World Meteorological Organization

COMMISSION FOR CLIMATOLOGY

Expert Meeting on WMO Climate Data Management System Strategy
Geneva, Switzerland, 12-14.XII.2018

EM-WCDMSS/Doc N(4)

Submitted by:
Steve Foreman

14.XII.2018

IMPLEMENTATION PLAN FOR THE REFERENCE CDMS TOOL SET

This document will become document 6.3(2) of the 2019 workshop.

REVIEW COMMENTS

Version / No	Reviewer	Issue	Notes	Response
1/1	Stuber	2.3 Management and resouces: "the "C" for "climate" in "OpenCDMS" might discourage other communities (such as hydrology) from adopting the tool set". That for sure. A CDMS well defined, that includes hydrological data, is very closed of a generic system dealing with managing generic sensor data. And this could interest some other communities. The "C" of climate could be replaced by example by a "E" for **E**nvironment. This possiblilty was discussed also with Bruce Bannerman, Peer Hecler, and Johannes Cullmann (head of Climate and Water Department of WMO).		Annotated
1/2	Stuber	Comment on strategy that could include proprietary systems.		New para 2.9.5.
1/3	Stuber	Stage 3: Observation metadata and climate-relevant products. "Observation metadata" is too dependant o the data management to pput its implementation in Stage 3. I will not give	This expands the scope of stage 2: is there scope for using OSCAR/surface as the interface from which the database aspects are populated?	Pending discussion

		here all the consequences of defining a Core system without the Metadata, but they are essential. Especially for data ingest, data ingest monitoring, provenance and data validity. They are necessary part of the Core implementation.		
1/4	Palmer	One comment on the draft Implementation Plan. In Para 29, this says "At this stage, OpenCDMS will link to Oscar/surface for information on the metadata associated with the observations (although a minimum set of metadata will need to be stored within OpenCDMS to allow processing). " This implies that OSCAR is the definitive source of station metadata, and OpenCDMS implementations are derivative from OSCAR. I don't think that OSCAR is good enough now or in the near future for this task. I have been looking at OSCAR for other reasons, and found 16 significant errors in a short look at the stations in the British Overseas Territories, including e.g. St Helena (where Napoleon was imprisoned after Waterloo in a rather nice villa just up the road from the observing station) is now part of the USA and managed by US-NWS	The intention is that the CDMS populates OSCAR/surface from delivery of Stage 3 onwards. Cleaning the data in OSCAR/surface is one way of getting the information published before Stage 2 is even developed. The errors in OSCAR/surface are because the owners of the metadata have not checked the automatic uploads – the upload may have mis-interpreted data, or the source may have been incorrect or out of data.	Pending discussion

		(a surprise to me, as I arrange the funding…), and that the CTBTO https://www.ctbto.org/ station there is managed by the USA (that could also lead to an international incident!). The station on Pitcairn is in the middle of the sea, I think the OSCAR map has the island in the wrong place (not surprising, the Bounty mutineers went there because no-one else could find it, it was another 30 years before another Royal Navy ship got there). Someone (not the British!) seems to have altered the entry for Falkland Islands. And a lot of "open since" information from Pub9 VolA seems to have been lost or corrupted. It also seems over-ambitious to expect OSCAR to store global metadata about 3rd party stations which may be available to NMHSs on a restricted basis, or may be "informal" sites.		

		So I would suggest that the national CDMS operated under the responsibility of the NMHS ("H" is "Hydrological") is the primary reference for station metadata, and there is a responsibility on the NMHS to ensure that OSCAR is kept up-to-date. This fits with the "Role and operations of NMHSs" definition of national responsibilities. Picking up on another thread from WIGOS, there is an interesting issue at St Helena – the CTBTO station is located within the obs enclosure of the GCOS station, but has a different 5-digit station number…..		
1/5	Stuber	**The passionate** A possibility also is to include, in the different phases of the project, the passionate of the environment. You have quoted the associations (ACRE, IEDRO, ..), and there are also some "open data provider" that be willing to participate to		Added to para 3.1.5.

