

A Coding Protocol for Labeling Scientific Literature on Carbon Dioxide Removal to Train Machine Learning Models

Sarah Lück*¹, Christiane Hamann¹, Doménica Michelle Jaramillo Sánchez¹, Ronja Kelch¹, Fariha Mawla¹, Fabian Metz¹, Leon Stephan¹, David Verdugo-Raab¹

¹ Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany

Carbon dioxide removal (CDR) plays an important role in any strategy to limit global warming to well below 2°C. As research on CDR rapidly expands, keeping pace with the scientific evidence through rigorous and transparent evidence synthesis is essential for informing sustainable deployment. To support this, we use artificial intelligence (AI) to develop the first comprehensive systematic map of CDR research. A critical first step in this process is the creation of a high-quality training dataset to enable machine learning models to accurately identify relevant scientific literature. This coding protocol provides a structured methodology for generating such a dataset, including the development of a coding scheme, inclusion criteria, and annotation workflow. The protocol ensures consistency and reproducibility, facilitating scalable AI-assisted evidence mapping in the CDR domain.

Introduction

To comply with the Paris Agreement and limit global warming to well below 2°C, rapid and deep reductions in greenhouse gas (GHG) emissions are essential. However, these efforts must be complemented by Carbon Dioxide Removal (CDR)—potentially at the gigaton scale by mid-century and beyond (IPCC 2022; Smith et al. 2023).

CDR serves three distinct roles in climate change mitigation (IPCC 2022). First, in the near term, it can reduce net CO₂ and GHG emissions, particularly in the land sector. Second, in the medium term, it can offset residual emissions from "hard-to-mitigate" sectors such as industry, long-distance transport, and agriculture (Buck et al. 2023; Luderer et al. 2018). Third, in the long term, CDR can support sustained net-negative emissions, helping to lower global temperatures in overshoot scenarios and stabilize warming at or below 1.5°C (IPCC 2022).

That said, CDR cannot substitute for stringent emissions reductions, which must remain the priority—even in sectors that are difficult to decarbonize (Fuhrman et al. 2024). Moreover, there are significant uncertainties about how rapidly and sustainably CDR technologies can be scaled, and whether temperature overshoots can be safely reversed (Schleussner et al. 2024). These uncertainties underscore the urgent need to reduce emissions as quickly as possible, while also ensuring robust policy support to enable CDR to achieve gigaton-scale removals in the second half of the 21st century (Smith et al. 2023; Nemet et al. 2018).

As scientific and technological research on CDR continues to accelerate (Minx et al. 2017), it is increasingly important to systematically track, evaluate, and synthesize emerging evidence. Conducting this synthesis in a rigorous, transparent, and methodologically sound manner ensures that policymakers, stakeholders, and practitioners have access to reliable information. This, in turn, is

crucial for the responsible and sustainable deployment of CDR, minimizing environmental risks while maximizing potential social and economic co-benefits.

The project step by step

The main aim for this project is to identify what research exists on Carbon Dioxide Removal (CDR) technologies in the scientific community. Technologies, in scope of this project, are defined broadly as means to an end, and include devices, practices and behaviours. The main challenge is that this literature is spread across multiple disciplines, having large corpora, using diverse terminology.

We address this problem by operationalising and developing inclusion criteria - that serve as system boundaries - and definitions used for categorizations of the records.

In order to deal with the large corpora, machine learning is used to scale-up the repetitive task of analysing abstracts and categorising the articles accordingly thereby staying within resource boundaries. We use Machine learning to both filter out relevant literature and to identify different categories. The whole process can be roughly divided into three steps:

1. Develop a search query. This query tries to find articles that are about Carbon Dioxide Removal and specific known CDR Technologies; however, it uses broad search terms for all its components by design. In essence, we are trying to gather as many documents as seems feasible and then use the subsequent step to filter out the irrelevant articles. For each CDR technology we created a separated query which are listed here: [link](#)
2. Train machine learning algorithms. We will mostly make use of so-called “supervised machine learning”. This means that a computer tries to mimic human decisions based on a training set. The training set consists of documents that are hand-coded by researchers.

This document is focussed on step 2. It contains the instructions for coding—in essence, the decisions researchers are trying to teach to a computer algorithm. Machine learning algorithms can only perform well if their training data is consistent; however, in practice, scientific publishing on CDR is heterogeneous and includes grey areas and edge-cases. This means that coders will have to continuously talk to each other to ensure that we get as close as possible to consistent coding.

