

Global Tracking of Climate Change Adaptation Policy Using Machine Learning: a Systematic Map Protocol

Anne J. Sietsma (✉ eeajs@leeds.ac.uk)

Priestley International Centre for Climate, University of Leeds, Leeds, United Kingdom

<https://orcid.org/0000-0003-0239-152X>

Max Callaghan

Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany

<https://orcid.org/0000-0001-8292-8758>

Robbert Biesbroek

Public Administration and Policy Group, Wageningen University & Research, Wageningen, the Netherlands <https://orcid.org/0000-0002-2906-1419>

Emily Theokritoff

Climate Analytics, Berlin, Germany; Integrative Research Institute on Transformations of Human-Environment Systems, Humboldt University, Berlin, Germany <https://orcid.org/0000-0003-0632-9862>

Adelle Thomas

Climate Analytics, Berlin, Germany; Climate Change Adaptation and Resilience Research Centre, University of The Bahamas, Nassau, Bahamas <https://orcid.org/0000-0002-0407-2891>

Iván Villaverde Canosa

: Priestley International Centre for Climate, University of Leeds, Leeds, United Kingdom

<https://orcid.org/0000-0002-9344-6452>

Carl-Friedrich Schleussner

Climate Analytics, Berlin, Germany; Integrative Research Institute on Transformations of Human-Environment Systems, Humboldt University, Berlin, Germany <https://orcid.org/0000-0001-8471-848X>

James D. Ford

Priestley International Centre for Climate, University of Leeds, Leeds, United Kingdom

<https://orcid.org/0000-0002-2066-3456>

Jan C. Minx

Priestley International Centre for Climate, University of Leeds, Leeds, United Kingdom; Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany <https://orcid.org/0000-0002-2862-0178>

Keywords: climate change, adaptation, policy tools, NLP, machine learning, evidence synthesis, evidence map

Posted Date: February 25th, 2022

DOI: <https://doi.org/10.21203/rs.3.pex-1836/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background – Countries around the globe have started implementing policies to respond to the current and future risks of climate change. The scientific literature on these adaptation policies is fragmented and no central typology is generally accepted, making tracking of global adaptation policy progress difficult.

Methods – In this protocol, we describe how we use machine learning methods to classify scientific literature on adaptation policies following the ROSES guidelines. We use a broad search query in Scopus, MEDLINE and Web of Science (up to November 2021). We manually classify a subset of the documents and use this to train multiple supervised machine learning algorithms, including a state-of-the-art algorithm based on BERT. The classification scheme is aimed at providing a multi-functional database: we classify first based on a newly created typology, which is based around the well-established NATO categories of policy instruments; this is supplemented with categories on the types of impacts, evidence on maladaptation, constraints, evidence type, governance level and geographic location.

Expected results – Using the typology and categories, as well as topic modelling, we create an overview of scientific literature on adaptation policies. This describes the breath of policy options, their geographic distribution, developments over time, and under-explored areas. If successful, this would result in the most comprehensive evidence map of adaptation policies to date; building on this, the machine learning algorithms and underlying data can serve as a basis for a living evidence map, moving towards the real-time tracking of adaptation progress.

Introduction

NB: our protocol is more extensive than required by Protocol Exchange and not all parts fit in the Protocol Exchange's pre-determined categories. In addition, tables and figures cannot be inserted in-text. We therefore highly recommend that you read the formatted document instead, which can be found under the Supplementary Files section.

1. Introduction

With the impacts of climate change becoming increasingly clear and pressing (IPCC, 2018, IPCC, 2021), adaptation to these impacts has become an important focus of climate policy, especially since the Paris Agreement (Lesnikowski et al., 2017). In the Agreement, countries commit to additional reporting, including a Global Stocktake on Adaptation, with the goal of assessing global progress and the adequacy and effectiveness of adaptation measures. For mitigation, such efforts are perhaps best summed up by the assessment of the “emissions gap” (United Nations Environment Programme, 2021b) which provides clear guidance on the progress of countries in reducing emissions and the size of the remaining

challenge. Quantifying the analogous “adaptation gap” (United Nations Environment Programme, 2021a) however has proven to be more difficult.

These difficulties are in part caused by fundamental disagreements on the meaning of adaptation and whether it can be meaningfully measured (Arnott et al., 2016, Runhaar et al., 2018, Craft and Fisher, 2018). More practically, adaptation tracking systems are not present in the majority of countries, even if they have a national adaptation plan (United Nations Environment Programme, 2021a, Leiter, 2021) and national tracking systems are often not designed for to support global comparisons (Magnan and Chalastani, 2019). This despite Global Stocktake’s reliance on country reports (Christiansen et al., 2020, Craft and Fisher, 2018). Independent efforts to create an overview of adaptation progress face a complex task as policies range from the international to the municipal level (Berrang-Ford et al. 2019; Olazabal et al. 2019); Persson and Dzebo (2019) even state that global or transnational adaptation governance is so ill-defined that it “can become an impossibly broad phenomenon to study” (p. 358). Adding to the challenge is the lack of a central database to record policies – though some efforts do exist (Nachmany et al., 2019, Fussel and Almond, 2021). Systematic assessments of the scientific literature meanwhile provided some insight in the past (Berrang-Ford et al., 2011), but since then, the sheer volume and heterogeneity of adaptation-relevant research has increased markedly, complicating synthesis efforts (Sietsma et al., 2021).

These challenges are not unique to the adaptation community. Spreading out from computer science, the term “Big Data” has been widely adopted, taking on various meanings along the way (Kitchin and McArdle 2016), all centred around the use of computers to extract meaningful insights from data with a high volume, velocity, and variety (Laney 2001). In recent years, such approaches have also made their way into climate change research (Rolnick et al., 2019). Some climate change researchers have used machine learning approaches to build and analyse large text-based datasets (Hsu and Rauber, 2021, Lesnikowski et al., 2019a, Biesbroek et al., 2020, Callaghan et al., 2020, Lamb et al., 2019). Relatedly, the phenomenon of “Big Literature” (Nunez-Mir et al., 2016), a rapid increase in the number of scientific publications, has led some within the evidence synthesis community to integrate machine learning methods into a more traditional systematic review process (Haddaway et al., 2020, Nakagawa et al., 2019, Van de Schoot et al., 2021). This includes the use of machine learning to create “evidence maps”; for example, to identify trends and research gaps on health impacts of climate change (Berrang-Ford et al., 2021c), to assess progress within climate change adaptation research (Sietsma et al., 2021) or to categorise and estimate the size of climate change impacts worldwide (Callaghan et al., 2021).