		such interesting Open Source project (Infoclimat, Oldweather, Wunderground, Weather Detective, etc.)		
1/6	ET-DRM	A project officer is needed to coordinate the work manage the interactions between contributors, plan the finances,…		
1/7	ET-DRM	Clarify that any commercial organization could join the project (whether HMEI or not)		Added to paras 3.1.4, 3.2.2
1/8	ET-DRM	Need WMO funding to support stages 1 & 2		Added para 2.5.2
1/9	ET-DRM	Communications: station owners (through the WIGOS focal points) are responsible for maintaining the OSCAR metadata for their station – and the CDMS can be a tool to achieve this.		Added to para 3.1.8
1/10	ET-DRM	Interaction with other activities: need to have standard data models to support both data and metadata.		Added para 2.8.5

DISCLAIMER

Regulation 43

Recommendations of working groups shall have no status within the Organization until they have been approved by the responsible constituent body. In the case of joint working groups the recommendations must be concurred with by the presidents of the constituent bodies concerned before being submitted to the designated constituent body.

Regulation 44

In the case of a recommendation made by a working group between sessions of the responsible constituent body, either in a session of a working group or by correspondence, the president of the body may, as an exceptional measure, approve the recommendation on behalf of the constituent body when the matter is, in his opinion, urgent, and does not appear to imply new obligations for Members. He may then submit this recommendation for adoption by the Executive Council or to the President of the Organization for action in accordance with Regulation 9(5).

Chairperson, Publications Board
World Meteorological Organization (WMO)
7 bis, avenue de la Paix
P.O. Box No. 2300
CH-1211 Geneva 2, Switzerland

Tel.: +41 (0)22 730 84 03
Fax: +41 (0)22 730 80 40
E-mail: Publications@wmo.int

NOTE:
The designations employed in WMO publications and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of WMO concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The mention of specific companies or products does not imply that they are endorsed or recommended by WMO in preference to others of a similar nature which are not mentioned or advertised.

This document (or report) is not an official publication of WMO and has not been subjected to its standard editorial procedures. The views expressed herein do not necessarily reflect those of WMO or its Members.

IMPLEMENTATION PLAN FOR THE REFERENCE CDMS TOOL SET

1 PURPOSE OF THIS DOCUMENT

1.1 This plan implements the strategy for the reference Climate Data Management (CDMS) tool set. The plan is built around an "agile programming" approach so that after the first two stages (in which the project and its management infrastructure are initiated and the core underpinning infrastructure is created), the priorities for further development are decided at each stage of the project with the aim of generating benefits to the users at each stage.

1.2 The overall aim of the strategy for the reference CDMS tool set is to provide all WMO Members with a software solution based on an electronic database to assist with management of climate-relevant information. In its initial stages of development, the reference CDMS tool set in itself will not provide sufficient functionality to allow Members to implement it as their core CDMS tool set. This plan therefore seeks to allow Members to inter-operate their existing systems with their current software, and to allow Members without current access to an

electronic CDMS tool set to implement one of the existing open source solutions (that do not have a sustainable model for maintenance and support) with the confidence that they can evolve to using the reference CDMS tool set.

1.3 The first stage of the project is concerned with setting up the management infrastructure for making decisions about the project, and a collaboration infrastructure to allow software developers around the world to work together. The second stage is to set up an initial version of the database upon which the software will be built together with other software components without which the data cannot be entered or accessed. This stage will also define the interfaces that allow these components to inter-operate with other software systems. Following the second stage, the development priorities will be determined as the start of the next stage is being planned. So far as is possible, the interfaces to a component of software developed at a stage in the project will also work with components of the tool set that are developed at later stages.

1.4 This plan concentrates on technical development of the reference CDMS tool set. The technical development includes providing mechanisms to interface the reference CDMS tool set with earlier CDMS systems. Implementing the reference CMDS tool set within organizations is not included in the plan.

1.5 Although the name of the project to implement the reference CDMS tool set will be agreed during the initiation stage, this plan assumes that the tool set will be called OpenCDMS.

2 STAGE 1: INITIATION

2.0.1 The purpose of the initiation stage is to set up the management practices for the project, agree the governance of the project, and publish an initial architecture for OpenCMS.

2.1 COMMUNICATIONS

2.1.1 Communications form an important element of Stage 1. The prime purpose of the communications plan for this stage is to obtain agreement from Cg-18 that OpenCDMS will underpin climate data management activities, to elicit sufficient financial contributions to the Special Account or in-kind contributions to allow the second stage of the project to be implemented, and to seek participation in the project from suppliers of meteorological, hydrological and oceanographic equipment.

