



Australian Government
Bureau of Meteorology

Open-CDMS Roadmap

Also covering high level plans for CliDE and ClimSoft



Published by the Bureau of Meteorology

© Commonwealth of Australia 2014

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced without prior written permission from the Bureau of Meteorology. Requests and inquiries concerning reproduction and rights should be addressed to the Communication Section, Bureau of Meteorology, GPO Box 1289, Melbourne 3001. Requests for reproduction of material from the Bureau website should be addressed to [AMDISS Bureau of Meteorology](#) at the same address.

Revision history

Date	Version	Description	Author
20 Feb 2015	0.2	First Draft of Open-CDMS Roadmap to Open-CDMS version 1.0, based on outcomes of several internal Bureau workshops	Bruce Bannerman
18 May 2015	0.4	Inclusion of feedback from WMO ET-CDMS. Addition of ClimSoft roadmap. NB: Timings are indicative only and will be dependent on start date.	Bruce Bannerman Steve Palmer (UK MetOffice)

Release history

Date	Version	Status	Audience	Approval
20 Feb 2015	0.2	For comment	CLiDE Steering Committee	

Document management register

Document	File reference

Review Status

Reviewer	Date Reviewed	Version Reviewed

Release Signatories

Approval	Name	Signature	Date

Contact details

Bruce Bannerman

Manager, Geospatial Data and Access

Bureau of Meteorology

GPO Box 1289 MELBOURNE VIC 3001

Phone: +61 3 9669 4093

Email: b.bannerman@bom.gov.au

Table of Contents

1. Introduction.....	1
2. Planned future versions of Open-CDMS.....	2
2.1. Preliminary work.....	2
2.2. Open-CDMS version 0.1.....	3
2.3. Open-CDMS version 0.2.....	4
2.4. Open-CDMS version 0.3.....	5
2.5. Open-CDMS version 0.4.....	6
2.6. Open-CDMS version 0.5.....	7
2.7. Open-CDMS version 0.6.....	8
2.8. Open-CDMS version 1.0.....	9
2.9. Summary of proposed Open-CDMS v1.0 functionality.....	10
A 1. Planned future versions of CliDE.....	2
A 1.1. CliDE Version 4.0.....	2
A 1.2. CliDE Version 4.1.....	3
B 1. DRAFT – Climsoft into Open-CDMS.....	2
B 1.1. Background.....	2
B 1.2. Assumption.....	2
B 1.3. Timescale.....	2
B 1.4. Governance.....	2
B 1.5. Tools and Techniques.....	3
B 1.6. Barriers.....	3

2. ~~Planned~~ future versions of Open-CDMS

This is a high level roadmap that outlines the current high level plans for the development of Open-CDMS. It also includes discussion on CliDE and ClimSoft in Annexes A and B respectively.

The document assumes that the development of Open-CDMS is conducted with an open source community, possibly within the governance of the not-for-profit organisation, the OpenWIS Association.

At the date of writing, no substantial commitment of funding has been made beyond funding of initial planning activities by both the WMO and the Bureau. In addition, WMO are funding maintenance activities on CliDE for fy 2014/15.

Therefore, any timings that are mentioned in this document are merely illustrative and should not be relied upon.

To date, expressions of interest in contributing resources to participate in the development of Open-CDMS have been received from:

- Bureau of Meteorology, Australia
- NIWA, New Zealand
- National Meteorological Directorate, Morocco

This version of the Open-CDMS Roadmap only considers work required up until version 1.0. Additional work will be required to extend the Roadmap.

The implementation dates of Open-CDMS are indicative only and will be constrained by the level of funding and resourcing that is committed to the open source project.

NB: While open source products may be mentioned in this roadmap, changes may occur once the high level Solution Architecture and Design have been completed.

1.1. Preliminary work

Establish the infrastructure to support an open source community to develop and maintain Open-CDMS. This will occur concurrent with initial design activities.

Estimated completion date: July 2016

There are a number of tasks that will need to be undertaken to provide the Open-CDMS community with the best chance of success. Some examples are:

- Establish:
 - Open-CDMS as a project within the OpenWIS Association. This organisation can be thought of as a not-for-profit open source foundation similar to OSGeo, Mozilla and Apache.