Despite the promise of machine learning methods, they are rarely applied in an adaptation policy context (notable exceptions: Biesbroek et al., 2022, Biesbroek et al., 2020, Lesnikowski et al., 2019a, Ulibarri et al., 2022). Moreover, the discourse around the Global Stocktake is mostly devoid of any substantive discussion on computer-based methods – even the Adaptation Committee (2020) limits their discussion of Big Data to one text box (p. 28), which is largely filled with explanations of definitions and refers to a few data initiatives outside of climate change. This does not do justice to the proven track record of

machine learning methods in other closely related contexts and applications – in fact, it is *less* substantive than earlier calls for Big Data applications for adaptation (Ford et al., 2016).

The lack of progress may simply be a reflection of the lack of overlap between the people and expertise in both research communities (Lesnikowski et al., 2019a, Haddaway et al., 2020); generally, adaptation researchers appear to favour in-depth qualitative approaches (Ford and Berrang-Ford, 2016). However, given the rapidly increasing pace of adaptation publications (Wang et al., 2018, Haunschild et al., 2016), adaptation researchers may ultimately be forced to use machine learning tools to synthesise research (Sietsma et al., 2021).

In the research outlined in this protocol, we adapt methods from other climate-relevant evidence maps (chiefly, Callaghan et al., 2021, Sietsma et al., 2021) to address the challenge of adaptation tracking. We do so by: a) training a machine learning algorithm to recognise when scientific papers are discussing adaptation policy; b) using machine learning to categorise these papers to connect them to both established policy research and to active debates among adaptation researchers; and c) using further Natural Language Processing (NLP) and evidence mapping techniques to visualise developments over time and geographic distribution of the evidence.

Following best practice for evidence mapping in the environmental sciences, we design our protocol around the ROSES standard (Haddaway et al., 2018). We deviate from this standard in three key instances to better fit the machine learning approach of this project: first, we will not retrieve full texts of our documents, relying instead on title and abstracts, as frequently done in machine-learning assisted systematic maps, because they condense key information and avoid excessive type-II errors (Marshall and Wallace, 2019, Callaghan et al., 2021); secondly, not all documents will be screened by hand, so we will describe instead how the typology was operationalised for human coders to label a subset of the documents (the “training set”), from which a supervised machine learning algorithm will then learn to categorise the full dataset (Marshall and Wallace, 2019, Callaghan et al., 2021, Berrang-Ford et al., 2021a); and third, as machine learning allows for a more quantitative estimate of both comprehensiveness and errors, we will discuss how these relate to the size of the training set and the interpretation of results.

2. Stakeholder Engagement

There are three main phases of stakeholder engagement within this project: writing this protocol, gathering feedback on our methods, and disseminating the results. We rely mostly on the breadth of experience of our team, but will aim to engage stakeholders from both the adaptation community *and* the machine learning/evidence synthesis community.

The project team includes researchers from a wide array of backgrounds. This includes university-based researchers on adaptation (JDF, IVC), public policy (RB, JCM), and machine learning methods for climate

change research (MC, JCM, AJS). Further expertise comes from the researchers based at Climate Analytics (CS, ET, AT), which, as an institute, specialises in bringing cutting-edge climate science into the policy debate.

All team members contributed to writing this protocol. The protocol will be publicly accessible and open for feedback.

Once complete, the research will again be made public, including the data and code. The hope is that both the dataset and the trained algorithms can provide a steppingstone for further research into climate change adaptation policy. To this end, we hope to present the results at scientific conferences, though the current pandemic makes that we cannot make our plans as concrete as we would wish. The quality of the final dataset will obviously also influence uptake by other stakeholders; if the data has sufficient detail, it may be worthwhile to create an easily accessible platform, which would require a new round of stakeholder engagement to determine their usability needs.

3 Objective

The overall objective is to systematically map evidence on climate change adaptation policies in the global academic literature, with a focus on the geographic distribution of different policy types and their development over time.

3.1 Research questions

We aim to create an evidence map of the scientific literature on climate change adaptation policy with global coverage. To link this to established policy research, our analysis will centre around the NATO typology – Nodality, Authority, Treasure and Organisation (Hood, 1986); we will also address issues of contention within the scientific literature on adaptation, specifically:

- Developments over time – is there an increase in adaptation policy?
- Geographic locations – how are adaptation policies distributed worldwide?
- Alignment to policy culture – do countries use similar policy tools for their adaptation policy as they are for their non-adaptation policies?
- Alignment to impacts – do the adaptation policies respond to the predicted or experienced impacts of climate change within a region or area?
- Constraints and limits – do we see increased attention to the constraints and limits of climate change adaptation, and if so, what types of constraints are most prevalent?

- Maladaptation – some adaptation actions may have unintended negative consequences, which can end up increasing vulnerability to climate change; how prevalent is this issue?

3.2 Problem scope

In general, for this project, we aim to be as comprehensive as possible: we wish to capture the full breadth of scientific research that has elements of both climate change adaptation and some form of policy- or policy-enabled response. The use of machine learning is crucial to taking such a broad view.

At the same time, we must take care not to cast too wide a net because even with the help of machine learning, we cannot meaningfully analyse too large a dataset. For this reason, we will limit our analysis to those papers which explicitly mention both elements – i.e. they must include a substantive analysis of both adaptation *and* policy. More details on the scope of the evidence considered is given in Table 1, along the Population, Intervention, Context, Timeframe (PICoST) framework for systematic reviews.

3.2.1 Adaptation

Adaptation here is defined in line with the IPCC to include all actions which aim to adjust to the actual or expected climate and its effects. We limit ourselves to adaptation in human systems, where adaptations seek to moderate or avoid harm or exploit beneficial opportunities caused by climate change. Crucially, this includes literature where neither the words “climate change” nor “adaptation” are used by the authors. Climate change has many different impacts, but the link of climate change may not always be known to the authors, or it may be omitted for political expediency– e.g. literature on improved irrigation may write about prolonged droughts without using the words climate change. To determine whether an impact is attributable to climate change, we will follow the latest report by the Intergovernmental Panel on Climate Change (IPCC, 2021).

3.2.2 Policy

For the purposes of this review, we refer to (public) policies as the outputs public actors that are designed to achieve defined goals and solutions to societal problems (Knill and Tosun, 2020). These outputs can take different forms, including programs, measures, decisions, legislation, strategies and other courses of action. Although there are different ways to study public policy, we focus on the underlying governance principles and instruments used by governments to achieve their goals.

Several typologies exist to classify the types of tools used by governments, see Capano and Howlett (2020). Here we adopt the frequently used NATO-typology: Nodality, Authority, Treasure, and Organization (Hood, 1986). This typology has also been used in an adaptation context, for example by Henstra (2017)

in a Canadian case study, by Lesnikowski et al. (2019b) to analyse the adaptation policy instrument mixes of 125 local governments, and by Biesbroek and Delaney (2020) to systematically map 183 adaptation policy studies in Europe.

