split('\n\n-----\n[Source')[0]
    if (newt != oldt):
      with open(cleandir + scraped.fn[i], 'w', encoding='utf-8') as f:
        f.write(scraped.text[i])
  else:
    if (scraped.text[i] == scraped0.text[i]):
      # preserve original creation date
      os.system('cp -pu ' + scrapeddir + scraped.fn[i] + ' ' + cleandir + scraped.fn[i])
    else:
      with open(cleandir + scraped.fn[i], 'w', encoding='utf-8') as f:
        f.write(scraped.text[i])
```

E PILO

Figure E.1 is a high-level overview of continuant classes, as displayed by the Protégé Desktop software. Figure E.2 is the corresponding overview of occurrents (processes and events). PILO classes are printed in bold face. An excerpt of the source code is included below, and the full version is available in the thesis code archive.
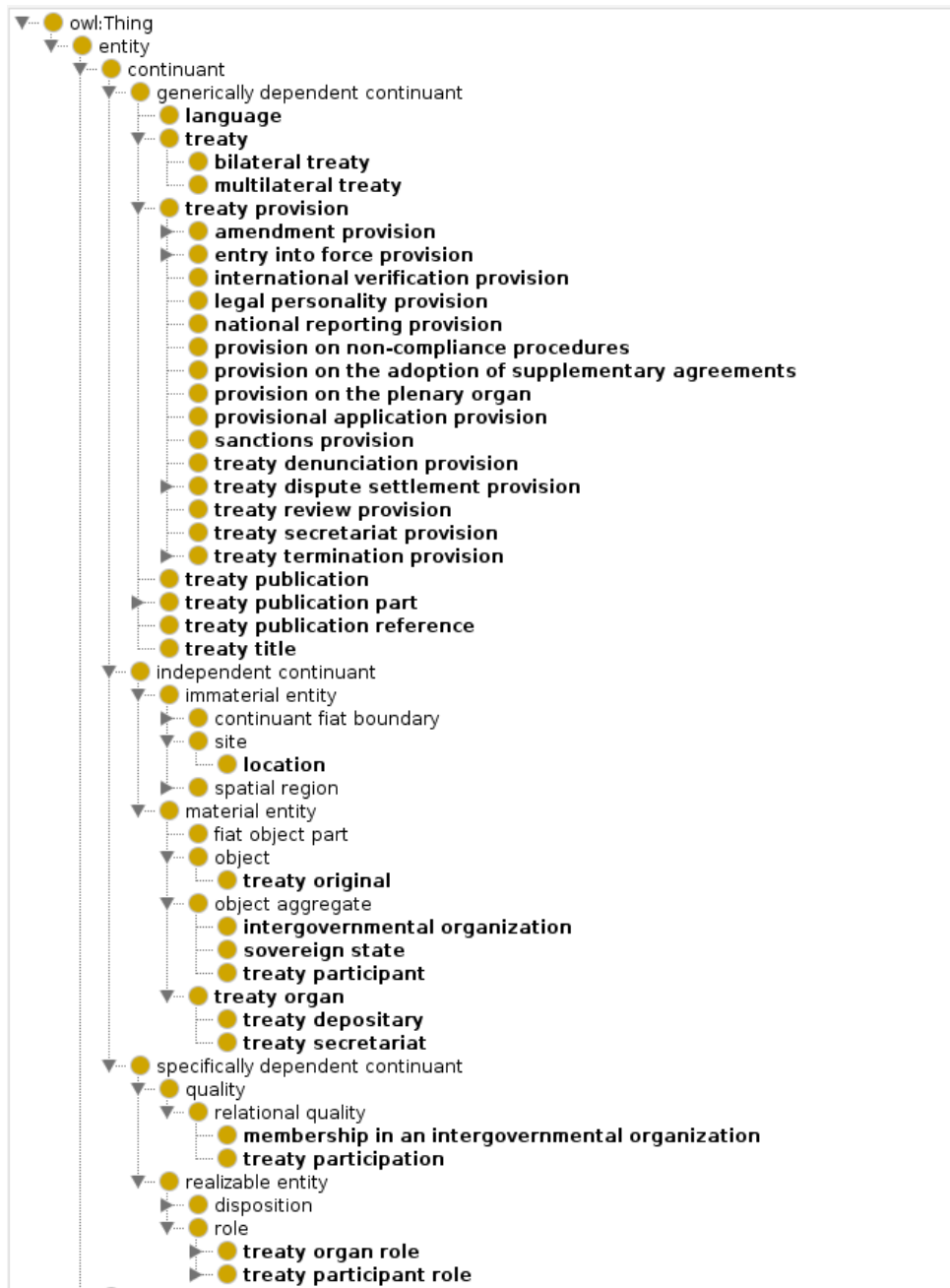
Figure E.1: PILO continuant classes
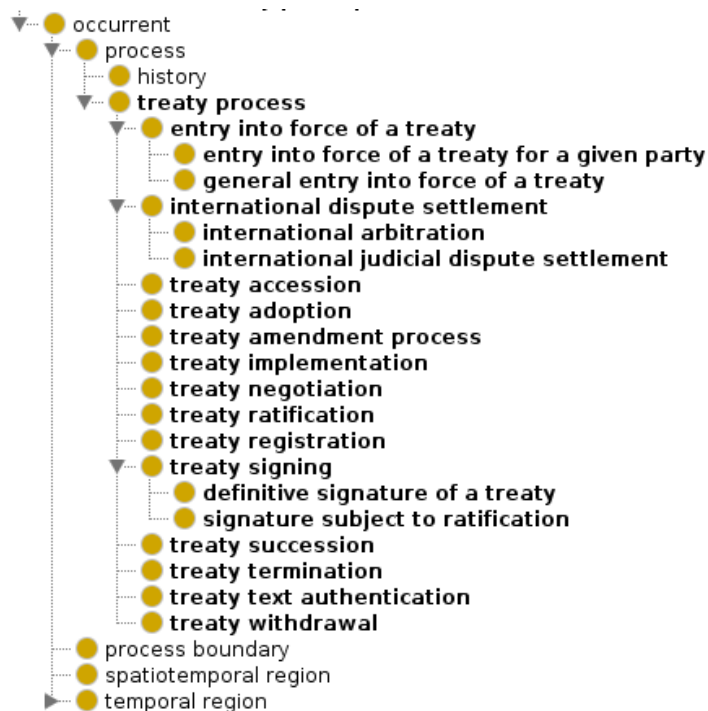
Figure E.2: PILO occurrent classes

Source code excerpt (in Turtle format)

```turtle
@prefix : <https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix obo: <http://purl.obolibrary.org/obo/> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix terms: <http://purl.org/dc/terms/> .
@base <https://gitlab.com/legalinformatics/pilo/core/pilo.ttl> .


<https://gitlab.com/legalinformatics/pilo/core/pilo.ttl> rdf:type owl:Ontology ;
                                    owl:imports <http://purl.obolibrary.org/obo/bfo/2020/bfo.owl> .


#################################################################
#    Classes
#################################################################


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#treaty
:treaty rdf:type owl:Class ;
        rdfs:subClassOf obo:BFO_0000031 ;
        rdfs:label "treaty"@en ;
        skos:altLabel "international agreement"@en ;
        skos:definition "An international agreement concluded between states or intergovernmental
            ↪ organizations in written form and governed by international law, whether embodied in
            ↪ a single instrument or in two or more related instruments and whatever its particular
            ↪  designation."@en .
```

```
###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#bilateralTreaty
:bilateralTreaty rdf:type owl:Class ;
                 rdfs:subClassOf :treaty ;
                 rdfs:label "bilateral treaty"@en ;
                 skos:altLabel "bilateral agreement"@en ;
                 skos:definition "A treaty which is open to participation by only two parties."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#multilateralTreaty
:multilateralTreaty rdf:type owl:Class ;
                 rdfs:subClassOf :treaty ;
                 rdfs:label "multilateral treaty"@en ;
                 skos:definition "A treaty which is open to participation by three or more parties
                     ↪ ."@en ;
                 rdfs:comment "Multilateral treaties may initially enter into force with only two
                         ↪ contracting parties, e.g. conventions adopted under the auspices of the
                         ↪ ILO require only two instruments of consent to be bound, hence the
                         ↪ openness criterion in the definition."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#provisAppProv
:provisAppProv rdf:type owl:Class ;
             rdfs:subClassOf :treatyProvision ;
             rdfs:label "provisional application provision"@en ;
             skos:definition "A treaty provision which governs the application of a part or the
                 ↪ entirety of a treaty pending its entry into force."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#amendProv
:amendProv rdf:type owl:Class ;
         rdfs:subClassOf :treatyProvision ;
         rdfs:label "amendment provision"@en ;
         skos:definition "A treaty provision which stipulates how the content of an agreement can
             ↪ be formally altered by its contracting parties."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#denunProv
:denunProv rdf:type owl:Class ;
         rdfs:subClassOf :treatyProvision ;
         rdfs:label "treaty denunciation provision"@en ;
         skos:altLabel "treaty withdrawal provision"@en ;
         skos:definition "A treaty provision which governs whether, when and how a party may end
             ↪ its participation in a treaty."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#termProv
:termProv rdf:type owl:Class ;
        rdfs:subClassOf :treatyProvision ;
        rdfs:label "treaty termination provision"@en ;
        skos:definition "A treaty provision which governs when and how participation of all parties
            ↪  in a treaty ends and a treaty ceases to be in force entirely."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#autoExpiryProv
:autoExpiryProv rdf:type owl:Class ;
                rdfs:subClassOf :termProv ;
                rdfs:label "auto-expiration provision"@en ;
                skos:definition "A treaty termination provision which stipulates that the agreement
                    ↪ automatically terminates at a certain date or after a certain period of time.
                    ↪ "@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#autoTermMemThreshProv
```

```
:autoTermMemThreshProv rdf:type owl:Class ;
                       rdfs:subClassOf :termProv ;
                       rdfs:label "auto-termination provision based on membership threshold"@en ;
                       skos:definition "A treaty termination provision which stipulates that the
                           ↪ agreement automatically terminates if the number of its parties falls
                           ↪ below a certain threshold."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#treatyAccession
:treatyAccession rdf:type owl:Class ;
                 rdfs:subClassOf :treatyProcess ;
                 rdfs:label "treaty accession"@en ;
                 skos:definition "A treaty process by which an eligible prospective contracting party
                     ↪  which is not a signatory becomes bound by a treaty."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#treatyAmendmentProcess
:treatyAmendmentProcess rdf:type owl:Class ;
                        rdfs:subClassOf :treatyProcess ;
                        rdfs:label "treaty amendment process"@en ;
                        skos:altLabel "the process of amending a treaty"@en ;
                        skos:definition "A treaty process by which contracting parties formally alter
                            ↪  the provisions of the treaty."@en ;
                        rdfs:comment "'Process' is added to the label clarify that this class stands
                            ↪ for the process, not its output, as the term 'amendment' is used for
                            ↪ both."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#treatyWithdrawal
:treatyWithdrawal rdf:type owl:Class ;
                  rdfs:subClassOf :treatyProcess ;
                  rdfs:label "treaty withdrawal"@en ;
                  skos:altLabel "treaty denunciation"@en ;
                  skos:definition "A treaty process by which a party ends its participation in a
                      ↪ treaty and at the end of which it becomes a third party."@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#treatyTermination
:treatyTermination rdf:type owl:Class ;
                   rdfs:subClassOf :treatyProcess ;
                   rdfs:label "treaty termination"@en ;
                   skos:definition "A treaty process by which all parties terminate their
                       ↪ participation in a treaty and at the end of which the treaty ceases to be
                       ↪ in force entirely."@en .


#################################################################
#    Individuals
#################################################################

###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#UN
:UN rdf:type owl:NamedIndividual ,
            :IGO ;
    rdfs:label "United Nations"@en .