- A contributor's agreement that assigns all intellectual property that is developed in Open-CDMS to the OpenWIS Association, together with functionality to manage these agreements.
- Project domain name, registered trade marks etc.
- Open source license(s) that the project is to utilise for software and libraries (e.g. GPL version 3 and L/GPL v3).
- Open data license that the project is to utilise for documentation, e.g. Creative Commons By Attribution, Share Alike.
- Initial Project Steering Committee
- Develop:
 - Project governance processes.
 - Source code styling guides, developer documentation etc.
 - Documentation that describes the project.
 - High level Solution Architecture, Design and project plan.
- Implement:
 - Git Source Code repository;
 - Wiki
 - Issue Tracker
 - Project email lists for Developers and Users to facilitate open communication channels.
 - Automated software build and smoke test systems
- Publicise the Open-CDMS project.

1.2. Open-CDMS version 0.1

Estimated completion date: July 2016

Open-CDMS component: 1.0 Climate-Database

This version will establish the initial database data model to be used by Open-CDMS and deploy an empty PostgreSQL database environment.

CDMS Components addressed:

- 3.1.3 Intellectual property (part)
- 3.1.5 Third party data (part)
- 3.1.2.4 Access to data (part)
- 3.1.6.1 Climate metadata (part)
- 3.1.6.2 Data lineage (part)
- 3.1.6.6 Quality Assurance (part)
- 3.2.1.1 Controlled access to data and systems (part)
- 4.1.1.1 Climate observations - Atmospheric (part)
- 4.2.1 Climate database (part)
- 4.2.2.1 Observations and Measurements (part)
- 4.3.1 Observations metadata (WIGOS Metadata Standard) (part)
- 4.3.3 Data provenance metadata (PROV-O) (part)

- 5.4.1.3 Multi-layer quality flags (part)
- 9.3.3 Security (part)

Functionality to be implemented:

- Establish the Open-CDMS data model that caters for:
 - Time series atmospheric meteorological observations;
 - Multiple sensors, per phenomena, per station;
 - Multiple tiers of quality assurance, from initial observations through to homogenised data;
 - WIGOS Observations Metadata;
 - Data Provenance Metadata, including point in time recovery;
 - Spatial enablement constrained to point observation based on ISO 19156 Observations and Measurements;
 - Intellectual Property inherent in data, including issues such as third party data, commercial data, data licensing, data usage constraints etc.; and
 - Implementing a robust security model.
- Establish a Debian Package to deploy an empty Open-CDMS database environment.

1.3. Open-CDMS version 0.2

Nominal completion date: December 2016

Open-CDMS component: 2.0 Ingest-and-extract

This version will establish the initial Ingest and Extract capability for Open-CDMS. This functionality is expected to **extend** the capability found in existing open source Extract Transform and Load (ETL) software such as GDAL/OGR, Talend or similar.

Open-CDMS Dependencies: 1.0 Climate Database

CDMS Components addressed:

- 3.1.3 Intellectual property (part)
- 3.1.5 Third party data (part)
- 3.1.6.1 Climate metadata (part)
- 3.1.6.2 Data lineage (part)
- 5.4.1.3 Multi-layer quality flags (part)
- 5.1.1.1 Business rules (part)
- 5.1.1.2 WMO Messages
- 5.1.1.5 Other formats (part)
- 5.1.1.6 Status log (part)
- 5.1.1.7 Automated with self recovery (part)
- 5.1.1.8 Transformation
- 5.1.2.1 Data extraction
- 9.3.3 Security (part)

Functionality to be implemented:

- Implement one or more open source ETL components to provide ETL capability.
 - Due diligence will be required to find the best technology fit, together with the open source community most receptive to working with Open-CDMS community for required additional functionality.
- Extend the ETL for specific CDMS requirements, e.g.:
 - WMO formats (TDCF and TAC)
 - AWS
- Establish ability to import existing CDMS Data:
 - CliDE
 - Climsoft
 - MCH
- Ingest according to defined business rules. Support for:
 - Data provenance
 - Observations Metadata
 - Data Intellectual Property
 - A range of climatological and technical business rules
- Extract data to:
 - selected WMO Product formats (e.g. CLIMAT, WWR, BUFR)
 - selected emerging formats (JSON, CF-netCDF)
 - comma delimited text

1.4. Open-CDMS version 0.3

Nominal completion date: December 2016

Open-CDMS component: 3.0 GUI

This version will establish the initial User Interface and Climate data presentation capability for Open-CDMS. This functionality is expected to utilise the capability found in existing open source products such as:

- Business Intelligence (BI) software such as Birt, Jasper or similar.
- Existing JavaScript components such as Highcharts may also be suitable.
- The Desktop GIS application (QGIS).
- The statistical analysis application (R).