2.1.2 The outcome of the communications is intended to be:

8.1 Resolution of Cg-18 that Congress endorses an open source reference CDMS tool set (OpenCDMS) and authorizes a Special Account to support its development, maintenance and operation;

2.1.2.2 A declaration by Members and members of HMEI that they will contribute to the development and maintenance of OpenCDMS;

2.1.2.3 Members not operating a CDMS based on an electronic database know that they can implement a franchise CDMS in the knowledge that their system will be migrated to OpenCDMS as OpenCDMS is developed.

2.3 MANAGEMENT AND RESOURCES

2.3.1 Although OpenCDMS will be open source, it has to be owned by a legal entity. That entity is the one that issues the software licence, enters into agreements with contributors that ensure that users of the software are allowed to use the software as intended by the project, and oversees management of the project. A key threat to open source projects in general is the concept of software patents that is recognized by some jurisdictions. Many open source projects, such as the OpenWIS Association, have chosen to set up an entity that is registered in a jurisdiction that does not recognize software patents in order to manage the risks associated with software patents. OpenCDMS may wish to do this, either by joining an existing open source project that is already owned by such an entity, or by setting up its own entity. An early decision of the project has to be the name; the "C" for "climate" in "OpenCDMS" might

discourage other communities (such as hydrology) from adopting the tool set. One suggestion is OpenEDMS (Environmental Data Management System)

2.3.2 The OpenCDMS project also needs to decide on the software licence it will use. The GPL 3 licence is commonly used, as it permits modification of the software by users but limits full integration of the software within another software package that is provided under a more restrictive licence. A frequently used method of ensuring that users do not confuse the trunk version of the software (that maintained by the project) with modified versions is to register the name of the project as a trademark and limit its use to the version of the software maintained by the project. The first stage of the project will set up the legal and licensing arrangements for the project.

2.3.3 A key component of the first stage of the OpenCDMS project is to set up the management structures. There are four components to this:

2.3.3.1 Management of the project development;

2.3.3.2 Management of funds supporting the project;

2.3.3.3 Management of the infrastructure supporting the project;

2.3.3.4 Management of regulatory and associated decisions needed by the project.

2.4 MANAGEMENT OF THE PROJECT DEVELOPMENT.

2.4.1 For OpenCDMS to thrive, it needs to be seen to be agile and managed by consensus among those contributing to its development. WMO's governance infrastructure is likely to be seen as too slow and restrictive by contributors to the project for them to maintain their interest in it. A formal management structure, set up as part of the entity that owns OpenCDMS, is needed to coordinate the developments and to assist the development team to come to consensus on the contents of the tool set. WMO could provide one of the members of the management team, but others would have to come from the community of developers, from users of CDMS and from other organizations with an interest in developing OpenCDMS. HMEI should be invited to nominate a representative to encourage commercial companies to contribute to the project.

2.4.2 The management team for the OpenCDMS project will need to be supported by a full time project officer whose responsibilities are provided in more detail in Annex 1. In addition to identifying where the management team needs to make decisions and summarizing the information needed by the team to make those decisions, the project officer will also take day to day responsibility for project finances, communications and community building.

2.5 MANAGEMENT OF FUNDS SUPPORTING THE PROJECT.

2.5.1 Not all those interested in developing OpenCDMS will be capable of providing development effort directly, and some may be reluctant to hire developers themselves. WMO may choose to set up a funding mechanism, such as a Special Account (the development of OpenCDMS will last more than one financial period, see Standing Instruction 5.6.32), to allow financial contributions from a number of donors to be aggregated to allow developers to be paid to develop specific components of the system. Management procedures will need to be set up to decide on how the funds should be spent and to monitor the effectiveness of the expenditure. The team managing the fund should include representatives of the project as well as representatives of WMO and the donors. The Special Account should complement developments provided directly by contributors, and is most likely to be used for infrastructure aspects of the project that ease its use by Members, such as installation tools or documentation. Note that Decision 19 (EC-69) decided to implement the I-DARE initiative, the resource plan for which ("A proposed long term resource plan for climate data management and rescue activities") also referred to the resources needed to implement the initial stages of an open source reference CDMS (€4M for the initial implementation).

2.5.2 Initial development of OpenCDMS (stages 1 and 2) is likely to need paid-for development in order to make progress in an acceptable time scale. This was foreseen in the resource plan for climate data management ??This makes setting up the Special Account, and the contributions to it, a high priority during the preparation for the project.