###  https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#UNTS
:UNTS rdf:type owl:NamedIndividual ,
              :treatyPublication ;
      rdfs:label "UNTS"@en ;
      :hasPublisher :UN .
```

F   GATE pipelines

There are two GATE apps developed for this project, one for UNTS treaty pages (UNTS-treatyrecords-IE) and the other one for treaty texts (treatytexts-IE), because each GATE app can only work with one corpus. Both have one or more Java Regexp Annotators that generate annotations, which are then used in an ontology-aware Jape-Plus Transducer to populate the PILO ontology with instances. Data from different sources are merged using a mapping between the source URLs (in the file `urlMap.tsv`), as treaty titles are not sufficiently consistent and unique to serve as identifiers.

F.1   UNTS annotation and ontology population

F.1.1   Java Regexp Annotator for UNTS treaty pages

This code block is an excerpt of the file `UNTSregexps.java` which specifies the rules for the Java Regexp Annotator plugin of GATE Developer. It has to be used as part of a GATE pipeline. An executable version of this pattern file, as well as the GATE app it belongs to, can be downloaded from the thesis code archive. It takes a GATE corpus of UNTS treaty pages gathered by the UNTS crawler as input and returns an annotated document corpus as output.

```
// Simplified date macro
date=(?:[0-9]{2}/[0-9]{2}/[0-9]{4})
date2=(?:[0-9]+ [A-z]+ [0-9]{4})


// Full treaty record
|^\[(https.+)\]\n\[(\d+.+)\]\n\n\n(Registration Number(?:.*\n?)+)
1 => UNTStreatyRecordURL
2 => LastRetrieved
3 => UNTStreatyRecord class="treatyPublication",isAbout="treaty",hasURL=$1,lastRetrieved=$2


// Treaty metadata
|Registration Number\t(.+)\t
|Registration Date(?:.+?)?(<<date2>>)\t
1 => TreatyRegNb kind="treatyRegNb",isOutputOf="treatyRegistration"
2 => TreatyRegDate kind="treatyRegDate",hasDateFormat="d MMMM yyyy",isDateOf="treatyRegistration"


|Title\t(.+)\t
1 => TreatyTitle class="treatyTitle",denotes="treaty",isPartOf="treatyPublication"


// If there are multiple dates, this only captures the first one
|Place\tDate\t(.+?)\t(<<date>>)\t([^\t\n]+)?\t?<<date>>?\t?([^\t\n]+)?
1 => AdoptionPlace class="location",isOfficialSiteOf="treatyAdoption"
2 => AdoptionDate kind="treatyAdoptionDate",hasDateFormat="dd/MM/yyyy"
3 => AdoptionPlace class="location",isOfficialSiteOf="treatyAdoption"
4 => AdoptionPlace class="location",isOfficialSiteOf="treatyAdoption"
```

```
|EIF information\t(?:.*?definitively on |with retroactive effect from )?(<<date2>>) ?,? ?(in
    ↪ accordance with article (\w+).*\s*.*?\s*.*?)?\t\nAuthentic
1 => TreatyEIFdate kind="treatyEIFdate",hasDateFormat="d MMMM yyyy",isDateOf="generalTreatyEIF"
2 => TreatyEIFartInfo class="treatyEIFprovInfo",isPartOf="treatyPublication"
3 => TreatyEIFartNb kind="treatyEIFartNb",isPartOf="treatyEIFprovInfo",denotes="generalTreatyEIFprov"


|Authentic texts\t((?:\w+\t)+)\nAttachments
1 => Language class="language",isOfficialLanguageOf="treaty"


|(?<=Depositary.{0,500}?\t)([^\t\n]+)?(?=.*?\t\nRegistration Date)
0 => TreatyDepositary class="treatyDepositary",isDepositaryOf="treaty"


|(?<=Subject terms.{0,500}?\t)([^\t\n]+)?(?=.*?\t\nAgreement type)
0 => SubjectTerm class="treatyDescriptor",isPartOf="treatyPublication",isAbout="treaty"


|Agreement type\t(.+?)\t
1 => TreatyType kind="treatyType"


|UNTS Volume Number\s+((\d+).*)\t
1 => UNTSvolRef class="treatyPublicationRef",isPartOf="treatyPublication"
2 => UNTSvolNb kind="UNTSvolNb"


|Certificate Of Registration\t(.+?)\t
1 => TreatyRegCert isCertificateOf="treatyRegistration"


// If there are multiple URLs, this only captures the first one
|Text.document..?s.\t.+?\[(https.+?pdf)\]\t
1 => TreatyTextURL isURLof="treatyText"


|(?i)Volume In Pdf\t.+?\[(https.+pdf)\]\t
1 => UNTSvolURL isURLof="UNTSvolume"


// Participation data
|^([^\t]+?) \[.+$
0 => TreatyAction kind="treatyAction"
1 => TreatyParticipant class="treatyParticipant"


|^(.+?) \[([^\]]+)\]\t(Signature)\t(<<date>>)
0 => TreatySignature class="treatySimpleSig",hasParticipant=$1,hasURL=$2,hasDate=$4,involves="treaty"


|^(?i)(.+?) \[([^\]]+)\]\t(?:Acceptance by )?Definitive signature[^\t]*?\t(<<date>>)\t(<<date>>)?
0 => TreatyDefSig class="treatyDefSig",hasParticipant=$1,hasURL=$2,hasNotifDate=$3,hasEffectDate=$4,
    ↪ involves="treaty"


|^(.+?) \[([^\]]+)\]\t(?:Ratification|Approval|Acceptance)\t(<<date>>)\t(<<date>>)?
0 => TreatyRatification class="treatyRatification",hasParticipant=$1,hasURL=$2,hasNotifDate=$3,
    ↪ hasEffectDate=$4,involves="treaty"


|^(.+?) \[([^\]]+)\]\t(?:Accession)\t(<<date>>)\t(<<date>>)?
0 => TreatyAccession class="treatyAccession",hasParticipant=$1,hasURL=$2,hasNotifDate=$3,
    ↪ hasEffectDate=$4,involves="treaty"


|^(.+?) \[([^\]]+)\]\t(?:Succession)\t(<<date>>)\t(<<date>>)?
0 => TreatySuccession class="treatySuccession",hasParticipant=$1,hasURL=$2,hasNotifDate=$3,
    ↪ hasEffectDate=$4,involves="treaty"
```

F.1.2   Populating PILO with UNTS data

This part of the GATE UNTS pipeline takes each annotated UNTS treaty page in turn
and populates the PILO database with instances. The file name in the GATE app is
`UNTSontopop.jape`.

```
Imports: {
import static gate.Utils.*;
import java.time.*;
import java.time.temporal.ChronoUnit;
import java.time.format.DateTimeFormatter;
import java.time.format.DateTimeParseException;
}


Phase: PopulateOntologyWithUNTSdata
Input: TreatyTitle AdoptionDate
Options: control = appelt


Rule: CreateInstancesAndProperties
({TreatyTitle}
{AdoptionDate}):inst
-->
{
  AnnotationSet instanceAnnots = inputAS.get();
  if (instanceAnnots == null || instanceAnnots.size() == 0) {
    System.err.println("Error: this document does not have any annotations!");
    return;
  }
  // generate treaty lookup map
  HashMap<String,String> URLmap = new HashMap<String,String>();
  String line = null;
  String dir = System.getProperty("user.home");
  String fn = dir + "/git/UNTS-treatyrecords-IE/examples/urlMap.tsv";
  try (BufferedReader csvReader = new BufferedReader(new FileReader(fn))) {
    while ((line = csvReader.readLine()) != null) {
      String[] arr = line.split("\t");
      URLmap.put(arr[0], arr[1]);
    }
  }
  catch (IOException err) {
    throw new JapeException(err);
  }
  // create a HashMap to look up instances by their className
  Map<OInstance,String> docInstMap = new HashMap<OInstance,String>();
  for (Annotation instanceAnn : instanceAnnots) {
    if (!instanceAnn.getType().matches("TreatyAction")) {
      String className = (String)instanceAnn.getFeatures()
        .get(gate.creole.ANNIEConstants.LOOKUP_CLASS_FEATURE_NAME);
      if (className != null) {
        OClass aClass = ontology.getOClass(ontology.createOURIForName(className));
        if (aClass == null) {
          System.err.println("Error class \""+ className +"\" does not exist! instanceAnn: " +
                             instanceAnn.getType());
          return;
        }
```

```java
      String theInstanceText = gate.Utils.stringFor(doc, instanceAnn);
      OURI instanceURI = ontology.createOURIForName(className + "_" + theInstanceText.hashCode());
      instanceAnn.getFeatures().put("instanceURI", instanceURI.toString());
      if (!ontology.containsOInstance(instanceURI)) {
        OInstance inst = ontology.addOInstance(instanceURI, aClass);
        // add (truncated) label
        if (!className.matches("treatyPublication|treatySimpleSig|treatyDefSig|" +
                               "treatyRatification|treatyAccession|treatySuccession")) {
          int StrLen = theInstanceText.length();
          if (StrLen > 100) {
            inst.setLabel(theInstanceText.substring(0,100), OConstants.ENGLISH);
          } else {
            inst.setLabel(theInstanceText.substring(0,StrLen), OConstants.ENGLISH);
          }
        }
        // add instance string (where appropriate)
        if (!className.matches("treatyPublication|location|treatyParticipant|treatySimpleSig|" +
                        "treatyDefSig|treatyRatification|treatyAccession|treatySuccession")) {
          DatatypeProperty dataProp =
            ontology.getDatatypeProperty(ontology.createOURIForName("hasInstanceString"));
          try {
            inst.addDatatypePropertyValue(dataProp, new Literal(theInstanceText,
                                                OConstants.ENGLISH));
          }
          catch (InvalidValueException err) {
            throw new JapeException(err);
          }
        }
        docInstMap.put(inst, className);
      } else {
        OInstance inst = ontology.getOInstance(instanceURI);
        docInstMap.put(inst, className);
      }
    }// if (has class feature)
  }// if (filter)
}// for each annotation in doc

// generate hash for instance URIs
String UNTSurl = stringFor(doc, instanceAnnots.get("UNTStreatyRecordURL"));
String titlenad = URLmap.get(UNTSurl);
String hash = "" + titlenad.hashCode();
// prepare datetime formatters
DateTimeFormatter textDate = DateTimeFormatter.ofPattern("d MMMM yyyy", Locale.ENGLISH);
DateTimeFormatter tableDate = DateTimeFormatter.ofPattern("dd/MM/yyyy");
// commonly used properties and instances
ObjectProperty involves = ontology.getObjectProperty(ontology.createOURIForName("involves"));
ObjectProperty hasParticipant = ontology.getObjectProperty(ontology.createOURIForName("
    ↪ hasParticipant"));
DatatypeProperty dateDP = ontology.getDatatypeProperty(ontology.createOURIForName("hasDate"));
String treatyTitleStr = stringFor(doc, instanceAnnots.get("TreatyTitle"));
String treatyAdoptionDateStr = stringFor(doc, instanceAnnots.get("AdoptionDate"));
LocalDate treatyAdoptionDate = LocalDate.parse(treatyAdoptionDateStr, tableDate);
String xsdtreatyAdoptionDateStr = treatyAdoptionDate.format(DateTimeFormatter.ISO_LOCAL_DATE);
OInstance UN = ontology.getOInstance(ontology.createOURIForName("UN"));
OInstance UNTS = ontology.getOInstance(ontology.createOURIForName("UNTS"));
// create OInstances that have no annotations
```

```
// create or get treaty instance
OInstance treatyInst = null;
if (!docInstMap.containsValue("treaty")) {
  OClass treatyClass = ontology.getOClass(ontology.createOURIForName("treaty"));
  OURI treatyInstURI = ontology.createOURIForName("treaty_" + hash);
  if (!ontology.containsOInstance(treatyInstURI)) {
    treatyInst = ontology.addOInstance(treatyInstURI, treatyClass);
  } else {
    treatyInst = ontology.getOInstance(treatyInstURI);
  }
  docInstMap.put(treatyInst, "treaty");
}
// create treatyAdoption instance
try {
  OClass treatyAdoptionClass = ontology.getOClass(ontology.createOURIForName("treatyAdoption"));
  OURI treatyAdoptionInstURI = ontology.createOURIForName("treatyAdoption_" + hash);
  OInstance treatyAdoptionInst = ontology.addOInstance(treatyAdoptionInstURI, treatyAdoptionClass);
  docInstMap.put(treatyAdoptionInst,"treatyAdoption");
  // add treatyAdoption properties
  treatyAdoptionInst.addObjectPropertyValue(involves, treatyInst);
  treatyAdoptionInst.addDatatypePropertyValue(dateDP, new Literal(xsdtreatyAdoptionDateStr,
                                              dateDP.getDataType()));