Platform Instructions

The coding for this project will take place on the NACSOS platform. The platform is hosted by the MCC in Berlin, but all elements of it are also provided publicly.

Coders should all have received login details for this platform. Once inside, you should be able to click on “scoping”, which takes you to the overview of projects you are a part of. This project has the ID 309 which can be accessed by the [the project overview page](#).

At the top, you should see your username. Clicking this takes you to the *review overview page*.

Scoping Review Helper

User metz@mcc-berlin.net Overview

User Credentials

Enter credentials here that can be used when submitting queries. N.B Must be UK Federation, only tested for University of Leeds. Security implications vague... At your own risk.

Organisation:

Username:

Password:

Scoping queries

Query ID	Query Title	Tag Title	! To review	Reviewed	Yes	No	Maybe	Yes buts
10343	CDR papers from AU - all docs	all	0	100 (84% relevant)	78	15	7	0
10342	CDR papers from AU - all docs	all	0	50 (74% relevant)	35	12	3	0

[Log out](#)
[Change password](#)

Click your username to get to this review overview page.

Click the "to review" number to get to the reviewing screen.

Once you are in the review page, you see an overview of all documents that are assigned to you for review. The column "to review" shows the number of documents that have not been coded yet. Click this to start the coding. If for any reason you need to go back, you can also click the numbers in the other columns to see documents that you have already coded.

The review screen presents you with a screen that is divided into four major areas:

1. At the top, there is a progress bar. This shows you how many documents you have already rated. You can also click the blocks in this progress bar to move between documents, for example if you want to revise a rating.
2. Below the progress bar on the left, you will see all the information of the current document. In most cases, this will be a title at the top, followed by an abstract, where certain keywords are highlighted to help you find relevant passages. Below that is bibliography information, including the authors, keywords, and publication year.
3. Below this is a notes section. This may be especially helpful for double-coded documents or documents you wish to discuss with the rest of the team. Remember to click the "Add note" button to save your note before moving on.
4. Finally, the actual coding categories are presented on the right. You should first decide whether an exclusion criterion applies to the record, rendering it not relevant. (see picture on the next page). If an article does not deal with a CDR-related topic you only click "Non-CDR" and immediately "No" for relevance. In all other cases you should click all categories that apply to your document. Hovering over a category option will show additional info on what is meant by the category. You should assign relevance last, as this will save your choices and move you to the next document

automatically.

The screenshot shows the 'Query Screener' interface for query number 10429. At the top, there is a progress bar and a welcome message: 'Query Screener (Query no. 10429) - Welcome, metz@mcc-berlin.net, your progress:'. Below this, the 'Query criteria' section is currently empty. The main content area displays a document titled 'Globalizing results from ocean in situ iron fertilization studies' with a snippet of text: '[1] Despite the growing number of in situ iron fertilization experiments, the efficiency of such fertilization to sequester atmospheric CO₂ remains largely unknown. For the first time, a global ocean biogeochemical model has been evaluated against those experiments and then used to estimate the effect of a long-term and large-scale iron addition on atmospheric CO₂. The model reproduces the observed timing and amplitude in chlorophyll, the shift in ecosystem composition, and the pCO₂ drawdown; it also proves to be of utility in interpreting the observations. However, a full ocean fertilization during 100 years results in a 33 mu atm decrease in atmospheric CO₂ that is 2 to 3 times smaller than found previously.' Below the text, there is a 'GLOBAL BIOGEOCHEMICAL CYCLES (2006) 10.1029/2005GB002591' citation and a list of authors: 'Aumont, O. [i]; Aumont, O. [i]; Bopp, L. [i]; Bopp, L. [i]'. The document type is 'Article' and the document ID is '29709'. On the right side, there are several sections for 'Which level 1 categories is this document relevant to?' (hover for more info), 'Which type of study categories is this document relevant to?' (hover for more info), 'Which content categories is this document relevant to?' (hover for more info), 'Which assessment categories is this document relevant to?' (hover for more info), 'Which exclusion categories is this document relevant to?' (hover for more info), and 'Is this document relevant according to the criteria shown?' (Click buttons or use number keys). The 'Exclusion labels' section includes categories like 'AR', 'BECCS', 'Biochar', 'Blue Carbon', 'CCS', 'CCUS', 'DAC/CCS', 'Enhanced Weathering (land based)', 'General Literature on CD/NET', 'Microalgae', 'Ocean alkalinity enhancement', 'Ocean fertilization & artificial upwelling', 'Other technologies', 'Restoration of landscapes/peats', and 'Soil Carbon Sequestration'. The 'Relevance' section has buttons for 'Yes (1)', 'No (2)', and 'Maybe (3)'. Red circles and arrows highlight the progress bar, the 'Exclusion labels' section, and the 'Relevance' section.