2.6 MANAGEMENT OF THE INFRASTRUCTURE SUPPORTING THE PROJECT.

2.6.1 Open source projects need an infrastructure of collaboration tools to support them. The project needs to choose which tools it intends to use to support collaboration and who will have the responsibility of maintaining and administering the tools. It also needs to set up hosting for the source code and any common infrastructure needed for development and testing.

2.7 MANAGEMENT OF REGULATORY AND ASSOCIATED DECISIONS NEEDED BY THE PROJECT.

2.7.1 Although the project will make its own decisions on the software and the software components that it delivers, those components will implement processes and procedures that will become informal standards. The project will comply with standards and specifications from WMO (including those in the CDMS Specifications, WMO-No. 1131) and the Manual on the High Quality Global Data Management Framework for Climate. There will be many topics, though, where there are no global standards or specifications for the project to adopt (such as how many missing values are permitted when calculating monthly means or climate normal). There needs to be a rapid procedure to help the project choose between competing options for algorithms and processes. Although there would not be enough time for the proposals to go through a formal WMO decision making process (even the Fast Track is too slow), there has to be a mechanism for representatives of a WMO constituent body to advise on which approach is likely to be more acceptable as a standard practice (probably members of expert teams or of a Community of Practice). This approach would allow the proposal to be tested in practice using the CDMS tool set, and would allow the proposal to be submitted to the formal approach when it was proven. Using the CDMS tool set to implement the candidate for WMO approval has the advantage that if the proposal is modified during its approval process, only one set of software has to be modified and there would be a standard roll-out procedure for users. The management procedures to allow this will need to be set up.

2.7.2 Request participation in the project from NMHS, HMEI and experts known by WMO and IOC to be working on CDMS tools.

2.8 BUILDING THE CRITICAL CORE

2.8.1 Publish a draft system architecture for the reference CDMS tool set so that the key interfaces can be identified and the owners of other systems can judge what they might be able to contribute to the OpenCDMS.

2.8.2 Experts familiar with ClimSoft, MCH and CliDE will identify which components of those systems might be re-used within the OpenCDMS tool set, and what would be needed for those systems to use the OpenCDMS database as their main source of information.

2.8.3 In addition to the "franchise" CDMS, the project should invite the owners of other CDMS to offer all or part of their software for potential use in the project. Provided that software is provided under an appropriate open source licence, this approach should allow the project to proceed more rapidly than if all components of the software had to be designed and coded from scratch.

2.8.4 Decide on the core elements of OpenCDMS that have to be in place in order that other CDMS may implement interfaces to it, and so that a stand-alone implementation can be used to receive, store, and share observations made by the organization operating OpenCDMS. These will form OpenCDMS version 1.

2.8.5 Not only is the architecture and modular design of OpenCDMS a critical delivery to allow the project to proceed, clear definitions of the standard data models to be used by OpenCDMS for storing and exchanging observations and metadata are essential to allow OpenCDMS to be used alongside with other software, including that of the Climate Services Information System toolkit. These data models have to be produced in the first stage of the project and their evolution subject to strict version control.

2.9 MAINTAINING CURRENT CAPABILITY

2.9.1 One of the objectives of OpenCDMS is to provide a sustainable replacement for the current "franchise" CDMS. During this first stage of the project, the following activities will be undertaken to allow a smooth transition.

2.9.2 **ClimSoft**. Source code for ClimSoft is stored under the OpenCDMS github project. The software is licensed under the GPL 3 licence that is expected to be the licence used by OpenCDMS. The project will negotiate with the owners of ClimSoft to transfer management of the software to the OpenCDMS project.

2.9.3 **MCH** (Meteorological, Climatological and Hydrological Database Management System). Ownership of MCH was transferred to WMO in 2011. MCH is released under a licence that requires the permission of a Permanent Representative to pass the software to a user. This is not compatible with an open source project. The project will need to negotiate with WMO to give OpenCDMS a software licence for MCH that is compatible with that for OpenCDMS and to publish the source code for MCH in the OpenCDMS github repository. The negotiation should also seek to pass management responsibility for developing MCH to the OpenCDMS project.

