Document information

Full project name:	Linking Climate and Development Policies – Leveraging International Networks and Knowledge Sharing				
Grant Agreement number:	642147	**Acronym:**	CD-LINKS		
Project URL:	www.cd-links.org				
Project start date:	01/09/2015	**Project end date:**	31/08/2019	**Project duration:**	48 months

Deliverable	**Number:**	7.2	**Title:**	Data Management Plan
Work Package	**Number:**	7	**Title:**	Project Management
Date of delivery	**Planned:**	Original: 01 March 2016 Update: 28 April 2017	**Actual:**	Original D7.2: 01 March 2016 Update: 28 April 2017
Dissemination level:	☒Public ☐Restricted			

Lead institute	International Institute for Applied Systems Analysis – IIASA		
Responsible author	**Name:**	Johanna Zilliacus	**E-mail:**	zilliacu@iiasa.ac.at
Reviewer	**Name:**	Volker Krey	**E-mail:**	krey@iiasa.ac.at

Version Log

Date	Uploaded by	Description	Summary of changes
01.03 2016	IIASA	First version	n/a
28.04.2017	IIASA	Updated version for Periodic Report 1	1) For WP1, the data management of the conducted case studies was updated based on the practice undertaken for the case studies. No common database of case studies was created due to the different types of information available for each case study, and instead the institutes in charge of each case study are managing the information related to the respective case studies. The two econometric studies of the case studies relied on a publicly available (CMCC) and a commercial database (UEA), which will be archived. 2) For WP3, the stocktaking database was expanded to include the policy scenario runs of both the national and global modelling teams (Tasks 3.2 and 3.3). This way the teams did not have to re-register for a new database with their models, facilitating the uploading process.

CD-LINKS

Linking Climate and Development Policies – Leveraging International Networks and Knowledge Sharing

GA number: 642147, Funding type: RIA

Deliverable number (relative in WP)	**D7.2**
Deliverable name:	Data Management Plan
WP / WP number:	WP7
Delivery due date:	Project month 6 (01/03/2016)
Actual date of submission:	01/03/2016
Dissemination level:	Public
Lead beneficiary:	IIASA
Responsible scientist/administrator:	Volker Krey, IIASA
Estimated effort (PM):	0.25
Contributor(s):	Data product related Work Package leaders (PIK, COPPE, PBL, WU, CMCC, IIASA) provided input to this deliverable (D7.2). Further, all project partners have had the opportunity to comment and provide feedback the Data Management Plan.
Estimated effort contributor(s) (PM):	
Internal reviewer:	

Changes with respect to the DoA

No changes were made with respect to the DoA.

Dissemination and uptake

The Data Management Plan (DMP) is developed for monitoring purposes. The DMP helps to plan and implement the specific activities of the project that will result in scientific output. It also gives more detailed information on the planned activities, and is available for the whole consortium in the project's intranet for document sharing (SharePoint).

Short Summary of results (<250 words)

The CD-LINKS project will produce a number of data sets for internal as well as public use, which are summarized in the CD-LINKS Data Management Plan (DMP). The list below provides an overview of the data sets to be generated within CD-LINKS, along with the work packages that are responsible for these data sets (in parenthesis).

- *Empirical data sets for the evaluation of existing climate and development policies (WP1)*
- *Inventory of existing climate and development policies (WP2)*
- *CD-LINKS stock taking scenario database (WP3/4)*
- *CD-LINKS diagnostic scenario database (WP3)*
- *CD-LINKS national and global low-carbon and sustainable development scenario database (WP3)*
- *CD-LINKS multiple policy objective scenario database (WP4)*
- *Global map of hot spots of criticalities, challenges, policy priorities, and non-aligned incentives (WP5)*

In the data sets section of the DMP, a more detailed description of these data sets can be found, including information regarding standards and metadata, data sharing and archiving and preservation.

The CD-LINKS consortium is taking measures to ensure that the scientific research data produced within the project will satisfy the relevant criteria defined in guidelines on data management in Horizon 2020.

The DMP will be updated for the months 24 and 42 of the project, as there will be more detailed information available on the different data sets when the project advances.

