European Commission

JRC TECHNICAL REPORTS

Modelling Inventory and Knowledge Management System of the European Commission (MIDAS)

Ostlaender, N., Acs, S., Listorti, G., Hardy, M., Ghirimoldi, G., Hradec, J. Smits, P.

2019

EUR 29729 EN

EU Science Hub
https://ec.europa.eu/jrc

JRC116242

EUR 29729 EN

PDF ISBN 978-92-76-02852-9 ISSN 1831-9424 doi:10.2760/900056

Luxembourg: Publications Office of the European Union, 2019

How to cite this report: Ostlaender, N., Acs, S., Listorti, G., Hardy, M., Ghirimoldi, G., Hradec, J. and Smits, P., *Modelling Inventory and Knowledge Management System of the European Commission (MIDAS)*, EUR 29729 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-02852-9, doi:10.2760/900056, JRC116242.

Contents

Acknowledgements

The *Modelling Inventory and Knowledge Management System of the European Commission* (MIDAS) is a Commission-wide activity that would have been impossible to achieve without the input of the *European Commission's Community of Practice (CoP) on Modelling*, and the members of the *Interservice Group on Modelling* (ISG-MOD). We would therefore like to thank the European Commission's community of modellers, as well as the modellers and model users in other services of the European Commission for providing the model descriptions as the central part of the MIDAS knowledge base; the members of the ISG-MOD for their input and feedback concerning the design of the MIDAS information system; the entire CC-MOD team and our external advisor Leen Hordijk for feedback and discussions on the subject.

Authors

Nicole Ostlaender, Szvetlana Acs, Giulia Listorti, Matthew Hardy, Gabriele Ghirimoldi, Jiri Hradec, Paul Smits

Abstract

The *Modelling Inventory and Knowledge Management System of the European Commission* (MIDAS) is a Commission-wide knowledge management tool for modelling, enabling enhanced transparency and traceability of models in use for EC policy making. It forms an integral part of the Competence Centre on Modelling (CC-MOD) of the Joint Research Centre of the European Commission (JRC). This document describes MIDAS, by providing a bird's-eye view of the MIDAS content, architecture, and functionality, and identifying the benefits of the system for the organisation and in the context of the Better Regulation Agenda.

1 Introduction

The European Commission (EC) is committed to a transparent and sound use of evidence for designing and evaluating EU policy throughout the policy cycle. This approach is reflected in the main regulatory framework of the EU, the EC Better Regulation policy. The Better Regulation Agenda, adopted in 2015 (European Commission 2015), is welcomed as a step forward for regulatory reform (Broughel 2015; Radaelli 2018; Renda 2015).

The *Better Regulation Guidelines*, that provide concrete guidance on how to implement the *Better Regulation Agenda*, recommend quantifying costs and benefits of policy reform to the extent possible (European Commission 2017). In this respect, we note that the use of simulation models has become increasingly important to support EU policy making. A model can be defined as an analytical representation or quantification of a real-world system, used to make projections or to assess the behaviour of the system under specified conditions. Models are increasingly used to assess the environmental, economic, and social impacts of policies.

The *European Commission's Competence Centre on Modelling (CC-MOD)* promotes a responsible, coherent and transparent use of modelling to underpin the evidence base for EU policies.

Maintaining an overview of ongoing modelling activities, by documenting the models and model combinations in use across the EC, is an elementary first step for a more transparent and coherent use of models in the policy cycle. It is, however, also a major challenge: for example, in the EC more than 150 models are currently being used. The list of domains ranges from greenhouse gas emissions, energy consumption and economy, to agriculture and structural integrity assessment, to name but a few. In addition, the majority of these models are run in combination with other models. Thus, they form networks of interaction, together with the related input datasets and assumptions. Capturing this knowledge, and communicating it in an understandable manner, will not only foster reproducibility and transparency of model use, but enable collaborative, interdisciplinary research that serves cross-policy issues.

In order to address these challenges, CC-MOD is responsible for the development and management of MIDAS, the Modelling Inventory and Knowledge Management System of the European Commission. This system today is used by the entire EC, containing descriptions of models previously or currently in use by the EC in support of the policy cycle. Since 2019, parts of the system have been opened to the European Parliament (EP).

This publication aims at providing a comprehensive overview of MIDAS. First, we present MIDAS in the context of CC-MOD activities (Chapter 2), and then outline its history (Chapter 3) and main approach (Chapter 4). This is followed by a brief introduction to the MIDAS content (Chapter 5), a technical chapter on the MIDAS high level architecture (Chapter 6), as well as an overview of its main functionalities (Chapter 7). We conclude by identifying the benefits of the system for the organisation and in the context of the Better Regulation Agenda (Chapter 8).

2 Background: The Competence Centre on Modelling

The *European Commission's Competence Centre on Modelling* (CC-MOD)([1]) of the JRC, launched in 2017, directly contributes to the *Better Regulation Agenda*, the *Interinstitutional Agreement on Better Law Making*, and the Communication on *Data, Information and Knowledge Management at the European Commission* (European Commission 2016). CC-MOD promotes a responsible, coherent and transparent use of modelling to underpin the evidence base for EU policies, by pooling the EC's competencies and best practices in developing and using models.

Starting with *MIDAS, the Commission-wide modelling inventory* (which will be extensively addressed in the next section), CC-MOD supports a proper documentation, use, and reuse of models. In addition, CC-MOD helps identifying common approaches to quality and transparency of model use. In a context characterised by increasing overlap between different policy areas, CC-MOD promotes the use of multi-disciplinary and integrated modelling approaches.

CC-MOD established and manages the *Community of Practice (CoP) on Modelling* as the forum for the exchange of modelling-related knowledge and best practices in support of the EU policy cycle. It brings together modellers and policymakers to:

- promote a responsible and coherent use of simulation models, underpinning the evidence base for EU policies;
- facilitate collaboration and dialogue both between the modelling teams and with policymakers;
- act as a Think Tank for modelling related issues;
- offer a point of contact for the users of modelling services, as well as a link to other modelling-related networks.

The CoP also acts as a facilitator for the organisation of working groups, seminars and trainings. The *1st Gathering of the CoP on modelling* took place in December 2018, while a scientific open EU conference on modelling for policy support is scheduled to take place in November 2019. The nucleus and steering group of the activities of CC-MOD and of the CoP on modelling is the *Interservice Group on Modelling* (ISG-MOD), which brings together EC Services involved or interested in modelling activities.

CC-MOD aims at improving *quality, transparency and sensitivity analysis of models* used for EU policy making. This involves the organisation of peer-reviews by external experts of the model quality and transparency. In addition, CC-MOD also provides services and tools related to sensitivity analysis and uncertainty quantification. These tools are increasingly employed worldwide in any field of science for the assessment, validation and verification of simulation models and computer codes in the presence of uncertainty. CC-MOD contributes to the methodological development and the diffusion of the discipline worldwide, by supporting and training interested modelling teams in the setting up and execution of uncertainty and sensitivity analyses. CC-MOD also further provides ad-hoc support to the policy *Directorate-Generals of the European Commission* (DGs) on model quality assurance.

CC-MOD also aims at increasing the *transparency and coherence of baseline scenarios* used for policy modelling at the EC. In close collaboration with the modelling teams and the relevant policy DGs, CC-MOD has mapped and compared across models the baseline scenarios currently used in the modelling in support for policy. Based on these insights, a number of recommendations will be put forward to facilitate the documentation of baseline assumption, increasing transparency and improving the quantification of baseline assumptions in EU impact assessments. The latter also includes efforts to further integrate foresight methodologies and baseline development.

([1]) Information on the Competence Centre on Modelling is available at https://ec.europa.eu/jrc/en/modelling

The Competence Centre offers training on *Social Multi-Criteria Evaluation of Policy Options* (SMCE) and ad-hoc support to policy DGs for impact assessments. SMCE builds on formal modelling techniques serving the purposes of decision and policy making in the policy formulation step of the policy cycle. SMCE is useful for helping the EC to integrate a plurality of technical aspects and social views into its impact assessment in a coherent and transparent manner. SMCE can provide a methodology which is:

— inter/multi-disciplinary, since the various criterion scores can assess a wide range of impacts (for example, by using results of economic, environmental, energy, and other simulation models);

— participatory, since fairness in the policy process is seen as an ethical obligation to take a plurality of social values, perspectives and interests into account;

— transparent, since all criteria are presented in their original form without any transformations in money, energy or whatever common measurement rod.