2.9.4 **CliDE**. CliDE is published under the GPL 3 licence (http://www.bom.gov.au/climate/pacific/docs/CliDE_user_manual_3.3.pdf) that is the same as expected for OpenCDMS. Another open source project, run by the British Geological Service, is also known as CliDE, so references within the OpenCDMS project should use a different name (such as CliDE-PCCS) to avoid confusion. The source code should be transferred to the OpenCDMS repository. OpenCDMS should negotiate with the Bureau of Meteorology on whether management of the source code should also transfer to OpenCDMS, or whether the OpenCDMS copy should be treated as a fork from the original and managed separately.

2.9.5 **Proprietary systems**. WMO Members use CDMS other than the "franchise" systems. Some of these (45% of Members) were developed by the Member using the system, and others use proprietary systems. The owners of those systems may wish to contribute code to OpenCDMS or may wish to contribute to development of tools that assist user migration from those systems to OpenCDMS.

2.10 INCLUDING THE CDMS SPECIFICATIONS

2.10.1 Review the CDMS specifications (WMO-No. 1131) to identify requirements that have to be built into the core architecture of OpenCDMS. Some aspects of ICT systems are difficult to add on to systems that have already been implemented, so it is important that these requirements are known to the developers during the initial design. Examples of requirements that are likely to have a pervasive impact on the design of OpenCDMS include: authentication and access controls, management and implementation of data policies (particularly those that may limit the use and distribution of data), logging of transactions and processing for audit purposes. Other pervasive requirements will be identified during the review of the WMO-No. 1131, and it is likely that reviewing WMO-No. 1131 suggests additional constraints that are not explicit in WMO-No. 1131.

2.10.2 In addition to pervasive constraints within WMO-No. 1131 the objectives of OpenCDMS will introduce others. OpenCDMS seeks to handle all the climate-relevant data for the organizations using it, and so will need to be able to be extended in a simple way to store and process additional types of data and metadata; not only will these extra data need to be stored, but they will also need to be processed and quality controlled. Handling of extra data types is also one example of the need for OpenCDMS to support additional or alternative processing of information held within it, and so has to provide suitable interfaces to allow such extensions to work smoothly – this has to be designed-in from the start of the project (designing for extensibility will make it easier to implement new core functionality in later releases). There will be other aspects of the design that have to be implemented in the first version of OpenCDMS in order to allow the initial version to work with components developed in later versions (OpenCDMS should be both backwards and forwards compatible except at infrequent major version changes; a strong policy on version numbering is essential for the success of OpenCDMS as it will help reduce the maintenance effort required by users).

2.11 ADDING NATIONAL CAPABILITIES

2.11.1 Addition of national capabilities is an example of providing extensions or alternatives to OpenCDMS. In future releases, it may be appropriate for some of the processing to be made flexible so that national practices can be controlled by control files rather than requiring alternative software to be written. This option should not be ruled out by the initial architecture.

2.12 SUPPORT FOR IMPLEMENTATION

2.12.1 OpenCDMS will not be implemented during the first stage. The OpenCDMS project may, however, recommend which franchise CDMS may be appropriate for users to install while OpenCDMS is being developed.

3 STAGE 2: CORE IMPLEMENTATION

3.0.1 The principal objectives of Stage 2 are to provide data ingestion, storage, exchange and quality control to allow a country without an existing CDMS based on an electronic database to use OpenCDMS to store and manage its observations and to produce the minimum set of climate-relevant products that are required by WMO standards and recommended practices. This includes storage of data obtained through data rescue activities that are either entered by hand or imported using a csv file. At this stage, OpenCDMS will link to Oscar/surface for information on the metadata associated with the observations (although a minimum set of metadata will need to be stored within OpenCDMS to allow processing).

3.1 COMMUNICATIONS

3.1.1 By the end of Stage 2, there will be a working version of OpenCDMS that provides a minimum capability for managing climate-relevant observations. With that capability available, the formal launch of OpenCDMS as an open source project will be appropriate, and Members without a CDMS based on an electronic database will be able to use OpenCDMS to manage their observations.

3.1.2 Communication will have different aims for each stakeholder group.

3.1.1 **Permanent Representatives of countries expected to use the first version of OpenCDMS** (and their climate data managers). The communication needs to focus on progress with OpenCDMS, ease of implementation of OpenCDMS, the benefits of its first version and the development path to provide additional support for their operations, how to implement OpenCDMS (climate data managers).

3.1.2 **Senior managers of organizations contributing to the first version of OpenCDMS**. The communication focusses on: the value being delivered from their contribution; progress; users wishing to implement the system; the total effort being invested (to show the leverage of the collaborative project); decision making process for the next stage of delivery.