Open-CDMS Dependencies:

- 1.0 Climate-Database
- 2.0 Ingest-and-extract

CDMS Components addressed:

- 7.1 Graphical User Interface (part)
- 7.1.1 Tables and Charts
- 7.1.3.1 Cartography (part)
- 6.1.2.1 Spatial Analysis (part)
- 6.1.2.2 Image Analysis (part)
- 6.1.2.3 Time series analysis (part)
- 9.3.3 Security (part)

Functionality to be implemented:

- Implement a BI application or similar to provide the functionality to a range of generate graphs, tabular reports, charts etc.
 - Due diligence will be required to find the best technology fit, together with the open source community most receptive to working with Open-CDMS community for required additional functionality.
- Implement QGIS to provide initial Desktop GIS, cartographic, spatial and image analysis capability.
- Implement R to provide initial statistical analysis capability.
- Establish the initial user interface, derived from CliDE.

1.5. Open-CDMS version 0.4

Nominal completion date: June 2017

Open-CDMS component: 4.0 Climate-metadata

This version will establish the initial climate metadata for Open-CDMS. Priority is aimed at implementing support for Observations metadata and Data Provenance metadata only.

Open-CDMS Dependencies:

- 1.0 Climate-Database
- 2.0 Ingest-and-extract
- 3.0 GUI

CDMS Components addressed:

- 7.1 Graphical User Interface (part)
- 5.5.1 Manage climate metadata (part)
 - 5.5.1.1 Create climate metadata (part)
 - 5.5.1.2 Maintain climate metadata (part)
 - 5.5.1.3 Quality control climate metadata (part)
 - 5.5.1.4 Metrics of climate metadata (part)
- 9.3.3 Security (part)

Functionality to be implemented:

- Implement functionality to manage observations metadata:
 - If required, extend the Data Model used by Open-CDMS Component 1.0 Climate-database to support the Observations Metadata capability that is implemented by this component.
 - If required, extend the Open-CDMS 2.0 Ingest-and-extract Component to support the WIGOS Observations Metadata Standard. This includes
 - Importing Observations Metadata.
 - Exporting Observations Metadata for inclusion within WMO's OSCAR Observations Metadata database.
 - Extend the Open-CDMS GUI to facilitate the effective management of Observations Metadata.
- Implement functionality to manage data provenance metadata:
 - If required, extend the Data Model used by Open-CDMS Component 1.0 Climate-database to support the Data Provenance Metadata capability that is implemented by this component. This is to be based on the W3C PROV standard.
 - Extend the Open-CDMS GUI to facilitate the effective management of Data Provenance Metadata.

1.6. Open-CDMS version 0.5

Nominal completion date: June 2017

Open-CDMS component: 5.0 Data-Rescue

This version will establish the initial Data rescue capability for Open-CDMS.

Open-CDMS Dependencies:

- 1.0 Climate-Database
- 2.0 Ingest-and-extract
- 3.0 GUI

CDMS Components addressed:

- 7.1 Graphical User Interface (part)
- 5.2.2.1 Data rescue metrics
- 5.2.3 Data entry
- 5.2.3.1 Forms
- 5.2.3.2 Key entry
- 5.2.3.3 Computation
- 9.3.3 Security (part)

Functionality to be implemented:

- Implement Data Rescue capability.
- Ensure that graphical forms can be easily customised while retaining the underpinning capability for:
 - Key entry
 - Computation.
- Implement functionality and if required, data model changes to support Data rescue metrics capability.
- Ensure that the following requirements are observed and maintained by the Data Rescue components:
 - data provenance of entered records
 - observation metadata for the sensor and station is either present in the database, or captured prior to input of the data.
 - data IP details are either present in the database, or captured prior to input of the data.
 - Initial quality assurance flags are appropriately established.
 - An image of each data entry form is stored in the Climate database with appropriate linkages to the relevant station and sensor.

1.7. Open-CDMS version 0.6

Nominal completion date: December 2017

Open-CDMS component: 6.0 Quality-Management

This version will establish the initial climate data quality management capability for Open-CDMS.