Mathematical approaches in SMCE furthermore allow a consistent aggregation of the diverse information.

Finally, CC-MOD promotes good practices in data management and facilitates the use of the EU Open Data Portal and the JRC Data Catalogue by modelling teams.

3 The MIDAS History

MIDAS ([2]) is the Corporate Modelling Inventory and Knowledge Management System of the EC. It is situated on the EC Network and accessible to EC staff. MIDAS acts as an element implementing the EC knowledge sharing strategy, by offering a platform where data, models, scientific publications and policy actions can be easily correlated, and where these connections can be browsed and better understood.

Like Rome, also *MIDAS wasn't built in a day*, and followed an evolution over several years as outlined below.

MIDAS was first launched in 2013 as an inventory of models run by the JRC to directly or indirectly support EU policies. At the time, MIDAS was built to enhance understanding of models by describing their context (how they are made, what they support, etc.), to facilitate sharing and preservation of knowledge across domains, disciplines, and time, and to enhance transparency and the consistency of model use by the JRC for better policy making (Ostlaender *et al*. 2015).

After two successful first years of use, in 2015 the *High Level Reflection Group of Information Management* expressed the need to extend MIDAS to the entire EC and to develop it into a corporate tool ([3]). In the following years, under the lead of a dedicated interservice group of the EC, and within the activities of CC-MOD, the scope and audience of MIDAS was thus extended to include models used by any EC services to support the policy cycle. This now also included models run by external organisations, where the EC used the results e.g. in the context of an impact assessment.

MIDAS was officially launched as an EC tool during the launch of CC-MOD on 26 October 2017 in Brussels, and can now be used by all EC services, containing descriptions of models previously or currently in use by the EC in support of the policy cycle.

Starting in 2017, MIDAS is integrated in the workflow for EC impact assessments (see Box 1). *The Better Regulation Toolbox version 2015* (European Commission 2015[2]) ([4]) already asked to describe any model used in a mandatory annex attached to the impact assessment report. The 2017 revision of the Toolbox (European Commission 2017[2]) goes one step further, and requests that any model used in an impact assessment should be described in MIDAS. This makes MIDAS an important corporate tool to use, reuse and document models in a proper way, leading to the propagation of sound methodology underpinning the EC's *Better Regulation Agenda* and potentially to significant efficiency gains in terms of financial and personnel outlays.

Box 1. Impact assessments (IA)

According to the Better Regulation guidelines, the 'impact assessment process is about gathering and analysing evidence to support policymaking' […](European Commission 2017). IAs refer to the ex ante phase of the policy cycle and are in use since 2002. With the *Better Regulation Agenda* (European Commission 2015), their use has been further extended and consolidated (EPRS 2015).

([2]) MIDAS can be accessed by EC services at http://midas.jrc.cec.eu.int. From 2019 onwards parts of the system are open to the European Parliamentary Research Service under the umbrella of the *Interinstitutional Agreement on Better Law-Making* (European Parliament, the Council of the European Union and the European Commission (2016). This version is accessible to all EC services and the Parliament at https://webgate.ec.testa.eu/midas-ii/. A version of MIDAS accessible to the general public is currently under discussion.

([3]) Ares(2015)2475458, Reflection paper on information management

([4]) *The Better Regulation Toolbox version 2015* (European Commission 2015[2]) complements the *Better Regulation Guidelines version 2015* (European Commission 2015[1])

In 2018 CC-MOD started the preparations to open parts of the system to the *European Parliamentary Research Service* (EPRS) under the umbrella of the *Interinstitutional Agreement on Better Law-Making* (European Parliament, the Council of the European Union and the European Commission (2016)), to support the EPRS during their evaluation of IAs prepared by the EC. This interinstitutional version of MIDAS was launched on the 7th February 2019 during the *Science Meets Parliament* Event.

4 The MIDAS Approach

When formulating the MIDAS vision and scope back in 2012, we were aware that this was not just about designing a technical solution. Instead, great institutional, cultural and resource challenges lay ahead of us: scientists and policy makers had to invest time and resources to share their knowledge in an understandable manner and across different domains, with us and with their peers, and be ready to maintain the information updated as long as the models are in use.

In the following sections we will provide a brief overview how we addressed the various challenges by the means of governance and user involvement (section 4.1), quality and consistency of data (section 4.2), and analytics and data representation (section 4.3).

4.1 Governance and user involvement

The institutional challenges have been tackled through a solid governance structure. First established in 2012, this structure changed over time following the extension of the system to its new audiences.

Today, MIDAS is governed by the Interservice Group on Modelling (ISG-MOD). This group, which meets around three times a year, is responsible for strategic decisions and high level user requirements. Some examples of decisions taken in this context are the role of MIDAS in the Better Regulation Toolbox, the opening of MIDAS to the EPRS, and a possible future opening to the public in general. The collaboration with the EPRS is managed under the umbrella of the Interinstitutional Agreement on Better Law-Making. A regular exchange with related *Better Regulation Bodies* like the *Impact Assessment Working Group* (IAWG) and the *Regulatory Scrutiny Board* (RSB) is also ensured.

To work on the cultural challenge of sharing a common understanding of cross-domain issues, the continuous involvement of the community of modellers and policy makers is fostered through the *CoP on Modelling*, and its related workshops and seminars. A particular focus is put on understanding barriers to and incentives for creating and sharing knowledge, and to maximise the relevance for the users.

4.2 High quality, consistent and updated data

In order to be of use for modellers and policy analysts, the information in MIDAS has to be relevant for the users, of high quality, and written in such a way that both experts and non-experts can benefit from it. This has been achieved thanks to the key principles presented below:

Capture the relevant: All information collected in MIDAS is based on user needs captured through interviews with modellers and policy makers, through questionnaires and discussions, and through analysis of use cases and other existing Knowledge Management Systems dealing with similar domains (see Ostlaender et al 2015). To enhance readability for experts and non-experts, we developed metadata profiles that distinguish between elements for overview information, and elements where technical details are described, while we developed best practices to guide users on how to write easy-to-read and relevant model descriptions.

Update & check regularly: The database is always open for updates, but dedicated updating campaigns are run at least once a year to ensure that the content remains up-to-date. During these campaigns a quality assurance process is in place where every model description is checked for completeness, consistency, and style. The updating campaigns have been recently supported by the use of a dedicated word embedding neural network SeTA (see Box 2, and Hradec et al 2019). This network is used for completeness checks to identify the use of models in impact assessments (see Chapter 5.5, and *Acs et al 2019*).

Box 2. Semantic Text Analysis Tool SeTA

The Semantic Text Analysis Tool SeTA is an Artificial Intelligence (AI) digital assistant to policy analysts in the form of a web application. It has been developed by the JRC based on a word embedding neural network, which has been trained on the European legislation and technical/scientific reports in order to identify similarity patterns between terms. Terms are single words like 'fairness' or combination of words like 'impact assessment' or model acronyms like 'GEM-E3' and 'PRIMES'. The neural network learns, in an unsupervised manner, the relations between these terms in a sentence, within a moving window, which means that it learns how the context in which the term is used can lead to its meaning, both in terms of synonyms and area of application or context in which it is used.

This particular neural network takes into consideration approximately 100 million sentences with 3.2 million unique words, and 700.000 phrases from 500.000 documents coming from the following sources: EU Bookshop, EUR-Lex, CORDIS, EU OpenDataPortal and JRC PUBSY. The documents cover a wide range of time (from 1956 until today), with an ever increasing volume per year. This was used to form a language model, in this case the model of the European Commission.

The system including various use cases is described in detail in (Hradec et.al. 2019).

Retrieve information stored and maintained elsewhere: We wanted to ensure that the effort for the policy analysts and modellers who feed MIDAS with information is minimised, while the collected information is as precise and up-to-date as possible. To achieve this MIDAS was designed to be placed in a *Service Oriented Architecture* (SOA), to make the most of existing EC knowledge infrastructure and other relevant information providers, in order to retrieve information that is already stored and maintained elsewhere (see Chapter 5.5).