Evidence of accomplishment
(report, manuscript, web-link, other)
The Data Management Plan has been uploaded to the project intranet platform for document sharing (SharePoint), and can be found as an annex to this document.

CD-LINKS

Initial Data Management Plan (DMP)

ADMIN DETAILS

Project Name: CD-LINKS
Grant Title: 642147
Funder: European Commission (Horizon 2020)
Deliverable: D7.2

Overview

The CD-LINKS project will produce a number of data sets for internal as well as public use. The list below provides an overview of the data sets to be generated within CD-LINKS, along with the work packages that are responsible for these data sets (in parenthesis).

- Empirical data sets for the evaluation of existing climate and development policies (WP1)
- Inventory of existing climate and development policies (WP2)
- CD-LINKS stock taking scenario database (WP3/4)
- CD-LINKS diagnostic scenario database (WP3)
- CD-LINKS national and global low-carbon and sustainable development scenario database (WP3)
- CD-LINKS multiple policy objective scenario database (WP4)
- CD-LINKS scenario database with focus on investment needs, financial flows, trade, cross-border effects and carbon lock-in (WP5)
- Global map of hot spots of criticalities, challenges, policy priorities, and non-aligned incentives (WP5)

In the data sets section below, a more detailed description of these data sets can be found, including information regarding standards and metadata, data sharing and archiving and preservation.

Further, the CD-LINKS consortium is taking measures to ensure that the scientific research data produced within the project will satisfy the relevant criteria defined in guidelines on data management in Horizon 2020.

1. Discoverability
   To aid discoverability of the data sets, it is aimed to obtain digital object identifiers (DOIs) where possible. In addition, the underlying scientific publications (reports and peer-reviewed journal articles) will cross-reference these data sets.

2. Accessibility

To ensure accessibility beyond the duration of the project, the use of proprietary data formats will be avoided. In addition, wherever web-interfaces will be applied to display and visualize the data in a convenient way, the underlying data will also be made accessible in numerical form to encourage re-use.

3. Assessability and intelligibility

The intention is that data sets will be made available to reviewers of the resulting publications to aid transparency in the review process.

4. Usability beyond the original purpose for which it was collected

All data sets produced in CD-LINKS are intended to be used also outside of the project. Experience in particular with existing scenario databases (e.g., AMPERE[1] and LIMITS[2] scenario database) has shown that re-analysis and use of the data beyond the original purpose is pursued heavily if provided on-line (as planned in CD-LINKS).

5. Interoperability to specific quality standards

The scenario data produced in the CD-LINKS project will use data formats and metadata standards established by the Integrated Assessment Modeling Consortium (IAMC). Applying community standards will allow easy combination with data sets from other European or international projects that rely on energy-economy and integrated assessment models. For the policy database we will add to an existing database that relies on an open source product, a semantic MediaWiki.

---

[1] https://tntcat.iiasa.ac.at/AMPEREDB
[2] https://tntcat.iiasa.ac.at/LIMITSDB

Data sets

Data set reference and name

Empirical data sets for the evaluation of existing climate and development policies (WP1)

Data set description

The data set will comprise unsystematic data for key indicators of development and climate policies analyzed in the case studies in WP1. Data used in the case studies will mostly be from third parties, i.e. from policy reports or statistical offices. In some cases also modelling data generated by the involved teams might be provided.

Standards and metadata

Data will be collected in a single spreadsheet file (Excel format), with a separate sheet for each case study. Each sheet will contain the used data sets as well as metadata, including a short description of source/origins and usage rights.

Data sharing

In face of possible limitations due to third party usage rights, the data will only be made available internally within the CD-LINKS consortium. In case no such limitations occur, the data can be made available to the public once the case studies are finished.

Archiving and preservation (including storage and backup)

The spreadsheet file with the data underlying the case studies will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

Inventory of existing climate and development policies (WP2)

Data set description

A systematic inventory of key policies in G20 countries that could significantly influence future GHG emissions will be developed. This includes the most important energy and climate policies, adaptation and development strategies, land use measures, and energy security and air pollution legislation.