Open-CDMS Dependencies:

- 1.0 Climate-Database
- 2.0 Ingest-and-extract
- 3.0 GUI
- 4.0 Climate-metadata

CDMS Components addressed:

- 7.1 Graphical User Interface (part)
- 3.1.6.6 Quality Assurance (part)
- 3.2.1.1 Controlled access to data and systems (part)
- 5.3.1.1 Consistency Checks
- 5.3.1.6 Data Recovery
- 5.3.2.1 Network Monitoring
- 5.4.1.1 Siting classification
- 5.4.1.3 Multi-layer quality flags (part)
- 9.3.3 Security (part)

Functionality to be implemented:

- Establish an extensible data quality management framework for Open-CDMS.
- Define the multi-layer quality flags to be used by Open-CDMS
 - If appropriate WMO defined quality flags are available, adopt them.
 - If not, define and clearly document what each quality flag means.
- Ensure capability for:
 - Site Administrator to select tests that appropriate for the climate networks being managed.
 - Consistent use of multi-layer quality flags.
 - Consistency Checks.
- Implement functionality to recover and maintain climate data, ensuring that:
 - Site Administrator can tightly control who has write access to the climate database.
 - Each edit has appropriate data provenance records covering the change.
 - Point in time recovery of data is maintained.
 - Metrics and details of changes for a given time period can be easily reviewed, and if required selected changes reversed.
- Implement Network Monitoring and Siting classification capability.

1.8. Open-CDMS version 1.0

Nominal completion date: December 2017

Open-CDMS component: 3.1 GUI

This version will establish the initial operational version of Open-CDMS.

Open-CDMS Dependencies:

- 1.0 Climate-Database
- 2.0 Ingest-and-extract
- 3.0 GUI
- 4.0 Climate-metadata
- 5.0 Data-Rescue
- 6.0 Quality-Management

CDMS Components addressed:

- 7.1 Graphical User Interface (part)
- 7.1.4.2 Integrated search of observations data and observations metadata
- 3.2.1.1 Controlled access to data and systems (part)
- 5.4.1.3 Multi-layer quality flags (part)
- 9.3.3 Security (part)

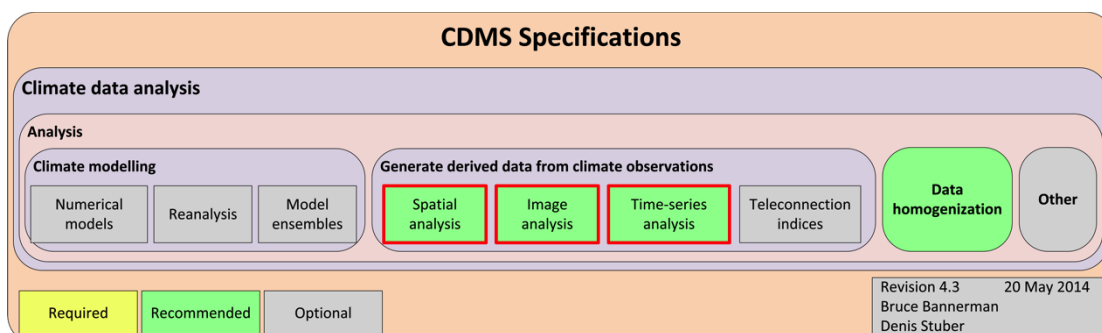
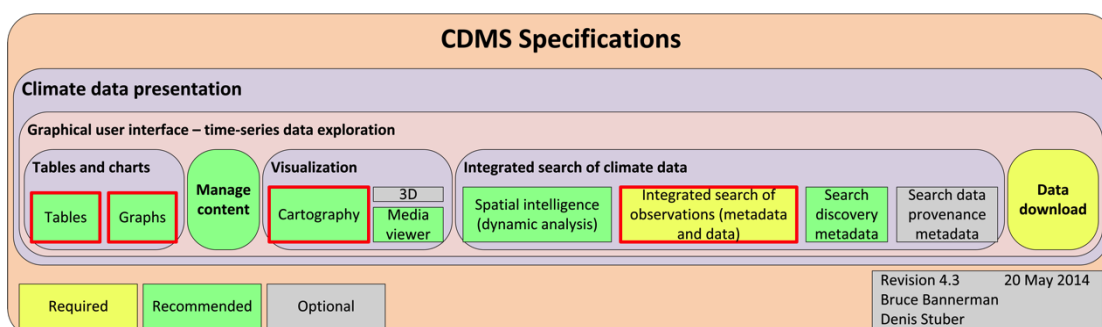
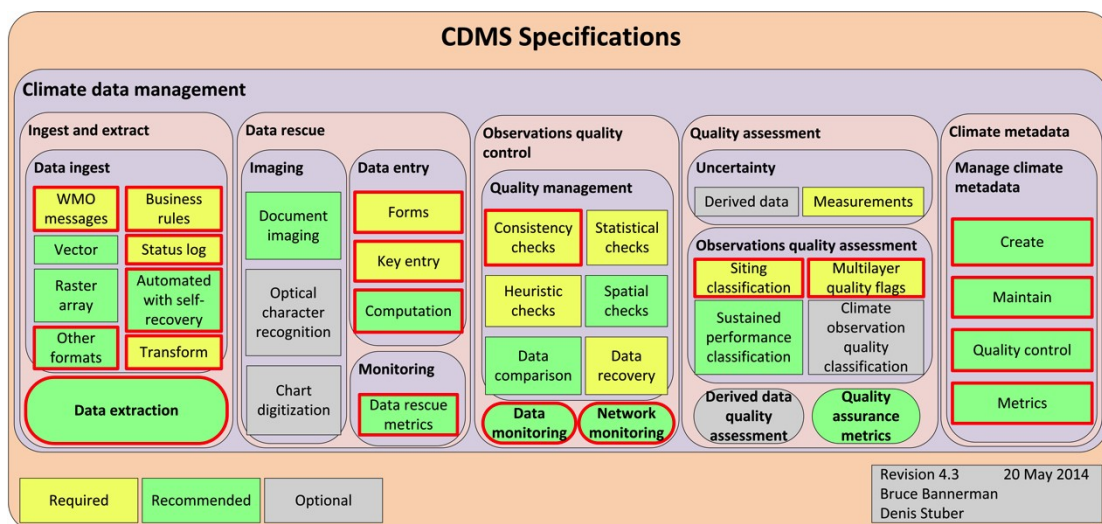
Functionality to be implemented:

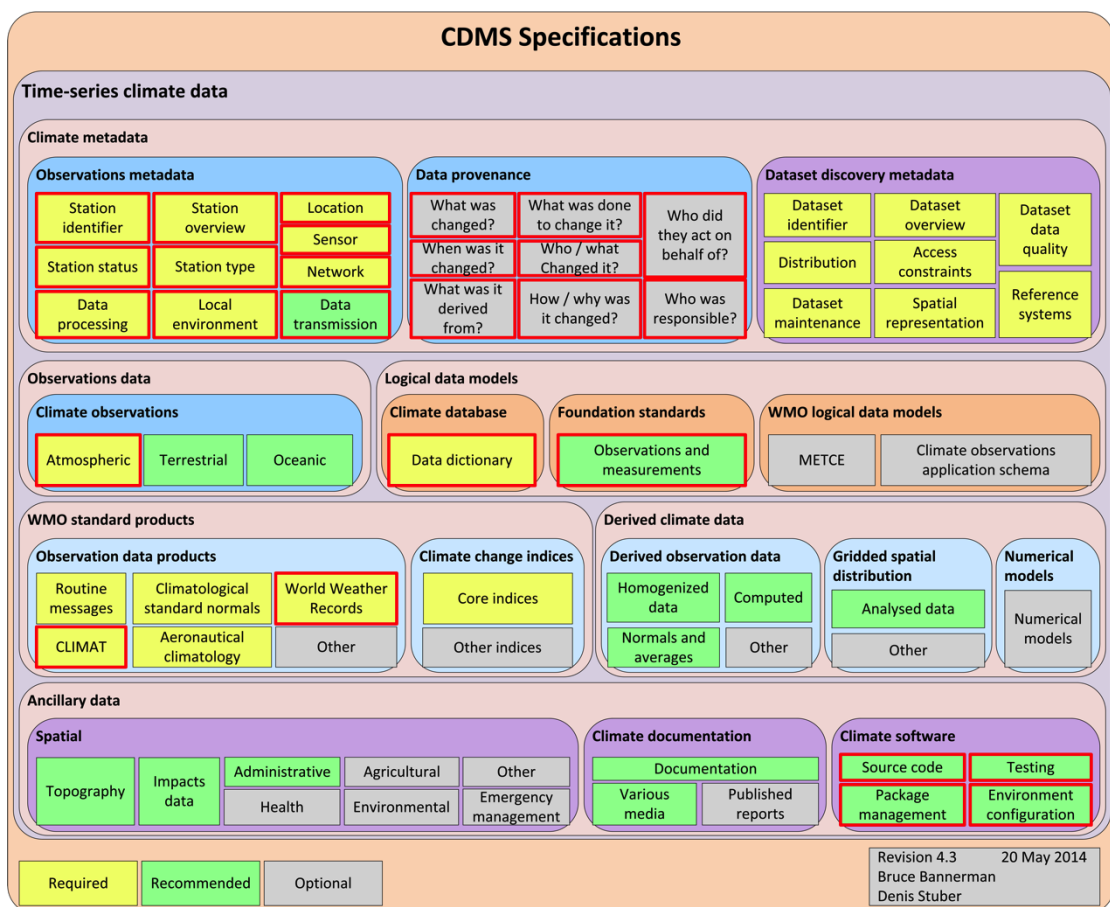
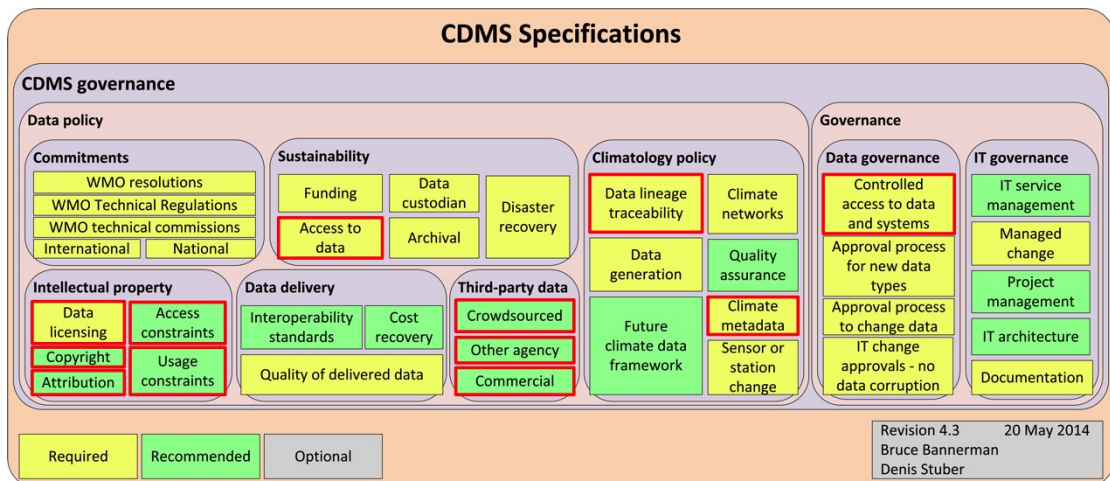
- Establish the initial operational release of Open-CDMS. Of note:
 - Extend the GUI to support integrated search of observations data and observations metadata (tabular, graphs and charts only).
 - Ensure that all components are well integrated.
 - Ensure that all components have automatic self-test capability to support remote Incident Management without the need for a site visit.
 - Ensure that Open-CDMS is thoroughly tested.
 - Implement a post installation wizard type interface to allow default systems administration settings to be over-written, e.g.:
 - Automated package updates
 - Backups
 - Ensure that Open-CDMS can be successfully deployed using standard Ubuntu package management processes, without expert intervention.