Use permanent identifiers for consistency and data access: Consistency of the information provided in MIDAS is also considered a major benefit. This includes the ability to identify that two models are related to the same entity, i.e. the same dataset, the same publication, or the same person. This is much easier said than done, and requires the use of *Permanent Identifiers* (PI) (see Box 3), which allow to universally identify the resource, independent of specific storage or access mechanism, or current ownership. Perhaps the best-known example of a PI is the *Digital Object Identifier* (DOI). Wherever possible we use PIs to identify the related resources and access them. For more information see Chapter 5.5.

Box 3. Persistent Identifiers

PIs are identifiers which reliably refer to resources regardless of those resources' physical location or current ownership (Tonkin 2008). Two characteristics are essential for PIs: ownership persistence and unambiguity (GBIF 2011). Persistence ensures that an identifier is permanently assigned to a resource. Unambiguity ensures that a resource can be uniquely identified. A PI only becomes useful if information about the resource it represents can be easily retrieved. For this, resolvable PIs can be used. In these cases, a resolver provides the mapping from the PI to the location of the corresponding digital object or related metadata.

Perhaps the best-known example of a PI is the DOI, which is assigned for identification of traditional publications using a well-established system. The DOI consists of a prefix which identifies the DOI registry and a suffix, a local identifier. In this way DOIs are easily resolvable by the DOI Proxy Server System. Based on DOIs, the CrossRef association (http://www.crossref.org/) enables the effective tracking and linking of citation, for enhanced access and transparency in scientific publications.

(cited from Ostlaender et al 2015)

4.3 Analytics & Data Representation

MIDAS puts all models into their scientific and policy context, by combining the knowledge that various modellers and model users in the EC have about a model in a single place, where it can be browsed and visualised.

Using data visualisation techniques helps to communicate the resulting complex network to a non-technical audience, revealing the bigger picture and allows users to identify patterns about model use they might not have been aware of. In recent years we added tools that allow users to generate their own graphs answering user-specific questions, e.g. '*Which models run by the JRC can be used for Climate Mitigation?*' or '*Which Impact Assessments used the EU Reference Scenario 2016?*' (see more on this in Chapter 7.1.1)

Report generators have been put in place to allow users to generate word documents on demand, containing always the latest information available in the knowledge base. Different purpose reports provide a modeller's or a policy maker's perspective on the available information.

5 What's in: An overview

MIDAS contains a description of models in use by the EC, which directly or indirectly support the policy cycle, independently of whether they were developed by the EC or by third parties. In more detail, a model should respect one of the following:

— It has policy relevance. This includes notably models used in IAs or policy relevant studies in any phase of the policy cycle (policy anticipation, formulation, implementation and evaluation) ([5]), but also models developed by the EC to support Member States or international organisations.

— It has scientific relevance in the interest of the EC.

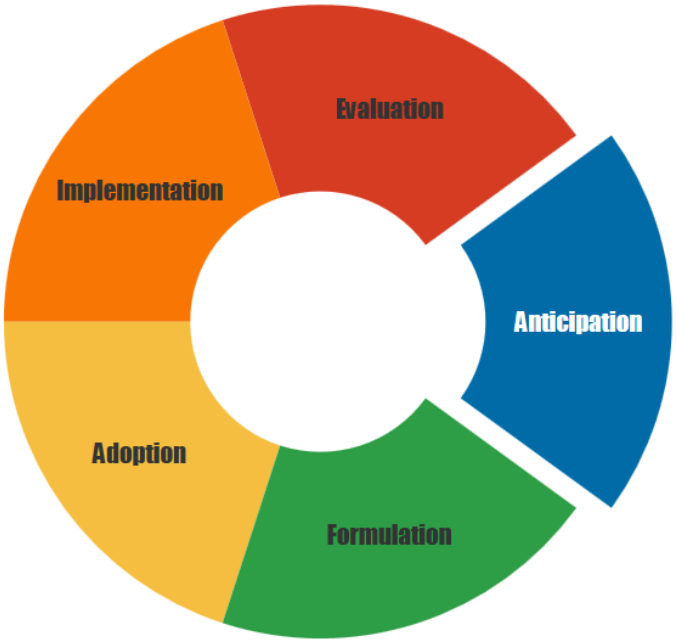

Figure 1. The policy making cycle, Source: Better Regulation, adapted for the use in MIDAS.

In addition, it describes how each model can be or has been used in the policy making context. In particular this covers the link to: other related models and datasets, policy relevant studies and impact assessments and related people (modellers and model users).

5.1 Models

The 'models' included in MIDAS are classified in one of the following three categories:

— **model**: analytical representation or quantification of a real-world system. A model can be used to make projections or to test the behaviour of the real-world system under certain conditions, like the presence or absence of a policy.

— **toolbox**: programming or user environment that provides a set of tools, for example implementations of mathematical equations that can be used to build specific models. The toolbox itself is more generic and multi-purpose, but the tools are often designed to be used in a certain domain like finance, fisheries, or land cover.

([5]) For IAs, models are included only if they contributed at least twice to IAs in the period 2005 – May 2017, and at least once since June 2017, regardless whether it was run internally by the EC or externally for the EC.

— **platform**: environment where models are run operationally for a certain (policy) purpose, often in an integrated manner, using a predefined configuration or a reference database.

In the remainder of the text, the term 'models' can refer to any of these three categories. Currently in MIDAS there are 230 entries, of which 201 are models, 19 toolboxes and 10 platforms. At the time of writing this report, 161 of these are actively in use.

A considerable set of metadata elements is collected:

— **Overview**: This includes general descriptive elements like the main purpose of the model, a summary of the models description, keywords, and model type, but also information on ownership and licensing.

— **Details**: Details about model/toolbox structure and approach, inputs and parametrisation, and outputs.

— **Quality, reliability & transparency**: Information on the quality, reliability and transparency of the model in the structured way. We are asking for example if the model has undergone sensitivity analysis, or if it has undergone external peer review. Transparency covers questions on whether equations (and database if applicable) are transparently documented and whether these documents are available to the general public.

— **Policy relevance and intended role in the policy cycle**: This describes the *potential use* of the model. It includes information on how the model can be used to assess the impact types listed in the Better Regulation Toolbox.

— **Use of the model by the EC**: This describes the *actual use* of the models, listing impact assessments and policy relevant studies (see below) the model contributed to.

— **Who runs it, why and how**: Details about the EC Departments and external institutes that are developing or actively running the model on behalf of the EC.

— **References**: Scientific articles and documents describing the model and its use.

MIDAS is work in progress and constantly extended. Also, if a model is no longer actively run by or on behalf of the EC, it is marked as 'no longer in use', but the information remains accessible. In this way MIDAS also acts as a corporate memory.

5.2 Data

Models use and produce data. Data thus form an integral part of the daily work of a modeller. MIDAS allows modellers to link models to the datasets that the models can use or require as inputs, or that they produce as outputs. MIDAS currently contains 615 datasets. The most represented are ESTAT, JRC, OECD, EEA and FAO datasets.

MIDAS offers the possibility to link datasets to models. The following information is provided:

— **Information about the data themselves**: This includes the dataset title, the Publisher (origin), and the access to the catalogue or website where the metadata is hosted and the data can be accessed ([6]).

— **Information about how models are linked to the datasets:** In the model descriptions, users can link the model to relevant datasets, by describing whether they are used as model inputs, or produced as model outputs.

([6]) Datasets are well documented in existing catalogues, therefore, they are an entity for which the majority of information is harvested from or kept at existing sources. More information in this can be found in the section describing the editor.

5.3 Studies

In MIDAS, 'studies' are analytical exercises that use modelling results to answer policy-relevant questions. The models might be run specifically for a study, or the study might use results from existing model runs, which are then interpreted or further processed for the new context.

Adding studies to MIDAS serves the purpose of increasing transparency on how (and in which combinations) models are used to answer policy questions. Currently MIDAS includes ongoing or completed policy relevant studies. An indicator for what is considered relevant is e.g. if the study was asked for by a policy DG and/or if the study was used in the policy context (e.g. an ex ante IA or an ex post evaluation).

Metadata elements are collected along two different axes:

- **Information about the study itself**: This includes a high level summary of the study, start and end-date, supported EU institutions, the policy question the study aims to answer, and links to documents describing the exercise and its results.
- **Information about the models contributing to the study**: All models that contributed to the study (including who has been running them), and ideally the year of the run. In addition, specific information on the contribution of each model to the study is also provided.