The task will build on existing databases, most notable the global climate policy database with more than 2400 policies, which was developed by the NewClimate Institute[3].

National partners from several G20 countries (EU, Brazil, India, China, Russia) will contribute to establishing this inventory jointly with the coordinators of WP2, PBL and WU.

Standards and metadata

The database is programmed in Semantic MediaWiki,[4] a free, open source extension to MediaWiki, the software that powers Wikipedia. It allows user inputs and tracking, moderating and undoing of changes. Data collection will be organized in steps of primary research, quality control and approval by different people.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the policy inventory data via www.climatepolicydatabase.org. All data is available online. Aspects of the data can be made available without any restriction or only accessible to registered users.

By the end of the project, or in part earlier, the quality controlled and approved data will be made publically available via a web-interface as well as in the form of downloadable spreadsheet files.

Archiving and preservation (including storage and backup)

All data can be downloaded as Excel or CSV file at any time. This will be used to archived earlier versions of the policy database. The CSV files will can be used independently from the online database even if the online not be available any longer. Relevant CSV files of the policy database, including the CD-LINKS contribution to it, will also be kept in a 10-year compliance archive at the project coordinator (IIASA).

---

[3] http://www.climatepolicydatabase.org
[4] https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki

Data set reference and name

CD-LINKS stock taking scenario database (WP3/4)

Data set description

This data set will consolidate data from existing global and national economy-energy-environment scenarios from all participating modelling teams participating in CD-LINKS (PIK, PBL, IIASA, CMCC, ICCS, COPPE, TERI, IIM, ERI, TU, HSE, NIES, RITE, external partner PNNL) in order to improve the understanding of differences across scenarios and models, as well as to identify key challenges and data requirements for the following scenario development in CD-LINKS.

To allow for comparison across modeling teams and countries, the scenarios will be grouped into the following four categories according to their climate policy assumptions:

- *(S0) No policy baseline* scenarios (no dedicated climate policies)
- *(S1) Pre-INDC Reference policy scenarios* reflecting existing and planned policies before the INDCs
- *(S2) INDC policy scenarios* reflecting the additional effect of the INDCs
- *(S3) Deep decarbonization scenarios*, reflecting 2°C-consistent climate policies (global models), or ambitious county-level climate change mitigation scenarios (national models)
- *(S4) Sustainable development scenarios*, accounting for other sustainability dimensions (e.g., air pollution, energy security, water, food, biodiversity) in addition to climate change mitigation.

As appropriate, other scenarios, such as existing technology roadmaps from the European Commission, national governments or scenarios developed by international institutions (e.g., the International Energy Agency, IEA) will be included in this database.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[5]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in

---

[5] https://tntcat.iiasa.ac.at/AR5DB/dsd?Action=htmlpage&page=about

the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

Given that this data set is consolidating information from other existing scenario data sets to inform the model and scenario development in CD-LINKS, it is not intended to publish this data set However, reports and journal publications that are a result of this analysis within CD-LINKS will obviously be made available.

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps. The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS diagnostic scenario database (WP3)

Data set description

Database with stylized diagnostic scenarios of national and global models participating in CD-LINKS. The experimental setup of diagnostic studies is dedicated to generate model output that can be used to estimate a set of diagnostic indicators of model response to carbon pricing policies. The ultimate goal is to better understand differences in model behavior, enable "fingerprinting" of model responses, and classify models along their "fingerprints". The CD-LINKS diagnostics work will adopt the scenario protocol developed in the ADVANCE diagnostics exercise, see modeling protocol[6] for details. The scenarios will be based on models from the institutions participating in CD-LINKS WP3, i.e. the global modeling teams PBL, IIASA, PIK, CMCC, ICCS and the national teams COPPE, PIK, ICCS, TERI, IIM, ERI, TU, HSE, NIES, RITE, and the external partner PNNL. To avoid duplication of efforts, diagnostic scenarios already submitted by modeling teams to the ADVANCE diagnostics exercise will be transferred to the CD-LINKS diagnostic database.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). To facilitate exchange within CD-LINKS, the new stock-taking CD-LINKS variable template will also be used for the diagnostics exercise. In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the ADVANCE diagnostics scenario database[7]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary

---

[6]
https://tntcat.iiasa.ac.at/ADVANCEWP1DB/static/download/ADVANCE_Diagnostic_scenarios_specs_final_2014_12_12r.pdf
[7] https://tntcat.iiasa.ac.at/ADVANCEWP1DB

scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

Given that this data set to some degree will draw upon diagnostic scenarios submitted to the ADVANCE diagnostics data and is primarily intended to inform the model and scenario development in CD-LINKS, it is not planned to publish this data set as such. However, given that some modeling teams have not been participating in the ADVANCE diagnostics exercise, an effort will be made to either transfer those scenarios within the CD-LINKS diagnostics database to the ADVANCE diagnostics database or they will be made publically available separately if scenario transfer is not an option.

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps. The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS coherent national and global low-carbon development scenario database (WP3)

Data set description

Database with coherent national and global low-carbon development scenarios. The main objective of this scenario set is to integrate national and global pathways analysis into coherent visions that can function as roadmaps for decision making. Based on storylines about future international climate policy frameworks, the integrated scenario set will be developed. The scenarios will be based on models from the institutions participating in CD-LINKS WP3, i.e. the global modeling teams PBL, IIASA, PIK, CMCC, ICCS and the national teams COPPE, PIK, ICCS, TERI, IIM, ERI, TU, HSE, NIES, RITE, and the external partner PNNL.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[5]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

After the end of the project, the quality controlled and approved data will be made publically available via a web-interface as well as in the form of consolidated snapshots in csv format, similar to the IPCC AR5 scenario database (see footnote 5).

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps. The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS multiple policy objective scenario database (WP4)

Data set description

Database with national and global scenarios including quantification of multiple policy objectives. In the scenario set, the synergies and trade-offs among different policy objectives will be explored by varying the emphasis placed on individual objectives. In particular, the scenarios will systematically explore the implications of different climate targets/policies for the attainability of other sustainability objectives, including economic development, air quality and health, energy poverty and income inequality, energy security, food security, biodiversity, water availability, climate resilience and adaptation. The scenarios will be based on models from the institutions participating in CD-LINKS WP4, i.e. IIASA, TERI, PIK, PBL, CMCC, ICCS, COPPE, ERI, TU, HSE, NIES, RITE, and the external partner PNNL.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[5]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

After the end of the project, the quality controlled and approved data will be made publically available via a web-interface as well as in the form of consolidated snapshots in csv format, similar to the IPCC AR5 scenario database (see footnote 5).

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps. The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS scenario database with focus on investment needs, financial flows, trade, cross-border effects and carbon lock-in (WP5)

Data set description

Database with scenarios analyzing investment needs, financial flows, trade, cross-border effects and carbon lock-in. This study will build on the scenarios developed in WP3 and WP4, but will include additional in-depth analyses performed in Tasks 5.1 to 5.3. Given the close relationship between the scenario set developed here and in WP3/4, the possibility of integrating the results of the additional analyses into the scenario databases of WP3 and WP4 will also be explored.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[5]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

After the end of the project, the quality controlled and approved data will be made publically available via a web-interface as well as in the form of consolidated snapshots in csv format, similar to the IPCC AR5 scenario database (see footnote 5).

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps.

The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name
Global map of hot spots of criticalities, challenges, policy priorities, and non-aligned incentives (WP5)

Data set description
The data set includes indicators representing the major implementation challenges, hot spots and criticalities for guiding climate and sustainable development-related policy making in G20 countries. The data set will be visualized in an interactive, open access mapping tool, which will be designed so as to maximize its uptake and further exploitation by others. The data set underlying the mapping tool will summarize research insights across WP1, WP2 and WP5. Examples of key national indicators that are envisioned to be covered by the data set include financial needs for climate mitigation/adaptation as a share of GDP; infrastructure requirements as a share of current infrastructure expenditures; level of corruption and/or lack of credible institutions; import and export of embedded emissions; current and projected capital committed to fossil fuels; fossil fuel resources underground; and fossil fuel exports as a share of national income.