Reagents

Equipment

Procedure

NB: as noted above, our protocol is more extensive than required by Protocol Exchange and not all parts fit in the Protocol Exchange's pre-determined categories. In addition, tables and figures cannot be inserted in-text. We therefore highly recommend that you read the formatted document instead, which can be found under the Supplementary Files section. Protocol exchange requests a list here, which we will therefore give first:

- 1) retrieve documents from scientific databases;
- 2) manually label a subset of these documents;
- 3) use the labelled documents to train a binary classifier which selects relevant literature;
- 4) assign categories to the documents, using a supervised multi-label classifier, topic modelling, and a geoparser;
- 5) combine the different results of the previous step to synthesise the data

4 Methods

4.1 A machine learning approach to evidence mapping

As described by Haddaway et al. (2020), a systematic map or evidence map aims to summarise an entire evidence base. Where a systematic review aims to combine the findings of individual studies (“what works where and how?”), an evidence map by contrast aims to describe what is known in an area (“what kinds of research exist?”). Both systematic reviews and maps share a focus on transparent and robust methods; they lay out the rationale for their main research questions in a protocol, which also describes the main methods, as we do here. The basis for the map is an often extensive search query; matching documents are then screened by hand, discarding irrelevant documents and dividing relevant documents into categories. Combined with meta-data, this allows researchers to describe developments in the field.

Results often take the form of a searchable or interactive database, as well as visualisations of the data, a list of knowledge gaps and clusters, and a report noting key findings.

Systematic maps in particular may benefit from incorporating computational methods, given their relatively descriptive nature and considerable evidence base (Haddaway et al., 2020). In this project, we use the ability of computers to handle large amounts of data to create an evidence map that is as inclusive as possible. This means that the documents cannot reasonably be screened by hand, which is why we make use of so-called “supervised machine learning” to both select and classify documents. Although we cannot provide a full introduction to machine learning here (we will instead refer the interested reader to the introduction provided by Marshall and Wallace, 2019), it is important to understand that supervised machine learning in essence attempts to mimic human decision making. It does so based on the so-called “training set”, which contains data labelled by humans. For this project, no appropriate pre-labelled dataset exists. A large proportion of this protocol is therefore dedicated to describing how we will create our training set, from which the algorithms will “learn” how to select relevant documents and which categories these documents belong to. We make use of additional machine learning methods to extract further information from our dataset and combine the different layers of information for our final assessment.

Altogether, our approach can be divided into 5 stages, given at the outset – see also Figure 1. Conceptually, this is similar to earlier computer-assisted evidence maps (including work by members of the project team, e.g. Callaghan et al., 2021, Callaghan et al., 2020, Sietsma et al., 2021).

In the next section, we will detail the search strategy for step 1. We then describe the criteria used to create the training set used in steps 3 and 4 in section *4.3 Article selection and classification*. More details on our chosen machine learning methods are in *4.5 Machine learning considerations*. The other machine learning methods used for categorisation – i.e. the geoparser and topic modelling – do not make use of the pre-labelled data. More details on their use are found in section *4.6 Data extraction*. The final section, *4.7 Data Synthesis*, describes how we create our evidence map with knowledge gaps and clusters.

4.2 Search strategy

4.2.1 Source

The search will be carried out on the Web of Science Core Collection^[1], which is a publisher-independent scientific database with a wide coverage. In addition, we will also search the MEDLINE database, which covers life sciences and biomedical research, as well as Scopus, which is maintained by the publisher Elsevier and contains a variety of high-quality journals. Note that these databases have a bias towards natural sciences, as well as towards English-speaking countries and the Global North (Mongeon and Paul-Hus, 2016, Vera-Baceta et al., 2019), which limits the representativeness of our results.

The search will be carried out on title, abstract and keywords for all databases, as well as Keywords Plus for Web of Science. We will not limit the search by date, but will limit our search to articles and reviews, in line with our problem scope.

As is common for a machine learning based screening process (Marshall and Wallace, 2019, Haddaway et al., 2020), we will not assess any full texts. Full text screening would be time-intensive (Haddaway and Westgate, 2019), whereas we will need a substantial number of hand-coded documents for our machine learning algorithms to function. We will therefore retrieve only abstracts, titles and meta-data (publication year, authors, author affiliations, keywords, references, field of research for Web of Science) for each article matching our search string; this information per article is what is meant by the word “document”.

[1]Web of Science Core Collection here includes:

- Science Citation Index Expanded (SCI-EXPANDED) –1900-present
- Social Sciences Citation Index (SSCI) –1900-present
- Arts & Humanities Citation Index (A&HCI) –1975-present
- Conference Proceedings Citation Index- Science (CPCI-S) –1990-present
- Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) –1990-present
- Emerging Sources Citation Index (ESCI) –2015-present

4.2.2 *Search string*

Our search string has three main components: climate change, adaptation and policy. Each of these parts in turn has multiple sub-components for which a query was constructed. Documents need to match at least one keyword from all strings – i.e. they are linked by a boolean AND. The majority of sub-components are internally linked by a boolean OR.

Our broad climate search string is a modified version of the query used by (Callaghan et al., 2020), which in turn is based on Grieneisen and Zhang (2011). We remove the majority of mitigation-related terms, but expand the general climate change part, and add keywords on impacts, vulnerability and risk. Note this general part of the query captures all literature which explicitly mentions “climate change” already. The added terms around climate impacts are based on the IPCC’s AR6 Table 12.2. This table describes changes in natural systems and impacts on human and natural systems that can be at least partially attributed to climate change. By including this, we capture literature which does not mention climate change explicitly, but does contain responses to impacts that are primarily driven by climatic changes.

In the adaptation component of our query, we take a similar approach: in line with our Problem Definition above, we need to strike a balance, such that we also capture the wider literature where climate change is a recognised contributor to the subject of interest, without capturing the much larger literature that investigates weather phenomena, rather than climate change. To do so, we split this part of the query in three parts: 1) recognised changes in natural systems based on the aforementioned table in AR6; 2) recognised impacts of these changes based on this same table; and 3) recognised adaptive responses based on the Final Government Draft of the IPCC AR6 WG2's Cross Chapter Box FEASIB, which lists responses to climate change; given that the full WG2 report was not available to us at the time of publishing this protocol, the response options mentioned in AR 5 table 14.1 are added to this. The second and third part here link general weather terms with a boolean AND to the impact and response keywords respectively. These general weather terms are a wider version of the same impacts covered under the first part.

Finally, in the policy part of our components, we again take a broad view of what could be considered relevant. We include terms around policy and governance, including key moments for adaptation in the UNFCCC process. Some governance-related terms (e.g. framework, management/managing) are so widely used that they no longer proved sufficiently selective on their own; similarly, keywords around governance levels (e.g. national/international, cities) also proved too general. We therefore combine these two types of keywords with the NEAR operator. This means that for example articles mentioning a "national framework" are still captured.

