The CMIP5 Model Documentation Questionnaire
## Development of a Metadata Retrieval System for the Metafor Common Information Model

Charlotte Pascoe[1], Bryan Lawrence[1], Marie-Pierre Moine[2], Rupert Ford[5], Gerry Devine[4] and the Metafor Project Team
[1] NCAS-BADC ; [2] CERFACS ; [3] University of Manchester ; [4] NCAS-Climate

metafor

7 CAPACITIES

The CMIP5 Questionnaire development life-cycle allows working scientists to be part of the development process

Mind Maps
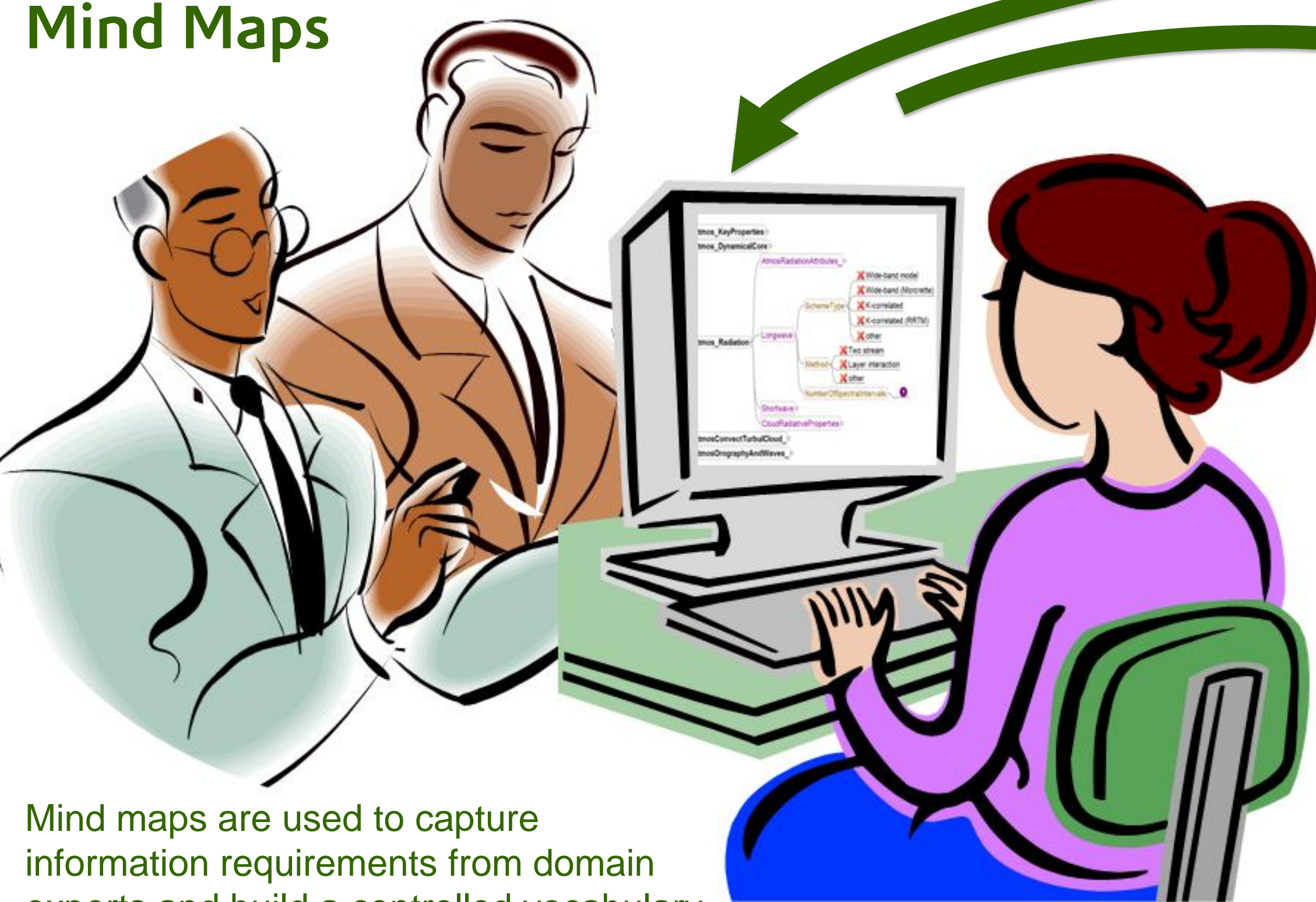
Mind maps are used to capture information requirements from domain experts and build a controlled vocabulary.

