Characterisation of metadata to enable high quality climate applications and services

Deliverable 300.3

Gap Analysis

CHARMe is funded by the EC under its FP7 Research Programme

Document Control

Contributors

Person	Role	Organisation	Contribution
R. Alegre	Project Researcher	UREAD	Initial Draft, contributions to all sections of the document.
P. Kershaw	Project Researcher	STFC	Contributions to all sections of the document, guidance, comments and corrections.
J. Blower	Project Researcher	UREAD	Document redistribution, guidance, comments and corrections.

Document Approval

Person	Role	Organisation
R. Phipps	Document Reviewer	UREAD

List of Acronyms

CHARMe	CHARacterisation of Metadata
CLI	Command-Line Interface
CMIP5	Coupled Model Intercomparison Project Phase 5
ESGF	Earth System Grid Federation
ncWMS	netCDF Web Map Service
RDF	Resource Description Framework
REST	Representational state transfer
RSS	Rich Site Summary
SPARQL	SPARQL Protocol And RDF Query Language
SRD	Software Requirements Document
TBD	To Be Defined
UC	Use Case
URD	User Requirements Document
WP	Work Package

References

ID	Author	Document Title	Date
[1]	STFC	Analysis of Existing Technical Solutions	14th May 2013
[2]	UREAD	User Requirements Document	16th July 2013

[3]	CGI	Software Requirements Document	TBD
[4]	CHARMe Consortium	CHARMe: Description of Work	2nd Oct 2012
[5]	R. Phipps	CHARMe Risk Register	17th Sep 2013

Revision History

Issue	Author	Date	Description
0.1	R. Alegre	22th July 2013	Initial Draft.
0.2	R. Alegre P. Kershaw	28th Aug 2013	Implementation of comments by STFC. Sections 2.1.7 to 2.1.15 reworded and extended.
1.0	R. Alegre J. Blower	17th Oct 2013	Mayor changes in the document: Document layout following Jon's comments: division of gaps by technical areas, not use cases. Both gaps and risks listed for each of the technical areas. Glossary of Technical Terms added as asked by the project coordinator.

Table of Contents

1 Introduction

An analysis of the results of the User Requirements Document [2] and the Analysis of the Existent Technical Solutions Document [1] may reveal use cases or requirements that cannot be fulfilled by existing technical solutions, and use cases for which no clear best-practice solution exists.

The purpose of this document is to guide the CHARMe project towards **areas that need particular attention and innovation**: these areas also present the greatest risks to the project and so will feed into the project's risk register.

This document will help identify which user requirements are being met and which might require additional effort. It is an overall assessment of the extent of the gaps or unknown areas that may pose a problem or cause delays during the course of the project.

The gap analysis provides key information to several stakeholders in the CHARMe project:

- The **scientific coordinating stakeholders**, specially the project coordinator, are provided with improved **awareness of the strength, limits and vulnerable areas** of the project development, including an input to update the project's risk register. The final list of gaps will be reviewed at the WP200's Scientific Coordination Meeting.
- **Software developers** are provided with further information on **areas where special attention and possible improvements** are needed to comply with the User Requirements. It will also help identifying specific areas for enhancement to guide the development of later versions of the software, and thus should be taken into consideration when preparing CHARMe's Software Requirements Document [3].

1.1 Gaps and Risks

The "**gaps**" in the context of this document represent the inability of the existing technical solutions considered in the Analysis of Existing Technical Solutions Document [1] to meet the user requirements detailed in the User Requirements Document [2]. These can be categorized as gaps in implementation, knowledge or expertise and we'll also assess on their impact in terms of:

- **Low (L)**, when no significant additional work needs to be employed to cover the gap or no major consequences that may affect the functioning of the technical area are foreseen.

- **Medium (M)**, when further research is needed and several components of the projects may be affected.
- **High (H)**, when the gap could pose a risk affecting several parts of the system's development and needs to be raised at a coordination meeting.