Standards and metadata
The interactive website will be built upon open-source frameworks, such as rails, symfony, django or d3.js. A sub set of the data from the other work packages, relevant to WP5 focus, will be directly imported into a local database from the website and displayed through interactive maps and graphs. Raw data (in comma separated values or json formats) will be also available for download after registration, in order to grant users full access.

Data sharing
During the project, a development version of the website will be available for testing purposes and for demonstrating the possibilities of the website as a communication tool. All partner institutions will have access to the development version. This will help to identify the key indicators that will summarize the findings from WP1, WP2 and WP5.
By the end of the project, data of controlled and approved quality will be made publically available via the website.

Archiving and preservation (including storage and backup)
During the project, data will be stored in a local database that will be dumped every week by a backup storage system at CMCC. After the end of the project, the website and tool will be maintained by the CMCC Foundation.

CD-LINKS

Data Management Plan (DMP) – Update Month 18

ADMIN DETAILS

Project Name: CD-LINKS
Grant Title: 642147
Funder: European Commission (Horizon 2020)
Deliverable: D7.2 (update)

Overview

The CD-LINKS project will produce a number of data sets for internal as well as public use. The list below provides an overview of the data sets to be generated within CD-LINKS, along with the work packages that are responsible for these data sets (in parenthesis).

- Empirical data sets for the evaluation of existing climate and development policies (WP1)
- Inventory of existing climate and development policies (WP2)
- CD-LINKS stock taking scenario database (WP3/4)
- CD-LINKS diagnostic scenario database (WP3)
- CD-LINKS national and global low-carbon and sustainable development scenario database (WP3)
- CD-LINKS multiple policy objective scenario database (WP4)
- Global map of hot spots of criticalities, challenges, policy priorities, and non-aligned incentives (WP5)

In the data sets section below, a more detailed description of these data sets can be found, including information regarding standards and metadata, data sharing and archiving and preservation.

Further, the CD-LINKS consortium is taking measures to ensure that the scientific research data produced within the project will satisfy the relevant criteria defined in guidelines on data management in Horizon 2020.

6. Discoverability
   To aid discoverability of the data sets, it is aimed to obtain digital object identifiers (DOIs) where possible. In addition, the underlying scientific publications (reports and peer-reviewed journal articles) will cross-reference these data sets.
7. Accessibility
   To ensure accessibility beyond the duration of the project, the use of proprietary data formats will be avoided. In addition, wherever web-interfaces will be applied to display and visualize the data

in a convenient way, the underlying data will also be made accessible in numerical form to encourage re-use.

8. Assessability and intelligibility

   The intention is that data sets will be made available to reviewers of the resulting publications to aid transparency in the review process.

9. Usability beyond the original purpose for which it was collected

   All data sets produced in CD-LINKS are intended to be used also outside of the project. Experience in particular with existing scenario databases (e.g., AMPERE[1] and LIMITS[2] scenario database) has shown that re-analysis and use of the data beyond the original purpose is pursued heavily if provided on-line (as planned in CD-LINKS).

10. Interoperability to specific quality standards

    The scenario data produced in the CD-LINKS project will use data formats and metadata standards established by the Integrated Assessment Modeling Consortium (IAMC). Applying community standards will allow easy combination with data sets from other European or international projects that rely on energy-economy and integrated assessment models. For the policy database we will add to an existing database that relies on an open source product, a semantic MediaWiki.

---

[1] https://tntcat.iiasa.ac.at/AMPEREDB
[2] https://tntcat.iiasa.ac.at/LIMITSDB

Data sets

Data set reference and name
Empirical data sets for the evaluation of existing climate and development policies (WP1)

Data set description
Two of the policy case studies data conducted within WP1 used econometric methods to analyze structured data sets. These two econometric case studies relied on a public available and a commercial database respectively while the other case studies reviewed reviewed documents, such as reports by ministries or government agencies, which typically provided data in forms of tables and figures.

Standards and metadata
Data is collected in a single spreadsheet file (Excel format), with a separate sheet for each of the relevant case study. Each sheet contains the used data sets as well as metadata, including a short description of source/origins and usage rights.