Metafor has been charged by the Working Group on Coupled Modelling (WGCM) via the Coupled Model Inter-comparison Project (CMIP) panel to define and collect model and experiment metadata for the CMIP Phase 5 (CMIP5) project. The core archive of model data produced by CMIP5, will be used for the next Intergovernmental Panel on Climate Change (IPCC) assessment.

To this end a web-based questionnaire is collecting information and metadata from the CMIP5 modelling groups on the details of the models used, how the simulations were carried out, how the models conformed to the CMIP5 experiment requirements and details of the hardware used to perform the simulations. The questionnaire is primarily a tool to document models in sufficient detail so that the CMIP5 data can be compared in a scientifically meaningful way.

The CMIP5 model documentation questionnaire is an ambitious metadata collection tool and will provide the most comprehensive metadata of any climate model inter-comparison project. However, completing the questionnaire will require a substantial effort from climate scientists. We hope the sense of ownership engendered by our interactive consultation strategy will motivate climate modellers to invest the time to ensure that the information gathered in the questionnaire becomes an invaluable resource for the whole community.

The controlled vocabulary captured in the mind map interviews is used to populate drop-down lists in the questionnaire and the structure of the mind maps generates the tree navigator. The tree navigator shown on the left hand frame of the questionnaire screen shot , it allows users to get to any component of the model with ease.
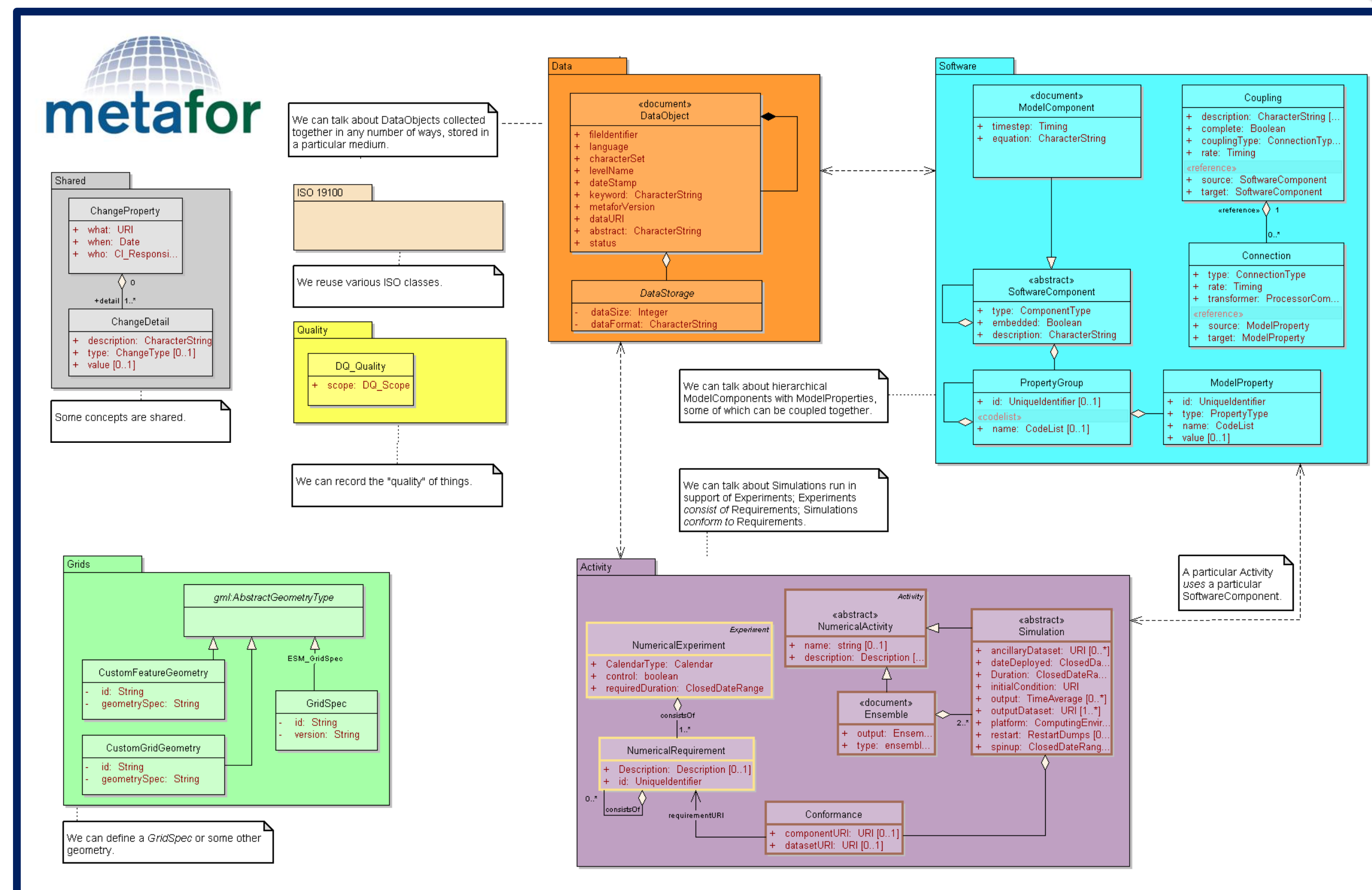
The questionnaire also allows users to enter descriptions of components which are not covered by the mind maps. The mind map driven sections of the questionnaire ensure that we collect a standardised set of metadata about each of the CMIP5 models but if users wish to describe their models in more detail then the questionnaire is flexible enough to allow them to do it. Additional terms entered by users will be used to populate the externally governed controlled vocabularies used by the Metafor Common Information Model (CIM).