We will also consider in this document "**risks**" that may affect the repercussion caused by the gaps discovered during this analysis. This document will just raise them and present them for them to be included in the CHARMe Risk Register [5] for future reference, discussion and assessment.

In the next section of this document, we'll explore in more depth which are the gaps and risks foreseen so far, and the different impacts the gaps can have in the development of the project.

2 Technical Areas

The Use Cases detailed in [2], imply that the CHARMe System consists of a group of **technical areas**. The exact characteristics of these areas are not fully defined at this stage of the project, but they will be specified in the Software Requirements Document [3], that takes this document as input.

In this section, we'll go through the list of broad categories of technical areas, detailing which technologies will be considered for their implementation. The Use Cases listed in [2] that may be related with each of these areas are depicted in the following table:

Technical Areas	UC-100	UC-101	UC-102	UC-103	UC-104	UC-105	UC-106	UC-107	UC-200	UC-201	UC-202	UC-203
Data model	X	X	X	X	X	X	X	X	X	X	X	X
Storage	X	X	X	X	X	X	X	X	X	X	X	X
Geospatial Referencing			X						X			X
Network Protocol		X	X	X	X	X	X					
User Interface	X	X	X	X	X	X	X	X	X	X	X	X
Authentication	X				X	X	X	X				X
Faceted Search Tool									X	X	X	X
Significant Events' Viewer									X	X	X	X
Inter-Comparison Tool									X	X	X	X
Fine-Grained Commentary Tool									X	X	X	X
Other: Initial population of CHARMe's DB		X	X	X					X	X	X	X
Other: Need for on-going moderation of user input					X	X	X					

2.1 Data model for annotations

We have decided on a **Linked Data** approach based on **RDF** and **Open Annotation**. We are confident [1] that Open Annotation is a suitable data model at a high level, although we need more detailed data models for particular types of annotation bodies (e.g. to record citations, or fine-grained commentary). These more detailed data models do not generally exist and are therefore gaps that we must fill with R&D in this project. In some cases we may need to develop particular ontologies.

This is an essential component in this project and thus needs special attention, as it will affect all use cases.

ID	Category	Description	Impact
Gap 2.1 - 1	Expertise	Need to develop specific custom ontologies to complement general data model.	H

ID	Description
Risk	None foreseen.

2.2 Storage for annotations

Two different technical solutions have been considered to implement this component:

- We can use a **native triplestore**, which has been defined in [1]. This would be a very flexible solution, but it might compromise performance as we haven't foreseen how exhaustive CHARMe users will interact and add new data to it. As the triplestore gets larger, the performance may get slower.
- We have also considered the implementation of a **relational database**, which would be much less flexible than a triplestore, but it is more mature and more scalable. This solution would also need a **translation layer for RDF**, which will need the developers to research on its exact capabilities. Also, the maturity of this translation layer solution could be a possible gap.

Although there are no major software gaps here, as both solutions are readily available, there are gaps in the designers and developers' knowledge and experience. Also, malfunctioning of this component would affect all use cases, so special attention is needed when designing this component.

The trade-off solution is to select a native triplestore implementation for flexibility in the development stage, but be mindful of the possible need to find alternative solutions as the CHARMe system grows beyond estimated use in the duration of the CHARMe project and which may compromise the performance.

ID	Category	Description	Impact
Gap 2.2 - 1	Knowledge Expertise	Further research on storage solutions by CHARMe's development team needed.	H

ID	Description
Risk 2.2 - 1	The amount of data in the CHARMe triplestore escalates to an extent that the system's performance is compromised and/or causes unexpected problems.

2.3 Geospatial referencing for annotations

CHARMe's Fine-Grained Commentary Tool requires annotations to be geospatially and temporally referenced, as indicated in [2].

In addition to the need to develop a data model for the annotations detailed in the previous section, we may need a means of storing geospatial information in an efficient way that makes it easy to do geospatial and temporal filtering.

There are solutions available, but not yet tested by CHARMe's development team. One example could be the **Strabon database**.

An alternative solution is to run a triplestore in parallel with a geospatial database such as **PostGIS**, which offers more maturity and it's been widely used.

