ES-DOC PI Report

Assumptions:

- Questionnaire:
  - Will not require any branding for CMIP6 (authority = Feb 2016 F2F).
  - Will be used to fill in Specializations, Responsible Parties, Citations, and to edit documents produced via other methods (authority=Feb 2016 F2F).
- The UKMO will use pyesdoc to produce CIM documents.
- Some groups may use the IPython notebook to generate model descriptions.
- Modeling groups will fill realms piecemeal rather than all at once. The same goes for the various other documents.
- The CIM will be changing iteratively as tools are developed. We would prefer a peer review of each class to consider it "stable enough not to be a blocker".
- There will be a great deal of work to do beyond just getting the CIM and generators ready. For instance, authentication, security, and the overall user experience will need to be touched on.
- Community review, documentation and training are essential and will be planned and organised as an integral part of the process.

Resources:

- Testing Plan with Status: https://github.com/ES-DOC/esdoc-testing
- Search Site: http://search.es-doc.org/
- CoG Site: https://www.earthsystemcog.org/projects/es-doc-models/
- Splash Page: http://es-doc.org/
- Wordpress Site: http://wordpress.es-doc.org/
- Questionnaire Releases: https://earthsystemcog.org/projects/es-doc-models/releases_questionnaire
- Github: https://github.com/ES-DOC

STATUS

as of

28 November 2016

Legend: 🟢 Completed and tested; 🟡 In Progress; 🔴🔴 Needs attention; 🔴 Not Started

Experiments and Experimental Requirements

How it will work: Experiment information, including Requirements, Temporal and Forcing Constraints was put into a spreadsheet by Charlotte. Mark wrote code to parse this and convert it to CIM documents, which were then ingested into the Archive.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Aug 2016	🟢	

Generated via a spreadsheet	Aug 2016	●	Charlotte did a fantastic job with this.
Archive ingestion	Aug 2016	●	Mark had to do lots of one-off coding to make this work since we were dealing with a spreadsheet.
Config files	Aug 2016	●	Mark wrote generator to output simple JSON config files for use by MOHC and others to initialise their internal systems.
D3 config files	Nov 2016	●	Mark wrote generator to output D3 visualization compatible JSON config files.
Visible in Viewer	Aug 2016	●	Mark E looked at this and liked it.

Conformances

How *it* *will* *work*: Users will be presented with a table, by MIP, of conformances. They will edit only those that don't conform. The table will be autogenerated from the Experiment spreadsheet using rules created by Charlotte. Mark will take these rules and pull the information from the spreadsheet and generate the table. How users will then edit this has yet to be determined. It is probable that the method will be a google spreadsheet.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Nov 2016	●	David and Charlotte adapted CIM to handle modification to data input e.g. regridding.
Rules for table generation	Nov 2016	●	Charlotte has drafted a rules document. Need a systematic method to identify forcings that require further input from the modelling groups.
Information is editable	Need Prioritizing	●	Mark to do this. Need to consider how user's will edit.. Is a web interface needed or just a google spreadsheet.
Edited information converted to CIM documents		●	Will need special code.
Visible in Viewer		●	

Responsible Party

How *it* *will* *work*: RPs will be linked to many documents. They will be generated through the Questionnaire or pyesdoc (group choice). They will be stand alone documents. How we will be enable groups to link to RPs from a potential list of hundreds is unclear.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Oct 2016	●	Has this been peer reviewed yet?
Published via the Questionnaire	Oct 2016	●	Test documents out in atom feed.

Lifecycle	Target	Status	Notes
Published via pyesdoc	Oct 2016	🟢	Real documents created during the Experiment publication process.
Ingested into the archive from pyesdoc	Oct 2016	🟢	Real documents ingested during the Experiment publication process.
Ingesting into the archive from the Questionnaire.	Feb 2016	🔴	Mark will need code to read the atom feed.
Visible in Viewer	Oct 2016	🟢	Those linked via other documents (e.g. Experiments).
Can be linked to other documents		🟡	pyesdoc works. Questionnaire may require some web-based mechanism to sort through thousands of published documents.

Performance

How it will work: We don't how this information will be collected or how it will be connected to other documents. There will be one Performance document per model, per Experiment.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Nov 2016	🟡	Minor changes recommended from peer review by Bryan Lawrence have yet to be implemented.
Published via the Questionnaire	Dec 2016	🟡	Once CIM updated, just needs testing. Should just work.
Published via pyesdoc	Dec 2016	🟢	Initial tests passed. Published to local disk.
Generated via the IPython Notebook		🔴	It is possible to seed template IPython notebooks for all possible model combinations and provide them on a Jupyter Hub server as an alternative to the Questionnaire.
Ingested into the archive		🔴	
Visible in Viewer		🔴	
Can be linked to other documents		🔴	

Machine

How it will work: We don't how this information will be collected or how it will be connected to other documents. This will be a small number of documents. One per institute.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Nov 2016	🟡	Minor changes recommended from peer review by Bryan

			Lawrence have yet to be implemented.
Generated via the Questionnaire	Dec 2016	🟡	Once CIM updated, just needs testing. Should just work.
Generated via pyesdoc	Dec 2016	🟢	Initial tests passed published to local disk.
Generated via the IPython Notebook		🔴	It is possible to seed template IPython notebooks for all possible model combinations and provide them on a Jupyter Hub server as an alternative to the Questionnaire.
Ingested into the archive		🔴	
Visible in Viewer		🔴	
Can be linked to other documents		🔴	

Simulations, Ensemble, Ensemble Axis, Ensemble Members

How it will work: David has written a script (cdf2CIM) that harvests select pieces of information from the netCDF files directly. This script will dump the information out. Mark will write a script to take this dump and convert it CIM documents. We anticipate the dump script to be integrated somehow into the ESGF software stack.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Aug 2016	🟢	
cdf2CIM code written	Sep 2016	🟢	
Test dump with CMIP5 data	Nov 2016	🟡	Dump completed but need to compare to actual documents to determine accuracy.
Web-service endpoint exposed	Jan 2017	🔴	Mark G. to implement , expect to finish during the Jan 2017 coding sprint.
Conversion to CIM Document written	Jan 2017	🔴	Mark G. to implement, expect to finish during the Jan 2017 coding sprint.
Integration into ESGF	Jan 2017	🟡	Initial meetings with ESGF held in Nov 2016.
Further_url connected	Mar 2017	🔴	This will be an ES-DOC web service. Coding up the rendering of CIM2 documents will begin in early 2017.