The search terms are given in Web of Science syntax in table in the full document. The plain-text search strings are given in Supplementary Materials 2.

4.23 *Comprehensiveness of the search*

Two main factors limit the comprehensiveness of our search strategy. Both are common limitations for works of this kind (Konno and Pullin, 2020). First, keywords are English only, so we will only capture non-English literature that is indexed in English. Combined with the language bias of the databases (Vera-Baceta et al., 2019), this will lead to a relative over-representation of literature from English-speaking countries. Second, our database does not include grey literature. Although it seems likely that a substantial amount of such literature is relevant to adaptation policy, it is more difficult to access in a reproducible manner and often will not include an abstract. The latter poses a problem for us, as our machine learning will take place at the abstract level. In addition to the different format, the lack of a comprehensive database of grey literature also would make its inclusion more complex and time-intensive. As such, this is left to future projects.

The above two caveats notwithstanding, we will strive for search terms and document selection that are as close to comprehensive as possible. The following factors should help ensure this:

- The project team consists of a diverse range of adaptation researchers, with backgrounds ranging from engineering to policy research. This team has all been involved in the formulation of the search query and coding guidelines. A diverse subset will be involved in the coding itself.
- This protocol will be made public prior to completing the work.
- We will cross-check the resulting documents against the references of the Special Report on 1.5C (IPCC, 2018) and the latest Adaptation Gap Report (United Nations Environment Programme, 2021a). We will then scan the list of missing articles for titles that appear adaptation relevant to identify which keywords might be missing.

We will not update the search beyond this within the current project. However, we wish to stress that, once the algorithm is trained, it can be used to filter future searches too. This means that, given proper support, this project could be used as the basis for a so-called “living evidence map” which automatically updates as new scientific papers are published.

4.2 Article selection and classification

4.2.1 Screening strategy

The above search strategy is purposefully kept broad, meaning it includes both relevant and irrelevant articles. We will make use of supervised machine learning to first select the subset that is relevant to climate change adaptation, and then to classify these relevant articles into different categories.

In practice, this leads us to a three-tiered approach for article selection.

In the first step, relevant articles are selected. We first check the articles basic bibliographic data against our criteria set out in the PICoST framework outlined earlier, selecting only articles and reviews and filtering out articles where the abstract is missing. After this, a group of coders will determine whether the document is relevant. There are only two possible labels here: relevant or irrelevant. To be considered relevant, the document should meet two content-based criteria:

The document must include a substantial focus on a response to climate change or to a weather phenomenon wherefore changes can confidently be attributed to climate change, as determined by the IPCC.

The adjustment must be either enabled by, supported by, or a direct result of at least one policy.

These inclusion/exclusion criteria alongside the others following from our PICoST criteria are given in more detail in Table 3 in the formatted text.

In the second step, for the relevant documents only, the type of policy in the document will also be labelled. The categories for these labels are outlined in the subsequent section. If a document contains multiple policies, or one policy fits multiple categories, it will also gain multiple labels. If the document does not contain sufficient information to determine if a given label is appropriate, it will be left blank. In practice, these first two steps are done concurrently.

In the third step, the labelled data from the first two steps is taken to train multiple machine learning algorithms. The documents selected by this algorithm, along with their categories, will form the basis of our analysis.

A separate, more extensive guide for coders, which also contains additional details on the different categories is available in Supplementary Materials 3. Note that this is a first version; as coding progresses, this guide will be amended with additional information and examples to ensure that the coding guidelines are clear and consistently followed by all.

4.3.2 Consistency & independence

Coding will be conducted by multiple researchers from different backgrounds. To ensure that their coding is consistent, 15% of the documents will be coded by two or more researchers. This allows us to find disagreements between researchers; such disagreements will be discussed with the wider team until consensus is reached.

Furthermore, for a conventional systematic review, it is not unusual to address issues of procedural independence – in practice, mostly ensuring that researchers do not include or exclude work they have themselves authored. Given that the majority of the inclusion/exclusion decisions in this project will be made by an algorithm, this is less of an issue. However, should a researcher encounter their own work during the coding process, they will refrain from coding it.

4.4 Category definitions

4.4.1 Translating the NATO typology to adaptation

As stated, a major part of our policy categorisation is based on the NATO typology. These four types – Nodality, Authority, Treasure, Organisation – form the first level of our categorisation. We also add a more detailed second and third level. This categorization scheme was developed collectively by the research team. See the table in the formatted document given in the supplementary materials.

4.4.2 Maladaptation

As adaptation policies are implemented, there is increased attention for the, often unintended, negative consequences of these policies, known as maladaptation. Although there is both a political and scientific debate on the exact meaning and proper use of the term, (Juhola et al., 2016, Glover and Granberg, 2021), we take maladaptation to mean situations where “exposure and sensitivity to climate change impacts are increased as a result of action taken” (Schipper, 2020 p. 409). The maladaptive effects do not need to impact the target group of the policy – in other words, if an adaptation policy shifts the vulnerability to another group, this is still considered maladaptive. Likewise, policies which negatively impact the general welfare of a population are considered maladaptive, as are actions which increase greenhouse gas emissions.

We do not record the type of maladaptation. Instead, we only record if the document provides any evidence of maladaptation or not.

4.4.3 *Constraints and limits*

Documents will also be marked according to whether constraints, limits or any synonyms describing “factors making it harder to plan and implement adaptation” are mentioned in the abstract. If the answer is ‘yes’, coders will be asked to specify the type of constraint, choosing from the options in table 5 below. Definitions and categories are based on the AR5 and AR6 (WG2) from the IPCC.

4.4.4 *Governance Level*

Documents will be marked according to the level at which the policy is implemented (in theory or in practice), choosing between:

- International (including supranational and regional such as the EU or ASEAN),
- National,
- Subnational level (including local, state, province, region, municipal, city).

4.4.5 *Type of impact responded to*

Some documents specify what kind of climate change effect is being responded to. This includes both observed impacts and potential hazards. These are categorised according to IPCC AR5 SPM2 (p. 7), with the addition of the more specific terms drought, heat waves and storms to reflect the increased confidence since that assessment. This results in:

- Glaciers, snow, ice and permafrost

- Rivers, lakes and floods;
- Drought;
- Extreme heat;
- Food production;
- Wildfire;
- Coastal erosion and/or sea level effects;
- Storms and hurricanes;
- Terrestrial ecosystems;
- Marine ecosystems;
- Livelihoods, health or economics.

If no specific impact is mentioned, this is left blank. This includes policies which respond to climate change in general.

In general, the most specific category will be the one selected (e.g. forest fires are influenced by droughts but will only be classified under Wildfire; agriculture generally depends on terrestrial ecosystems to some extent, but will be classified under Food production). However, the categories are not mutually exclusive, so in case multiple specific impacts are mentioned, all will be recorded.