Inclusion and Exclusion

The first choice for all documents is whether it is relevant or not. In essence, a document is relevant if it discusses Carbon Dioxide Removal (CDR) Technology in any way. This includes discussing the application of a CDR - Technology, its side-effects or societal implications, such as governance or public perception. Note that the technology is not always mentioned explicitly, sometimes rather a process is described that amounts to removing CO₂ from the atmosphere and storing it, e.g. Pyrolysis of Biomass resulting in the sequestration of CO₂. Long-term storage¹ (*at least 5 years* is important in this context. We are not interested in technology pathways that re-release the carbon after short periods of time. Within this definition, CO₂ from direct air capture used in sparkling water or used for the production of synthetic car fuels are not considered as CDR. Although some of the technologies are not regarded as CDR, such as CCS, algae farming or CCUS we include them anyway. We might want to analyse them later..

More detailed inclusion/exclusion criteria are given in the table below. If a document meets all of the inclusion criteria, you click "Yes" on the question for relevance. If it does not meet one or more of the below criteria, click "No" to move to the next document. If you are unsure, click "Maybe". Articles rated as "Maybe" will be discussed with your colleagues. (Note again that the assignment of relevance should be the last step in coding the record as it immediately saves changes and presents the next record.)

¹ Current working definition: At least one accounting period: one year.

<p>Any study where CO₂ Storage by active and conscious human intervention is discussed and is (potentially) accomplished over a longer period (more than 1 year). This includes any direct discussion of CDR Technologies (explicitly and implicitly) or of concrete implications for the application of CDR Technologies. Furthermore, it includes any process of C-fixation, where the end-use is framed in the context of CCS or remains open.</p> <p>Edge cases to include:</p> <ul style="list-style-type: none"> - biochar is produced but it is not clear what happens to the biochar 	<p>Explicitly excluded are</p> <ul style="list-style-type: none"> - studies discussing CCS in the production of biofuels and other short-lived products (e.g. cosmetics etc.) where no net-negative emissions are achieved. - studies where biochar is burned - studies not dealing with a CDR Technology - natural processes where there is no human influence or the data can not be used to influence human actions, such as Example1, Example2
<p>The abstract must be available in English. <i>(If no abstract is presented see point 2 below)</i></p>	<p>The abstract is only available in languages other than English.</p>

General Remarks on Coding Conventions

1. We code very inclusively, meaning:

All records where the abstract mentions a CDR technology are included **unless** it is only mentioned in the context of concluding **remarks for further research, as a rhetorical figure OR** a pathway **only involves short-term CO₂ storage**. Furthermore, if we cannot exclude a paper on the abstract base, we include that record.
2. We use the title, abstract and - if available - the author keywords to judge the study. If an abstract is not available, judge from the title if possible. If the information does not suffice for **an inclusion** click “maybe”. And add the note “No abstract”. A title clearly mentioning Carbon Dioxide Removal, a CDR technology or CDR technologies in general is sufficient for inclusion.
3. We code what has been stated in the abstract and can be regarded as primary research and the objective of the study. If the title/abstract deals with more than one technologies/scientific methods all are coded.
4. Since we want to use machine learning methods to judge the remaining articles, please do not use information besides the title/abstract/keywords to come to an inclusion judgement, rather use the “Maybe” label. You should infer as little as possible to come to a judgement.

5. You may and should refer to author keywords.
6. We do not use the Web of Science Keywords for judgements, as they are apparently automatically generated and often misleading.
7. It is not sufficient to include a study if only one vague reference to CDR or a CDR Technology is mentioned. Include the record if multiple references are given.

Categories

For all relevant articles we further sort them into 4 different categories, namely CDR technology, research method, main focus of study, side-effects. Please click all categories that apply. This can mean **the categories are not mutually exclusive**. We try to fix coding boundaries in the tables below.

It is possible that none of the options in one or multiple of the categories apply. If the information in the abstract itself is not sufficient to decide what option within a category applies, please do not make a selection. But because all the documents are research articles on CDR we expect for each article that it deals with a CDR technology and uses a research method. Those two categories must be filled for each included article. Do not use outside sources or make speculative inferences.

The tables below present all categories, their labels and label descriptions, furthermore the third column on rules/application provides for some cases additional information on how a given label is applied and edge cases are treated.