Specializations

How it will work: Specializations are wrappers/extenders of the base Model Process classes in the CIM. They contain the additional "scientific properties", now identified as "processes" and "sub-processes" specific to that realm. There will be one per realm. These will be ingested directly by the Questionnaire and IPython notebooks to form the basis of the questions asked. As of November 2016, the format, structure, and content of these files was still being decided.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Dec 2016	(two pink dots)	Format seems to be stabilizing as of 22 Nov 2016. **Everything hinges on this.**
Ocean	Dec 2016	(yellow dot)	Multiple drafts checked in. Now adapting to latest changes in the CIM. POC is Eric.
Sea Ice	Dec 2016	(yellow dot)	Initial draft checked in on 24 Nov 2016. POC is Ruth.
Atmosphere	Dec 2016	(yellow dot)	Almost there, updating syntax POC is Charlotte.
Land Surface		(red dot)	Need the first three realms done; the generation, search, and view done before soliciting a realm expert.
Atmospheric chemistry		(red dot)	Need the first three realms done; the generation, search, and view done before soliciting a realm expert.
Ocean biogeochemistry		(red dot)	Need the first three realms done; the generation, search, and view done before soliciting a realm expert.
Land Ice		(red dot)	Need the first three realms done; the generation, search, and view done before soliciting a realm expert.
Aerosols		(red dot)	Need the first three realms done; the generation, search, and view done before soliciting a realm expert.
Generated via the IPython Notebook	Jan 2016	(yellow dot)	Initial tests pass. will need to double check once the Specializations are relatively stable. Some extra coding may be required to deal with changes.
Generated via the Questionnaire	Jan 2016	(two pink dots)	Need stable CIM, treeview is complicated and make take extra time.
Generated via pyesdoc	Jan 2016	(yellow dot)	Initial tests pass. will need to double check once the Specializations are relatively stable. Some extra coding may be required to deal with changes.
Seeded from CMIP5 document	Jan 2016	(two pink dots)	Need mapping from CMIP5 to CIM2 for EACH realm, which has to be done by realm experts. There also has to be a model identifier (model A to model D) mapping, which in fact may be tricker. Then Mark has to write code to do the conversion.
Previous document editable via the Questionnaire	Mar 2016	(red dot)	

Lifecycle	Target	Status	Notes
Previous document editable by pyesdoc		🔴	
Short model tables for papers		🔴	Auto-generated tables describing/comparing model document details, suitable for publishing papers.
Previous document edited via the IPython notebook		🔴	
Ingested into the archive		🔴	
Visible in Viewer		🔴	
Visible in the Comparator	2018	🔴	This may be the last thing we do for CMIP6.

Citation

How *it* *will* *work*: CIM2 content is undeveloped. How they will be generated and linked is still unknown. How they related to the further_info_url, is unknown.

Lifecycle	Target	Status	Notes
CIM2 stable enough not to be a blocker	Aug 2016	🟢	Has not been peer reviewed but current structure used by Experiments.
Generated via pyesdoc	Aug 2016	🟢	Was done as part of the Experiment spreadsheet to CIM publication process.
Ingested into the archive from pyesdoc	Aug 2016	🟢	Was done as part of the Experiment spreadsheet to CIM publication process. Questionnaire untested.
Can be linked to other documents via pyesdoc	Aug 2016	🟢	Was done as part of the Experiment spreadsheet to CIM publication process.
Visible in Viewer	Oct 2016	🟢	Can be viewed linked with Experiment documents.
Generated via the Questionnaire	Dec 2016	🟡	Just need to test this. Should just work.
Ingested into the archive from the Questionnaire	Jan 2016	🔴	Mark will need code to read the atom feed.
Can be linked to other documents via the Questionnaire		🔴	A much more complicated process. There could be thousands of these to chose from.

Documentation

How *it* *will* *work:* The es-doc splash page is migrating to a new url (http://wordpress.es-doc.org/) that is a wiki so it can be edited by everyone outside of the es-doc stack. Must be ready when we initially email the world on behalf of the WGCM.

Lifecycle	Target	Status	Notes
Wordpress site created	Jan 2016	🟢	http://wordpress.es-doc.org/

Structure finalized	Dec 2016	🟡	Some initial content is there but we need an outline. Dedicated telco planned for the near term.
How to choose CIM generator documented	Jan 2017	🔴	Need to guide users to either the Questionnaire or pyesdoc.
Questionnaire documented		🔴	
Pyesdoc documented		🔴	
Viewer documented		🔴	
Comparator documented		🔴	
URL transferred		🔴	From current splash page to wordpress site.

Community review and training

How it will work: TODO

Appendix A: Document creation and linking

ES-DOC DOCUMENTS - creation and linking

*Conformance table created by ES-DOC, groups edit for non-conformance, ES-DOC then create all Conformance documents

Entry point to published documentation (view and edit) via further_info-URL