Hence the software solutions for this component already exist in the market, although there is a current gap in our knowledge that will be investigated during the design and development of WP740.

ID	Category	Description	Impact
Gap 2.3 - 1	Expertise	CHARMe's developer team has little experience on this subject and further research is needed.	M

ID	Description
Risk	None Foreseen

2.4 Network protocol for exchanging annotations

The **SPARQL** protocol is provided "for free" with most triplestores, although for convenience we may require a REST interface too.

This **REST** interface may pose a possible gap in our expertise, as this option has not been yet studied for CHARMe's particular case. However, some solutions exist that will require further testing, as in the case of **Linked Data API**

We may also need geospatial network capabilities for WP740 (CHARME's Fine-Grained

Commentary Tool) and this may be provided by **GeoSPARQL**, which is an OGC standard, or **stSPARQL**, created by the developers of Strabon.

We can assume that once CHARMe is released, different communities of climate data users as well as data providers and external automated systems will be able to perform concurrent queries and submit information in the CHARMe Nodes. The network's efficiency will be tested during the development stages, however it is difficult to gauge exact network requirements at this stage.

ID	Category	Description	Impact
Gap 2.4 - 1	Expertise	Further research on the REST interface needed.	L

ID	Description
Risk 2.4 - 1	The CHARMe system's efficiency is compromised when trying to cope with several parallel queries or interactions with the data stored in the CHARMe Nodes.

2.5 User interfaces

CHARMe uses will interact with the systems via different user interfaces:

- **CHARMe Plug-In's GUI:** this tool will feature most of the use cases (UC-100 to UC-107) using web development technologies and a REST interface with the CHARMe repository. This would be similar to the **Javascript** plug-in concept used in Metafor. However, data providers may be concerned with the installation of a Javascript plugin at their servers because it **might compromise the security** of their site and expose information to unknown or malicious users, as CHARMe registration is not needed to just view and filter data. To mitigate this, the data providers in the CHARMe consortium have been asked to provide a list of constraints on system integration that will be taken into account during the design of the CHARMe Plug-In. These constraints will be included in the Software Requirements Document [3].

- **Command-Line Interface:** advanced CHARMe users that need a more exhaustive interaction with the data stored in the CHARMe nodes may prefer doing so via a CLI, which will provide an interface for users to submit queries to the system using **RDF, Turtle** or **ldJSON**.

- Advanced user interfaces for the **WP700 tools**: These will be explained in more detailed in Section 2.7 to 2.10 of this document.

There is a major gap in this technical area: the interfaces will largely have to be developed from scratch, and that could raise unexpected gaps and risks that might translate into delays in the software deliveries or results below expectations. Although there are precedents we could follow, e.g. ES-DOC, in most cases, significant custom development would be needed. To reduce the impact of this gap, the user interface developers will devote some time to do a detailed design of their applications prior to development.

ID	Category	Description	Impact
Gap 2.5 - 1	Expertise	Most of the user interfaces to be developed from scratch.	H

ID	Description
Risk 2.5 -1	The CHARMe Plug-In may compromise the Data Provider's site's security and inadvertently expose information to unknown or malicious users.

2.6 Authentication/Authorization system

User registration and add/modify/delete operations raise the issue of authentication and authorisation. To make changes to the repository, the user needs to be registered so that they can authenticate and also, they must be assigned **access privileges** so that an authorisation policy can be enforced determining which actions they can and cannot execute. For example, a System Moderator is likely to have a more full set of access rights than a CHARMe User.

It is still unclear at this stage of the project how to manage the user registration for actors [2] other than CHARMe Users, i.e.: system moderators (including third parties subject to approve or discard user requests quarantined by other system moderators), system administrators and remote systems. What their privileges exactly are and who will be responsible for granting them will be decided in later stages of the project.