3.1.3 **WMO constituent bodies and the Secretariat**. The communication should concentrate on: the benefits being enabled by the first version; Members intending to implement the first version; ease of migration from the franchise CDMS; rate of progress with OpenCDMS; compliance offered by OpenCDMS; data preservation offered by OpenCDMS.

3.1.4 **Senior managers of organizations and projects not contributing to the first version of OpenCDMS**. Aims of OpenCDMS and benefits to contributors and recipients; leverage provided by collaborative working; standardization enabled by OpenCDMS; for HMEI members and other for-profit or not-for-profit providers of CDMS and related software: benefits of collaborative development of infrastructure that is not a differentiator thus freeing resource for investment in differentiation; benefits of being interoperable with OpenCDMS.

3.1.5 **Individual developers who could contribute to OpenCDMS**. In addition to software developers, this group of stakeholders includes weather and climate enthusiast such as those behind Infoclimat, Oldweather, Wunderground, Weather Detective and

others. Aims of OpenCDMS; benefits to society from OpenCDMS; working practices of OpenCDMS; progress of first stage; how to contribute to OpenCDMS.

3.1.6 **Climate scientists using observations and/or generating products**. Benefits of OpenCDMS; functionality and standardization within OpenCDMS; auditing, metadata and q/c supported by OpenCDMS; future objectives for OpenCDMS; benefits of being interoperable with OpenCDMS.

3.1.7 **Data rescue organizations and international projects**. Objectives of OpenCDMS; OpenCDMS as a repository for managing data as they are rescued; using OpenCDMS to transfer rescued data to the NMHS responsible for the data; benefits of being interoperable with OpenCDMS.

3.1.8 **Observation suppliers**. Aims of OpenCDMS; ability of OpenCDMS to implement data policies; OpenCDMS as a way of collecting their data and distributing them to a national centre; extensibility so could use with other types of data handled by the supplier; reminder that organizations managing observing stations are responsible for managing the metadata in OSCAR/surface and not the OpenCDMS will make it simpler to do so as well as being able to ingest metadata from OSCAR/surface.

3.2 MANAGEMENT AND RESOURCES

3.2.1 Development at this stage will be managed using the open source committee set up at stage 1.

3.2.2 Resources will be drawn from software developers and other experts volunteered by Members and members of HMEI and other for-profit or not-for-profit providers of CDMS and related software, supplemented if necessary by contracts with individual developers supported by funding from the Special Account.

3.3.3 A key output for the management theme in this stage of the project is a definition of the detailed management and governance procedures that will be used to manage development of the project in later stages.

3.3 BUILDING THE CRITICAL CORE

3.3.1 This stage builds the first version of OpenCDMS (OpenCDMS v1). The contents of the version were determined during the first stage.

3.3.2 During this stage, the essential contents of the next version (OpenCDMS v2) have to be specified. The contents should be limited to what can be delivered within a time scale of at most a year (preferably six months), allowing time for additional functionality to be added through consensus among the volunteer developers. In addition to the list of the components to be added, the stage should specify the interfaces, standards and mechanisms for those components to interact with each other and with external software so that the components can be developed in parallel.

3.4 MAINTAINING CURRENT CAPABILITY

3.4.1 Fundamental to any CDMS is the database that supports it. Any organization migrating from one of the franchise databases to OpenCDMS will need to migrate their data to OpenCDMS, so at this stage utilities will be provided to allow the franchise CDMS code to access the OpenCDMS database and the OpenCDMS code to access that franchise CDMS databases. This functionality will be used in the development of stage 3.

3.5 INCLUDING THE CDMS SPECIFICATIONS

3.5.1 The functionality agreed during stage 1 will be implemented during this stage.

3.5.2 The functionality that is considered to be high priority for implementation during stage 3 will be identified, and any clarification on applicable standards and specifications will be sought in order to avoid delays to stage 3.

3.6 ADDING NATIONAL CAPABILITIES

3.7.1 Stage 2 will concentrate on core functionality.

3.7 SUPPORT FOR IMPLEMENTATION

3.7.1 Experts supporting installation of the franchise CDMS should feed back their experiences to the development team for OpenCDMS so that issues that might affect OpenCDMS can be addressed during the development of OpenCDMS.