0 Technology (CDR Technologies)

Every included study has to be assigned to a technology. Not mentioning a specific technology gets a “General literature on CDR” label, a technology which cannot be assigned to another technology a “Other/New technology” label.

0 Technology (CDR Technologies)		
Bioenergy with Carbon Capture and Storage (BECCS)	The application of Carbon Dioxide Capture and Storage (CCS) technology to bioenergy conversion processes.	<p>This label includes carbon negative Biofuel production</p> <p>If BECCS is applied, we do not code CCS separately. Unless BECCS and CCS are also discussed separately.</p> <p>BECCS usually refers to Energy or Heat production in a power plant , where the flue-gases are filtered and stored using CCS.</p> <p>If co-firing (using Biomass and fossils together) is discussed, we code both CCS and BECCS</p> <p>Mentioning bioenergy alone in the context of mitigation does not suffice for an inclusion.</p>
Afforestation and reforestation (AR)	Planting of new forests on lands that historically have not contained forests and replanting of forests that have been cleared. Planting forests per se reduces CO2. Studies mentioning AR are therefore included very inclusively.	
Agroforestry	<i>Agroforestry</i> is a land use management system in which trees or shrubs are grown around or	

	among crops or pastureland. Agroforestry is one of the CDR technologies mentioned in the latest IPCC report and as such it is always included.	
Forest management	Every aspect of management which leads to a sustained carbon dioxide removal. Forest management per se does not increase carbon storage, the already existing storage is rather maintained. Because of that the study must refer to carbon sequestration, e.g. soil carbon, biomass in the forest etc, to be included into the map.	
Direct Air Carbon Capture and Storage (DACCS)	Chemical process by which CO ₂ is captured directly from the ambient air, with subsequent storage. Also known as direct air capture and storage (DACs).	<p>If DACCS is applied, we do not code CCS separately. Unless DACCS and CCS are also discussed separately.</p> <p>We also include DAC-only papers here that focus on the process of dissection C from ambient air.</p> <p>Included are also research on DACCS indoors although their capacity is small.</p>
Enhanced Weathering (terrestrial)	Enhancing the removal of carbon dioxide (CO ₂) from the atmosphere through dissolution of silicate and carbonate rocks by grinding these minerals to small particles and actively applying them to soils.	<p>Examples of silicate and carbonate rocks are basalt and feldspar and lime and dolostone. Sometimes components of limestone are mentioned as well e.g. CaCo₃.</p> <p>Liming is explicitly included.</p> <p>We are interested in active human mediated</p>

		<p>weathering. Excluded are:</p> <p>Descriptive analyses of natural weathering processes</p> <ul style="list-style-type: none"> - Non active anthropogenic effects: <ul style="list-style-type: none"> o Analyses of aerosol effects <p>If it is unclear whether the EW process is applied on land or in the ocean use Enhanced Weathering as default.</p> <p>If both usages (on land and in the ocean) are described use both labels.</p> <p>Liming in agriculture seems to have controversially effects on CO₂, it might sequester CO₂ or increase CO₂ emissions. Because we are not sure about the overall outcome we will include papers on liming in agriculture.</p> <p>Ocean liming is a one way to in enhance ocean alkalinity and is included as "Ocean alkalinity enhancement".</p>
<p>Ocean alkalinity enhancement (enhanced ocean weathering)</p>	<p>Enhancing the removal of carbon dioxide from the atmosphere through dissolution of silicate and carbonate rocks by grinding these minerals to small particles and actively applying them to coasts and oceans.</p>	<p>Is the opposite of acidification.</p> <p>We are interested in active human mediated weathering. Excluded are:</p> <ul style="list-style-type: none"> - Descriptive analyses of natural weathering processes - No active anthropogenic effects: <ul style="list-style-type: none"> o Analyses of aerosol effects