}
catch (InvalidValueException err) {
  throw new JapeException(err);
}
// create treatyEIF instance
if (!instanceAnnots.get("TreatyEIFdate").isEmpty()) {
  try {
    OClass treatyEIFClass = ontology.getOClass(ontology.createOURIForName("generalTreatyEIF"));
    OURI treatyEIFInstURI = ontology.createOURIForName("generalTreatyEIF_" + hash);
    OInstance treatyEIFInst = ontology.addOInstance(treatyEIFInstURI, treatyEIFClass);
    docInstMap.put(treatyEIFInst, "generalTreatyEIF");
    // add treatyEIF properties
    treatyEIFInst.addObjectPropertyValue(involves, treatyInst);
    String dateStr = stringFor(doc, instanceAnnots.get("TreatyEIFdate"));
    LocalDate d = LocalDate.parse(dateStr, textDate);
    String xsdDateStr = d.format(DateTimeFormatter.ISO_LOCAL_DATE);
    treatyEIFInst.addDatatypePropertyValue(dateDP, new Literal(xsdDateStr,
                                                dateDP.getDataType()));
    // add nb of days elapsed since adoption
    DatatypeProperty daysDP = ontology.getDatatypeProperty(ontology.createOURIForName("incubDays"))
        ↪ ;
    Long daysdelta = ChronoUnit.DAYS.between(treatyAdoptionDate, d);
    treatyEIFInst.addDatatypePropertyValue(daysDP, new Literal(daysdelta.toString(),
                                                daysDP.getDataType()));

  }
  catch (InvalidValueException | DateTimeParseException err) {
    throw new JapeException(err);
  }
}
// create treatyRegistration instance
if (!instanceAnnots.get("TreatyRegDate").isEmpty()) {
  try {
    String className = "treatyRegistration";
    OClass aClass = ontology.getOClass(ontology.createOURIForName(className));
```

```java
      OURI instanceURI = ontology.createOURIForName(className + "_" + hash);
      OInstance inst = ontology.addOInstance(instanceURI, aClass);
      docInstMap.put(inst, className);
      // link registration instance to treaty instance and to UN
      inst.addObjectPropertyValue(involves, treatyInst);
      inst.addObjectPropertyValue(hasParticipant, UN);
      // add date
      String dateStr = stringFor(doc, instanceAnnots.get("TreatyRegDate"));
      LocalDate d = LocalDate.parse(dateStr, textDate);
      String xsdDateStr = d.format(DateTimeFormatter.ISO_LOCAL_DATE);
      inst.addDatatypePropertyValue(dateDP, new Literal(xsdDateStr, dateDP.getDataType()));
    }
  catch (InvalidValueException | DateTimeParseException err) {
    throw new JapeException(err);
  }
}
// automatically create properties/relations
for (Annotation instanceAnn : instanceAnnots) {
  if (!instanceAnn.getType().matches("TreatyAction")) {
    String sourceClassName = (String)instanceAnn.getFeatures()
      .get(gate.creole.ANNIEConstants.LOOKUP_CLASS_FEATURE_NAME);
    if (sourceClassName != null) {
      OURI sourceInstOURI = ontology.createOURI((String)instanceAnn.getFeatures()
                                    .get("instanceURI"));
      OInstance sourceInst = ontology.getOInstance(sourceInstOURI);
      // create object properties
      for (Map.Entry<Object, Object> e : instanceAnn.getFeatures().entrySet()) {
        if (!e.getKey().toString().matches("class|instanceURI|hasParticipant|lastRetrieved|" +
                              "hasDate|hasNotifDate|hasEffectDate|hasURL")) {
          String objPropName = e.getKey().toString();
          ObjectProperty objProp = ontology.getObjectProperty(ontology.createOURIForName(
              ↪ objPropName));
          String targetInstClassName = e.getValue().toString();
          OInstance targetInst = null;
          // number of keys for this value should be 1
          if (docInstMap.containsValue(targetInstClassName)) {
            for (Map.Entry<OInstance,String> en : docInstMap.entrySet()) {
              if (en.getValue().equals(targetInstClassName)) {targetInst = en.getKey();}
            }
            try {
              sourceInst.addObjectPropertyValue(objProp, targetInst);
            }
            catch (InvalidValueException | NullPointerException err) {
              throw new JapeException(err);
            }
          }
        }
      }
      // create treaty participant instances
      if (e.getKey().toString().matches("hasParticipant")) {
        try {
          String objPropName = e.getKey().toString();
          ObjectProperty objProp = ontology.getObjectProperty(ontology.createOURIForName(
              ↪ objPropName));
          String targetInstName = e.getValue().toString();
          OURI targetInstURI = ontology.createOURIForName("treatyParticipant_" +
                                    targetInstName.hashCode());
```

```java
          OInstance targetInst = ontology.getOInstance(targetInstURI);
          sourceInst.addObjectPropertyValue(objProp, targetInst);
        }
        catch (InvalidValueException | NullPointerException err) {
          throw new JapeException(err);
        }
      }
      // add data properties for dates and datetimes
      if (e.getKey().toString().matches("lastRetrieved|hasDate|hasNotifDate|hasEffectDate")) {
        String dataPropName = e.getKey().toString();
        DatatypeProperty dataProp = ontology.getDatatypeProperty(ontology.createOURIForName(
            ↪ dataPropName));
        String dateStr = e.getValue().toString();
        try {
          if (dataPropName.matches("lastRetrieved")) {
            String xsdDateTimeStr = dateStr;
            sourceInst.addDatatypePropertyValue(dataProp, new Literal(xsdDateTimeStr,
                                                          dataProp.getDataType()));

          } else {
            LocalDate d = LocalDate.parse(dateStr, tableDate);
            String xsdDateStr = d.format(DateTimeFormatter.ISO_LOCAL_DATE);
            sourceInst.addDatatypePropertyValue(dataProp, new Literal(xsdDateStr,
                                                          dataProp.getDataType()));

          }
        }
        catch (InvalidValueException | DateTimeParseException err) {
          throw new JapeException(err);
        }
      }
      // add data property for URLs
      if (e.getKey().toString().matches("hasURL")) {
        String dataPropName = e.getKey().toString();
        DatatypeProperty dataProp = ontology.getDatatypeProperty(ontology.createOURIForName(
            ↪ dataPropName));
        String dataPropStr = e.getValue().toString();
        try {
          sourceInst.addDatatypePropertyValue(dataProp, new Literal(dataPropStr));
        }
        catch (InvalidValueException err) {
          throw new JapeException(err);
        }
      }
    }//for Map.Entry
  }// if has class feature
}//if relevant instanceAnn
}//for instanceAnn


// create UNTSvolume instance (at the end to avoid trouble with isPartOf treatyPublication
    ↪ properties)
if (!instanceAnnots.get("UNTSvolRef").isEmpty()) {
  try {
    String className = "treatyPublication";
    OClass aClass = ontology.getOClass(ontology.createOURIForName(className));
    String instStr = stringFor(doc, instanceAnnots.get("UNTSvolNb"));
    OURI instanceURI = ontology.createOURIForName("UNTSvol" + instStr);
    OInstance inst = ontology.addOInstance(instanceURI, aClass);
```

```java
      inst.setLabel("UNTS volume " + instStr, OConstants.ENGLISH);
      docInstMap.put(inst, className);
      // link it to UNTS
      ObjectProperty partOf = ontology.getObjectProperty(ontology.createOURIForName("isPartOf"));
      inst.addObjectPropertyValue(partOf, UNTS);
      // add UNTSvolume pdf version URL
      if (!instanceAnnots.get("UNTSvolURL").isEmpty()) {
        String url = stringFor(doc, instanceAnnots.get("UNTSvolURL"));
        DatatypeProperty dataProp = ontology.getDatatypeProperty(ontology.createOURIForName("hasURL")
            ↪ );
        inst.addDatatypePropertyValue(dataProp, new Literal(url));
      }
    }
    catch (InvalidValueException err) {
      throw new JapeException(err);
    }
  }
}
}
```

F.1.3   SPARQL query for UNTS metadata

This sample query can be used to retrieve UNTS metadata from a PILO database. GATE and Protégé have their own data query and export functionalities, but SPARQL has the advantage of being a standardised query language (like SQL) and can be used in many different software applications, online and offline. See https://www.w3.org/TR/sparql11-query/ for the specification.

```sparql
PREFIX :     <https://gitlab.com/legalinformatics/pilo/core/pilo.ttl#>
PREFIX rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT DISTINCT ?treaty ?treatyTitle ?treatyAdoptionDate ?treatyAdoptionPlace ?treatyEIFdate ?
    incubDays ?treatyEIFprovInfo ?treatyLangs ?treatySubjectTerms ?UNTSvolRef ?UNTStreatyRecordURL ?
    UNTStrLastR ?treatyRegDate ?treatyDepositary

WHERE {
  ?treaty               rdf:type                :treaty .
  ?UNTStreatyRecord     rdf:type                :treatyPublication ;
                        :isAbout                ?treaty ;
                        :hasURL                 ?UNTStreatyRecordURL ;
                        :lastRetrieved          ?UNTStrLastR .
  ?treatyTitleInst      :denotes                ?treaty ;
                        :hasInstanceString      ?treatyTitle .
  ?treatyAdoption       rdf:type                :treatyAdoption ;
                        :involves               ?treaty ;
                        :hasDate                ?treatyAdoptionDate .
  ?AdoptionLoc          :isOfficialSiteOf       ?treatyAdoption ;
                        rdfs:label              ?treatyAdoptionPlace .
  OPTIONAL { ?treatyEIF         :involves               ?treaty ;
                        :hasDate                ?treatyEIFdate ;
                        :incubDays              ?incubDays } .
  OPTIONAL { ?treatyRegistration  rdf:type              :treatyRegistration ;
                        :involves               ?treaty ;
```

```
                                    :hasDate                 ?treatyRegDate } .
  OPTIONAL { ?EIFprovInfo           rdf:type                 :treatyEIFprovInfo ;
                                    :isPartOf                ?UNTStreatyRecord ;
                                    :hasInstanceString       ?treatyEIFprovInfo } .
  OPTIONAL { ?Lang                  :isOfficialLanguageOf    ?treaty ;
                                    :hasInstanceString       ?treatyLangs } .
  OPTIONAL { ?Subject               :isAbout                 ?treaty ;
                                    rdf:type                 :treatyDescriptor ;
                                    :hasInstanceString       ?treatySubjectTerms } .
  OPTIONAL { ?Depositary            :isDepositaryOf          ?treaty ;
                                    :hasInstanceString       ?treatyDepositary } .
  OPTIONAL { ?UNTSvolRefInst        rdf:type                 :treatyPublicationRef ;
                                    :isPartOf                ?UNTStreatyRecord ;
                                    rdfs:label               ?UNTSvolRef } .
}
```

F.2   GATE treaty texts IE

The GATE treatytextsIE app is a conditional corpus pipeline that can be used to annotate
and extract treaty text metadata, provisions and mentions, and populate a PILO database
with some or all of this information. Only some excerpts of the provisions and mentions
regex pattern files are included here due to space constraints. The full version of the app
is available for download from the thesis code archive.

F.2.1   Treaty provisions annotation

```
// end of section
eos=\n\n(?:Article|Regulation|Chapter|Part|In Witness Whereof|Annex|Appendix|[A-Z]{2,} ?\d? ?[.-] [A-
    ↪ Z]{3,})


////////////// Treaty life cycle //////////////
//// Entry into force ////
// named arts
|(?ism)^(Article (\w+)\s*.?\s*\b([^\n]*Entry into force(?:,? \w+){0,5})\n+.+?)(?=<<eos>>)
1 => EntryIntoForceProvision class="EIFprov",governs="treatyEIF",hasArtNb=$2,hasArtTitle=$3