Studies are also linked to those IAs which use their results. If results from at least one model from a certain study were used in support of a certain IA, the link is established and the role the model played is added in the IA description as well.

Encoding studies adds transparency to how models are used to answer policy questions (and in which combinations), and foster traceability and re-use of model results.

5.4 Impact Assessments

MIDAS includes ongoing or completed ex ante impact assessments carried out in the context of the EU policy cycle, where modelling results were used for any of the following aspects: problem definition, evaluation of existing policy, baseline scenarios and assessment of policy options (for more details see also Acs at al 2019).

At the time of writing this report, there are 151 Impact Assessments registered in MIDAS, dating back to 2005, which made use of results coming from 74 different models. The Impact Assessments cover 51 different subjects, with Environment, Transport and Energy being the most prominent ones (note that IAs related to climate change belong to the Environment subject area). Information about the subjects is taken from EUR-Lex, where each impact assessment is annotated with a 'subject matter' ([7]).

Similar to the studies, MIDAS collects information about IAs along two different axes:

- **Information about the IA itself**: This includes the IA title, related legal documents, responsible EC Services, year of publication of the IA and more ([8]).
- **Information about the models contributing to the IA:** This includes all models that contributed to this impact assessment in the context of the problem definition, the evaluation of existing policies, the baseline scenario, or the assessment of policy options (i.e. role of the model in the IA). In addition, it also contains detailed information notably on the organisation that was running the model.

In MIDAS, in IAs, there are two ways to add the information on related models:

([7]) More information about 'subject matter' can be found at https://publications.europa.eu/en/web/eu-vocabularies/at-dataset/-/resource/dataset/subject-matter

([8]) Impact assessments are well documented in both the Inception Roadmap of the SG and the EUR-Lex system of the OP. Therefore a large number of impact assessments are harvested from these sources.

— by describing the model contributions directly, or

— by referring to a study that describes the modelling exercise used by the IA (i.e., establish a link to the relevant study in MIDAS, select the relevant models that contributed from the study and identify their role).

The latter is particularly important for studies that are frequently re-used in various impact assessments, such as the *EU reference scenario 2016 Energy, transport and GHG emissions: trends to 2050*.

5.5 Modellers and model users

MIDAS also connects people. Currently, in MIDAS there are more than 260 contact points related to models. MIDAS currently has two main user groups:

— **Modellers**, who describe their models in MIDAS and analyse them by using the MIDAS tools. The model descriptions that the input providers deliver are the core of MIDAS.

— **Model users and policy makers** (e.g. staff from DGs, or the EPRS), who are looking for models in order to assess the use of models for impact assessment and other policy support, and to access related datasets, model descriptions and documents.

MIDAS also forms a good base for initial discussions, workshops and conferences related to modelling within the EC.

6 MIDAS high level architecture

In this chapter we describe the high level architecture of the MIDAS application. This includes an overview of the modular architecture of the system, the embedding of MIDAS in a Service Oriented Architecture (SOA), and interdependencies with other *Knowledge Management Systems*.

6.1 Modular architecture

MIDAS is conceptually built following a SOA in order to provide a flexible, scalable and modular application. MIDAS mainly focuses on services in an interoperability vision and approach.

It includes:

— **Front End Infrastructure**: The *Front End Infrastructure*, i.e. the end user interface and the model metadata cache, consists of a two-tier application layer. The first layer follows a conventional web application approach and is built on top of JavaScript technologies ([9]). This interface includes MIDAS visualisation engines (developed with d3.js libraries ([10])). The second layer is a standalone PHP ([11]) block service layer providing an interface between the MIDAS database and the metadata cache.

— **Back End Infrastructure**: The *Back End Infrastructure* of MIDAS is a multi-tier PHP application built on top of Zend architectures ([12]). The MIDAS back end infrastructure relies on 3 main components: A multi-tier *Model-View-Controller* (MVC) ([13]) PHP application built on top of a Zend Framework and implementing the abstraction layer and data mapping patterns by *Object-Relational Mapping* (ORM) ([14]) technologies.

([9]) https://it.wikipedia.org/wiki/JavaScript
([10]) https://d3js.org/
([11]) http://php.net/
([12]) http://www.zend.com/
([13]) https://it.wikipedia.org/wiki/Model-view-controller
([14]) https://it.wikipedia.org/wiki/Object-relational_mapping

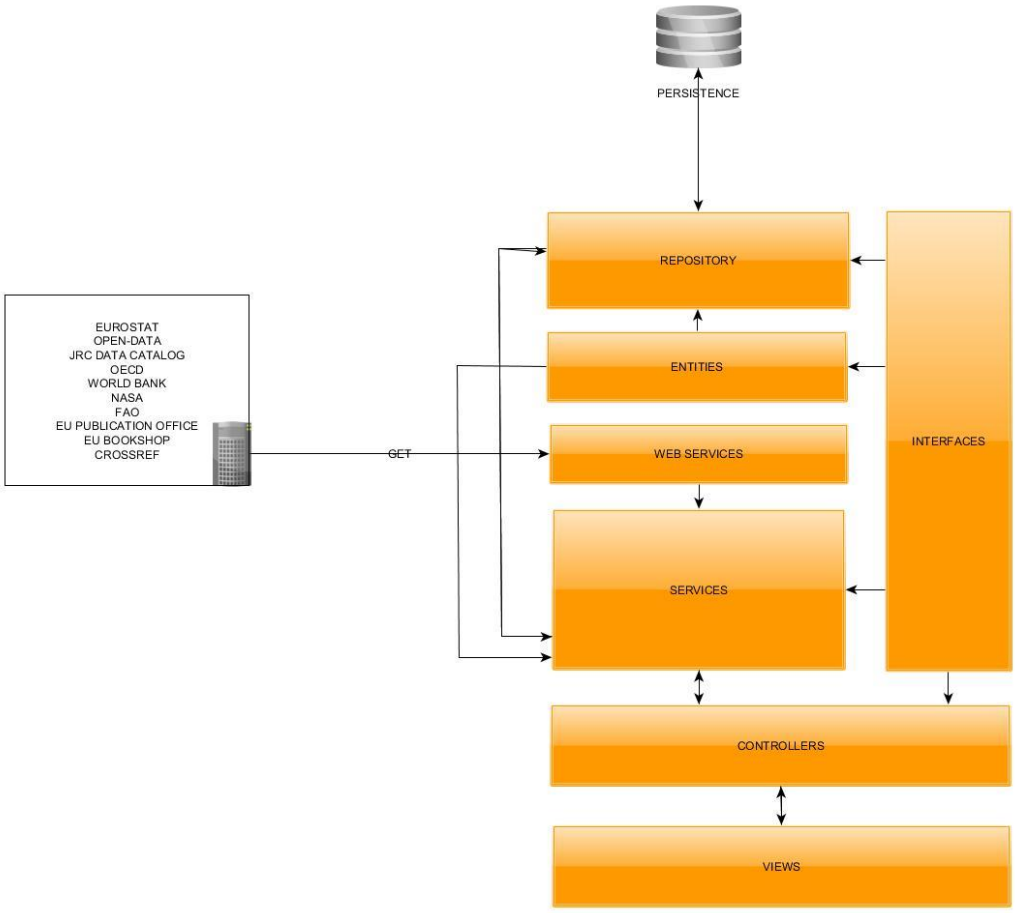

Figure 2. MIDAS back end infrastructure diagram

MIDAS back end infrastructure includes an ad-hoc data input editor, open to all registered EC modellers and policy makers to insert and maintain the metadata elements describing a specific model, impact assessment, or study. The editor is by design separated from the end-user interface, and requires *European Central Authentication Service* (ECAS) ([15]) authentication and local authorisation.

6.2 Interdependencies with other Knowledge Management Systems

All MIDAS tools and components are arranged around the MIDAS database, which forms the core of the application hosting the descriptions of the models, and related entities. For models the system provides a complete metadata profile, while for related entities the system is situated within a SOA to query existing inventories of related entities. The DB only hosts minimal information to avoid duplication and to minimise maintenance of the stored information. To ensure synergies and reduce maintenance, preference is given to authority services provided e.g. by the *Publications Office of the European Union* (OP) ([16]). Wherever possible, MIDAS stores persistent (or at least unique) identifiers. If a PI and resolving engine are not available, we use URLs to resolve into the location of the entity itself or a metadata entry of the entity.