<p>Ocean fertilization & Artificial upwelling</p>	<p>Deliberate increase of nutrient supply to the near-surface ocean in order to enhance biological production through which additional carbon dioxide from the <i>atmosphere</i> is sequestered. This can be achieved by the addition of micro-nutrients or macro-nutrients.</p> <p>Upwelling refers to the use of pipes or other methods to pump nutrient-rich deep ocean water to the surface where it has a fertilizing effect.</p>	
<p>Biochar</p>	<p>Stable, carbon-rich material produced by heating <i>biomass</i> in an oxygen-limited environment (pyrolysis) .</p>	<p>< (e.g. Sorption etc.) and studies discussing the co-production of biochar and biofuel are included.</p> <p>We include Hydrochar under this label.</p> <p>Uses of Biochar as fuel (where it is burned and CO₂ released) are excluded:</p> <ul style="list-style-type: none"> - E.g. Direct fuel cell applications
<p>Soil Carbon Sequestration (SCS)</p>	<p>Land management changes which increase the soil organic carbon content, resulting in a net removal of CO₂ from the atmosphere.</p>	<p>Applications of biochar as soil amendment are included if the effects on the C content of the soil is discussed.</p> <p>“crop residue retention” is practice used for SCS</p> <p>We include records dealing with cover crops only if their application is framed in a CDR context.</p>

		<p>It seems cultivation of biofuel crops is beneficial for long term storage of soil carbon. We include those cases and label them with SCS and bioenergy. Example1, Example2</p> <p>Sediment carbon, i.e. carbon stored in material which sank to the ground in water, is not soil carbon. We do not code SCS here.</p> <p>Carbon Farming is a practice where farming serves to sequester carbon in soil. We will code it as SCS but it can include other technologies, such as biochar.</p>
Blue carbon	<p>Blue carbon is the carbon captured by living organisms in coastal (e.g., mangroves, salt marshes, seagrass beds, macro algae and seaweed) and marine ecosystems, and stored in biomass and sediments.</p>	<p>We include all discussion of restoration efforts of Blue Carbon (often dealing with mangrove forests and seagrass beds), either before restoration or evaluation of already existing restored forests/beds, Example.</p> <p>Included is conservation of marine ecosystems (mangroves etc.) if it is done with a focus on carbon sequestration. Example</p> <p>Coral reefs are not a clear Blue carbon technology as their potential of sequestering carbon is low. If they are discussed with a focus on Blue Carbon we follow the restoration/conservation rules.</p> <p>Mentioning “management” without mentioning carbon sequestration or looking at the sequestered carbon will not be included, cp. forest management.</p> <p>Blue carbon often manifests in sediment carbon.</p>

		Sediment consists of material which sank to the ground. It is no soil carbon.
Restoration of landscapes and peats	A process by which formerly destroyed landscapes that sequester CO ₂ are restored to their natural conditions.	Is only applied to land-based activities. Mere rewetting for peats is not sufficient, a record needs to state that CO ₂ is sequestered. Should only be double coded with AR and SCS if either AR or SCS are discussed additionally.
Carbon dioxide Capture Utilization and Storage (CCUS)	A process in which CO ₂ is captured and then used to produce a new product. Only if the CO ₂ is stored in a product for a climate-relevant time horizon, this is referred to as carbon dioxide capture, utilisation, and storage. Only then, and only combined with CO ₂ recently removed from the atmosphere, can CCUS lead to carbon dioxide removal. <i>Examples would be usage for furniture, as building material,</i>	Included are in general products which are solid, excluded are products which are highly flammable. More detailed inclusion/exclusion list: link Enhanced oil recovery (EOR) is excluded as study and does not count as CCUS. Exceptions are if EOR is discussed as part of BECCS/DACCS etc. We observe that many studies talk about CCUS but actually mean CCS. We code here CCS, eg. 1826959 . If several long lived and short lived products are discussed we include the study
Other/New Technology	Any technology that removes CO ₂ from the Atmosphere and is not listed in this overview.	Applied to any new technology that comes up. Be reminded of the wide definition technology employed here. example
General Literature on CDR and CDR	Any literature that discusses the process or	If a record compares multiple specific CDR Technologies all single labels are applied plus

Technologies.	impact of CDR- Technologies, without a limited focus on one or multiple explicit technologies.	<p>“General literature”.</p> <p>If a record deals with a class of multiple technologies - e.g. terrestrial or marine CDR – use general literature. Also, if a record deals with a (partly) unknown portfolio of CDR technologies, use the labels for the known technologies and “general literature” additionally.</p> <p>Included are papers that explicitly deal with scenarios excluding CDR Technologies and discussing that exclusion.</p>
Direct Ocean Capture	<p>Electrochemical processes for oceanic carbon removal</p> <p>Note: Only coded for a few documents as part of an additional labelling for OAE, to be found in NACSOS2 in the project “CDR to find OAE”</p>	<p>Since apparently the terminology is not fixed yet and some abstracts use Ocean Alkalinity Enhancement and Direct Ocean Capture synonymously we label for processes using electrochemical processes to capture carbon in the ocean both labels: OAE and Direct Ocean Capture</p>

Additional Technology

Additional technology labels are applied to either further qualify a CDR Technology or to distinguish literature that deals with key technology components of a CDR Technology. On NACSOS we do not differentiate Technology from Additional Technology.