// EIFdelay and EIFnConsent extraction
|(?i)\n\n(Article (\w+)\b[^.]*.?\s?Th\w+(?: present)? (?!amendment)(?:\w+ )?\w+ shall \w+ into force
    ↪ (?:on the )?([^.]+? (?:day|month|year)[s]?) (?:after|following) (?:the \w+ \w+ which (?:the )
    ↪ ?(?:Governments of )?)?(?:not less than|at least|the)? ?([^.]{0,20}?) (?:States|Governments|
    ↪ Parties)[^.]*? have (?:become parties|deposited \w* ?instrument|[^.]*?(?:ratif|consent to be
    ↪ bound))[^.]+[.])
1 => GeneralEIFprovision class="generalTreatyEIFprov",governs="generalTreatyEIF",hasArtNb=$2,EIFdelay
    ↪ =$3,EIFnConsent=$4


|(?:Subject to the prov\w+ of Art\w+ \d+.\d, t|T)h\w+(?: present)? (?!(?:amendment|governing))(?:\w+
    ↪ )?\w+ (?:shall \w+|enters) into force (?:on the same day as the \w+ or )?(?:on the )?([^.]+?
    ↪ (?:day|month|year)[s]?) (?:after|following|from) (?:the (?:date|day) of )?(?:the )?deposit
    ↪ [^.]{0,50}? of (?!its )(?:the )?([^.]{0,20}?) ?(?<!an )instrument[s]? of ratification[^.]*[.]
```

```
0 => GeneralEIFprovision class="generalTreatyEIFprov",governs="generalTreatyEIF",EIFdelay=$1,
    ↪ EIFnConsent=$2


|(?i)\n(.*This \w+ shall enter into force(?:, with respect to [^.]{0,100}?)?(?: on the)? (\w+ (?:day|
    ↪ month)[s]?(?: of the \w+ month)?) after (?:consent to be bound has been expressed by|(?:the \
    ↪ w+ .n which )?the)? ?(\w+) (?:States|instrum\w+ of ratif[^.]+)[.])
1 => GeneralEIFprovision class="generalTreatyEIFprov",governs="generalTreatyEIF",EIFdelay=$2,
    ↪ EIFnConsent=$3


|(?i)\n(Article (\w+)\n+\d.?\s?This (?!amendment)(?:\w+ )?\w+ shall be binding only upon[^.]+[.]\n+\d
    ↪ .\s?It shall \w+ into force (\w+ (?:day|month)[s]?) (?:after|following) the date on which the
    ↪ \w+ of (\w+) Members[^.]+[.])
1 => GeneralEIFprovision class="generalTreatyEIFprov",governs="generalTreatyEIF",hasArtNb=$2,EIFdelay
    ↪ =$3,EIFnConsent=$4


////////////////// Termination //////////////////
// named arts
|(?is)\n(Article (\w+)\s*.?\s*\b((?:\w*.? ?\w+ (?:and|or) )?(?:Termination|Liquidation)(?!(?: of
    ↪ assistance|\n\nThe conciliation proc))(?:,? \w+){0,5})\n+.+?)(?=<<eos>>)
1 => TerminationProvision class="termProv",governs="treatyTermination",hasArtNb=$2,hasArtTitle=$3


// unnamed arts
|(?i)\n(Article (\w+)\n\n.*(?:This \w+ shall cease to be in force|Notwithstanding the termination of
    ↪ this Conv).+)
1 => TerminationProvision class="termProv",governs="treatyTermination",hasArtNb=$2


// autoExpiry period
|(?i)\n([^\n]*shall remain in force (?:for a period of|until the end of the) (\w+ (?:\w* ?\w+ )?year
    ↪ .?) (?:\w+ ){0,5}(?:entry into force|unless extended)[^.]+[.])
1 => autoExpiry class="autoExpiryProv",governs="treatyTermination",autoExpiryPeriod=$2


// autoExpiry date
|(?i)\n([^\n]*shall remain in force until (\w+ \w+ \d+),? unless (?:it is )?(?:extended|prolonged)
    ↪ [^.]*[.])
1 => autoExpiry class="autoExpiryProv",governs="treatyTermination",autoExpiryDate=$2


/////////// Secretariat /////////////
// named arts
|(?is)\n(Article (\w+)\s*.?\s*\b([^\n]*(?:Secretariat|Secretary|Bureau|The Registry|Staff|(?:Staff|
    ↪ Office|Duties) of the [^DP]\w+)(?: and \w+-?\w+)?)\n+.+?)(?=<<eos>>)
1 => SecretariatProvision class="secretariatProv",governs="treatySecretariat",hasArtNb=$2,hasArtTitle
    ↪ =$3


// named arts (title before art)
|(?is)\n((Secretariat)\n+Article (\w+)\n+.+?)(?=<<eos>>)
1 => SecretariatProvision class="secretariatProv",governs="treatySecretariat",hasArtNb=$3,hasArtTitle
    ↪ =$2


/////////////// Dispute settlement //////////////////
// named arts
|(?is)\n(Article (\w+)\s*.?\s*\b([^\n]*(?:Dispute Settlement|(?<!Prevention of )Dispute[s]?|
    ↪ Resolution of disagreements(?=\n+[^\n]+arbitrat)|Interpret\w+ \w+ (?:arbitration|Constitution
    ↪ ))(?!\n+[^\n]+(?:Committee shall seek to resolve|(?:referred|bring) to the Council|\n+[^\n]+
    ↪ committee of experts)))\n.+?)(?=<<eos>>)
1 => DisputeSettlementProvision class="disputeProv",governs="intlDisputeSett",hasArtNb=$2,hasArtTitle
    ↪ =$3
```

```
// unnamed arts
|(?is)\n(Article (\w+)\n+[^\n]*(?:settlement of disputes under this|dispute[^.]+ interpretation or
    ↪ application[^.]+ be (?:referred|submitted)|settlement of the dispute by|procedure for
    ↪ settlement of dispute|controversy between the Parties[^\n]+arbitration|establish a Claims
    ↪ Commission).+?)(?=<<eos>>)
1 => DisputeSettlementProvision class="disputeProv",governs="intlDisputeSett",hasArtNb=$2
```

F.2.2   Treaty mentions annotation

The regex patterns for environmental problems and risks are included in Chapter 4.1 and thus not repeated here.

```
//////////////////// Human activities ////////////////////
// Fishing/hunting
|(?i)(?:fishing|fishery|fisheries)
|\bto fish\b
|(?:fish|bird[s]?) kill\w*
|(?i)(?:whaling|poaching|hunt\w*)
|(?i)(?:allowable catch|introduction from the sea)
0 => FishHunt class="mention",isAbout="fishingHunting"


agricprod=(?:crop|timber|coffee|dairy|olive|cocoa|sugar|cereal|rice|grain)
// Farming & forestry
|(?i)\bagricult(?!ure organi.ation)\w+
|(?i)\b(?:horticult|aquacult|farm|\w+ husbandry|forestry)\w*
|(?i)\b(?:grower|breeder)[s]?|bred in captivity|pastoral\w*
|(?i)<<agricprod>>[s]?.(?:prod|grow(?!er)|cultiv)\w*
|grow(?!er)\w* <<agricprod>>
|(?<!-)producing,[^.]+?(?<!grows )(?:coffee|sugar)
0 => Agriculture class="mention",isAbout="agriculture"


// Production/manufacturing
|(?i)\b(?:manufactur|fabrication|factor[yi])\w*
|(?i)\b(?<!(?:oxygen|rural|crop|land resource) )produc(?:er|tion)(?! of \w* ?(?:evidence|document|
    ↪ report|record|statistics))
|(?i)\b(?:produces, processes|packaging,? \w* ?processing|products or \w+ produced)
|(?i)\b(?:prod\w+ energy|energy prod|industrial \w* ?\w* ?process|domestically produced)\w*
0 => Manufacturing class="mention",isAbout="manufacturing"


// Storage/stocks
|(?i)\b(?:storage|storing|store\W|stores|stored|stockpil\w+|warehousing|stow(?!away)|container\w*)
|(?i)\b(?:(?:imports|supply|verified|obsolete),? stocks|stock level[s]?)
|(?i)\b(?:stocks of (?:sugar|cocoa|coffee|mercury|\w* ?products)|(?:sugar|cocoa|coffee|mercury)
    ↪ stocks)
0 => Storage class="mention",isAbout="storage"


// Transport/freight
|(?i)(?<!(?:atmospheric|environmental(?: cycle,)?|bioaccumulate and are|long-range|horizontal|animal
    ↪ which normally)\W)\btransport\w*
|(?i)\b(?:trans?)?ship(?:ping|ment)[s]?
|(?i)\b(?:carriage(?!way)|freight|safe transfer|delivery(?= to))
|(?i)(?:launch\w* into outer space|return\w* to earth|launching(?= (?:State|authority)))
0 => Transport class="mention",isAbout="transport"
```

```
// Trade
|(?i)\b(?<!department of foreign affairs and )trade\W(?!(?:name|mark|union|secret))
|(?i)\b(?:trading|import(?!a)\w*|export\w*|international commerce|traffick\w*)
|(?i)\b(?:transfer to .+ any (?:non-nuclear-weapon )?State|acquire nuclear weapon)[s]?
0 => Trade class="mention",isAbout="internationalTrade"


// Weapons of mass destruction (any related activity)
|(?i)weapon[s]? of mass destruction
|(?i)(?:nuclear|radiological|chemical|biological|toxi.|poisoned) weapon\w*
|(?i)nuclear (?:arm|disarm|test)\w*
|(?i)weapon.+biological agents
0 => WMDs class="mention",isAbout="weaponOfMassDestruction"


///////////////// Means of intervention //////////////////
// Quantitative restrictions
|(?i)\b(?:quota|quantitative restriction|fixed allowance|quantified \w* ?limit\w*|assigned amount)[s
      ↪ ]?\b
|(?i)\b(?:emission reduction (?:or limit\w+ )?target|(?:emission|release) limit value)[s]?\b
|(?i)\b(?:maximum \w* ?(?:amount|quantit\w+|limit|level|catch|dose)|ceiling value)[s]?\b
|(?i)\b(?:limited quantities|(?:exposure|production) limit|subject to \w* ?restrict\w+ as to number)[
      ↪ s]?\b
|(?i)\b(?:level of (?:consumption|production)[^.]+? (?:not exceed|limit)|quantity per day)[s]?\b
|(?i)\b(?:levels of radioact\w+ greater than de minimis .exempt. \w+|within limits \w+ ?\w* by)\b
|(?i)\b(?:permits for (?:not allow|remove from use).+(?:above|greater than) \d*.?\d+ per.?cent)\b
|(?i)\b(?:exact quantit\w+|(?:reduc|decreas)\w+ \w* ?the (?:amount|quantity) of|quantit[^.]+?no
      ↪ justif\w+|permit[^.]+quantities)[s]?\b
0 => QuantRestriction class="mention",isAbout="quantitativeRestriction"


// Spatial regulation/zoning
|(?i)(?:protect|sanctuary|delineat|designat|clearly define|speci|given|certain|limited|endangered|
      ↪ pest.free|fishing|dumping|mined|booby-trapped|affected|inspect|populated)\w* (?:natural|
      ↪ geographical)? ?area[s]?
|(?i)exclusive \w+ zone(?!.+\n*.+nothing in this \w+ shall)
|(?i)(?:sea-bed|tropical) zone
|(?i)(?:natural heritage|(?:land|water|nature) reserve[s]?)
|(?i)(?:international scientific preserves|areas of special importance)
|(?i)(?:nuclear-weapon-free|demilitari.ed|combat|safety) zone
|(?i)zone[s]? (?:under special protection|susceptible to)
|(?i)\b(?:spatial planning|siting policy)
|(?i)(?<=\n)Article.+\Wsiting
|(?i)(?:site-limited|(?:identif|managing).+sites)
0 => SpatialReg class="mention",isAbout="spatialRegulation"


// Risk/impact assessment
evaltype=(?:impact|risk|hazard|chemical|safety|environmental)


|(?i)<<evaltype>> (?:assessment|analysis|evaluation)[s]?
|(?i)(?<!<<evaltype>> )(?:assess|eval|(?<!evaluat\w{1,3} and )analy|investig|determin\w+ (?:of|
      ↪ whether))[^.;\n]+?(?:risk|impact|hazard|effects)\w*
|(?i)(?:examin.+?chemical.+?risk|risk profile)[s]?
|assess\w* the conseq\w+
|medical examination[s]?(?=[^.]+?(?:hazard|at \w* ?intervals|periodical))
0 => RiskImpactAssessment class="mention",isAbout="riskImpactAssessment"


// Inspections
```

```
|(?i)(?<!public )inspect\w*
|(?i)advance notice of a projected visit
|(?:Atomic Energy Agency|IAEA)[^.]+? safeguards|safeguards[^.]+? (?:Atomic Energy Agency|IAEA)
|(?i)(?:investig[^.]+in loco|on-the-spot investigation[s]?)
|(?i)(?:fact-finding mission|visit the scene)[s]?
0 => Inspection class="mention",isAbout="inspection"