([15]) https://en.wikipedia.org/wiki/Central_Authentication_Service

([16]) The Publications Office of the European Union (OP) publishes and disseminates the publications of the institutions and other bodies of the European Union, see https://publications.europa.eu

All these information items, such as datasets, parameters, models, studies, impact assessments, publications and laws are heavily interconnected, creating interdependencies and relations between the information systems which are storing and managing this information.

The following systems and repositories are under the responsibility of the OP:

— EUR-Lex ([17]): The EUR-Lex website provides free access to *The Official Journal of the European Union*, EU case law and other resources for EU law. The database dates back until 1951.

— EU Publications ([18]): The *EU Publications* website provides access to reports, studies, information booklets, magazines and other publications from the EU institutions and other bodies. This was formerly known as the EU Bookshop.

— EU Vocabularies ([19]): The *EU Vocabularies* website provides access to vocabularies managed by the EU institutions and bodies. This includes controlled vocabularies, schemas, ontologies, data models, etc. The *EU Vocabularies* website replaces the *EuroVoc* and the *Metadata Registry* (MDR) sites. We are using in particular: *EuroVoc*, *Corporate Bodies*, and *Subject Matter*.

— EU ODP ([20]): EU ODP stands for *EU Open Data Portal*, a catalogue of datasets from the EU institutions and other EU bodies like EUROSTAT for example.

Some of the content can be accessed through CELLAR, which is the common repository of content managed by the OP. CELLAR contains the files and the metadata of various collections of documents, which can be fetched using the machine-readable SPARQL ([21]) endpoint facility or the HTTP RESTful web services. CELLAR is powered by semantic technology and enables direct access to information stored as Linked Data.

Another resource of the EC we use is the *Publications Repository of the Joint Research Centre*, named *PUBSY* ([22]). PUBSY is an online service giving access to data about research publications produced by the JRC. It was established to assist with central storage, management and search to our research publications that go beyond the official publications stored in the EU Bookshop. PUBSY has a native application program interface (API).

Other systems are:

— DOI ([23]) and Crossref ([24]) services: The DOI service (see also Box 3) gives the possibility to retrieve publications metadata information based on the unique publication DOI number. This, in conjunction with the Crossref service, provides MIDAS with full publication's metadata and citation information.

— NASA ([25]), FAO ([26]), World Bank ([27]): MIDAS makes use of the web services provided by these important institutions to harvest metadata information about public datasets on different topics of interest.

([17]) https://eur-lex.europa.eu
([18]) https://publications.europa.eu/en/web/general-publications/publications
([19]) https://publications.europa.eu/en/web/eu-vocabularies/
([20]) http://data.europa.eu/euodp/en/home
([21]) http://www.w3.org/TR/sparql11-query/
([22]) http://publications.jrc.ec.europa.eu/
([23]) http://dx.doi.org/
([24]) https://www.crossref.org/
([25]) https://neo.sci.gsfc.nasa.gov/
([26]) http://www.fao.org/faostat/en/
([27]) https://data.worldbank.org/

7 Functionality – a brief overview

In this chapter we provide you with a brief overview of the MIDAS functionality and content types. Please note that the shown functionality refers to the corporate version of MIDAS (the MIDAS portal accessible from within the EC Network ([28]) from all Commission Sites). A dedicated version has also been made available to the European Parliament ([29]).

Figure 3. The MIDAS Portal, Corporate version

The *Front End* is organised into a set of modules.

In this document, we will focus on the modules called *Search*, *Impact Assessments*, *Model Support to the EU Policy Cycle* and *Reports*, as these are the end-user applications providing access to the *MIDAS Fact Sheets* and dedicated reports.

([28]) This version is accessible only from the Commission Network, under the address http://midas.jrc.cec.eu.int

([29]) The Parliament version is accessible from the TESTA Network under the address https://webgate.ec.testa.eu/midas-ii/

The *Editor* module and the related workflows will play a main role in the integration of MIDAS into the relevant business processes from 2019 onwards for the creation of impact assessments in the EC.

7.1 The Fact Sheets

Fact Sheets are online data sheets (or fiches) that summarise all information that the MIDAS system contains about its five main entities: models, impact assessments, studies, data, and people (modellers and model users).

The metadata profiles of these entities or content types are of varying complexity, as highlighted in the chapter 5: The main metadata profile is that of models, being the main content type of the inventory. For the remaining types we only store the most relevant information. Additional information is loaded from or referred to if existing in other EC or external organisations where the information is maintained.

The fact sheets bring together the knowledge of anyone that interacted with it: i.e. for models you will find information on who is running them in or on behalf of the EC, which other models can connect to them to use the model outputs, and which impact assessments and studies used the model results. This means we bring together the perspective of all modellers and model users in the EC about each model and its use in the EC.

7.1.1 Search & Analyse

The main aim of this module is to allow users to find and export relevant information from MIDAS. The model offers a search engine, the possibility to query the MIDAS Knowledge Base, and to visually analyse the connections made.

7.1.2 The Search engine

The aim of the *Search* engine (Figure 4) is to give the user as much flexibility as possible to extract the desired content. Users must first select the correct tab to indicate which of the content types they are searching for. An Excel-like structure allows filtering and sorting by content type (columns). Filters can be combined so that it is possible, for example, to look for any models run inside the JRC that support the policy area 'Agriculture'. Due to space limitations, not all searchable columns are immediately visible so users may have to show/hide columns to display the required fields. A full screen mode has been provided to allow a larger number of search fields to be displayed at the same time.

All results can be exported in various formats (such as Excel or PDF), or they can be selected and further analysed using the '*View Selection as Graph*' function, further referred to as the '*Analyse*' function, as it allows a visual analysis of the selected content.

Figure 4. Search and Analyse - Switch between Search and Analysis Graph

The user can switch between the *Search* and the *Analyse* function, keeping the selection made in the search. The analysis graph (Figure 5) will show all the selected elements, and the connections between them. The user can extend the selection using either the search function or the graph itself, which allows adding new elements related to those already visible on the canvas.

It is possible to s*how, hide and weight* (based on the number of incoming links) any of the links in the analysis graph. This gives the users the flexibility required to design a graph that fits their needs. An example of question that can now be answered is, for example: *for the given models and studies, which of the models was most involved?*