Data sharing
In face of possible limitations due to third party usage rights, the data will only be made available internally within the CD-LINKS consortium. In case no such limitations occur, the data can be made available to the public once the case studies are finished (currently in progress).

Archiving and preservation (including storage and backup)
The spreadsheet file with the data underlying the case studies will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

Inventory of existing climate and development policies (WP2)

Data set description

A systematic inventory of key policies in G20 countries that could significantly influence future GHG emissions has been developed. This includes the most important energy and climate policies, adaptation and development strategies, land use measures, and energy security and air pollution legislation.

The task builds on existing databases, most notable the global climate policy database with more than 2400 policies, which was developed by the NewClimate Institute[3].

National partners from several G20 countries (EU, Brazil, India, China, Russia) have contributed to establishing this inventory jointly with the coordinators of WP2, PBL and WU.

Standards and metadata

The database is programmed in Semantic MediaWiki,[4] a free, open source extension to MediaWiki, the software that powers Wikipedia. It allows user inputs and tracking, moderating and undoing of changes. Data collection will be organized in steps of primary research, quality control and approval by different people.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the policy inventory data via www.climatepolicydatabase.org which makes all data available online. Aspects of the data can be made available without any restriction or only accessible to registered users. By the end of the project, all quality controlled and approved data will be made publically available via a web-interface as well as in the form of downloadable spreadsheet files. However, a large portion of the data has already been included into the public policy inventory described above with a specific view allowing to separately look at the CD-LINKS contribution to the database[5].

Archiving and preservation (including storage and backup)

All data can be downloaded as Excel or CSV file at any time. This will be used to archived earlier versions of the policy database. The CSV files will can be used independently from the online database even if the online not be available any longer. Relevant CSV files of the policy database, including the CD-LINKS contribution to it, will also be kept in a 10-year compliance archive at the project coordinator (IIASA).

---

[3] http://www.climatepolicydatabase.org
[4] https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki
[5] http://www.climatepolicydatabase.org/index.php?title=CDlinks_policy_inventory

Data set reference and name

CD-LINKS stock-taking and fast-track scenario database (WP3/4)

Data set description

This data set consolidates data from existing global and national economy-energy-environment scenarios from all participating modelling teams participating in CD-LINKS (PIK, PBL, IIASA, CMCC, ICCS, COPPE, TERI, IIM, ERI, TU, HSE, NIES, RITE, external partner PNNL) that was collected during the first year of the project in order to improve the understanding of differences across scenarios and models, as well as to identify key challenges and data requirements for the following scenario development in CD-LINKS. To allow for comparison across modeling teams and countries, the scenarios will be grouped into the following four categories according to their climate policy assumptions:

- *(S0) No policy baseline* scenarios (no dedicated climate policies)
- *(S1) Pre-INDC Reference policy scenarios* reflecting existing and planned policies before the INDCs
- *(S2) INDC policy scenarios* reflecting the additional effect of the INDCs
- *(S3) Deep decarbonization scenarios*, reflecting 2°C-consistent climate policies (global models), or ambitious county-level climate change mitigation scenarios (national models)
- *(S4) Sustainable development scenarios*, accounting for other sustainability dimensions (e.g., air pollution, energy security, water, food, biodiversity) in addition to climate change mitigation.

In addition to the stock-taking exercise of existing scenarios, the database also hosts a set of policy scenarios that are being developed in support of publications that should feed into the IPCC Special Report on 1.5° C (so called "fast-track" scenarios within CD-LINKS). This scenario set includes global and national pathways that follow implemented (and planned) policies through 2020 or implement the Intended Nationally Determined Contributions (INDCs) goals submitted by countries to the Unitef Nations Framework Climate Change Convention (UNFCCC) under the Paris agreement and thereafter transition to CO_2 emissions budgets that are consistent with the 2 and 1.5° C targets.

Standards and metadata

The scenario data has been and is continued to be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/).

In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[6]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format are generated.