4.4.6 *Evidence type*

Documents will be marked according to whether they provide ex-post or ex-ante evidence on policies. For the purposes of this project, ex-post refers to all studies which analyse the effects of a policy which has already been enacted. Ex-ante refers to all studies which analyse potential effects of a policy once the policy has started being implemented.

4.4.7 *Countries mentioned*

We will use a pre-trained algorithm (Halterman, 2017) to identify geographic locations in the full search. As this algorithm has been trained on non-academic texts however, countries for the labelled data will be noted so we can estimate the accuracy of this algorithm for our particular dataset.

Where documents mention locations or geographical entities, annotators will record the country or countries which contain those geographical entities. For example, if a paper mentions “Berlin”, the annotator will enter “Germany”. Where geographical entities are supranational or non-national (e.g. the European Union, or the Atlantic Ocean) this field shall be left blank.

4.5 Machine learning considerations

4.5.1 Algorithm

To briefly reiterate, in this project we will employ supervised learning of two different types: first, a binary classifier is used for study identification (relevant vs not relevant); on the documents identified as relevant, we will then use a multi-label classifier for each coding level. In both cases, we are using the studies coded by hand as training and validation sets.

Choosing the appropriate algorithm for both these classification tasks is crucial to ensure that the machine learning predictions are fit for purpose. We will therefore test a variety of models from the Scikit Learn package (Pedregosa et al., 2011). Prior work (Berrang-Ford et al., 2021b, Sietsma et al., 2021) had positive results especially using Support Vector Machines (Chang and Lin, 2011); as this does not natively support multi-label predictions, we will use a one-VS-rest set-up for the category predictions. Following Callaghan et al. (2021) we will also test more state-of-the-art deep learning approaches based on BERT (Devlin et al., 2018), including a BERT model that has undergone additional pre-training for a corpus of documents related to climate change (Webersinke et al., 2021).

We will use a nested cross-validation procedure (Cawley and Talbot, 2010, see also Callaghan et al., 2021) to optimize hyperparameters and measure the accuracy of our classifiers. In simple terms, this entails dividing the labelled data up in several subsets. All but one of the subsets are then used to train an algorithm which is used to make predictions on the remaining subset – this is known as the “test set”. The procedure is repeated using another subset as the test set until each of the subsets has functioned as test set. If, for example, we divide in ten subsets, this means we have predictions for all labelled documents based on 10 algorithms that were each trained on 90% of the total labelled dataset. Comparing these predictions against the manually created labels provides an estimate of performance with quantified uncertainty. The process can then be repeated with different hyper-parameters and different algorithms. This allows us to choose the algorithm and hyper-parameters that are most appropriate for our dataset.

4.5.2 Random and non-random samples

Ensuring sufficient numbers of true positives may be a challenge as our initial dataset will be large and relatively unfocussed. With tens of thousands of documents in total and an evidence base that for some

of the more specific categories will not be larger than a few dozen documents in total, we will need a mixture of strategies to provide the machine learning algorithm with sufficient examples to learn from, without biasing results. In practice, we will make use of a mixture of the following types of samples:

- *Random* – the majority of the coded documents will be selected at random. These documents will also be used to estimate the performance of our classifier.
- *Preliminary machine learning* – some documents will be selected based on preliminary results from our classifier. This serves two purposes: first, to identify early on which areas the classifier is struggling with; second, to increase the number of positive examples.
- *Keyword-based* – if a particular area is lacking positive examples, some samples will be drawn based on keywords. Care must be taken here not to bias the results, which in practice means choosing keywords that will be used not just by a large majority of the positive examples we are looking to identify, but which are also still used by a substantial body of other literature.
- *From literature* – if there is a need to further increase the number of positive examples, we may choose to create a sample based on the references of key literature (e.g. IPCC reports). This will ensure that the classifier performs well on highly-regarded documents.

The degree to which non-random samples will be used is dependent on the performance of the classifier. The type of sample will be recorded for all documents assigned to any given reviewer so that we can ensure a sufficiently large random set is used to evaluate the classifier performance and to prevent bias.

4.5.3 Accuracy targets & size of the training set

When setting accuracy targets for supervised machine learning algorithms, a key consideration is the size of the training set (i.e. the number of hand-coded documents), as this to a large degree determines the performance of the classifier. In theory, it may be possible to set a given level of accuracy (e.g. 95%) and keep increasing the size of the training set until this accuracy is reached. In practice however, using an accuracy target in such a way is often impractical for two reasons: first, hand-coding documents is time-intensive (Haddaway and Westgate, 2019); second, the performance of the classifier cannot increase beyond the quality of the input data. In other words, there is likely to be a substantial grey area on what is considered “relevant” which will lead to inconsistency among coders and therefore inconsistent training data. As a consequence, the algorithm will not be able to accurately distinguish between documents in this grey area either, no matter the size of the training dataset.

The extent to which machine learning classifiers can accurately make predictions is unknown *a priori*. Previous efforts using a similar strategy to the one outlined here have found that, even with thousands of documents coded, the accuracy of the classifier remains lower than what would be ordinarily expected in science, which is to say, less than 90% of true positives (Sietsma et al., 2021, Hsu and Rauber, 2021).

Note however that traditional systematic reviews likely also suffer the same issue – it simply remains unreported as the “performance” of the human coders there is never quantified. Using the performance metrics obtained through cross-validation, as described earlier, we can provide estimates on the accuracy algorithm and will report these results as well as their implications for uncertainty bands.

Although a simple accuracy target cannot be set, we can use these accuracy scores to estimate when the classifier is reaching its maximum potential. More concretely, for this project, the training dataset will consist of at minimum 1 500 hand-coded documents, of which at least 1000 are from a random sample. This will then be used to train a first version of the classifier, which will be used to estimate if performance is still increasing with increases in the size of the training set. If this is the case, document screening will continue until the increase in accuracy of the classifier has not increased meaningfully for a minimum of 500 documents added.

The performance for most categories in the multi-label classifier is likely to be lower than that for the binary inclusion/exclusion classifier, given that the same number of documents is used here to provide information on multiple options within the category. Since we further expect the data to be unbalanced (i.e. some categories will have relatively few positive examples while others have many), we may use additional targeted samples to increase performance of the multi-label classifier. Foregrounded results will be limited to the category types where the classifier achieves consistently high accuracy. Recall also that the NATO-typology is hierarchical, which is useful here: if we do not have enough examples to make positive predictions at a lower level, we may still get usable predictions at the higher level.

4.6 Data extraction

Our evidence map is primarily based on the categories predicted for all the relevant documents. The criteria for these categories have been described above, though it is worth repeating that depending on data availability and associated performance of the classifier, some categories may later be merged. As stated also, the geographic location will be based on a pre-trained geoparser, namely Mordecai (Halterman, 2017). The hand-coded locations will only be used to estimate its accuracy.