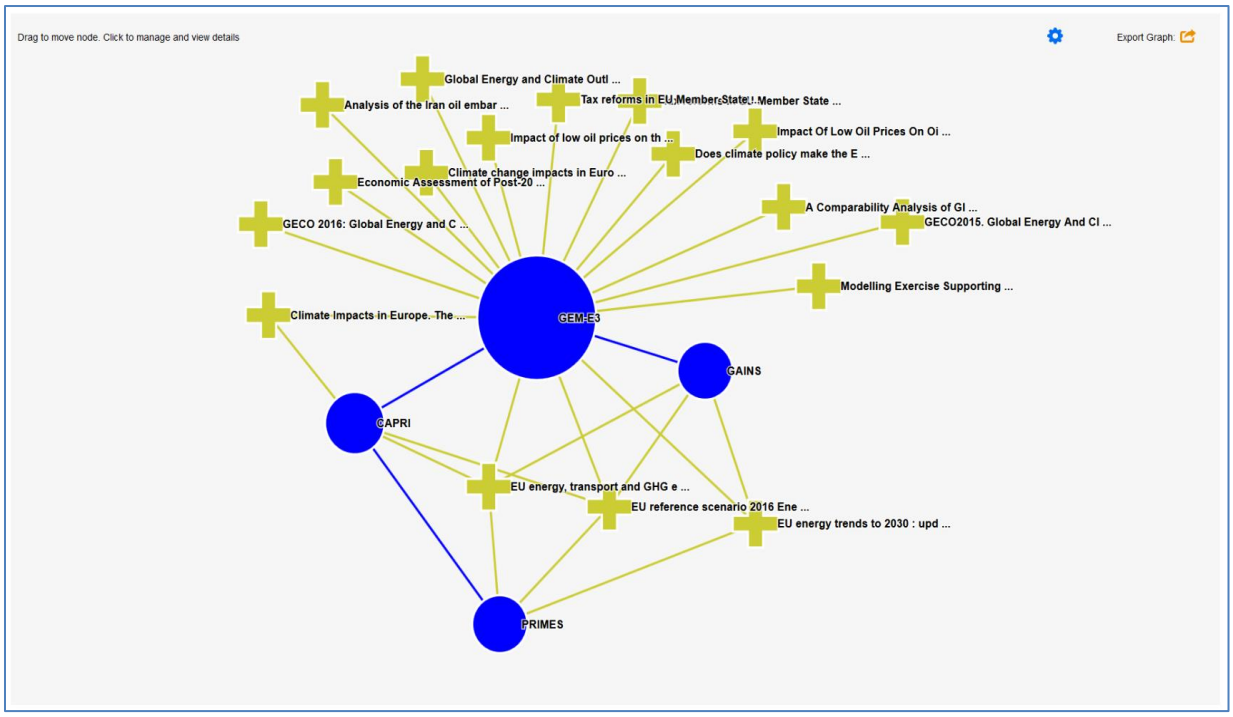

Figure 5. Analysis Graph: For the given set of models and studies: which of the models was used in which studies? Models are weighted by number of studies they were involved in.

The graph and all related tables showing the selection and resulting connections can then be exported by the user for further use.

7.2 Model Support to the EU Policy Cycle

In MIDAS, all models actively in use by the EC have been associated to one or more phases of the policy cycle by the people responsible for the description. The module on *Model Support to the EU policy Cycle* gives access to this information (see Figure 6).

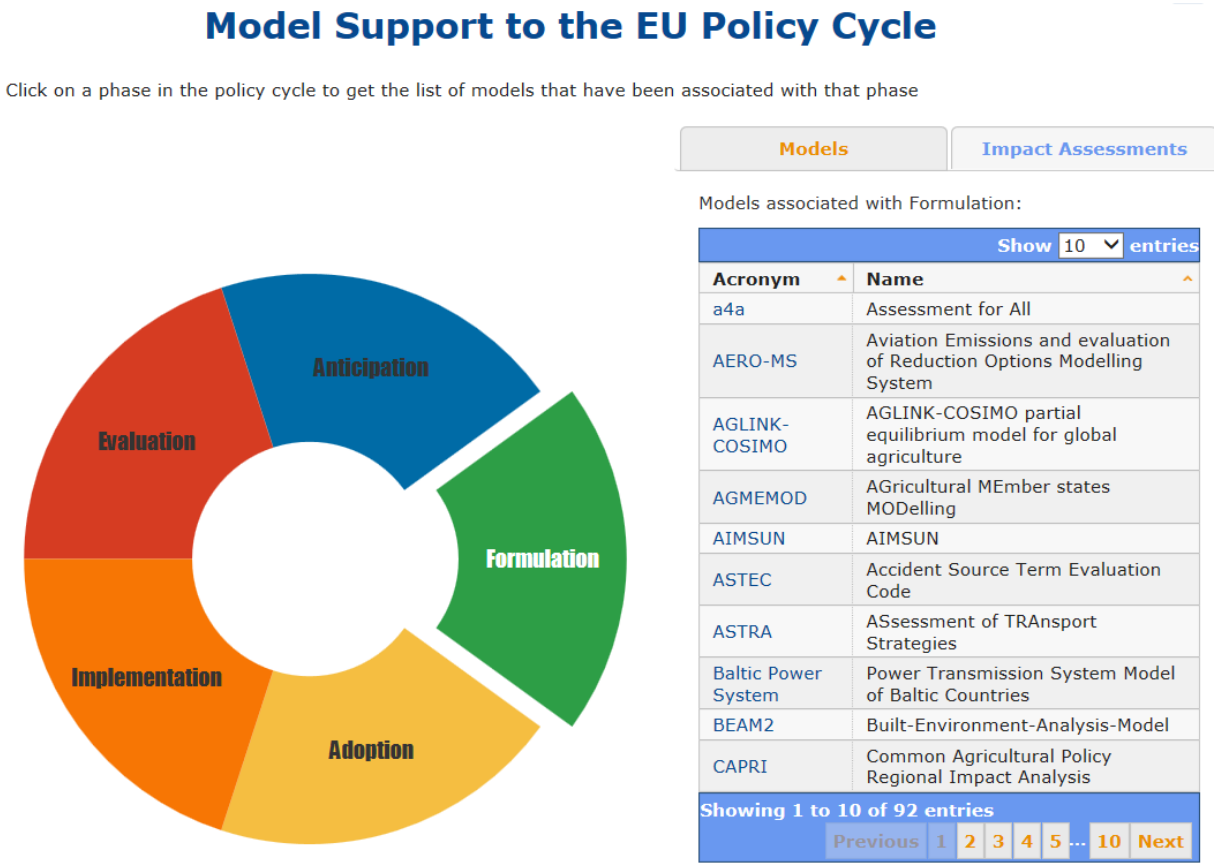

Figure 6. Model Support to the EU Policy Cycle

The model description itself contains a section explaining how the model can be used to support the associated phase(s) of the policy cycle. If the model has been designed to support a specific policy or group of policies, then this is added as well. We are currently in the process of also linking studies to the specific phase of the policy cycle they supported.

7.3 Impact Assessments

Modelling is important for *ex ante* impact assessments, and modelling results can be used for all aspects of an impact assessment: problem definition, evaluation of existing policy, baseline scenarios and assessment of policy options. The aim of the *Impact Assessments* module in MIDAS is to explain all the concepts, related information and tools that MIDAS hosts to support policy analysts and modellers to fulfil the obligations by the *Better Regulation Guidelines*, and to find the right models to support impact assessments.

In addition to the *Search*, which allows users to find impact assessments according to year of publication, models used, subject matter, lead DG, title, etc., MIDAS also hosts the following additional tools to support users in their activities related to impact assessments:

— **Impact Assessment Models Summary:** To get an overview of previous use of models for Impact Assessments, an interactive graph for exploring the use of models by subject matter, and by their specific role in the Impact Assessment.

— **MIDAS Impact Area Browser**: This tool supports users who are looking for suitable models that can assess certain types of impacts, required by the *Better Regulation Toolbox*. Browsing by impact category, from the most general to the most specific

impact types, this tool allows to find which models can be used to assess these impacts, and how.

— **MIDAS IA Annex Generator**: Once a model has been used for an impact assessment, this tool helps users to create the mandatory annex describing the model, which should be attached to the impact assessment report.

All tools are described below in some detail:

7.3.1 Impact Assessment Models Summary

If an impact assessment has been registered in EUR-Lex, we can capture information on the relevant subject matter this impact assessment covered, which is specified in a code list managed by the OP (see chapter 6). Examples of subject matter are *Environment*, *Transport*, *Pollution* etc.

As MIDAS also contains information on the role of models in supporting an impact assessment, we can combine the subject matter, the models used, and the role of these models into a single graph (see Figure 7).

By using this interactive graph, users can explore for example, which models have been used in a given range of years for 'evaluation of existing policy' in impact assessments covering '*Transport*'.