Another pending issue is that of **security domains**. This is twofold. Consider firstly, the deployment of a CHARMe plugin at a Data Provider site. The plugin is deployed within the security domain of the Data Provider. Resources under the authority of the Data Provider need to be secured with the Data Provider's access policy. In many cases, this may already be in place. For example, for ESGF, CMIP5 data is secured using ESGF's security system and particular access policy for this data. However, the CHARMe plugin accesses resources at the CHARMe node it links to. This node may be under an independent administrative domain to the Data Provider site where the plugin is deployed. When a user interacts with the plugin their actions will need to be secured by the access control system at the CHARMe node.

The second issue is that of the deployment of the CHARMe infrastructure. In the proposal document a system of multiple CHARMe nodes was envisaged. In this scenario, each node may be under an independent administrative domain. To enable users to use such a system, some means of linking the authorisation policy across these domains is desirable. Likewise, a means of **single sign-on** would enable the user to use the same authentication credentials to access resources in all the nodes. Systems to manage access policy in a distributed environment have been implemented in many Grid-based infrastructures using the concept of a **Virtual Organisation** to centrally manage access privileges across a number of distributed nodes. ESGF uses such a system. A number of different single sign-on technologies are also available such as **OpenID** and **Shibboleth**.

We also need to consider in this section **UC-107**, which allows CHARMe Users interested in a particular dataset to subscribe to it and be updated when the CHARMe metadata available for that dataset changes. This could be done by registering the user's email address or registering the user's CHARMe ID, using **RSS**, **Atom**, etc. Which method is preferred to accomplish this will be decide during the Software Requirements analysis. There is a pending issue however on **data privacy**. The user may need to release their e-mail address to CHARMe and therefore will need to trust the system.

ID	Category	Description	Impact
Gap 2.6 - 1	Knowledge	Access control required to secure access to the repository by different types of user profiles not yet conceived.	M
Gap 2.6 – 2	Knowledge	Management of security administrative domains.	M
Gap 2.6 – 3	Implementation	User privacy: it should be possible to formulate statement of terms of use of system and how user's data will be used.	L

ID	Description
Risk 2.6 – 1	Risk that user registration authentication and security domains not clarified at an early enough stage in the design process.

Risk 2.6 – 2	Risk that CHARMe plugin will not provide the required security and that a full CHARMe node will have to be implemented per site.

2.7 Faceted Search Facility

The Faceted Search Tool will allow users to find Commentary metadata stored in the CHARMe Nodes using facets. There are many ways in which this search can be enabled, but many methods do not scale well to large varieties of metadata holdings.

The CHARMe metadata will be saved in a **triplestore** along with a **NoSQL database** to index the elements stored in a triple store for efficient search.

The Faceted Search Tool will be based upon a successful facility deployed for the ESGF using **Apache SOLR** for searching and indexing textual data available in the triple store, which has proven to work well in well-known sites like the BBC World Service or the ENVILOD project.

The developers of this facility have extensive experience in this matter, and no major gaps are foreseen. However, some of the tools in WP700 have common areas of development and there is a need to study how this tool will be integrated with the others in this WP. This analysis will take place during the elaboration of a detailed design of the WP700 tools.

ID	Category	Description	Impact
Gap 2.7 - 1	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L

ID	Description
Risk	None foreseen.

2.8 Significant Events Viewer

The Significant Events Viewer will provide users with a visualisation tool allowing interactive browsing and visualization of time series of feedback data with their associated events.

This advanced tool will use different technologies for its implementation. Some of them have been widely used and it's not foreseen that there will pose any gap in the future. These are **Django**, **JQuery** and **relational databases**, specifically **MySQL**.
Other technologies developed at ECMWF and successfully used in other projects are the MARS archive, ODB and "magics".

As with the Faceted Search Tool, there are no difficulties foreseen that may affect the development of the Significant Events Viewer. The mentioned technologies have been successfully used for the ECMWF's Forecaster and other internal projects. The only possible gap would be the integration of this tool with the others in WP700.

ID	Category	Description	Impact
Gap 2.8 - 1	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L

ID	Description
Risk	None foreseen.

2.9 Inter-Comparison Tool

A simple web-based tool will be developed that allows side-by-side visual comparison of datasets. The user will be able to select one or more datasets from a CHARMe search interface such as the Faceted Search Tool.