Meta data will also be retrieved for all articles. This includes the publication year, allowing for a temporal analysis. Further, the author affiliations often include an address or place name. This field can also be fed through the geoparser to identify the geographic location of the authors.

In addition to these categories, we will also make use of topic modelling. This is a so-called “unsupervised” machine learning method. In contrast to the supervised methods described earlier, unsupervised learning does not make use of a training set. Rather, the algorithm searches for structures in the data itself – in the case of topic modelling specifically, the algorithm will find clusters of words which frequently occur together for a set number of topics. Each topic will then be named by the researchers and topics can be grouped together into overarching topic groups. These topic names and

groups will be determined inductively and in combination with the findings from the classifier; as such, we cannot provide additional details on their content before the final dataset of relevant documents has been compiled.

4.7 Data Synthesis

4.7.1 Synthesis strategy

Ordinarily, a systematic review or map will result in a narrative summary where vote-counting is especially discouraged (Haddaway et al., 2018). For this machine learning-based project, more quantitative measures however are all but inevitable. Indeed, determining the size of the evidence base for our various categories is among the core objectives for this research and is necessary to provide context to further findings. As such, we expect to start our evidence synthesis with a description of our final dataset and its development over time, as well as its geographic distribution. These basic descriptors may already point towards biases in the evidence base.

By combining different layers within the final dataset, we can then investigate more complex questions. For example, the NATO categories can be combined with the results of the topic model to investigate what types of tools are most prevalent in different subject areas. We can also further investigate geographic biases, if there are any, by quantifying the prevalence of different topics per region.

4.7.2 Knowledge gaps and clusters

Since topic modelling entails the identification of clusters within a document set, this tool is ideally suited to identify knowledge clusters within academic literature. Topics within the topic model can also overlap, which can be used to identify larger topic groups. Moreover, since topic modelling assumes that each document consists of a mixture of topics, we can also investigate the co-occurrence of topics within documents. This can be used to further highlight knowledge clusters – e.g. if the topic model would include the categories typhoons and re-location, and these topics frequently occur together, this would be an indication that the evidence base here is strong.

To identify knowledge gaps, such a table with co-occurrences could also be useful – e.g. if the same typhoons topic would have little overlap with a coastal zone management topic, this would suggest a lack of evidence. In addition, the categorization of the selected documents should provide some insight into understudied areas – both in terms of subject areas and in terms of geographic distribution.

Lastly, we can compare our dataset against other current assessments of adaptation science generally and adaptation policy in particular. For this, the upcoming IPCC WG2 assessment as well as the Adaptation Gap Report (United Nations Environment Programme, 2021a) should form a solid basis. In the

case of the former, we would for example expect to see more evidence for topics highlighted in the Summary for Policymakers and its figures. Pending government approval, a figure comparing adaptation options is expected.

4.7.3 Critical appraisal

It should be noted that we do not control for study quality in any way, except by limiting our search to established databases of peer-reviewed research. This is especially important given the disparate communities of research from which we draw research. These communities have varying epistemological bases and standards, which despite the diversity of researchers involved in this project, the team may not always be equipped to fully appreciate, especially judging from the abstracts alone. Overall, we aim to be inclusive, which may at times result in the inclusion of documents of a scientific standard that would not be acceptable to all. In our view, this is inevitable given the scale of the project.

This same scale also means that a small number of papers are unlikely to significantly influence the results. On the one hand, this means that the point raised above about scientific quality only becomes a major concern if the general standard of the field is insufficient; on the other hand, there are indications that in some areas of research, the standard may indeed be low (e.g. Scheelbeek et al., 2021) and voices with more fundamental criticisms may get crowded out. Any narrative emerging from the data should therefore be assessed critically in light of the power structures that underpin the science of adaptation policy (Overland and Sovacool, 2020, Nightingale, 2017). More generally, during the analysis, the whole team should keep in mind that we are not *evaluating* adaptation policies, but rather *documenting* where research on adaptation policies is published (similar to the proposal by Tompkins et al., 2018). In short, we assess quantity, not quality of adaptation policy literature. Still, if successful, this would result in the most comprehensive evidence map of adaptation policies to date.

Troubleshooting

Time Taken

Anticipated Results

References

Adaptation Committee. Data for adaptation at different spatial and temporal scales. (UNFCCC Secretariat, Bonn, 2020).

Arnott, J. C., Moser, S. C. & Goodrich, K. A. Evaluation that counts: A review of climate change adaptation indicators & metrics using lessons from effective evaluation and science-practice interaction.

- Environmental Science & Policy* **66**, 383-392, doi:<https://doi.org/10.1016/j.envsci.2016.06.017> (2016).
- Berrang-Ford, L. *et al.* A systematic global stocktake of evidence on human adaptation to climate change. *Nature Climate Change* **11**, 989-1000, doi:10.1038/s41558-021-01170-y (2021).
- Berrang-Ford, L. *et al.* Mapping global research on climate and health using machine learning (a systematic evidence map) [version 1; peer review: awaiting peer review]. *Wellcome Open Research* **6**, doi:10.12688/wellcomeopenres.16415.1 (2021).
- Berrang-Ford, L. *et al.* Systematic mapping of global research on climate and health: a machine learning review. *The Lancet Planetary Health* **5**, e514-e525, doi:[https://doi.org/10.1016/S2542-5196\(21\)00179-0](https://doi.org/10.1016/S2542-5196(21)00179-0) (2021).
- Berrang-Ford, L., Ford, J. D. & Paterson, J. Are we adapting to climate change? *Global Environmental Change* **21**, 25-33, doi:<https://doi.org/10.1016/j.gloenvcha.2010.09.012> (2011).
- Biesbroek, R. & Delaney, A. Mapping the evidence of climate change adaptation policy instruments in Europe. *Environmental Research Letters* **15**, 083005, doi:10.1088/1748-9326/ab8fd1 (2020).
- Biesbroek, R., Badloe, S. & Athanasiadis, I. N. Machine learning for research on climate change adaptation policy integration: an exploratory UK case study. *Regional Environmental Change* **20**, 85, doi:10.1007/s10113-020-01677-8 (2020).
- Biesbroek, R., Wright, S. J., Eguren, S. K., Bonotto, A. & Athanasiadis, I. N. Policy attention to climate change impacts, adaptation and vulnerability: a global assessment of National Communications (1994–2019). *Clim Policy* **22**, 97-111, doi:10.1080/14693062.2021.2018986 (2022).
- Callaghan, M. *et al.* Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. *Nature Climate Change* **11**, 966-972, doi:10.1038/s41558-021-01168-6 (2021).
- Callaghan, M. W., Minx, J. C. & Forster, P. M. A topography of climate change research. *Nature Climate Change* **10**, 118-123, doi:10.1038/s41558-019-0684-5 (2020).
- Capano, G. & Howlett, M. The Knowns and Unknowns of Policy Instrument Analysis: Policy Tools and the Current Research Agenda on Policy Mixes. *SAGE Open* **10**, 2158244019900568, doi:10.1177/2158244019900568 (2020).
- Cawley, G. C. & Talbot, N. L. On over-fitting in model selection and subsequent selection bias in performance evaluation. *The Journal of Machine Learning Research* **11**, 2079-2107 (2010).
- Chang, C.-C. & Lin, C.-J. LIBSVM: A library for support vector machines. **2**, Article 27, doi:10.1145/1961189.1961199 (2011).