Algae Farming		Applied if C is stored in algae. If they discuss harvesting or growing algae in general, there

	<p>Any managed processes that fixate C in macro or micro algae, in tanks or in the sea, to be harvested later.</p>	<p>need to be a framing in CO2 storage (<u>ex1</u>)</p> <p>Some papers discussing algae will be relevant:</p> <ul style="list-style-type: none"> - Algae Farming to gather Biomass for BECCS - as material for mid/long term stable products <p>Others will be irrelevant:</p> <ul style="list-style-type: none"> - Algae farmed for biofuel production. - Algae farmed as fodder for livestock
<p>Carbon Capture and Storage (CCS)</p>	<p>A process in which a relatively pure stream of carbon dioxide (CO2) from industrial and energy-related sources is separated (captured), conditioned, compressed and transported to a storage location for long term isolation from the atmosphere</p>	<p>This is a partial technology for BECCS and DACCS and EOR. If only BECCS/DACCS, EOR are discussed in an article we do not apply an additional CSS label.</p> <p>Only if the technologies are discussed separately apply multiple labels.</p> <p>If a paper focuses on the capture process also code CCS if it is capture from a highly concentrated CO2 stream.</p> <p>Wastewater from mining can be treated with CO2 to detoxify it. We include this as CCS, <u>example</u> , <u>example2</u></p> <p>We code CCS where an emulsion of CO2 and</p>

		water using CaCO ₃ is generated, <u>example</u>
--	--	--

Scientific Method

Please assign to each included article a scientific method. If you are unsure what method is used in the article use “Unknown method”.

Scientific Method		
Life Cycle Assessments (LCA)	Analysis of all environmental impacts of a product or service over its lifetime. By quantifying all inputs and outputs of material flows and assessing how these material flows affect the environment. In the context of CDR often assessments of C storage with regard to material flows.	<ul style="list-style-type: none"> - Is related to both statistical analysis/accounting and modelling - If the LCA label is applied do not double code <i>statistical analysis</i> or <i>modelling</i>
Modelling	(Mathematical) simulation of socio-economic systems, energy systems, technological systems, human processes	<p>We also code modelling if there are concrete quantified predictions of the future mentioned in the abstract.</p> <p>Aspen Plus is a software for simulating industrial processes and an indicator for “modelling”.</p> <p>Other words which indicate modelling:</p> <ul style="list-style-type: none"> - techno-economic assessment - IAM, integrative assessment modelling - scenario
Experiment	The manipulation of variables to establish cause and effect relationships.	For all experiments we do not label data analysis additionally if they do some basic calculations, e.g. mean, stdv etc.

		More complex analysis asks for an additional label in data analysis, e.g. an example with <u>spatial correlation</u>
<i>Experiment - Field study</i>	Testing a hypothesis outside of artificial and highly controlled settings. The idea is to include natural variation into the experimental setting.	Includes observational studies. Mesocosm experiment, e.g. a tank in the ocean, are labelled as field study.
<i>Experiment - Laboratory experiments</i>	Testing a hypothesis in an artificial and highly controlled setting of a laboratory.	Microcosm studies where an ecosystem is simulated are labelled as laboratory experiments. Includes pot experiments in greenhouses.
Review	Any study that relies on a literature review as scientific basis for its claims.	This is understood very broad in the context of this research and includes: <ul style="list-style-type: none"> - Overview pieces (e.g.in book chapters) - Perspective pieces that rely limited and sometimes skewed literature bases (often these do not provide methods and give an introduction into a topic/ overview of what research needs to be done) - Introductions to journal volumes - Proceedings paper are no reviews automatically but preliminary research papers, if possible identify a method of all available methods, including review commentary are included as reviews, <u>example</u>

Systematic review	A type of evidence synthesis, using repeatable analytical methods to collect secondary data and analyse it (quantitatively and qualitatively). It is often mentioned as method in the abstract.	Mentioning numbers of reviewed /synthesised articles in the abstract is an indication for a systematic review. Meta-analysis, i.e. the comparison and evaluation of several data sets, can be a part of a systematic review. But if the study only conducts a meta-analysis we code it as “data analysis”
Survey	Using any physical or digital tool to gather answers to a predefined set of questions.	also included here are interviews, e.g. with experts
Data analysis / Statistical analysis / Econometric	Data analysis, remote sensing analysis,, statistical models, e.g. panel analysis and accounting studies	Meta-analysis, i.e. the comparison and evaluation of several data sets, can be a part of a systematic review. But if the study only conducts a meta-analysis we code it as “data analysis”. Included here are: <ul style="list-style-type: none"> - machine learning approaches - multi-criteria analysis, <u>ex</u> (included in the author keywords) - statistical linear-mixed effect models, <u>ex</u>
Qualitative research	This label is attributed to any qualitative design not caught by other labels. This might include focus groups, case studies, developing frameworks for assessments etc.	Double check if the article does not qualify as perspective (review category) before applying this label.