Python Parser

A python parser processes the XML files generated by the mind maps

```
<component name="Radiation">
    <definition status="missing">Definition of component type Radiation required</definition>
    <parameter name="RadiativeTimeStep" choice="keyboard">
        <definition status="missing">Definition of property name RadiativeTimeStep required</definition>
        <value format="numerical" name="time step"/>
    </parameter>
    <parametergroup name="Longwave">
        <parameter name="SchemeType" choice="">
            <definition status="missing">Definition of p
            <value name="Wide-band model"/>
            <value name="Wide-band (Morcrette)"/>
            <value name="K-correlated"/>
            <value name="K-correlated (RRTM)"/>
            <value name="other"/>
        </parameter>
        <parameter name="Method" choice="XOR">
            <definition status="missing">Definition of p
            <value name="Two stream"/>
            <value name="Layer interaction"/>
            <value name="other"/>
        </parameter>
        <parameter name="NumberOfSpectralIntervals" choice="keyboard">
            <definition status="missing">Definition of property name NumberOfSpectralIntervals required</definition>
            <value format="numerical" name=""/>
        </parameter>
    </parametergroup>
```

django
http://www.djangoproject.com

Django is used to generate the dynamic structure and content of the web based questionnaire from processed xml and the METAFOR Common Information Model (CIM)