1.9. Summary of proposed Open-CDMS v1.0 functionality.

The diagrams below summarise the WMO CDMS Specification components that are addressed as at Open-CDMS version 1.0.

NB: In many cases, only part of the required functionality of components is actually implemented.





Annex A

CLiDE and Open-CDMS

A 1. Planned future versions of CliDE

Current version of CliDE is 3.3

A 1.1. CliDE Version 4.0

Estimated release date: July 2015

This version of CliDE will introduce some changes to the database data model and address ease of use issues, together with addressing priority issues that were identified from a user survey.

Additional Functionality to be implemented:

- Software fixes for greater system stability.
- Establish Debian Package Management capability to support future automated upgrades (provided that the product has not been customised).
- User Interface improvements to improve the ease of use of CliDE.
- Establish capability to summarise database data holdings for each station and sensor.

Issues:

- There are some differences in the version of both CliDE and its data model at several NMHS.
 - Manual intervention may be required to upgrade to version 3.4 of CliDE and its database.
- Several NMHS have made customisations to the version of CliDE that they have deployed.
 - NMHS that have customised CliDE will need to analyse the differences between their customisation and CliDE version 3.4 and upgrade the product themselves.
- Many Pacific Islands servers are at end of life with no funding to replace them.

A 1.2. CliDE Version 4.1

Estimated release date: July 2016

This version is expected to be the last major version of CliDE. An upgrade path will be provided to Open-CDMS when that product has reached sufficient functionality and maturity.

Additional Functionality to be implemented:

- Create additional data entry forms, if required.
- Generate WMO Standard Products, CLIMAT and World Weather Records.
- Create additional time series graphs and tabular reports if required.

B 1. DRAFT – Climsoft into Open-CDMS

Annex B ClimSoft and Open-CDMS

Steve Palmer

11 March 2015

B 1.1. Background

This paper explores how the Climsoft Project could be subsumed into the Open-CDMS Project being discussed under WMO CCI OPACE1 Expert Team on Climate Data Management Systems.

B 1.2. Assumption

It is the responsibility of all National Meteorological and Hydrological Services to maintain and promote use of their national climate data, including all available past and future observation information. This implies that all NMHS should be operating a CDMS which meets the required functionality defined by WMO and approved by its Members.

B 1.3. Timescale

The Climsoft Project has declared that it is moving to an open-source mode.

- The current Version 3.2 is now in use in all the active user NMHSs. This uses the MS-Jet RDBMS engine, with a background database, typically in MySQL, though other RDBMS engines can be used.
- Version 4 is being developed, with a target of end of 2015. This uses open-source tools and techniques, albeit that the core functionality is being developed in VB.NET. Version 4 will remove the MS Jet RDBMS engine and address the (former background) RDBMS directly, with minimal changes to the data model used in Version 3.2. Work under the Version 4 development will produce a first version of a data access layer which will support incorporation of techniques and products developed in R and other languages. Only the core functionality will remain in VB.NET.
- Version 5 will be developed during 2016, using fully open-source tools and techniques. Changes to the data model are envisaged.

Comparison with the draft Open-CDMS roadmap shows that convergence could be established with Version 5 of Climsoft in 2016, and full integration achieved in 2017.

B 1.4. Governance

The structure envisaged in the Open-CDMS Roadmap is similar to that already in progress for Climsoft. In particular:

- The Climsoft Project is established with a Steering Group for strategic management, and a Technical Advisory Group who are responsible for monitoring coding standards and technical changes, and signing off on all merges into the core of the Climsoft Project.
- The Climsoft Project has established roles for Contributors and Users.
- A draft Contributors Agreement has already been written, and is expected to be agreed at the next Steering Group meeting.
- The Climsoft Project is exploring how to vest ownership of the source code (as a development from the current agreement which licenses this to WMO Members) and is in discussion with the Open-WIS Consortium.

It will be seen that in most respects the Climsoft Project is already in progress to establish similar governance arrangements to those envisaged for the Open-CDMS project.

B 1.5. Tools and Techniques

The Climsoft Project is in discussion with other teams about closer integration with their developments. In particular, the Reading University Statistics In Applied Climatology course is being rewritten in R, and these modules will have a direct interface to a Climsoft database, with the advantage that SIAC students can work and learn directly on their own national data, and how to develop their own products and services.

The TAMSAT database is being rewritten, and these tools and techniques can be interfaced to Climsoft V4, raising the possibility of much improved gridded data for rainfall.

A group within the Met Office is currently evaluating tools for Extreme Value Analysis; there is an agreement that code resulting from this work should use a defined data access layer, which could be easily interfaced to Climsoft, thus enabling other NMHSs to use state-of-the-art techniques for EVA, which is especially important for risk analysis which is essential to the role of an NMHS in disaster warning.

Resources for the Climsoft Project include advice from people with experience in Open-WIS, Open-GIS and the IRIS Project.

Support for multiple languages for the user interface is essential. Climsoft already has this ability through use of a translation table; this should be mandatory on all code offered to the core of Open-CDMS.

B 1.6. Barriers

The main risk to the migration of the Climsoft Project into the Open-CDMS project is that decisions are made in the planning stages of Open-CDMS which are seen by the Climsoft Project Steering Group as barriers to entry. One barrier would be over technology which does not recognise existing skills and experience (e.g. a decision to exclude VB.NET from open-source developments). Another barrier would be participation in the governance of the Open-CDMS project, especially an imbalance between developed-country members and developing country members, when the majority of initial users of the Open-CDMS are expected to be developing country NMHSs.