// Data & research
|(?i)\b(?:research|science)\w*
|(?i)(?<!personal )\bdata\b
|(?i)(?:scientific|techn)\w* (?:eval|assess|info|evidence|co.?op|knowledge|studies|study|results|
    ↪ investig)\w*
|(?i)(?:studies|analy|assess)\w*(?=.+?(?:carr\w+ out|result|assess|criteria|cause|effect|impact|
    ↪ samples|monitor|hazard))
|(?i)(?:statistic|forecast)\w*
|(?i)(?:dev|applic|exchang)\w+ of scientific
|(?i)(?:measur\w+ the concentra|systematic observ)\w+
|(?i)monitoring (?:\w+ tracking|system|program)\w*
|(?i)(?:poll|env|radiat)\w+ monitoring
|(?i)(?:analysis|reporting) of \w* ?accidents
0 => DataResearch class="mention",isAbout="dataResearch"


// Education
|(?i)(?<!the United Nations )educat\w+
|(?i)teaching
|(?i)(?:public|consumer|rais\w+) awareness
|(?i)(?<!(?:public|consumer) )awareness.(?:raising|campaign|program)\w*
0 => Education class="mention",isAbout="education"


// Training/capacity-building
|(?i)\b(?:re.?)?training
|(?i)\b(?<!(?:production |her than a ))train(?!,? \w* ?(?:road vehicle|coach))\b
|(?i)human resource[s]? development
|(?i)(?:build\w* capacity|capacity..?building)
|(?i)(?:develop|strengthen)\w* (?:of )?(?:human resource|capa(?!city-building))\w*
|Secretariat [^.]+ assist in the development of its work
|workers [^.]+ shall be \w* ?instructed
|programs of military instruction
0 => Training class="mention",isAbout="training"


// Technology transfer
|(?i)technolog\w+(?: \w+,?){0,5} transfer\w*
|(?i)(?<!heat )(?:(?:facilitate|provide) access|transfer|exchange)\w*,?(?: \w+,?){0,6} \w*.?technolog
    ↪ \w+
|(?i)access to \w* ?technolog.+facilitat\w+
|(?i)technolog\w+ under fair and most favo.?rable terms
|(?i)strengthen\w+ \w+ scientific and technol\w+ capab\w+
|(?i)supply(?: of)? equipment \w+ the State[s]?
0 => TechTransfer class="mention",isAbout="techTransfer"


// Insurance
|(?i)\b(?:insurance|bond or other \w+ of security)
|(?i)financial (?:guarantee|securit)\w*
0 => Insurance class="mention",isAbout="insurance"


// Liability
```

```
|(?i)\b(?:liability|liabilities|liable for (?!punishment)|liable to pay)
|shall be liable(?! (?:to pay|for))
0 => Liability class="mention",isAbout="liability"

// Criminal responsibility
|(?i)(?:crime|criminal|penal(?=\W)|infraction|prosecution|extradition)[s]?
0 => criminalResp class="mention",isAbout="criminalResponsibility"
```

G Treaty profile sample

This is an overview of the main variables constructed for the *Paris Agreement*. The source code to generate this figure can be found in the thesis code archive. An online version of this profile page could have embedded text excerpts that appear upon click or hover on the relevant values.

Paris Agreement

Adopted: 2015-12-12; in force: 2016-11-04
Treaty label: ParisAgreement; filename: 2015-Paris-1992-ClimateChange.EN.txt

Overview

I. Treaty Subject Matter

Cluster	Variable	
Problems & Risks	Climate Change	
	Ozone Depletion	
	Pollution	
	Waste	
	Land Degradation	
	—Deforestation	
	—Desertification	
	Endangered Species	
	Pests & Weeds	
	Animal Welfare	
Human Stressors	Fishing & Hunting	
	Agriculture	
	Production	
	Storage & Stockpiling	
	Transport	
	Trade	
	Weapons of Mass Destr.	
Means of Intervention	Quant. Restrictions	
	Spatial Regulation	
	Risk/Impact Assessment	
	Inspections	
	Data & Research	
	Education	
	Training/Capacity-build.	
	Tech Transfer	
	Insurance	
	Liability	
	Criminal Responsibility	

II. Treaty Provisions

Cluster	Variable	
Start of Life	Provisional Application Provision	No
	Entry into Force Provision	Yes
	—Min. Nb. of Ratifs for EIF	55
	—Additional EIF Conditions	Yes
	—EIF Waiting Period (Years)	0.08
Change	Amendment Provision	Yes
	—Simplified Amendment	Yes
	Adoption of Annexes or Protocols	Yes
End of Life	Denunciation Provision	Yes
	—Denunciation Waiting Period (Years)	3
	—Other Denunciation Conditions	No
	—Denunciation Effect Delay (Years)	1
	—Other Denunciation Effect	No
	Termination Provision	No
	—Auto-Termination Memb. Threshold	-
	—Auto-Expiry	No
Treaty Organs	Legal Personality	No
	Organisational Auspices	UN
	Secretariat	Yes
	Conference of Parties	Yes
Compliance & Effectiveness	National Reporting	Yes
	International Verification	No
	Progress Review	Yes
	Non-Compliance Procedures	Yes
	Sanctions	No
	Dispute Settlement	Yes
	—Compulsory Dispute Settlement	opt-in

H Treaty EIF predictions

This Appendix contains the source code for the data analysis in Chapter 5.

H.1 Data loading & wrangling

```python
import os, re, time, random
from datetime import datetime
import numpy as np
import pandas as pd

wd = os.path.expanduser('~/workspaces/org-workspaces/EIFpreds/')
rs = 7
textdf = pd.read_csv(os.path.expanduser('~/git/phd/textdf2.csv'), encoding='utf-8') \
        .dropna(how='all').set_index('fn')
smdf = pd.read_csv(os.path.expanduser('~/git/phd/smdf2.csv'), encoding='utf-8') \
       .dropna(how='all').set_index('vars')
idcols = ['ieaid','SignatureDate','TreatyName','shortTitle','treatyLabel']
df = pd.read_csv(os.path.expanduser('~/git/phd/treatymeta.csv'), encoding='utf-8').dropna(how='all')
df['treatyLangLists'] = df.treatyLangs.str.split('\t').str[:-1]
df['nLangs'] = df.treatyLangLists.apply(lambda x: (len(x) if x is not np.nan else np.nan))
df['incubWeeks'] = df.incubDays/7
df['incubYears'] = df.incubDays/365.25
df['incubConsent'] = np.nan
df = pd.merge(df, textdf, on=idcols, how='outer').sort_values(by='treatyAdoptionDate')
smvars = list(smdf.index + 'B')
# binarise categorical variables
catcols = ['treatyAdoptionPlace','orgAuspices','compDispSett']
dummies = pd.get_dummies(df[catcols], prefix=catcols)
dummies = dummies[dummies.columns[dummies.sum() >=10]]
dummies.columns = dummies.columns.str.replace('-','')
df = pd.concat([df, dummies], axis=1)
# deal with missing values in numerical variables
df.EIFdelay.fillna(0, inplace=True)
df.denunNotifMinYrs.fillna(0, inplace=True)
df.denunEffectYrs.fillna(1, inplace=True)
df['incubConsent'] = df.incubYears - df.EIFdelay
cols = (['treatyAdoptionYear','nLangs'] + smvars +
        provdfdisp[provdfdisp.type!='str'].vars.values.tolist() + list(dummies.columns))
# exclude rows with missing target and key feature
dfX = df.loc[df.incubYears.notna() & df.EIFnConsent.notna(), cols]
# exclude cols with missing values
cols = [c for c in cols if dfX[c].count() == len(dfX)]
# exclude cols with variance = 0
cols = [c for c in cols if dfX[c].nunique() > 1]
dfX = dfX[cols]
# convert to float (for neural networks)
for c in cols:
  dfX[c] = dfX[c].astype(float)
# add random features
rng = np.random.default_rng(rs)
dfX['bin_var'] = rng.integers(low=0, high=2, size=len(dfX)).astype(float)
```

```
ncols = len(dfX.columns)
X = dfX.copy()
y = df.incubConsent[df.incubConsent.notna() & df.EIFnConsent.notna()].copy()
print(dfX.shape)
```

H.2   Comparison of model performance

```python
import random
from sklearn.model_selection import KFold, ParameterGrid
from sklearn.linear_model import LinearRegression, Ridge, Lasso
from sklearn.metrics import mean_absolute_error as mae
from sklearn.metrics import r2_score
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, regularizers


starttime = datetime.now()
resdf = pd.DataFrame(np.nan, columns=['rs','outer_k','inner_k','testSize','testIndex','regressor',
                                      'nInputs','numInputScale','binInputScale','hpGrid','hpRes',
                                      'bestHps','params','trainMAE','testMAE', 'trainR2','testR2',
                                      'testPreds','meanPredMAE','medianPredMAE'], index=[0])
resdf.index.name = 'datetime'
# convert list cols to object dtype
resdf.testIndex = resdf.testIndex.astype('object')
resdf.testPreds = resdf.testPreds.astype('object')
# number of random trials
n_repeats = 10
outer_k = inner_k = 5
# loop for each repeat
for i in range(n_repeats):
  random.seed(i)
  np.random.seed(i)
  tf.random.set_seed(i)
  # define cross-validation method for outer and inner loops
  outer_cv = KFold(n_splits=outer_k, shuffle=True, random_state=i)
  inner_cv = KFold(n_splits=inner_k, shuffle=True, random_state=i)
  for j, (outer_train_index, outer_test_index) in enumerate(outer_cv.split(X)):
    t = str(datetime.now())[:-4]
    resdf.loc[t, ['rs']] = i
    resdf.loc[t,'outer_k'] = j
    resdf.loc[t,'inner_k'] = -1
    test_ind = y.index[outer_test_index]
    resdf.loc[t,'testSize'] = len(test_ind)
    resdf.at[t,'testIndex'] = list(test_ind)
    X_train, X_test = X.drop(index=test_ind), X.loc[test_ind,:]
    y_train, y_test = y.drop(index=test_ind), y[test_ind]
    resdf.loc[t,'meanPredMAE'] = abs(y_test - y_train.mean()).mean()
    resdf.loc[t,'medianPredMAE'] = abs(y_test - np.median(y_train)).mean()
    # rescale continuous input vars (z-score)
    resdf.loc[t,'numInputScale'] = 'z-score'
    numcols = ['treatyAdoptionYear','EIFnConsent', 'EIFdelay', 'denunNotifMinYrs','denunEffectYrs']
    numcols = [c for c in numcols if c in X.columns]
    X_train_means = X_train[numcols].mean()
```