Django is a high-level Python Web framework which makes it possible to build web applications with minimal code.
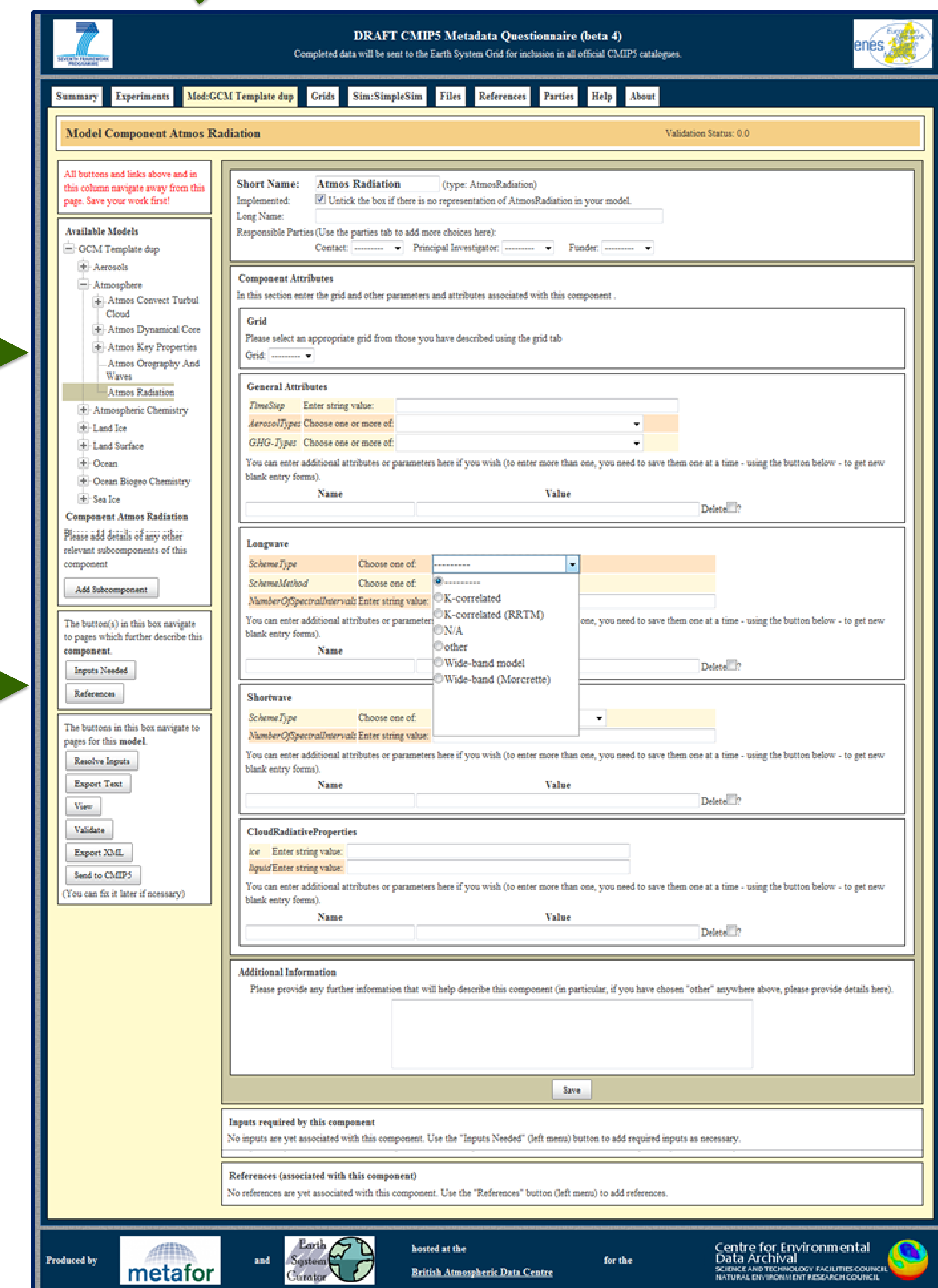
The CMIP5 questionnaire makes much use of Django templates. The templates allow us to separate design, content and Python code.

This separation of concerns has given us the flexibility to use the most appropriate tools for the different parts of the questionnaire life-cycle.

The careful selection of tools provides a framework that enables working scientist (who we would never ordinarily get to interact with UML and XML) to be part of the iterative development process and ensure that the CMIP5 questionnaire reflects what scientists want to know about climate models.

UML overview of the Metafor Common Information Model (CIM). CIM XML documents are used to generate experiment descriptions and conformance pages in the CMIP5 questionnaire.

CMIP5 Questionnaire

http://q.cmip5.ceda.ac.uk/

Screen shot from the Beta4 release of the CMIP5 model documentation questionnaire showing a metadata entry page for describing the Atmosphere Radiation component of a model.

Beyond CMIP5

The development framework described here can be applied to any model. The Metafor project team will be using this framework to document the models used by the ENSEMBLES project. .