- Christiansen, L., Olhoff, A. & Dale, T. Understanding adaptation in the Global Stocktake. (UNEP DTU ..., Copenhagen, 2020).
- Craft, B. & Fisher, S. Measuring the adaptation goal in the global stocktake of the Paris Agreement. *Clim Policy* **18**, 1203-1209, doi:10.1080/14693062.2018.1485546 (2018).
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).
- Ford, J. D. & Berrang-Ford, L. The 4Cs of adaptation tracking: consistency, comparability, comprehensiveness, coherency. *Mitigation and Adaptation Strategies for Global Change* **21**, 839-859, doi:10.1007/s11027-014-9627-7 (2016).
- Ford, J. D. *et al.* Big data has big potential for applications to climate change adaptation. *P Natl Acad Sci USA* **113**, 10729-10732, doi:10.1073/pnas.1614023113 (2016).
- Füssel, H.-M. & Almond, S. The European Climate Data Explorer-a new web portal providing interactive access to climate change information for Europe. (Copernicus Meetings, 2021).
- Glover, L. & Granberg, M. The Politics of Maladaptation. *Climate* **9**, 69 (2021).
- Grieneisen, M. L. & Zhang, M. The current status of climate change research. *Nature Climate Change* **1**, 72-73, doi:10.1038/nclimate1093 (2011).
- Haddaway, N. R. & Westgate, M. J. Predicting the time needed for environmental systematic reviews and systematic maps. *Conservation Biology* **33**, 434-443, doi:10.1111/cobi.13231 (2019).
- Haddaway, N. R. *et al.* On the use of computer-assistance to facilitate systematic mapping. *Campbell Systematic Reviews* **16**, e1129, doi:<https://doi.org/10.1002/cl2.1129> (2020).
- Haddaway, N. R., Macura, B., Whaley, P. & Pullin, A. S. ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environmental Evidence* **7**, 7, doi:10.1186/s13750-018-0121-7 (2018).
- Halterman, A. Mordecai: Full Text Geoparsing and Event Geocoding. *The Journal of Open Source Software* **2**, doi:10.21105/joss.00091 (2017).
- Haunschild, R., Bornmann, L. & Marx, W. Climate Change Research in View of Bibliometrics. *PLoS One* **11**, e0160393-e0160393, doi:10.1371/journal.pone.0160393 (2016).
- Henstra, D. Climate Adaptation in Canada: Governing a Complex Policy Regime. *Review of Policy Research* **34**, 378-399, doi:<https://doi.org/10.1111/ropr.12236> (2017).
- Hood, C. *The tools of government*. (Springer, 1986).

- Hsu, A. & Rauber, R. Diverse climate actors show limited coordination in a large-scale text analysis of strategy documents. *Communications Earth & Environment* **2**, 30, doi:10.1038/s43247-021-00098-7 (2021).
- IPCC. Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. (IPCC, Geneva, Switzerland, 2018).
- IPCC. Summary for Policymakers., (Oxford, 2021).
- Juhola, S., Glaas, E., Linnér, B.-O. & Neset, T.-S. Redefining maladaptation. *Environmental Science & Policy* **55**, 135-140, doi:<https://doi.org/10.1016/j.envsci.2015.09.014> (2016).
- Knill, C. & Tosun, J. *Public policy: A new introduction*. (Bloomsbury Publishing, 2020).
- Konno, K. & Pullin, A. S. Assessing the risk of bias in choice of search sources for environmental meta-analyses. *Research Synthesis Methods* **11**, 698-713, doi:<https://doi.org/10.1002/jrsm.1433> (2020).
- Lamb, W. F., Creutzig, F., Callaghan, M. W. & Minx, J. C. Learning about urban climate solutions from case studies. *Nature Climate Change* **9**, 279-287, doi:10.1038/s41558-019-0440-x (2019).
- Leiter, T. Do governments track the implementation of national climate change adaptation plans? An evidence-based global stocktake of monitoring and evaluation systems. *Environmental Science & Policy* **125**, 179-188, doi:<https://doi.org/10.1016/j.envsci.2021.08.017> (2021).
- Lesnikowski, A. *et al.* Frontiers in data analytics for adaptation research: Topic modeling. *Wiley Interdisciplinary Reviews: Climate Change* **10**, e576 (2019).
- Lesnikowski, A. *et al.* What does the Paris Agreement mean for adaptation? *Climate Policy* **17**, 825-831, doi:10.1080/14693062.2016.1248889 (2017).
- Lesnikowski, A., Ford, J. D., Biesbroek, R. & Berrang-Ford, L. A policy mixes approach to conceptualizing and measuring climate change adaptation policy. *Climatic Change* **156**, 447-469, doi:10.1007/s10584-019-02533-3 (2019).
- Magnan, A. K. & Chalastani, V. I. Towards a Global Adaptation Progress Tracker: first thoughts. *IPCC Working Paper N°1* (2019).
- Marshall, I. J. & Wallace, B. C. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Systematic reviews* **8**, 1-10 (2019).
- Mongeon, P. & Paul-Hus, A. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics* **106**, 213-228 (2016).