```python
X_train_stds = X_train[numcols].std()
for c in numcols:
  X_train[c] = (X_train[c] - X_train_means[c]) / X_train_stds[c]
  X_test[c] = (X_test[c] - X_train_means[c]) / X_train_stds[c]
resdf.loc[t,'binInputScale'] = '0|1'
# OLS regression
resdf.loc[t, 'regressor'] = 'OLS'
resdf.loc[t, 'nInputs'] = len(X.columns)
model = LinearRegression().fit(X_train, y_train)
resdf.loc[t,'trainR2'] = model.score(X_train, y_train)
resdf.loc[t,'testR2'] = model.score(X_test, y_test)
resdf.at[t,'testPreds'] = model.predict(X_test).round(4)
resdf.loc[t,'testMAE'] = mae(y_test, model.predict(X_test))
resdf.at[t,'params'] = [{'coefs': model.coef_, 'intercept': model.intercept_.round(5)}]
# hyperparameter tuning loop
for k, (inner_train_index, inner_dev_index) in enumerate(inner_cv.split(X_train)):
  t = str(datetime.now())[:-4]
  resdf.loc[t,'inner_k'] = k
  dev_ind = y_train.index[inner_dev_index]
  resdf.loc[t,'testSize'] = len(dev_ind)
  resdf.at[t,'testIndex'] = list(dev_ind)
  # get unscaled X vars and standardise with inner training data
  X_dev = X.loc[dev_ind,:]
  inner_X_train = X.loc[X_train.index.drop(dev_ind),:]
  inner_y_train, y_dev = y_train.drop(dev_ind), y_train[dev_ind]
  resdf.loc[t,'meanPredMAE'] = abs(y_dev - inner_y_train.mean()).mean()
  resdf.loc[t,'medianPredMAE'] = abs(y_dev - np.median(inner_y_train)).mean()
  inner_X_train_means = inner_X_train[numcols].mean()
  inner_X_train_stds = inner_X_train[numcols].std()
  for c in numcols:
    inner_X_train[c] = (inner_X_train[c] - inner_X_train_means[c]) / inner_X_train_stds[c]
    X_dev[c] = (X_dev[c] - inner_X_train_means[c]) / inner_X_train_stds[c]
  # Ridge (L2 regularisation)
  resdf.loc[t, 'regressor'] = 'Ridge'
  hpGrid = {'alpha': [i.round(1) for i in np.arange(0.1, 20.2, 0.2)]}
  resdf.at[t, 'hpGrid'] = [hpGrid]
  hpres = {}
  for a in hpGrid['alpha']:
    model = Ridge(alpha=a, fit_intercept=True, tol=1e-3, solver='auto').fit(inner_X_train,
                                                          inner_y_train)
    hpres[a] = mae(y_dev, model.predict(X_dev))
  resdf.at[t, 'hpRes'] = [hpres]
  bestHp = min(hpres, key=lambda x: hpres[x])
  resdf.at[t, 'bestHps'] = bestHp
  model = Ridge(alpha=bestHp, fit_intercept=True, tol=1e-3, solver='auto').fit(inner_X_train,
                                                          inner_y_train)
  resdf.loc[t,'trainR2'] = model.score(inner_X_train, inner_y_train)
  resdf.loc[t,'testR2'] = model.score(X_dev, y_dev)
  resdf.at[t,'testPreds'] = model.predict(X_dev).round(4)
  resdf.loc[t,'testMAE'] = mae(y_dev, model.predict(X_dev))
  resdf.at[t,'params'] = [{'coefs': model.coef_, 'intercept': model.intercept_.round(5)}]
  # Lasso (L1 regularisation)
  t = str(datetime.now())[:-4]
  resdf.loc[t, 'regressor'] = 'Lasso'
  hpGrid = {'alpha': [0.0001, 0.001, 0.005, 0.01, 0.05] + \
            [i.round(1) for i in np.arange(0.1, 2.1, 0.1)]}
```

```python
resdf.at[t, 'hpGrid'] = [hpGrid]
hpres = {}
for a in hpGrid['alpha']:
  model = Lasso(alpha=a, fit_intercept=True, max_iter=3000, tol=1e-3,
                warm_start=False, selection='cyclic').fit(inner_X_train, inner_y_train)
  hpres[a] = mae(y_dev, model.predict(X_dev))
resdf.at[t, 'hpRes'] = [hpres]
bestHp = min(hpres, key=lambda x: hpres[x])
resdf.at[t, 'bestHps'] = bestHp
model = Lasso(alpha=bestHp, fit_intercept=True, max_iter=3000, tol=1e-3,
              warm_start=False, selection='cyclic').fit(inner_X_train, inner_y_train)
resdf.loc[t,'trainR2'] = model.score(inner_X_train, inner_y_train)
resdf.loc[t,'testR2'] = model.score(X_dev, y_dev)
resdf.at[t,'testPreds'] = model.predict(X_dev).round(4)
resdf.loc[t,'testMAE'] = mae(y_dev, model.predict(X_dev))
resdf.at[t,'params'] = [{'coefs': model.coef_, 'intercept': model.intercept_.round(5),
                         'dualgap': model.dual_gap_.round(5)}]
# MLP regression
keras.utils.set_random_seed(i)
t = str(datetime.now())[:-4]
resdf.loc[t, 'regressor'] = 'MLP'
hpGrid = {'layers': [2,3,4], 'units': [64, 96, 128],
          'init_lr': np.arange(0.0002, 0.0008, 0.0001)}
resdf.at[t, 'hpGrid'] = [hpGrid]
hpres = []
for hps in ParameterGrid(hpGrid):
  model = initializer = lr_schedule = stop_early = history = None
  initializer = keras.initializers.HeUniform(seed=i)
  model = keras.Sequential()
  model.add(keras.Input(shape=(len(X_train.columns),)))
  for hl in range(hps['layers']):
    model.add(layers.Dense(units=hps['units'], activation='relu',
                           kernel_initializer=initializer))
  model.add(layers.Dense(1))
  lr_schedule = keras.optimizers.schedules.ExponentialDecay(hps['init_lr'], 20, 0.90,
                                                            staircase=True)
  model.compile(optimizer=keras.optimizers.Adam(learning_rate=lr_schedule),
                loss='mean_absolute_error')
  stop_early = keras.callbacks.EarlyStopping(monitor='val_loss', patience=50)
  history = model.fit(inner_X_train, inner_y_train, batch_size=32, epochs=300,
                      verbose=0, validation_data=(X_dev, y_dev), callbacks=[stop_early])
  hist = pd.DataFrame(history.history)
  hist['weighted_avg_loss'] = (hist.loss + hist.val_loss*2) / 3
  if hist[hist.loss<=hist.val_loss].size != 0:
    subdf = hist[hist.loss<=hist.val_loss]
    hps['weighted_min_loss'] = subdf.loc[subdf.val_loss<=subdf.val_loss.min()+0.02,
                                         'weighted_avg_loss'].min()
  else:
    hps['weighted_min_loss'] = hist.loc[hist.val_loss<=hist.val_loss.min()+0.02,
                                        'weighted_avg_loss'].min()
  hps['best_epoch'] = hist.index[hist.weighted_avg_loss==hps['weighted_min_loss']][0]
  hps['trainMAE'] = hist.loss[hps['best_epoch']]
  hps['testMAE'] = hist.val_loss[hps['best_epoch']]
  for floatval in ['init_lr','weighted_min_loss','trainMAE','testMAE']:
    hps[floatval] = round(hps[floatval], 5)
  hpres.append(hps)
```

```python
    hpresdf = pd.DataFrame(hpres)
    resdf.at[t, 'hpRes'] = hpres
    bestHps = hpresdf[hpresdf.trainMAE<=hpresdf.testMAE].sort_values(by='testMAE') \
                                            .head(1).to_dict('records')[0]
    resdf.at[t, 'bestHps'] = [bestHps]
    resdf.loc[t,'trainMAE'] = bestHps['trainMAE']
    resdf.loc[t,'testMAE'] = bestHps['testMAE']

# model evaluation with best performing hps from inner loop
# Ridge
hpresdf = pd.DataFrame([el[0][0] for el in resdf.loc[(resdf.regressor=='Ridge'),
                                            ['hpRes']].tail(inner_k).values])
bestHp = hpresdf.mean().sort_values(ascending=True).index[0]
t = str(datetime.now())[:-4]
resdf.loc[t,'inner_k'] = -1
resdf.loc[t, 'regressor'] = 'Ridge'
resdf.at[t, 'bestHps'] = bestHp
model = Ridge(alpha=bestHp, fit_intercept=True, tol=1e-3, solver='auto').fit(X_train, y_train)
resdf.loc[t,'trainR2'] = model.score(X_train, y_train)
resdf.loc[t,'testR2'] = model.score(X_test, y_test)
resdf.at[t,'testPreds'] = model.predict(X_test).round(4)
resdf.loc[t,'testMAE'] = mae(y_test, model.predict(X_test))
resdf.at[t,'params'] = [{'coefs': model.coef_, 'intercept': model.intercept_.round(5)}]
# wait to prevent this row from getting overwritten due to truncated datetime
time.sleep(0.1)
# Lasso
hpresdf = pd.DataFrame([el[0][0] for el in resdf.loc[(resdf.regressor=='Lasso'),
                                            ['hpRes']].tail(inner_k).values])
bestHp = hpresdf.mean().sort_values(ascending=True).index[0]
t = str(datetime.now())[:-4]
resdf.loc[t, 'regressor'] = 'Lasso'
resdf.at[t, 'bestHps'] = bestHp
model = Lasso(alpha=bestHp, fit_intercept=True, max_iter=3000, tol=1e-2,
              warm_start=False, selection='cyclic').fit(X_train, y_train)
resdf.loc[t,'trainR2'] = model.score(X_train, y_train)
resdf.loc[t,'testR2'] = model.score(X_test, y_test)
resdf.at[t,'testPreds'] = model.predict(X_test).round(5)
resdf.loc[t,'testMAE'] = mae(y_test, model.predict(X_test))
resdf.at[t,'params'] = [{'coefs': model.coef_, 'intercept': model.intercept_.round(5),
                        'dualgap': model.dual_gap_.round(5)}]
time.sleep(0.1)
# MLP
keras.utils.set_random_seed(i)
model = initializer = lr_schedule = stop_early = history = None
subdf = resdf.hpRes[(resdf.regressor=='MLP')].tail(inner_k).values
hpresdf = pd.DataFrame(np.nan, columns=range(len(subdf[0])), index=range(inner_k))
for combindex in hpresdf.columns:
    hpresdf[combindex] = [gridsearch[combindex]['testMAE'] for gridsearch in subdf]
bestHpsInd = hpresdf.mean().sort_values(ascending=True).index[0]
bestHps = pd.DataFrame([listel[bestHpsInd] for listel in subdf])
meanEpoch = round(bestHps.best_epoch.mean())
bestHps = bestHps[['layers','units','init_lr','best_epoch']].head(1).to_dict('records')[0]
bestHps['best_epoch'] = meanEpoch
t = str(datetime.now())[:-4]
resdf.loc[t, 'regressor'] = 'MLP'
resdf.at[t, 'bestHps'] = [bestHps]
```

```python
    initializer = keras.initializers.HeUniform(seed=i)
    model = keras.Sequential()
    model.add(keras.Input(shape=(len(X_train.columns),)))
    for hl in range(bestHps['layers']):
      model.add(layers.Dense(units=bestHps['units'],
                             activation='relu',
                             kernel_initializer=initializer))
    model.add(layers.Dense(1))
    lr_schedule = keras.optimizers.schedules.ExponentialDecay(bestHps['init_lr'], 20, 0.90,
                                                       staircase=True)
    model.compile(optimizer=keras.optimizers.Adam(learning_rate=lr_schedule),
                 loss='mean_absolute_error')
    history = model.fit(X_train, y_train, batch_size=32, epochs=bestHps['best_epoch']+1, verbose=0)
    hist = pd.DataFrame(history.history)
    resdf.loc[t, 'trainMAE'] = round(hist.loss.values[-1], 5)
    resdf.loc[t,'testMAE'] = round(model.evaluate(X_test, y_test), 5)
    resdf.at[t,'testPreds'] = model.predict(X_test).round(5).T[0]
    resdf.loc[t,'trainR2'] = r2_score(y_train, model.predict(X_train).T[0]).round(5)
    resdf.loc[t,'testR2'] = r2_score(y_test, model.predict(X_test).T[0]).round(5)
  print(resdf.tail().T)
# clean-up & fillnas
resdf.drop(index=0, inplace=True)
for col in ['rs','outer_k','inner_k','nInputs','numInputScale','binInputScale']:
  resdf[col].ffill(inplace=True)
for col in ['testSize','testIndex','meanPredMAE','medianPredMAE']:
  resdf.loc[resdf.inner_k==-1, col] = resdf.loc[resdf.inner_k==-1, col].ffill()
for col in ['testSize','testIndex','meanPredMAE','medianPredMAE']:
  resdf[col].ffill(inplace=True)
for col in ['rs','outer_k','inner_k','testSize','nInputs']:
  resdf[col] = resdf[col].astype(int)
resdf.to_csv(wd+'resdf.csv', encoding='utf-8', index=True)
print(resdf.shape)
print('Time elapsed:', str(datetime.now() - starttime)[:-4])
```

Map MLP testPreds to testIndex

```python
subdf = resdf[(resdf.regressor=='MLP')&(resdf.inner_k==-1)]
treatypreds = {}
for testi in y.index:
  treatypreds[testi] = []
for row in subdf.iterrows():
  for testi in row[1].testIndex:
    idx = row[1].testIndex.index(testi)
    treatypreds[testi].append(row[1].testPreds[idx])
treatypreds = pd.DataFrame(treatypreds)
treatypredsdf = treatypreds.describe().T
treatypredsdf['std'] = treatypreds.mad()
treatypredsdf.drop(columns=['count','25%','75%'], inplace=True)
treatypredsdf.rename(columns={'50%':'Median'}, inplace=True)
treatypredsdf.columns = treatypredsdf.columns.str.title() + 'Pred'
treatypredsdf.rename(columns={'StdPred':'PredMAD'}, inplace=True)
treatypredsdf.insert(0, 'treatyLabel', df.treatyLabel)
treatypredsdf.insert(1, 'incubConsent', y)
treatypredsdf['PredRange'] = treatypredsdf.MaxPred - treatypredsdf.MinPred.abs()
treatypredsdf['MeanPredErr'] = treatypredsdf.MeanPred - y
treatypredsdf['MAE'] = (treatypreds - y).abs().mean()
```