4 STAGE 3: OBSERVATION METADATA AND CLIMATE-RELEVANT PRODUCTS

4.0.1 The principal objective of OpenCDMS at Stage 3 is to expand the variety of observations that can be managed within OpenCDMS, to store and manage observation metadata within OpenCDMS, and to store climate-relevant products within OpenCDMS. At this stage, OpenCDMS will record metadata associated with observations and export to Oscar/surface the metadata required by that system (it will also import metadata from Oscar/surface; this will allow initial population of OpenCDMS and also allow metadata not originating from the organization operating the instance of OpenCDMS to be populated within that instance). OpenCDMS will be able to store scanned copies of rescued data and provide interfaces to tools commonly used in data rescue activities. Provisionally, this version will be known as OpenCDMS version 2 on the assumption that it was not possible to design sufficient flexibility into version 1 to allow components of version 1 to be used unchanged by components introduced in version 2. A project decision will be needed to determine whether stage 3 introduces version 1.*n* or version 2.

4.0.2 Whereas the aims of the second stage were determined by a limited team during the initiation phase of the project, the third stage has to encourage contributions from volunteers. In consequence, although some of the contents of version 2 of OpenCDMS will be determined by the team who specified version 1, a substantial part will be decided by a consensus-seeking process among the contributors. This may mean that some important functionality will be delayed as a result of other functionality being assigned higher priority by contributors. The delay, although frustrating to those accustomed to conventional software development practices, represents an investment that is necessary to encourage an active and sustainable open source community to support OpenCDMS.

4.1 COMMUNICATIONS

4.1.1 This is the first stage in which the project becomes fully open source. A major aspect of the communications activity during this stage is directed towards (potential) contributors to encourage them to feel fully part of the OpenCDMS community. Open source development progresses differently from conventional in-house software, so the emphasis of communication to donors, the technical commissions and climate scientists is to manage the perceptions of how the development is proceeding – particularly as the open source development methodology will highlight bugs, misunderstandings and confusions and not make the real progress as visible.

4.1.1.1 **Permanent Representatives of countries expected to use the first version of OpenCDMS.** The primary aim when communicating with these stakeholders is to reassure them about the quality and speed of development and to allow them to plan for implementing OpenCDMS in their organization.

4.1.1.2 **Senior managers of organizations and projects contributing to the first version of OpenCDMS**. These stakeholders will need to be reassured that the open source development is under control, delivering to an acceptable time scale and creating code of high quality. To encourage them to keep supporting the project, they also need to know how many users anticipate implementing the software and how this will impact on data availability from those users.

4.1.1.3 **WMO constituent bodies and the Secretariat**. These stakeholders will be concerned that OpenCDMS is addressing the problems of compliance and increasing the availability of information of known quality. Communications should concentrate on the rate of development, the expected uptake, and the success in uncovering problems with the software and standards.

4.1.1.4 **Senior managers of organizations and projects not contributing to the first version of OpenCDMS**. The project will hope that some of these will contribute towards the next stage of the project, so the aim of communication with this stakeholder group is to demonstrate that the open source approach is working, that users are intending to implement the software when it is available, and to highlight the intentions for the next stage and the benefits that additional contributors would add to the project in terms of time scale, scope and meeting the needs of users (both data suppliers and data users).

4.1.1.5 **Individual developers who could contribute to OpenCDMS**. Individual developers are likely to be motivated by the outcomes that will be enabled by OpenCDMS or by particular challenges involved in the implementation (and often a combination of each). Communications to developers should concentrate in the outcomes that OpenCDMS will deliver, the plans for the next stage (including scope of adding proposals from contributors) together with the technical challenges to be overcome, the approach being taken for development, how to watch what was happening and how to contribute.

4.1.1.6 **Climate scientists using observations and/or generating products**. This group of stakeholders has the potential to become contributors as well as users of OpenCDMS. Communication should concentrate on what OpenCDMS will deliver at this stage and the outline plans for the next stage, together with what that functionality will enable, particularly in terms of the availability and quality of observations from data-sparse regions. It should also explain how OpenCDMS might be useful in their own work.

4.1.1.7 **Data rescue organizations**. These stakeholders are keen that their outputs are made widely available and are demonstrably of high quality. Communication to these stakeholders should concentrate on what OpenCDMS can do for storing, managing and quality control of observations and subsequently exchanging them with others. Explaining how the open source development works may encourage them to contribute to the development by adding in functionality to handle scanned reports and interfaces to digitization tools.