---

[6] https://tntcat.iiasa.ac.at/AR5DB/dsd?Action=htmlpage&page=about

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

The stock-taking scenarios included in this data set essentially consolidate information from other existing scenario data sets to inform the model and scenario development in CD-LINKS, and thus it is not intended to publish this part of the data set. However, the quality controlled and approved "fast-track" scenarios developed within the CD-LINKS project as an intended input to the IPCC Special Report on 1.5° C, will be made publically after the end of the project via a web-interface as well as in the form of consolidated snapshots in csv format, similar to the IPCC AR5 scenario database (see footnote 6).

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database are archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database is backed up on a daily basis in the form of database dumps.

By the end of the project, the final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS diagnostic scenario database (WP3)

Data set description

Database with stylized diagnostic scenarios of national and global models participating in CD-LINKS. The experimental setup of diagnostic studies is dedicated to generate model output that can be used to estimate a set of diagnostic indicators of model response to carbon pricing policies. The ultimate goal is to better understand differences in model behavior, enable "fingerprinting" of model responses, and classify models along their "fingerprints". The CD-LINKS diagnostics work adopts the scenario protocol developed in the ADVANCE diagnostics exercise, see modeling protocol[7] for details. The scenarios are based on models from the institutions participating in CD-LINKS WP3, i.e. the global modeling teams PBL, IIASA, PIK, CMCC, and the national teams COPPE, PIK, ICCS, TERI, IIM, ERI, TU, HSE, NIES, RITE, and the external partner PNNL. To avoid duplication of efforts, diagnostic scenarios already submitted by modeling teams to the ADVANCE diagnostics exercise were be transferred to the CD-LINKS diagnostic database.

Standards and metadata

The scenario data has been collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). To facilitate exchange within CD-LINKS, the initial stock-taking CD-LINKS variable template was also used for the diagnostics exercise. In addition to the scenario data itself, also metadata about the scenarios and models were collected, using the same format as in the ADVANCE diagnostics scenario database[8]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format are generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data are available to partner institutions involved in the scenario analysis. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary

---

[7]

https://tntcat.iiasa.ac.at/ADVANCEWP1DB/static/download/ADVANCE_Diagnostic_scenarios_specs_final_2014_12_12r.pdf

[8] https://tntcat.iiasa.ac.at/ADVANCEWP1DB

scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

Given that this data set to some degree will draw upon diagnostic scenarios submitted to the ADVANCE diagnostics data and is primarily intended to inform the model and scenario development in CD-LINKS, it is not planned to publish this data set as such. However, given that some modeling teams have not been participating in the ADVANCE diagnostics exercise, an effort will be made to either transfer those scenarios within the CD-LINKS diagnostics database to the ADVANCE diagnostics database or they will be made publically available separately if scenario transfer is not an option.

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database are archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database is backed up on a daily basis in the form of database dumps.

By the end of the project, the final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS coherent national and global low-carbon development scenario database (WP3)

Data set description

Database with coherent national and global low-carbon development scenarios. The main objective of this scenario set is to integrate national and global pathways analysis into coherent visions that can function as roadmaps for decision making. Based on storylines about future international climate policy frameworks, the integrated scenario set will be developed. The scenarios will be based on models from the institutions participating in CD-LINKS WP3, i.e. the global modeling teams PBL, IIASA, PIK, CMCC, and the national teams COPPE, PIK, ICCS, TERI, IIM, ERI, TU, HSE, NIES, RITE, and the external partner PNNL.

It should be noted that the "fast-tack" scenarios included in the CD-LINKS stock-taking and fast-track database above, were originally intended to be collected in this database. However, to speed up the process of scenario collection on the timeline relevant for the IPCC Special Report on 1.5° C, it was decided to add them to the existing stock-taking database.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[6]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

After the end of the project, the quality controlled and approved data will be made publically available via a web-interface as well as in the form of consolidated snapshots in csv format, similar to the IPCC AR5 scenario database (see footnote 6).

