The CHARMe Project: Commentary Metadata for EO Datasets
Kevin Marsh1, Philip Kershaw2, Jon Blower2 and Paul Van Der Linden3
1NCEO Centre for Environmental Data Archival, STFC, RAL, Oxford, UK,
2University of Reading, UK, 3UK Met Office

Background
Accurate, long-term monitoring of the Earth is of vital importance for gathering information about the climate. This information, in turn, forms an important part of the evidence base for operational and policy decisions that have far reaching effects on society. Climate data are used by both public and private sectors for applications such as controlling greenhouse gas emissions, energy production, food security and flood prediction.

The CHARMe system will provide links between datasets and citations and other commentary from many systems developed so far.

Objectives
The main objectives of CHARMe are:
1) To develop an open-source system for recording “commentary metadata” that links with climate datasets and other sources of information such as descriptions of sensors and instruments.
2) To provide interfaces for commentary metadata to be entered, queried and displayed through existing community websites, and through machine-readable interfaces.
3) To identify, and engage with, key strategic stakeholders (including climate data users, producers and high-level global initiatives), and to ensure that the CHARMe concept is understood and supported by both providers and users of climate data.
4) To develop tools that demonstrate other ways in which commentary metadata can be produced and exploited in a variety of scenarios.

There will be a collection of CHARMe nodes to store C-metadata, which will have many different clients. The ESF METAFOR / ES-DOC system for climate model descriptions provides some ideas. It uses a lightweight plugin to existing sites, and CHARMe will use a similar approach. The CHARMe system will also include a Significant Events viewer, a Faceted Search tool and an Intercomparison tool.

Linked Data and Open Annotation
Linked data is now used by many organisations including data.gov.uk, Australian Bureau of Meteorology, and the UK Met Office to enable more context to be given to the information they hold. CHARMe will use Open Annotation (W3C standard) for modelling annotations, as it is an ontology which is well suited to the CHARMe project requirements. Using this, CHARMe will allow users to record the motivation behind an annotation.

Linked data is now used by many organisations including data.gov.uk, Australian Bureau of Meteorology, and the UK Met Office to enable more context to be given to the information they hold. CHARMe will use Open Annotation (W3C standard) for modelling annotations, as it is an ontology which is well suited to the CHARMe project requirements. Using this, CHARMe will allow users to record the motivation behind an annotation.

The CHARMe system will provide links between datasets and citations and other commentary information using open standards such as Open Annotation. This approach is highly flexible and can adapt to changing user needs as the system grows.

Commentary Metadata
Commentary metadata allows the annotation of datasets, and in CHARMe this includes:
- Post-fact annotations, e.g. citations, ad-hoc comments and notes;
- Results of assessments, e.g. validation campaigns, intercomparisons with models or other observations, reanalysis;
- Provenance, e.g. dependencies on other datasets, processing algorithms and chain, data source;
- Properties of data distribution, e.g. data policy and licensing, timeliness, reliability;
- External events that may affect the data, e.g. volcanic eruptions, El-Nino index, satellite operational changes to the orbit calculations.

Commentary metadata enables both data providers and data consumers to share commentary information linked to the data.

The CHARMe system will collect and share annotations associated with climate datasets, and information is called commentary metadata (“C-metadata”). Users will be able to add or view commentary metadata, to complement existing information from data providers. The term C-metadata comes from the ABCDE metadata taxonomy developed by Lawrence et al. Support for C-metadata has been lacking in existing community websites, and this information is called commentary metadata.

Commentary metadata enables both data providers and data consumers to share commentary information linked to the data.

The CHARMe project aims to connect users, and prospective users, of climate datasets with the previous expertise that has accumulated in the community, and enable them to contribute back information on their own experiences.

The CHARMe project aims to connect users, and prospective users, of climate datasets with the previous expertise that has accumulated in the community, and enable them to contribute back information on their own experiences.

The use of Linked Data technologies will also enable CHARMe information to be discovered and used by both public and private sectors for applications such as controlling greenhouse gas emissions, energy production, food security and flood prediction.

The CHARMe project aims to connect users, and prospective users, of climate datasets with the previous expertise that has accumulated in the community, and enable them to contribute back information on their own experiences.

The CHARMe system will also include a Significant Events viewer, a Faceted Search tool and an Intercomparison tool.

There will be a collection of CHARMe nodes to store C-metadata, which will have many different clients. The ESF METAFOR / ES-DOC system for climate model descriptions provides some ideas. It uses a lightweight plugin to existing sites, and CHARMe will use a similar approach. The CHARMe system will also include a Significant Events viewer, a Faceted Search tool and an Intercomparison tool.

The CHARMe system will provide links between datasets and citations and other commentary information using open standards such as Open Annotation. This approach is highly flexible and can adapt to changing user needs as the system grows.

The CHARMe system will provide links between datasets and citations and other commentary information using open standards such as Open Annotation. This approach is highly flexible and can adapt to changing user needs as the system grows.

Acknowledgments: With thanks to The CHARMe Consortium

Kevin Marsh, Philip Kershaw, Jon Blower and Paul Van Der Linden

1NCEO Centre for Environmental Data Archival, STFC, RAL, Oxford, UK,
2University of Reading, UK, 3UK Met Office

The role of CEDA
CEDA includes BADC and NEODC, and provides professional data management support to NCAS and the research community. CEDA also ensures that relevant NCAS and NCEO data are securely archived for both present and future use. Over 2PB of climate, ED and weather datasets (inc. ground-based, satellite and modelsynthesis data) are available for use by academic researchers. CEDA is directly involved in many of the CHARMe work packages.

Summary
CHARMs will create connected repositories of “commentary metadata” and help users tap into existing expert knowledge about climate datasets.
CHARMs will provide this information in existing archives and websites. However, nothing in the project is really specific to climate, and can be applied to other data archives.

The use of Linked Data technologies will also enable CHARMe information to be discovered and used in other systems.