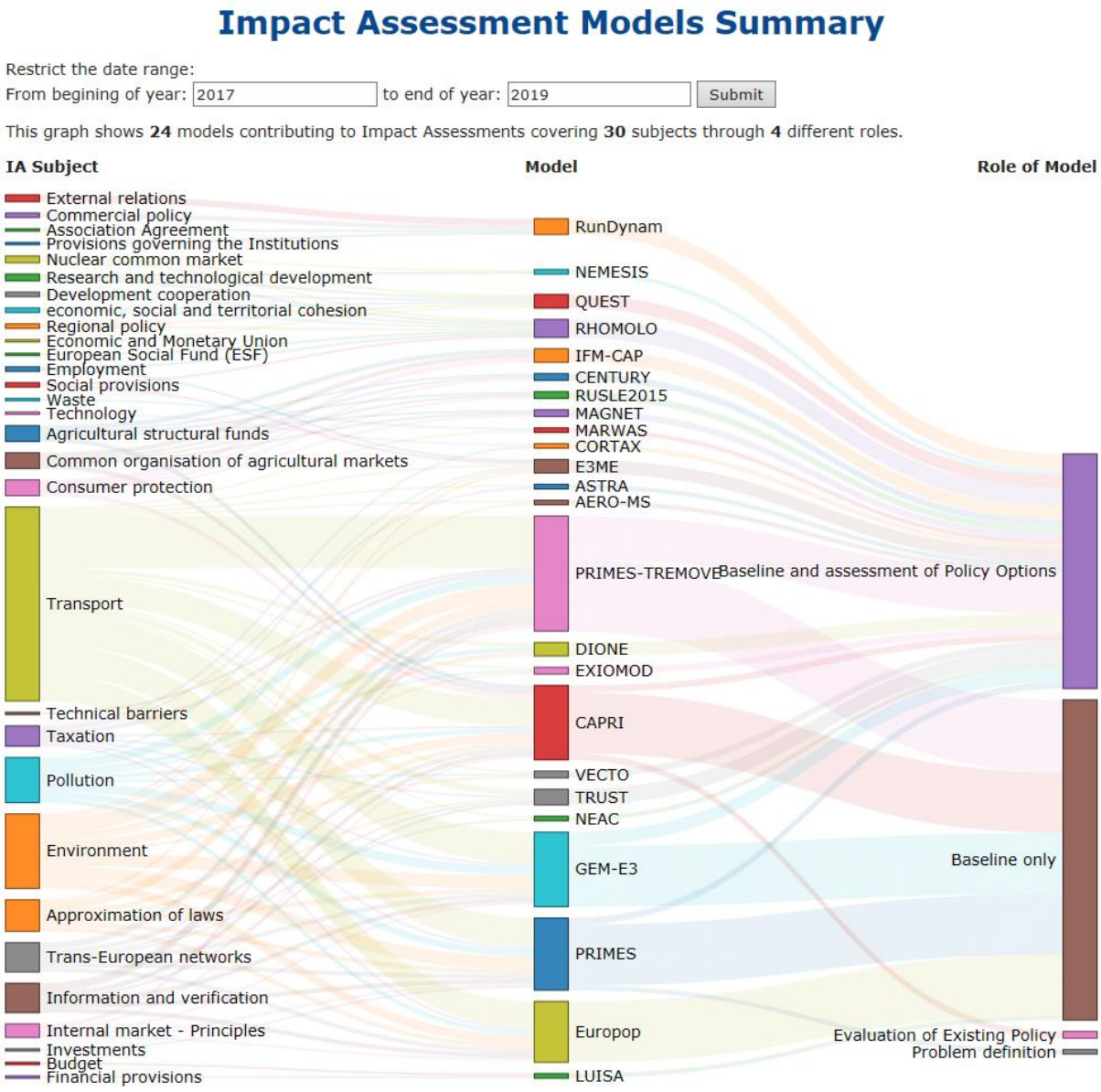

Figure 7. The Impact Assessment Models Summary (years: 2017 - 2019)

7.3.2 MIDAS Impact Area Browser

Choosing a model for an impact assessment is not trivial, and the technical description might not be sufficient to understand how a model can be used.

The *Better Regulation Toolbox* provides an entry point here, as it describes the kind of impacts that have to be assessed in an impact assessment. These impacts are structured into four impact areas: economic, social, environmental, and fundamental rights. They are further structured into *Impact Categories*, and *Subcategories* ([30]).

In MIDAS we associate models that can be used for impact assessments with these categories, including a description of how the model outputs can be used to assess these impacts. With the *MIDAS Impact Area Browser*, users can now browse these categories, to find e.g. all models that can assess economic impacts, and how. If users are interested in a very specific impact category, they can browse down to a specific subcategory.

([30]) Subcategories are represented in the *Better Regulation Toolbox* through leading questions.

This has been implemented for a subset of all models, namely those where we have in-house expertise. We plan to extend this in the future to all models used for impact assessments in the EC.

Branch slected : 'Effects on income, distribution and social inclusion'

Search: Show 10 entries

Impact category	Can be assessed by model	Through
Access to and quality of social protection benefits	EUROMOD	Analysing the extent to which actual and prospective policy measures can influence the household disposable income
Financing and organisation of social protection systems	EUROMOD	Analysing the extent to which actual and prospective policy measures can influence the household disposable income
Households income and at risk of poverty rates	POLES	Energy price impacts at country-level (from EU or third countries) or international energy and climate policies
Households income and at risk of poverty rates	E3ME	consumer prices and expenditures, and implied household distributional effects.
Households income and at risk of poverty rates	GEM-E3	Household income is a key outcome of any GEM-E3 analysis
Households income and at risk of poverty rates	EUROMOD	Analysing the extent to which actual and prospective policy measures can influence the household disposable income
Inequalities and the distribution of incomes and wealth	E3ME	consumer prices and expenditures, and implied household distributional effects.
Inequalities and the distribution of incomes and wealth	EUROMOD	Analysing the extent to which actual and prospective policy measures can influence the household disposable income
Inequalities and the distribution of incomes and wealth	FSSIM-DEV	Gross farm income, Net Farm Income, Farm-household income, Poverty gap
Inequalities and the distribution of incomes and wealth	IFM-CAP	Inequalities and the distribution of incomes and subsidies across EU farms

Showing 1 to 10 of 10 entries Previous 1 Next

Figure 8. MIDAS Impact Area Browser - sample results for: *Social impacts >> Effects on income, distribution and social inclusion*.

7.3.3 MIDAS IA Annex Generator

As mentioned previously, any model used in the impact assessment has to be described in a mandatory annex to the impact assessment report.

The *MIDAS IA Annex generator* is a tool that supports this specific demand: Based on the information stored in MIDAS, the *IA Annex Generator* creates a prefilled and annotated word-template, including information on all elements which are mandated in the *Better Regulation Guidelines*. The template can be generated for any model listed in MIDAS, and can then be used as a basis to draft the mandatory Annex.

Instructions for the use of the template are in cursive print and highlighted in yellow, and can be removed from the template once the editing is completed. These instructions are directly based on the *Better Regulation Toolbox #62, Section 5 'Transparency'* (European Commission 2017[2]).

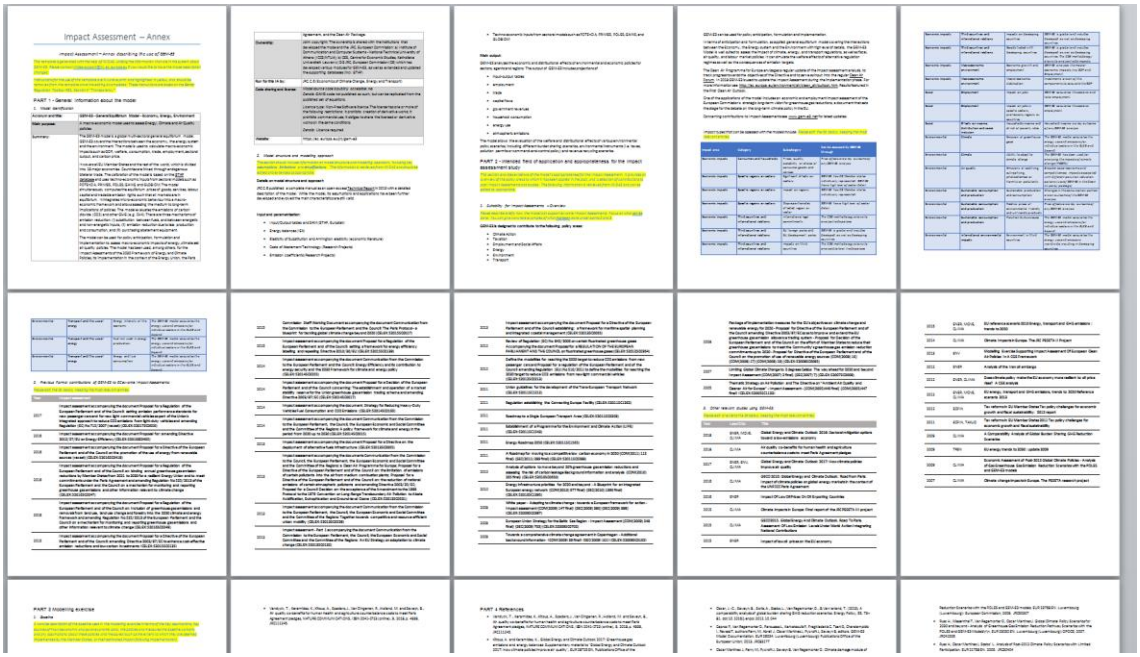

Figure 9. Overview sample of an IA ANNEX, generated using the MIDAS IA Annex Generator

7.4 Reporting

In this module, users can request the automatic generation of reports in form of word documents, both for specific items like models, impact assessments and studies, and for the inventory as a whole.