Unknown method	It is the default label if none of the other labels are applicable.	Double check if the article does not qualify as perspective (review category) before applying this label.
----------------	---	---

Main focus

This category aims at catching the wider context in which a technology or its application are discussed in an article.

Main focus		
Public perception & acceptance	Developments of opinions/stances in a given population and media coverage.	
Policy, politics and governance	We follow the understanding of “public policy” in order to be as broad as possible in this category: Public policy is an institutionalised proposal or a decided set of elements like laws, regulations, guidelines, and actions to guide CDR implementation.	
Earth system/ Climate Systems	Global ocean-, land- and air-systems or their interactions and respective outcomes with regard to Climate Change. This includes management, intervention through deploying CDR technologies and their estimated	This category is limited to: Quantitative assessments e.g. in the form of climate (consequences) modelling This label usually co-occurs with modelling as method.

	outcomes.	Hints are usage of following scenario models: MAgPIE, LPJ, LPJmL, JSBACH, the name carbon cycle-climate model (see http://climatemodels.uchicago.edu/)
Socio-economic pathways	Societal- and market decisions and trajectories. They elicit options for actions. A classical example would be to investigate with a model deployment strategies for different technologies under economic constraints. Here, we do not include optimal deployment of one technology, this would be labelled as technology specific.	<p>This includes:</p> <ul style="list-style-type: none"> - Market modelling - Emission pathways - modelling of scenarios - Energy system portfolios - Integrated assessments - Government investments - Land-use depending on the context it is discussed <p>This label usually co-occurs with modelling as method.</p> <p>Not included are optimal solutions for one technology, e.g. best suitable land for planting. 4377360. We code such documents as technology.</p> <p>Often these documents are framed in a way that they want to guide policies, e.g. If they just mention policy in an outlook, we do not code the policy label here.</p>
Technology	Features, processes, and the application of the technology.	This is the most vague/catch-all label in this category. As many articles revolve around technicalities of a given CDR Technology and do not contextualise further.

		If another context category (e.g. Earth Systems) is applied only code this label if, the record discusses variables of/ variations to the process of the CDR technology additionally.
Equity & Ethics	Analysis or discussions of the ethical/ normative dimension of CDR technologies and their application.	This might include, risk assessments, assessments of burden-sharing and the like. An <u>example</u>

Assessment

The "Assessment" category is used if an article analyses the application of CDR Technologies.

For that purpose, different assessment criteria are distinguished.

Economic considerations	Information about the economic viability of the Technology and its cost.	Only apply this label if considerations are either given quantitatively OR a result of a quantitative method, i.e.. model, statistical

		analysis, experimental measurement.
Storage/sequestration potentials	Information about the global and local CO ₂ – storage/capturing potentials of CDR Technologies.	<p>Only apply this label if the storage potentials are either given quantitatively OR a result of a quantitative method, , i.e.. model, statistical analysis, experimental measurement.</p> <p>Storage potentials are only assessed for CO₂/carbon.</p> <p>The capturing process involves ad/absorption of the CO₂ to a filtering material and the desorption process from this material. We include in this label all quantified stages of this process.</p>
Environmental side effects	Anything caused by deploying or introducing a CDR Technology on the (Ecological/Physical) environment which does not relate to carbon storage/removal. We want to code very inclusive on side-effects to capture as much as possible.	<p>This is a vague label to also capture qualitative effects that are mentioned in the abstracts.</p> <p>Apply this label (or the other side effect label) if results are not directly caused by the CDR technology, e.g. looking at algae in ocean fertilisation experiments are no side effects.</p>
Non-environmental side effects	Anything caused by deploying or introducing a CDR Technology on the social/economic surroundings which does not relate to carbon storage/removal. We want to code very inclusive on side-effects to capture as much as possible.	<p>This is a vague label to also capture qualitative effects that are mentioned in the abstracts.</p> <p>This includes potential economic trade-offs/ positive/negative externalities. (Quantified and non-quantified) For example fertilisation effects on crop / farm land or increased crop yield in general.</p> <p>No implicit double coding with the Economic</p>