This tool requires access to both metadata from the CHARMe repository and data. Climate model forecast data from the CMIP5 activity will be available via the ESGF.

In order to make inter-comparisons, the data in the CHARMe Nodes must be structured in a way that allows machine based inter-comparison, so it's essential that the data model component depicted in Section 2.1 has a well-defined schema. Thus, there is a risk that the CHARMe Nodes' structure is not compatible with the inter-comparison tool and other applications built on top of CHARMe. To deal with this risk, the designers of the data model should take into account the specific requirements of the CHARMe advanced tools, in particular the Inter-Comparison Tool.

Other general technologies broadly and successfully used that will be used to implement this tool are **Django**, **JQuery**, **Apache Jena**, and **Apache Lucene/Solr**.

However, it is being discussed at the moment how this tool will be developed differently than how it was detailed in the Document of Work [5]. There are certain similarities in the functionality of this tool and the Fine-Grained commentary tool that may require changes to the original design of this tool and affect its development.

ID	Category	Description	Impact
Gap 2.9 - 1	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L

ID	Description
Risk 2.9 - 1	The data model has a structure that doesn't allow metadata inter-comparison.

2.10 Fine-Grained Commentary Tool

CHARMe Users can use the Fine-Grained Commentary Tool to visualize and request the manipulation of a subset of a dataset. This tool will merge and extend the capabilities of the following existing technologies and projects:

- the **ncWMS,** which is a Web Map Service for geospatial data represented in different formats, such as NetCDF, GRIB or BUFR;
- the **Godiva2 data visualization system**, which provides a means for scientists to browse interactively through large environmental datasets using only a web browser; and
- the **BlogMyData project**, which allows scientists to explore data visually, making and discovering other users' comments about subsets of datasets.

This tool will allow visualization and manipulation of 4D annotations: temporal and georeferenced annotations. As explained in Section 2.3, the data model will need to be extended to cope with geospatial and temporal queries, for which GeoSPARQL is a good candidate. Other alternatives that complement the triple store in the CHARMe Nodes with a PostGIS database have also been considered, but this decision will be made at a later stage of the project.

As with the other tools to be developed as part of WP700, some similarities in functionality can be found that may require further discussion facilitated by the elaboration of a detailed design prior to development.

ID	Category	Description	Impact
Gap 2.10 - 1	Knowledge Expertise	Research needed on how to extend CHARMe's data model to cope with geospatial and temporal queries.	M
Gap 2.10 - 2	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L

ID	Description
Risk	None foreseen.

2.11 Other Technical Areas

In order to demonstrate the usability of the CHARMe system, as well as for testing purposes during earlier stages, we need a **populated CHARMe database** containing different kinds of annotations. This is a major gap, as it is not part specifically of any WP and it needs a comprehensive approach, trying to explore all possible use cases of annotation bodies (citations, geographically referred annotations, temporally referred annotations, …) and targets (datasets, other annotations, …). This also poses a risk that a well-populated CHARMe database is not ready in time for testing.

This document only explores the gaps related to use cases where the user is a CHARMe User [2]. There will be other kinds of users interacting with the system not considered in this document like Remote Systems, Administrators or Moderators [2]. There is currently a gap on how these users will interact with the system, although this may be subject of future versions of CHARMe.

At this stage, we do need to take into consideration **moderation of user commentary**. It will depend on how extensively the CHARMe system is used when released, but we need to provide the CHARMe User community with a trustworthy moderation system. At time of writing, there are no currently specific details on how this will be handled or who will be nominated to assume this role.

ID	Category	Description	Impact
Gap 2.11 - 1	Implementation	Initial population of the CHARMe's database of Commentary metadata.	H
Gap 2.11 - 2	Design	Role of CHARMe Administrators and Moderators	M

		and specification of their interaction with CHARMe.	

ID	Description
Risk 2.11 - 1	CHARMe Moderators not able to review all possible requests on metadata manipulation.
Risk 2.11 - 2	Initial CHARMe database not populated in time for testing.