4.1.1.8 **Observation suppliers**. An objective of WIGOS is to encourage organizations other than NMHSs to contribute observations to the global archive. Communication to these organizations should address how OpenCDMS will help them store and manage their observations and the importance of the quality management that OpenCDMS supports. These organizations are also likely to be concerned that usage constraints that they place on their data will be respected, so the communication has to confirm that OpenCDMS restricts the use of observations to the uses permitted by the supplier, thus giving the supplier reassurance that if they provide observations to an organization using OpenCDMS they can be confident about how their observations are used. These stakeholders are also potential users of OpenCDMS and, because they are likely to handle information that is not traditionally handled by the WMO community, they may wish to contribute extended functionality to OpenCDMS to make it suitable for use within their community.

4.2 MANAGEMENT AND RESOURCES

4.2.1 This stage will be managed fully under the agreed open source management and governance procedures. These may need to be updated using a consensus-driven approach to overcome issues identified during the stage.

4.2.2 A consensus approach should be used to decide on the contents of the release (OpenCDMS v1.*n* if possible). The project should be open to unsolicited contributions that forward the aims of the project and meet its development standards.

4.2.3 As this is the first stage of the project to be run as a truly open source project, it is likely that some of the functionality considered to be critical by WMO and IOC experts may not be contributed by volunteer developers. If this proves to be the case, funding from the Special Account could be used to commission development of the missing components.

4.2.4 The draft contents and design of the next stage should be developed by consensus.

4.3 BUILDING THE CRITICAL CORE

4.3.1 This stage should deliver sufficient functionality for users to capture, store, manage and quality control observations and observation-related metadata, and to store selected climate-relevant products. It will also provide reporting and auditing capability to allow the maturity of climate data management practices in the organization using the OpenCDMS instance.

4.3.2 User documentation is often lacking in open source systems, and where it is provided the developers often do not write it from the perspective of the users. It may be appropriate to seek contributions from data managers to write documentation, or it may be necessary to commission a technical write, funded from the Special Account.

4.4 MAINTAINING CURRENT CAPABILITY

4.4.1 This release of OpenCDMS should allow the key functionality of the franchise CDMS to be retired. Additional functionality that is not common to the majority of these systems may still need to be provided by the franchise CDMS, but the interfaces provided by OpenCDMS may allow this remaining capability to be delivered using OpenCDMS as the underpinning data store.

4.5 INCLUDING THE CDMS SPECIFICATIONS

4.5.1 This release of OpenCDMS will manage conventional climate observations (ECVs) and associated metadata. It will provide limited capability for managing products and product metadata. It is unlikely to address management of products generated outside the instande of OpenCDMS.

4.6 ADDING NATIONAL CAPABILITIES

4.6.1 This version of OpenCDMS will not address specific national needs.

4.7 SUPPORT FOR IMPLEMENTATION

4.7.1 This version of OpenCDMS has to have an automated installer that looks after configuring the software for the hardware and system environment in which it is being installed.

4.7.2 Developers who have had a broad exposure to the architecture may need to assist the first installations to identify and solve problems that arise from integrating the software with the working practices of organizations. Such experiences should be documented and the solutions shared so that subsequent installations do not need to overcome the same issues, and so that the developments can learn from the problems encountered and take account of them when designing later releases of the software.

5 STAGE 4 AND BEYOND

5.0.1 Stages 1-3 will have given OpenCDMS the basic functionality for Members to manage their observations and create products to the minimum standards required.

ANNEX 1. RESPONSIBILITIES OF THE PROJECT OFFICER

The responsibilities of the project officer will include the following.

a. Responsible for finance and budget.

b. Making sure licensing regimes, contributor agreements etc are managed actively.

c. Green Climate Fund, Regional Associations and other stakeholder communications, reporting and relationship management, with specific emphasis on ensuring Member institutions in the developing countries are fully involved.

d. Project Steering Committee – establish this and act as Secretary, including reporting on milestones and exceptions.

e. Technical Advisory Committee – act as Chair.

f. Developing the Open Source community for the project. Specific objective to increase the Contributor community based in developing country institutions.

g. Facilitating Open Source communications, peer review and ensuring pull requests etc are properly implemented.

h. Identify issues which are not sufficiently defined and seek advice from appropriate bodies, especially WMO Expert Teams and WMO Programmes. Link to related programmes, especially EUCOS.