		considerations. Apply this label (or the other side effect label) if results are not directly caused by the CDR technology, e.g. looking at algae in ocean fertilisation experiments are no side effects .
Ex-post	Ex-post analysis looks at results after they have occurred	Is given in the context of policy and model evaluations. Explicitly not used for any experimental or observational study
Ex-ante	Ex-ante analysis looks at predictions	Is given in the context of policy and model evaluations. Explicitly not used for any experimental study

Exclusion

If this category is applied, the record is judged as not relevant, based on the inclusion criteria.

It is used to catch some topics that might later be included or relevant for further investigation.

This allows us to set the limits to our living map more flexibly. Please apply an exclusion label if the study is excluded.

Bioenergy	Energy derived from any form of biomass or its metabolic by-products.	This includes all uses of biological material for heat or energy production including biofuel and burning biochar.
Enhanced Oil Recovery (EOR)	The practice of pumping CO ₂ into oil reservoirs (effectively sequestering it) to increase the oil yield.	

Short-lived CCUS	A process in which CO2 is captured and then used to produce a new product, which stores CO2 only for a short period of time.	This includes usages of CO2 produce cosmetics via algae or to produce fodder for livestock.
Solar Radiation Modification (SRM)	Sunlight would be reflected back to space to limit or reverse human-caused climate change.	Please also apply if the record is relevant based on the other technology labels applied.
Non-CDR	The paper clearly deals with a non-related topic	This is our “Catch all”- label, used if no other exclusion category applies. Meaning that the paper does not address CDR Technologies at all. If it is applied no other category needs to be applied. Fuel cells are included here if there is no relation to CDR.

References

- Buck, Holly Jean, Wim Carton, Jens Friis Lund, and Nils Markusson. 2023. “Why Residual Emissions Matter Right Now.” *Nature Climate Change* 13 (4): 351–58. <https://doi.org/10.1038/s41558-022-01592-2>.
- Fuhrman, Jay, Simone Speizer, Patrick O’Rourke, Glen P. Peters, Haewon McJeon, Seth Monteith, Laura Aldrete Lopez, and Frances M. Wang. 2024. “Ambitious Efforts on Residual Emissions Can Reduce CO2 Removal and Lower Peak Temperatures in a Net-Zero Future.” *Environmental Research Letters* 19 (6): 064012. <https://doi.org/10.1088/1748-9326/ad456d>.
- IPCC. 2022. *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, et al. Cambridge, UK and New York, NY, USA. 10.1017/9781009157926.
- Luderer, Gunnar, Zoi Vrontisi, Christoph Bertram, Oreane Y. Edelenbosch, Robert C. Pietzcker, Joeri Rogelj, Harmen Sytze De Boer, et al. 2018.

- “Residual Fossil CO₂ Emissions in 1.5–2 °C Pathways.” *Nature Climate Change* 8 (7): 626–33. <https://doi.org/10.1038/s41558-018-0198-6>.
- Minx, J.C., W.F. Lamb, M.W. Callaghan, L. Bornmann, and S. Fuss. 2017. “Fast Growing Research on Negative Emissions.” *Environmental Research Letters* 12 (3). <https://doi.org/10.1088/1748-9326/aa5ee5>.
- Nemet, Gregory F, Max W Callaghan, Felix Creutzig, Sabine Fuss, Jens Hartmann, Jérôme Hilaire, William F Lamb, Jan C Minx, Sophia Rogers, and Pete Smith. 2018. “Negative Emissions—Part 3: Innovation and Upscaling.” *Environmental Research Letters* 13 (6): 063003. <https://doi.org/10.1088/1748-9326/aabff4>.
- Schleussner, Carl-Friedrich, Gaurav Ganti, Quentin Lejeune, Biqing Zhu, Peter Pfleiderer, Ruben Prütz, Philippe Ciais, et al. 2024. “Overconfidence in Climate Overshoot.” *Nature* 634 (8033): 366–73. <https://doi.org/10.1038/s41586-024-08020-9>.
- Smith, S. M., O. Geden, G.F. Nemet, M J Gidden, W.F. Lamb, C. Powis, R. Bellamy, et al. 2023. “The State of Carbon Dioxide Removal - 1st Edition.” Available at: <https://www.stateofcdr.org>.