- Nachmany, M., Byrnes, R. & Surminski, S. Policy brief National laws and policies on climate change adaptation: a global review. (Grantham Research Institute on Climate Change, London, 2019).
- Nakagawa, S. *et al.* Research weaving: visualizing the future of research synthesis. *Trends in ecology & evolution* **34**, 224-238 (2019).
- Nightingale, A. J. Power and politics in climate change adaptation efforts: Struggles over authority and recognition in the context of political instability. *Geoforum* **84**, 11-20, doi:<https://doi.org/10.1016/j.geoforum.2017.05.011> (2017).
- Nunez-Mir, G. C., Iannone, B. V., Pijanowski, B. C., Kong, N. N. & Fei, S. L. Automated content analysis: addressing the big literature challenge in ecology and evolution. *Methods Ecol Evol* **7**, 1262-1272, doi:10.1111/2041-210x.12602 (2016).
- Overland, I. & Sovacool, B. K. The misallocation of climate research funding. *Energy Research & Social Science* **62**, 101349, doi:<https://doi.org/10.1016/j.erss.2019.101349> (2020).
- Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825-2830 (2011).
- Persson, Å. & Dzebo, A. Special issue: Exploring global and transnational governance of climate change adaptation. *International Environmental Agreements: Politics, Law and Economics* **19**, 357-367, doi:10.1007/s10784-019-09440-z (2019).
- Rolnick, D. *et al.* Tackling Climate Change with Machine Learning. doi:arXiv:1906.05433v2 (2019).
- Runhaar, H., Wilk, B., Persson, Å., Uittenbroek, C. & Wamsler, C. Mainstreaming climate adaptation: taking stock about “what works” from empirical research worldwide. *Regional Environmental Change* **18**, 1201-1210, doi:10.1007/s10113-017-1259-5 (2018).
- Scheelbeek, P. F. *et al.* The effects on public health of climate change adaptation responses: a systematic review of evidence from low-and middle-income countries. *Environ Res Lett* **16**, 073001, doi:<https://doi.org/10.1088/1748-9326/ac092c> (2021).
- Schipper, E. L. F. Maladaptation: When Adaptation to Climate Change Goes Very Wrong. *One Earth* **3**, 409-414, doi:<https://doi.org/10.1016/j.oneear.2020.09.014> (2020).
- Sietsma, A. J., Ford, J. D., Callaghan, M. W. & Minx, J. C. Progress in Climate Change Adaptation Research. *Environmental Research Letters* **In press**, doi:<https://doi.org/10.1088/1748-9326/abf7f> (2021).
- Tompkins, E. L., Vincent, K., Nicholls, R. J. & Suckall, N. Documenting the state of adaptation for the global stocktake of the Paris Agreement. *WIREs Climate Change* **9**, e545, doi:10.1002/wcc.545 (2018).

Ulibarri, N. *et al.* A global assessment of policy tools to support climate adaptation. *Clim Policy* **22**, 77-96, doi:10.1080/14693062.2021.2002251 (2022).

United Nations Environment Programme. Adaptation Gap Report 2021: The gathering storm – Adapting to climate change in a post-pandemic world. Report No. ISBN: 978-92-807-3834-6, (Nairobi, 2021).

United Nations Environment Programme. Emissions Gap Report 2020: The Heat Is On – A World of Climate Promises Not Yet Delivered. (Nairobi, 2021).

Van de Schoot, R. *et al.* An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence* **3**, 125-133 (2021).

Vera-Baceta, M.-A., Thelwall, M. & Kousha, K. Web of Science and Scopus language coverage. *Scientometrics* **121**, 1803-1813, doi:10.1007/s11192-019-03264-z (2019).

Wang, Z., Zhao, Y. & Wang, B. A bibliometric analysis of climate change adaptation based on massive research literature data. *Journal of Cleaner Production* **199**, 1072-1082, doi:10.1016/j.jclepro.2018.06.183 (2018).

Webersinke, N., Kraus, M., Bingler, J. A. & Leippold, M. Climatebert: A pretrained language model for climate-related text. *arXiv preprint arXiv:2110.12010* (2021).

Acknowledgements

The authors declare no competing interests.

Figures

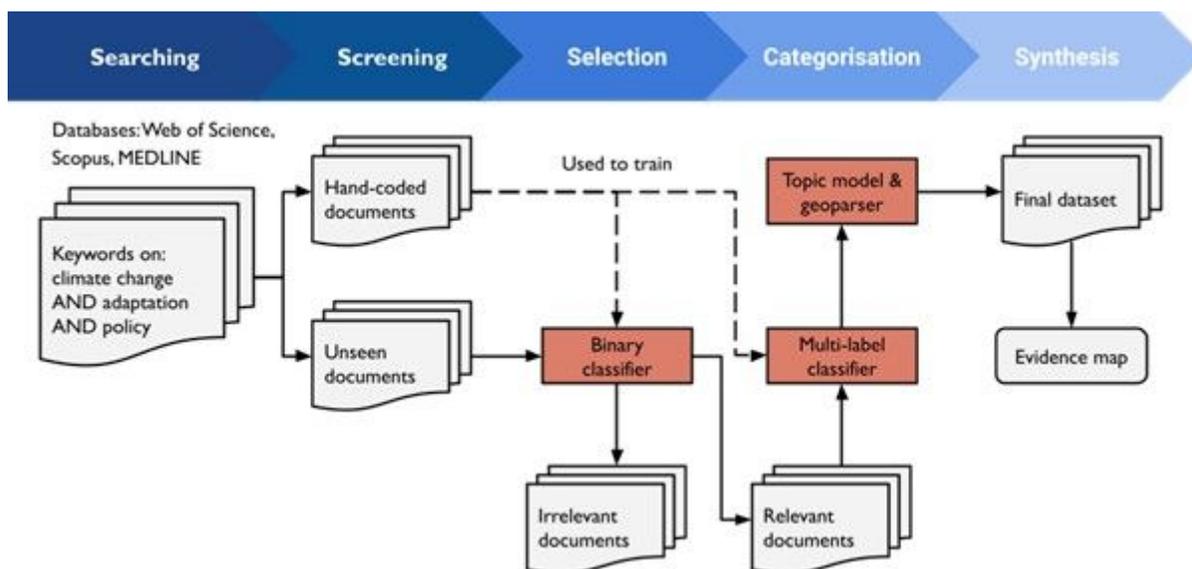


Figure 1

Figure 1: overview of the process. Machine learning components are highlighted in red.

Core objective	Systematically synthesise evidence on climate change adaptation policies in the global academic literature, tying it explicitly to the established NATO typology, with a focus on the geographic distribution of different policy types and their development over time, as well as any identified constraints and limits and climate impacts
Population	Humans experiencing or expecting to experience climate change



Figure 2

Table 1: the review objective and its key components according to the PICoST framework

Inclusion	Exclusion
<p>The document must include a substantial focus on a perceived, experienced, or observed adjustment (adaptation) to the actual or expected impacts of climate change or to a weather phenomenon wherefore changes can confidently be attributed to climate change, as determined by the IPCC including floods</p>	<p>No response is discussed, or the responses are only discussed at a superficial level. This includes papers focussed on assessing risks or impacts, where those assessments are not used to make concrete suggestions for adaptations.</p> <p>The response is not to a climate change</p>

Figure 3

Table 3: inclusion/exclusion criteria (papers must meet all to be included)

Category of constraint	Description
Economic	Existing livelihoods, economic structures and economic mobility.
Social/cultural	Social norms, identity, place attachment, beliefs, worldviews, values, awareness, education, social justice, and social support.
Human capacity	Individual, organisational and societal capabilities to set and achieve adaptation objectives over time. This includes training, education and skill development.

Figure 4

Table 5: The categories of constraints and limits used for categorizing documents

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [220208ClimateChangeAdaptationPolicyProtocolFINALclean.pdf](#)
- [InstructionsforCoding.pdf](#)
- [211125Query.txt](#)
- [ROSESforSystematicMapProtocols.xlsx](#)