```
treatypredsdf['ErrMAD'] = abs(((treatypreds - y).abs() - treatypredsdf.MAE)).mean()
treatypredsdf['MinAbsErr'] = (treatypreds - y).abs().min()
treatypredsdf['MaxAbsErr'] = (treatypreds - y).abs().max()
treatypredsdf['ErrRange'] = (treatypreds - y).max() - (treatypreds - y).min()
treatypredsdf['AbsErrRange'] = treatypredsdf.MaxAbsErr - treatypredsdf.MinAbsErr
```

H.3   Feature permutation

Feature permutation for MLPs

```
subdf = resdf[(resdf.regressor=='MLP')&(resdf.inner_k==-1)]
cols = X.columns[~X.columns.str.contains('(?:treatyAdoption|orgAuspices)', regex=True)]
binvars = [c for c in cols if dfX[c].nunique() == 2]
numvars = [c for c in cols if dfX[c].nunique() > 2]
# remove imbalanced features
floor = len(dfX) * 0.2
ceiling = len(dfX) * 0.8
binvars = [c for c in binvars if dfX[c].sum()>floor and dfX[c].sum()<ceiling]
testmlps = pd.DataFrame(np.nan, columns=['treatyInd','treatyLabel','incubConsent',
                                         'model_dt','model_id','origPred','origRank',
                                         'permutCol','permutVal','permutPred','predDiff'],
                        index=[0]).set_index(['treatyInd','treatyLabel','incubConsent',
                                              'model_dt','model_id','origPred','origRank',
                                              'permutCol','permutVal'])
for testi in subdf.index:
  model = initializer = lr_schedule = stop_early = history = None
  rs = int(resdf.rs[testi])
  outer_k = resdf.outer_k[testi]
  keras.utils.set_random_seed(rs)
  # reproduce datasets
  test_ind = resdf.testIndex[testi]
  X_train, X_test = X.drop(index=test_ind), X.loc[test_ind,:]
  y_train, y_test = y.drop(index=test_ind), y[test_ind]
  numcols = ['treatyAdoptionYear','EIFnConsent','EIFdelay','denunNotifMinYrs','denunEffectYrs']
  X_train_means = X_train[numcols].mean()
  X_train_stds = X_train[numcols].std()
  for c in numcols:
    X_train[c] = (X_train[c] - X_train_means[c]) / X_train_stds[c]
    X_test[c] = (X_test[c] - X_train_means[c]) / X_train_stds[c]
  # reproduce model
  bestHps = resdf.bestHps[testi][0]
  initializer = keras.initializers.HeUniform(seed=rs)
  model = keras.Sequential()
  model.add(keras.Input(shape=(len(X_train.columns),)))
  for hl in range(bestHps['layers']):
    model.add(layers.Dense(units=bestHps['units'],activation='relu',kernel_initializer=initializer))
  model.add(layers.Dense(1))
  lr_schedule = keras.optimizers.schedules.ExponentialDecay(bestHps['init_lr'], 20, 0.90,
                                                            staircase=True)
  model.compile(optimizer=keras.optimizers.Adam(learning_rate=lr_schedule),
                loss='mean_absolute_error')
  history = model.fit(X_train, y_train, batch_size=32, epochs=bestHps['best_epoch']+1,
                      verbose=0, validation_data=(X_test, y_test))
```

```python
  hist = pd.DataFrame(history.history)
  testpreds = model.predict(X_test, verbose=0)
  if (testpreds.round(5).T[0] != resdf.testPreds[testi]).any():
    print(f'Predictions of model {testi} not reproduced!')
  permdf = pd.DataFrame(np.nan, columns=binvars, index=X_test.index)
  X_test_perm = X_test.copy()
  for c in binvars:
    X_test_perm.loc[X_test_perm[c] == 0., c] = 1.
    permdf.loc[:,c] = model.predict(X_test_perm, verbose=0)
    X_test_perm = X_test.copy()
  for c in binvars:
    permdf.loc[X_test[c] == 1., c] = np.nan
  permdf['treatyInd'] = test_ind
  permdf['treatyLabel'] = df.treatyLabel[test_ind]
  permdf['incubConsent'] = y_test
  permdf['origPred'] = testpreds
  permdf['origRank'] = 0
  permdf = pd.melt(permdf,id_vars=['treatyInd','treatyLabel','incubConsent','origPred','origRank'],
                   var_name='permutCol',value_name='permutPred').dropna()
  permdf['model_dt'] = testi
  permdf['model_id'] = f'model_{rs}.{outer_k}'
  permdf['permutVal'] = 1
  permdf = permdf.set_index(['treatyInd','treatyLabel','incubConsent','model_dt','model_id',
                             'origPred','origRank','permutCol','permutVal'])
  testmlps = pd.concat([testmlps,permdf], axis=0, join='outer')
testmlps = testmlps.reset_index(drop=False)
testmlps['predDiff'] = testmlps.permutPred - testmlps.origPred
testmlps['origAbsErr'] = abs(testmlps.origPred - testmlps.incubConsent)
for t in testmlps.treatyLabel.unique():
  ranking = sorted(testmlps.loc[testmlps.treatyLabel==t, 'origAbsErr'].unique())
  testmlps.loc[testmlps.treatyLabel==t, 'origRank'] = \
    testmlps.origAbsErr[testmlps.treatyLabel==t].apply(lambda x: ranking.index(x))
testmlps.dropna(how='all', inplace=True)
for col in ['treatyInd','origRank']:
  testmlps[col] = testmlps[col].astype(int)
testmlps.to_csv(wd+'testmlps.csv', encoding='utf-8', index=False)
print(testmlps.shape)
```

Feature permutation for Ridge regressors

```python
subdf = resdf[(resdf.regressor=='Ridge')&(resdf.inner_k==-1)]
cols = X.columns[~X.columns.str.contains('(?:treatyAdoption|orgAuspices)', regex=True)]
binvars = [c for c in cols if dfX[c].nunique() == 2]
numvars = [c for c in cols if dfX[c].nunique() > 2]
# remove imbalanced features
floor = len(dfX) * 0.2
ceiling = len(dfX) * 0.8
binvars = [c for c in binvars if dfX[c].sum()>floor and dfX[c].sum()<ceiling]
ridgemodels = pd.DataFrame(np.nan, columns=['treatyInd','treatyLabel','incubConsent',
                                            'model_dt','model_id','origPred','origRank',
                                            'permutCol','permutVal','permutPred','predDiff'],
                           index=[0]).set_index(['treatyInd','treatyLabel','incubConsent',
                                                  'model_dt','model_id','origPred','origRank',
                                                  'permutCol','permutVal'])

for testi in subdf.index:
  model = None
```

```python
  rs = int(resdf.rs[testi])
  outer_k = resdf.outer_k[testi]
  keras.utils.set_random_seed(rs)
  # reproduce datasets
  test_ind = resdf.testIndex[testi]
  X_train, X_test = X.drop(index=test_ind), X.loc[test_ind,:]
  y_train, y_test = y.drop(index=test_ind), y[test_ind]
  numcols = ['treatyAdoptionYear','EIFnConsent','EIFdelay','denunNotifMinYrs','denunEffectYrs']
  X_train_means = X_train[numcols].mean()
  X_train_stds = X_train[numcols].std()
  for c in numcols:
    X_train[c] = (X_train[c] - X_train_means[c]) / X_train_stds[c]
    X_test[c] = (X_test[c] - X_train_means[c]) / X_train_stds[c]
  # reproduce model
  bestHp = resdf.bestHps[testi]
  model = Ridge(alpha=bestHp, fit_intercept=True, tol=1e-3, solver='auto').fit(X_train, y_train)
  testpreds = model.predict(X_test)
  if (testpreds.round(4) != resdf.testPreds[testi]).any():
    print(f'Predictions of model {testi} not reproduced!')
  permdf = pd.DataFrame(np.nan, columns=binvars, index=X_test.index)
  X_test_perm = X_test.copy()
  for c in binvars:
    X_test_perm.loc[X_test_perm[c] == 0., c] = 1.
    permdf.loc[:,c] = model.predict(X_test_perm)
    X_test_perm = X_test.copy()
  for c in binvars:
    permdf.loc[X_test[c] == 1., c] = np.nan
  permdf['treatyInd'] = test_ind
  permdf['treatyLabel'] = df.treatyLabel[test_ind]
  permdf['incubConsent'] = y_test
  permdf['origPred'] = testpreds
  permdf['origRank'] = 0
  permdf = pd.melt(permdf, id_vars=['treatyInd','treatyLabel','incubConsent','origPred','origRank'],
                   var_name='permutCol', value_name='permutPred').dropna()
  permdf['model_dt'] = testi
  permdf['model_id'] = f'model_{rs}.{outer_k}'
  permdf['permutVal'] = 1
  permdf = permdf.set_index(['treatyInd','treatyLabel','incubConsent','model_dt','model_id',
                             'origPred','origRank','permutCol','permutVal'])
  ridgemodels = pd.concat([ridgemodels,permdf], axis=0, join='outer')
ridgemodels = ridgemodels.reset_index(drop=False)
ridgemodels['predDiff'] = ridgemodels.permutPred - ridgemodels.origPred
ridgemodels['origAbsErr'] = abs(ridgemodels.origPred - ridgemodels.incubConsent)
for t in ridgemodels.treatyLabel.unique():
  ranking = sorted(ridgemodels.loc[ridgemodels.treatyLabel==t, 'origAbsErr'].unique())
  ridgemodels.loc[ridgemodels.treatyLabel==t, 'origRank'] = \
    ridgemodels.origAbsErr[ridgemodels.treatyLabel==t].apply(lambda x: ranking.index(x))
ridgemodels.dropna(how='all', inplace=True)
for col in ['treatyInd','origRank']:
  ridgemodels[col] = ridgemodels[col].astype(int)
ridgemodels.to_csv(wd+'RidgeTestModels.csv', encoding='utf-8', index=False)
print(ridgemodels.shape)
```