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps. The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name

CD-LINKS multiple policy objective scenario database (WP4)

Data set description

Database with national and global scenarios including quantification of multiple policy objectives. In the scenario set, the synergies and trade-offs among different policy objectives will be explored by varying the emphasis placed on individual objectives. In particular, the scenarios will systematically explore the implications of different climate targets for the attainability of other sustainability objectives, including economic development, air quality and health, energy poverty and income inequality, energy security, food security, biodiversity, water availability, climate resilience and adaptation. The scenarios will be based on models from the institutions participating in CD-LINKS WP4, i.e. IIASA, TERI, PIK, PBL, CMCC, ICCS, COPPE, ERI, TU, HSE, NIES, RITE, and the external partner PNNL.

Standards and metadata

The scenario data will be collected using the data template defined by the Integrated Assessment Modeling Consortium (IAMC, see http://www.globalchange.umd.edu/iamc/scientific-working-groups/data-protocols-and-management/iamc-time-series-data-template/). In addition to the scenario data itself, also metadata about the scenarios and models will be collected, using the same format as in the IPCC AR5 scenario database[6]. For the purpose of analysis, consolidated snapshots of scenario data in comma separated value (csv) file format will be generated.

Data sharing

During the duration of the CD-LINKS project, all partner institutions will have access to the scenario data via a web-interface. In addition, for the purpose of convenient scenario data analysis, consolidated snapshots of scenario data will be available to partner institutions. Therefore, when submitting data to a CD-LINKS scenario database, modeling teams agree to the internal use of their scenario data by all CD-LINKS partner institutions. The scenario data shall not be distributed to people that are not a partner in the CD-LINKS consortium. Credentials of the web-database need to be treated with care and shall not be shared with other people.

Individual modeling teams participating in the CD-LINKS project shall retain control of their preliminary scenario data with regard to external use outside the CD-LINKS project. External use of preliminary scenario data by other partners than the modeling team itself is not allowed without explicit permission from the modeling team.

After the end of the project, the quality controlled and approved data will be made publically available via a web-interface as well as in the form of consolidated snapshots in csv format, similar to the IPCC AR5 scenario database (see footnote 6).

Archiving and preservation (including storage and backup)

During the CD-LINKS project, all model registration forms and scenario data templates, submitted to the scenario database will be archived as individual files (Excel spreadsheets). In addition, the imported data in the underlying ORACLE database will be backed up on a daily basis in the form of database dumps. The final scenario database will be archived using the consolidated snapshot in csv format as well as a full database dump using the SIARD (Software Independent Archiving of Relational Databases) standard. The resulting files will be kept in a 10-year compliance archive at the project coordinator (IIASA).

Data set reference and name
Global map of hot spots of criticalities, challenges, policy priorities, and non-aligned incentives (WP5)

Data set description
The data set includes indicators representing the major implementation challenges, hot spots and criticalities for guiding climate and sustainable development-related policy making in G20 countries. The data set will be visualized in an interactive, open access mapping tool, which will be designed so as to maximize its uptake and further exploitation by others. The data set underlying the mapping tool will summarize research insights across WP1, WP2 and WP5. Examples of key national indicators that are envisioned to be covered by the data set include financial needs for climate mitigation/adaptation as a share of GDP; infrastructure requirements as a share of current infrastructure expenditures; level of corruption and/or lack of credible institutions; import and export of embedded emissions; current and projected capital committed to fossil fuels; fossil fuel resources underground; and fossil fuel exports as a share of national income.

Standards and metadata
The interactive website will be built upon open-source frameworks, such as rails, symfony, django or d3.js. A sub set of the data from the other work packages, relevant to WP5 focus, will be directly imported into a local database from the website and displayed through interactive maps and graphs. Raw data (in comma separated values or json formats) will be also available for download after registration, in order to grant users full access.

Data sharing
During the project, a development version of the website will be available for testing purposes and for demonstrating the possibilities of the website as a communication tool. All partner institutions will have access to the development version. This will help to identify the key indicators that will summarize the findings from WP1, WP2 and WP5.
By the end of the project, data of controlled and approved quality will be made publically available via the website.

Archiving and preservation (including storage and backup)
During the project, data will be stored in a local database that will be dumped every week by a backup storage system at CMCC. After the end of the project, the website and tool will be maintained by the CMCC Foundation.