3 Summary

This document has presented the components and areas we can identify at this early stage of design that will take part of the CHARMe system, discovering possible problems that may pose a risk during future implementation stages.

The following tables offer a summary of the gaps and risks found for each of the system technical areas depicted along this document:

3.1 Summary of Gaps

Technical Area	Gap ID	Category	Description	Impact
Data model for annotations	2.1 - 1	Expertise	Need to develop specific custom ontologies to complement general data model.	H
Storage for annotations	2.2 - 1	Knowledge Expertise	Further research on storage solutions by CHARMe's development team needed.	H
Geospatial referencing for annotations	2.3 - 1	Expertise	CHARMe's developer team has little experience on this subject and further research is needed.	M
Network Protocol for exchanging annotations	2.4 - 1	Expertise	Further research on the REST interface needed.	L
User Interfaces	2.5 - 1	Expertise	Most of the user interfaces to be developed from scratch.	H
Authentication / Authorization systems	2.6 - 1	Knowledge	Access control required to secure access to the repository by different types of user profiles not yet conceived.	M
	2.6 – 2	Knowledge	Management of security administrative domains.	M
	2.6 – 3	Implementation	User privacy: it should be possible to formulate statement of terms of use of system and how user's data will be used.	L
Faceted Search Facility	2.7 - 1	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L
Significant Events Viewer	2.8 - 1	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L
Inter-Comparison Tool	2.9 - 1	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L
Fine-Grained Commentary Tool	2.10 - 1	Knowledge Expertise	Research needed on how to extend CHARMe's data model to cope with geospatial and temporal queries.	M
	2.10 - 2	Implementation	Research and implementation of areas of collaboration with other advanced tools.	L

Other	2.11 - 1	Implementation	Initial population of the CHARMe's database of Commentary metadata.	H
	2.11 - 2	Design	Role of CHARMe Administrators and Moderators and specification of their interaction with CHARMe.	M

3.2 Summary of Risks

Technical Area	Risk ID	Description
Data model for annotations		No risks foreseen.
Storage for annotations	2.2 - 1	The amount of data in the CHARMe triplestore escalates to an extent that the system's performance is compromised and/or causes unexpected problems.
Geospatial referencing for annotations		No risks foreseen.
Network Protocol for exchanging annotations	2.4 - 1	The CHARMe system's efficiency is compromised when trying to cope with several parallel queries or interactions with the data stored in the CHARMe Nodes.
User Interfaces	2.5 – 1	The CHARMe Plug-In may compromise the Data Provider's site's security and inadvertently expose information to unknown or malicious users.
Authentication / Authorization systems	2.6 – 1	Risk that user registration authentication and security domains not clarified at an early enough stage in the design process.
	2.6 – 2	Risk that CHARMe plugin will not provide the required security and that a full CHARMe node will have to be implemented per site.

Faceted Search Facility		No risks foreseen.
Significant Events Viewer		No risks foreseen.
Inter-Comparison Tool	2.9 - 1	The data model has a structure that doesn't allow metadata inter-comparison.
Fine-Grained Commentary Tool		No risks foreseen.
Other	2.11 - 1	CHARMe Moderators not able to review all possible requests on metadata manipulation.
	2.11 - 2	Initial CHARMe database not populated in time for testing.

Appendix A – Glossary of technical terms

This appendix contains a glossary of technical terms used in this document. It just aims at creating a quick reference for non-technical readers. These definitions are widely available on the Web; a link to the reference with more available information has been included for each of the terms.