The MIDAS team also offers report generation for ad-hoc queries, usually representing a certain perspective or (policy) question. In cases repeated queries for the same information are needed, the reports are shared on the *Reporting* module of MIDAS.

Currently, MIDAS already offers the following reports:

— *Fact Sheets (item specific reports)*:

- *Models*: Word documents that include all information listed in MIDAS about a particular model.
- *Impact Assessments*: Word documents that include all information listed in MIDAS about a particular impact assessment.
- *Studies*: Word documents that include all information listed in MIDAS about a particular study.
- *Directorate-General (DG)*: Word documents summarising all modelling activity in or for a specific Policy DG or EC Service.

— *Inventory reports*:

- The inventory as a whole.
- A report that provides the inventory from a policy perspective, providing models with a particular support to policy making.

Figure 10. MIDAS Reporting Section

8 Conclusion & Outlook

MIDAS directly contributes to the *Better Regulation Agenda* by enhancing the *transparency* of the use of models in support to policy making and the *traceability* of their results.

The information and reports generated by MIDAS on models and model use can be used in practice in documents and reports, such as the compulsory annex introduced by the *Better Regulation Agenda* for impact assessments on '*Analytical models used in the preparation of the impact assessment*'. This could effectively contribute to the recommendation to enhance transparency on data, assumptions, methodology and results (RegWatchEurope cited in Golberg 2018; Impact Assessment Institute 2017) and to keep impact assessments understandable and useful for decision-makers, as suggested in the academic debate (Golberg 2018).

MIDAS also provides an *overview of the models used* in support to policies *within and across different policy domains*. Models are often used together to answer complex policy questions which involve various policies and related fields. Identifying which models are used and in which respect can allow for synergies, greater and more consistent results for policy making. This can also help achieving a more efficient use of resources.

Enhanced transparency in models and model use may help to understand better how models work and to validate their behaviour. This can encourage sound use of models, and thus contribute to more *sound quantitative evidence* in support to the policy making process and to tackle remaining difficulties in quantification in impact assessments (RSB SG and JRC Working Group 2018).

By providing information on the role of models in support to policy making, for example in ex ante impact assessment or ex post evaluation of related policies, MIDAS can also contribute to increased *coherence in the policy cycle* in linking ex ante with ex post assessment.

In addition, organised evidence gathering, like the one achieved thanks to the MIDAS modelling inventory, contributes to *learning* within the Institutions (Smismans 2015). This also facilitates the work of EU internal *scrutinizing* bodies like the RSB and the European Parliament.

Finally, having been opened to the European Parliament MIDAS constitutes a practical example of *Interinstitutional collaboration* and sharing of a common evidence base.

In conclusion, MIDAS enhances the transparency of models, the traceability of model results, and our understanding of the ongoing support within and across different domains. It is an important corporate and interinstitutional tool to use, reuse and document models in a proper way, directly contributing to the dissemination and use of sound methodology underpinning the EC's Better Regulation policy.

References

ACS, S., OSTLAENDER, N., LISTORTI, G., HRADEC, J., HARDY, M., Smits, P., Hordijk, L. (2019) *Modelling for EU Policy support: Impact Assessments*. JRC Technical Report. Forthcoming

Broughel, J. (2015) 'What the United States Can Learn from the European Commission's Better Regulation Initiative', *European Journal of Risk Regulation* 6(3): 380–381.

Hradec, J., Ostlaender, N., Macmillan, C., Acs, S., Listorti, G., Tomas, R. and Arnes Novau, X., Semantic Text Analysis tool: SeTA, EUR 29708 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-01518-5 (online), doi:10.2760/577814 (online), JRC116152.

EPRS (2015) 'Ex-ante impact assessment in the European Commission's new Better Regulation Guidelines', European Parliamentary Research Service.

European Commission (2015) 'Better Regulation for Better results - An EU agenda', *COM (2015) 215 final, Brussels, 19 May.*

European Commission (2015[1]) 'Better Regulation Guidelines' *SWD (2015) 111 final, Strasbourg, 19 May.*

European Commission (2015[2]) 'Better Regulation Toolbox', accessible at: http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf, last access: 20th March 2019.

European Commission (2016) 'Data, Information and Knowledge Management at the European Commission', *C (2016) 6626 final, Brussels, 18 October.*

European Commission (2017[1]) 'Better Regulation Guidelines', *SWD(2017) 350 final, Brussels, 7 July*.

European Commission (2017[2]) 'Better Regulation Toolbox', accessible at: https://ec.europa.eu/info/better-regulation-toolbox_en, last access: 20th March 2019.

European Parliament, the Council of the European Union and the European Commission (2016) '*Interinstitutional agreement of 13 April 2016 on Better Law-Making*' Official Journal of the European Union L 123, 12.5.2016, p. 1–14

GBIF (2011). A *Beginner's Guide to Persistent Identifiers, version 1.0*. Released on 9 February 2011. Authors Kevin Richards, Richard White, Nicola Nicolson, Richard Pyle, Copenhagen: Global Biodiversity Information Facility, 33 pp, accessible online at http://links.gbif.org/persistent_identifiers_guide_en_v1.pdf. [accessed 20 March 2019]

Golberg, E. (2018) '"Better Regulation": European Union Style', *Harvard Kennedy School, Mossavar-Rahmani Center for Business and Government, M-RCBG Associate Working Paper Series* (98).

Ostlaender, N., Bailly-Salins, T, Hardy, M., Perego, A., Friis-Christensen, A. dalla Costa, S., (2015) *Describing models in context – A step towards enhanced transparency of scientific processes underpinning policy making*, International Journal of Spatial Data Infrastructures Research, Vol.10, 27-54.

Radaelli, C. M. (2018) 'Halfway Through the Better Regulation Strategy of the Juncker Commission: What Does the Evidence Say?', *JCMS: Journal of Common Market Studies* 56: 85–95.

Renda, A. (2015) 'Too good to be true? A quick assessment of the European Commission's new Better Regulation Package', *CEPS Special Report* (108).

RSB SG and JRC Working Group (2018) *Quantification in Commission Impact*

Assessments and Evaluations, Brussels, available at https://ec.europa.eu/info/publications/report-rsb-sg-jrc-working-group-quantification-commission-impact-assessments-and-evaluations_en.

Smismans, S. (2015) 'Policy evaluation in the EU: The challenges of linking ex ante and ex post appraisal', *European Journal of Risk Regulation* 6(1): 6–26.

Tonkin, E.(2008).Persistent Identifiers: Considering the Options July 2008, Ariadne Issue 56 http://www.ariadne.ac.uk/issue56/tonkin/[accessed 20 March 2019]

List of abbreviations and definitions

AI	Artificial Intelligence
CC-MOD	European Commission's Competence Centre on Modelling
CoP	European Commission's Community of Practice
DGs	Directorate-Generals of the European Commission
DOI	Digital Object Identifier
EC	European Commission
ECAS	European Central Authentication Service
EP	European Parliament
EPRS	European Parliamentary Research Service
IA	Impact Assessment
IAWG	Impact Assessment Working Group
IGP	Information Providers Guide
ISG-MOD	Inter-Sservice Group on Modelling
JRC	Joint Research Centre of the European Commission
LISO	JRC Local Information Security Officer
MDR	Metadata RegistryOP
MIDAS	Modelling Inventory and Knowledge Management System of the European Commission
MVC	Model-View-Controller
OP	Publications Office of the European Union
ORM	Object-Relational Mapping
PI	Permanent Identifier
RSB	Regulatory Scrutiny Board
SMCE	Social Multi-Criteria Evaluation of Policy Options
SOA	Service Oriented Architecture

List of figures

List of boxes

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications
You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-29729-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub

ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/900056

ISBN 978-92-76-02852-9