```
(17570, 12)
```

Bibliography

Alschner W, Pauwelyn J, and Puig S, 'The Data-Driven Future of International Economic Law' (2017) 20(2) Journal of International Economic Law 217.

Alschner W, Seiermann J, and Skougarevskiy D, 'Text of Trade Agreements (ToTA)—A Structured Corpus for the Text-as-Data Analysis of Preferential Trade Agreements' (2018) 15(3) Journal of Empirical Legal Studies 648.

Alschner W and Skougarevskiy D, 'Mapping the Universe of International Investment Agreements' (2016) 19(3) Journal of International Economic Law 561.

Arbeitsgruppe EDV und Recht FUB, 'Juradat kämpft gegen die Informationslawine' (1970) 3(4) Kritische Justiz 463.

Arp R, Smith B, and Spear AD, *Building Ontologies with Basic Formal Ontology* (MIT Press August 2015).

Ashburner M and others, 'Gene Ontology: Tool for the Unification of Biology' (2000) 25(1) Nature Genetics 25.

Aust A, *Modern Treaty Law and Practice* (3rd ed., Cambridge University Press 2013).

Barrett S, *Environment and Statecraft: The Strategy of Environmental Treaty-Making* (Oxford University Press 2005).
— 'An Economic Theory of International Environmental Law', in D Bodansky, J Brunnée, and E Hey (eds), *The Oxford Handbook of International Environmental Law* (Oxford University Press 2007).

AA Batabyal (ed), *The Economics of International Environmental Agreements* (Ashgate 1999).

Bench-Capon T and others, 'A History of AI and Law in 50 Papers: 25 years of the International Conference on AI and Law' (2012) 20 Artificial Intelligence and Law 215.

Benjamini Y and Hochberg Y, 'Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing' (1995) 57(1) Journal of the Royal Statistical Society 289.

Berners-Lee T and Hendler J, 'Publishing on the Semantic Web' (2001) 410(6832) Nature 1023.

Berners-Lee T, Hendler J, Lassila O, and others, 'The Semantic Web' (2001) 284(5) Scientific American 28.

U Beyerlin (ed), *Ensuring Compliance with Multilateral Environmental Agreements: A Dialogue between Practitioners and Academia* (Studies on the Law of Treaties vol. 2, Martinus Nijhoff 2006).

Bird S, Klein E, and Loper E, *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit* (O'Reilly 2009).

Blei DM and Smyth P, 'Science and Data Science' (2017) 114(33) Proceedings of the National Academy of Sciences 8689.

Bontcheva K and others, 'Evolving GATE to Meet New Challenges in Language Engineering' (2004) 10(3/4) Natural Language Engineering 349.

Bowles MD, 'The Information Wars: Two Cultures and the Conflict in Information Retrieval, 1945-1999' (1999).

Breitmeier H, Young OR, and Zürn M, *Analyzing International Environmental Regimes: from Case Study to Database* (MIT Press 2006).

Brunnée J, 'COPing with Consent: Law-Making Under Multilateral Environmental Agreements' (2002) 15(01) Leiden Journal of International Law 1.

Carbon S and Mungall C *Gene Ontology Data Archive* version 2022-10-07 (Zenodo July 2018.

Ceballos G and others, 'Accelerated Modern Human-induced Species Losses: Entering the Sixth Mass Extinction' (2015) 1(5) Science Advances e1400253.

Chayes A and Chayes AH, *The New Sovereignty: Compliance with International Regulatory Agreements* (Harvard University Press 1995).

Choi S.-S, Cha S.-H, Tappert CC, and others, 'A Survey of Binary Similarity and Distance Measures' (2010) 8(1) Journal of Systemics, Cybernetics and Informatics 43.

Consortium TGO, 'The Gene Ontology Resource: 20 years and still GOing strong' (2018) 47(D1) Nucleic Acids Research D330.

consortium wwPDB, 'Protein Data Bank: the Single Global Archive for 3D Macromolecular Structure Data' (2018) 47(D1) Nucleic Acids Research D520.

Cordonier Segger M.-C and Khalfan A, *Sustainable Development Law: Principles, Practices, and Prospects* (Oxford University Press 2004).

Cunningham H and others, 'Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics' (2013) 9(2) PLOS Computational Biology 1.

Deeks A, 'High-Tech International Law' (2020) 88 George Washington Law Review 574.

Fellicious C, Weissgerber T, and Granitzer M, 'Effects of Random Seeds on the Accuracy of Convolutional Neural Networks' in *International Conference on Machine Learning, Optimization, and Data Science* (Lecture Notes in Computer Science, vol 12566, Springer 2020).

Finus M, *Game Theory and International Environmental Cooperation* (Edward Elgar 2001).

Fobbe S, 'Introducing Twin Corpora of Decisions for the International Court of Justice (ICJ) and the Permanent Court of International Justice (PCIJ)' (2022) 19(2) Journal of Empirical Legal Studies 491.

Franke TM, Ho T, and Christie CA, 'The Chi-Square Test: Often Used and More Often Misinterpreted' (2012) 33(3) American Journal of Evaluation 448.

Geslin A, *La mise en application provisoire des traités* (Edition A Pedone 2005).

Goodfellow I, Bengio Y, and Courville A, *Deep Learning* (http://www.deeplearningbook.org, MIT Press 2016).

Grimmer J, Roberts ME, and Stewart BM, *Text as Data: A new Framework for Machine Learning and the Social Sciences* (Princeton University Press 2022).

Grimmer J and Stewart BM, 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts' (2013) 21(3) Political Analysis 267.

Haas PM, Keohane RO, and Levy MA, *Institutions for the Earth: Sources of Effective International Environmental Protection* (MIT Press 1993).

Hafner G, 'Provisional Application of Treaties' (2021) 24(1) Austrian Review of International and European Law Online 67.

He K and others, 'Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification' in *International Conference on Computer Vision (ICCV)* (IEEE 2015).

Hearst MA, 'Untangling text data mining' (1999).

Helm C, *Economic Theories of International Environmental Cooperation* (Edward Elgar 2000).

Hirschberg J and Manning CD, 'Advances in Natural Language Processing' (2015) 349(6245) Science 261.

P Hitzler and others (eds), OWL 2 Web Ontology Language: Primer (Second Edition) (W3C Recommendation, 2012) ⟨https://www.w3.org/TR/2012/REC-owl2-primer-20121211/⟩ accessed 20 October 2016.

Hoffmann M and others, 'The Impact of Conservation on the Status of the World's Vertebrates' en (2010) 330(6010) Science 1503.

Impériali C and Kiss AC, *L'effectivité du droit international de l'environnement: contrôle de la mise en oeuvre des conventions internationales* (Economica 1998).

Jenks CW, 'Some Constitutional Problems of International Organizations' (1945) 22 British Yearbook of International Law 11.

Jordan MI and Mitchell TM, 'Machine learning: Trends, perspectives, and prospects' (2015) 349(6245) Science 255.

Jumper J and others, 'Highly Accurate Protein Structure Prediction with AlphaFold' (2021) 596(7873) Nature 583.

Keeley JF, 'Coding Treaties: An Example from Nuclear Cooperation' (1985) 29(1) International Studies Quarterly 103.

Klabbers J, *The Concept of Treaty in International Law* (Kluwer Law International 1996).
— *An Introduction to International Institutional Law* (Cambridge University Press 2002).
— 'The Concept of Legal Personality' (2005) 11 Ius Gentium 35.

Koskenniemi M, *From Apology to Utopia: the Structure of International Legal Argument* (Cambridge University Press 2005).

— 'The Fate of Public International Law: Between Technique and Politics' (2007) 70(1) Modern Law Review 1.

Kotzé LJ, 'Earth System Law for the Anthropocene' (2019) 11(23) Sustainability.

Lazer D and others, 'Computational Social Science' en (2009) 323(5915) Science 721.

Lazer DMJ and others, 'Computational Social Science: Obstacles and Opportunities' (2020) 369(6507) Science 1060.

LeCun Y, Bengio Y, and Hinton G, 'Deep Learning' (2015) 521(7553) Nature 436.

Leith P, 'The Rise and Fall of the Legal Expert System' (2016) 30(3) International Review of Law, Computers and Technology 94.

Libbrecht MW and Noble WS, 'Machine Learning Applications in Genetics and Genomics' (2015) 16(6) Nature Reviews Genetics 321.

Maljean-Dubois S and Rajamani L, *La mise en oeuvre du droit international de l'environnement* (Martinus Nijhoff 2011).

Marx V, 'The Big Challenges of Big Data' (2013) 498(7453) Nature 255.

Mason R and Swanson T, 'A Kuznets Curve Analysis of Ozone-Depleting Substances and the Impact of the Montreal Protocol' English (2003) 55(1) Oxford Economic Papers 1.

McDonald J, *Handbook of Biological Statistics* (3rd edition, Sparky House Publishing 2014).

Mertsch AQ, *Provisionally Applied Treaties: Their Binding Force and Legal Nature* (Brill 2012).

Miles EL, *Environmental Regime Effectiveness: Confronting Theory with Evidence* (MIT Press 2002).

Mitchell RB, 'International Environmental Agreements: A Survey of Their Features, Formation, and Effects' (2003) 28(1) Annual Review of Environment and Resources 429.

— *International Environmental Agreements Database Project* version 2020.1 ⟨http://iea.uoregon.edu/⟩.

Mitchell RB and others, 'What We Know (and Could Know) About International Environmental Agreements' (2020) 20(1) Global Environmental Politics 103.

Moore PM, 'The Decisions Bridging the GATT 1947 and the WTO Agreement' (1996) 90(2) American Journal of International Law 317.

Müller W, *WTO Agreement on Subsidies and Countervailing Measures: A Commentary* (Cambridge University Press 2017).

Murdoch JC and Sandler T, 'The Voluntary Provision of a Pure Public Good: The Case of Reduced CFC Emissions and the Montreal Protocol' English (1997) 63(03) Journal of Public Economics 331.

Musen MA, 'The Protégé Project: A Look Back and a Look Forward' (2015) 1(4) AI Matters 4.

Palmirani M and Vitali F, 'Akoma-Ntoso for Legal Documents' in G Sartor and others (eds), *Legislative XML for the Semantic Web: Principles, Models, Standards for Document Management* (Springer 2011).

Pedregosa F and others, 'Scikit-learn: Machine Learning in Python' (2011) 12 Journal of Machine Learning Research 2825.

Peroni S, Palmirani M, and Vitali F, 'UNDO: The United Nations System Document Ontology' (d'Amato C and others eds, Springer 2017).

Popper KR, *The Logic of Scientific Discovery* (Basic Books 1959).

Rogoff MA and Gauditz BE, 'The Provisional Application of International Agreements' (1987) 39 Maine Law Review 29.

Russell SJ and Norvig P, *Artificial Intelligence : A Modern Approach* (Pearson Education 2022).

PH Sand (ed), *The Effectiveness of International Environmental Agreements: A Survey of Existing Legal Instruments* (Grotius Publications 1992).

Sands P, *Principles of International Environmental Law* (Second Edition, Cambridge University Press 2003).

Schmidt C, *Designing International Environmental Agreements: Incentive Compatible Strategies for Cost-Effective Cooperation* (New horizons in environmental economics, Edward Elgar 2000).

Schweighofer E, *Legal Knowledge Representation: Automatic Text Analysis in Public International and European Law* (Kluwer Law International 1999).

Schweighofer E and Winiwarter W, 'Legal Expert System KONTERM - Automatic Representation of Document Structure and Contents' (DDEXA '93, Springer 1993).

Seabold S and Perktold J, 'Statsmodels: Econometric and Statistical Modeling with Python' (2010).

Simitis S, 'Die Informationskrise des internationalen Rechts und die Datenverarbeitung' (1969) 9 Zeitschrift für Rechtsvergleichung 276.

Smith B and others, 'The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration' en (2007) 25(11) Nature Biotechnology 1251.

Stähler F, *On the Economics of International Environmental Agreements* (Springer 1998).

Swanson TM and Johnston S, *Global Environmental Problems and International Environmental Agreements: The Economics of International Institution Building* (Edward Elgar: Published in association with UNCTAD 1999).

Trask A and others, 'Neural Arithmetic Logic Units' (Bengio S and others eds, NIPS, 2018) vol 31.

T Treves, A Tanzi, and L Pineschi (eds), *Non-Compliance Procedures and Mechanisms and the Effectiveness of International Environmental Agreements* (Asser Press 2009).

Tukey JW, 'The Future of Data Analysis' (1962) 33(1) The Annals of Mathematical Statistics 1.

Tversky A and Kahneman D, 'Judgment under Uncertainty: Heuristics and Biases' (1974) 185(4157) Science 1124.

Victor DG, Raustiala K, and Skolnikoff EB, *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice* (International Institute for Applied Systems Analysis; MIT Press 1998).

Vignes D.-H, 'Une notion ambiguë: l'application à titre provisoire des traités' (1972) 18(1) Annuaire Français de Droit International 181.

Viñuales JE, 'On Legal Inquiry' in D Alland and others (eds), *Unity and Diversity of International Law: Essays in Honour of Professor Pierre-Marie Dupuy* (Nijhoff 2014).

Wagner UJ, 'The Voluntary Provision of a Pure Public Good? Another look at CFC Emissions and the Montreal Protocol' (2009) 61(1) Oxford Economic Papers 183.

Weiss EB, 'Understanding Compliance with International Environmental Agreements: The Baker's Dozen Myths' (1999) 32 University of Richmond Law Review 1555.

EB Weiss and HK Jacobson (eds), *Engaging Countries: Strengthening Compliance with International Environmental Accords* (MIT Press 1998).

Wettestad J, *Designing Effective Environmental Regimes: the Key Conditions* (Edward Elgar 1999).

Wilkinson MD and others, 'The FAIR Guiding Principles for Scientific Data Management and Stewardship' (2016) 3(1) Scientific Data 1.

Wilks Y, 'Computational Linguistics: History' in K Brown (ed), *Encyclopedia of Language & Linguistics* (2nd Edition, Elsevier 2006).

Wolfrum R, 'Means of Ensuring Compliance with and Enforcement of International Environmental Law' English [1998] Recueil des cours de l'Académie de droit international de La Haye = Collected courses of the Hague Academy of International Law 9.

Young OR, *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms* (MIT Press 1999).