Annotation	Set of connected resources, typically including a body and target, where the body is somehow about the target. http://www.openannotation.org/spec/core/
Apache Jena	Java framework for building Semantic Web and Linked Data applications. http://jena.apache.org/
Apache Lucene	High-performance, full-featured text search engine library suitable for applications that require full-text search, especially cross-platform. http://lucene.apache.org/
Apache SOLR	Search platform for the Apache Lucene project. http://lucene.apache.org/solr/
Atom	Set of web standards describing an XML language used for web feeds and a protocol for creating and updating web resources. http://en.wikipedia.org/wiki/Atom_(standard)
BlogMyData	Project that allows scientists to explore data visually using Godiva2, then make comments about features in the data on a blog. http://www.blogmydata.org/
CMIP5	Framework for coordinated climate change experiments for the next five years that includes simulations for assessment in the AR5 and its extensions. http://cmip-pcmdi.llnl.gov/cmip5/
Django	Python Web framework that encourages rapid development and clean, pragmatic design. It eases the creation of complex, database-driven websites. https://www.djangoproject.com/
ESGF	The Earth System Grid Federation is a collaboration that develops, deploys and maintains software infrastructure for the management, dissemination, and analysis of model output and observational data. http://esgf.org/
GeoSPARQL	GeoSPARQL defines a vocabulary for representing geospatial data in RDF, and it defines an extension to the SPARQL query language for processing geospatial data. http://www.opengeospatial.org/projects/groups/geosparqlswg
Godiva2	Web client that allows users to visualize data from a ncWMS server interactively. http://www.resc.rdg.ac.uk/trac/ncWMS/wiki/GodivaTwoUserGuide

Javascript	Scripting language of the Web allowing addition of functionality, validation of user input, communication with web servers, etc. http://www.w3schools.com/js/
JQuery	Multi-browser JavaScript library designed to simplify the client-side scripting of HTML. http://jquery.com/
Linked Data	Recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF. http://linkeddata.org/
ldJSON	A standard for line delimiting JSON in stream protocols, JSON being a language for representing simple data structures and associative arrays. http://en.wikipedia.org/wiki/Line_Delimited_JSON
Magics	ECMWF's Meteorological plotting software. https://software.ecmwf.int/wiki/display/MAGP/Magics
MARS archive	The Meteorological Archival and Retrieval System is the main repository of meteorological data at ECMWF. It contains terabytes of operational and research data as well as data from Special Projects. http://www.ecmwf.int/services/archive/
MySQL	One of the most used relational database management systems along with PostgreSQL. http://www.mysql.com/
Native Triplestore	Purpose-built database for the storage and retrieval of triples, a triple being a data entity composed of subject-predicate-object, like "Data Provider A owns Dataset B”. http://en.wikipedia.org/wiki/Triplestore
ncWMS	Web Map Service for geospatial data that are stored in CF-compliant NetCDF files requiring minimal configuration. http://www.resc.rdg.ac.uk/trac/ncWMS/
NoSQL database	A NoSQL database provides a mechanism for storage and retrieval of data that employs less constrained consistency models than traditional relational databases. It offers simplicity of design, horizontal scaling and finer control over availability. http://en.wikipedia.org/wiki/NoSQL
ODB	Observational DataBase is a tailor made hierarchical database software developed at ECMWF to manage very large observational data volumes and to enable flexible post-processing of these data. http://www.ecmwf.int/services/odb/
Open Annotation	Framework for creating associations between related resources and annotations, using a methodology that conforms to the architecture of the W3C. It satisfies complex requirements while remaining simple enough to also allow for the most common use cases, such as attaching a piece of text to a single web resource.

	http://www.openannotation.org/spec/core/
OpenID	Allows users to be authenticated by certain co-operating sites using a third party service, eliminating the need for webmasters to provide their own ad hoc systems. http://openid.net/
PostGIS	PostGIS is a spatial database extender for PostgreSQL relational database. It adds support for geographic objects allowing location queries to be run in SQL. http://postgis.net/
RDF	The Resource Description Framework is a family of W3C specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modelling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats. http://www.w3.org/RDF/
Relational Database	Database that has a collection of tables of data items, all of which is formally described and organized according to the relational model. In typical solutions, tables may have additionally defined relationships with each other. http://en.wikipedia.org/wiki/Relational_database
REST interface	Representational State Transfer (REST) is an architectural style that abstracts the architectural elements within a distributed hypermedia system. REST ignores the details of component implementation and protocol syntax in order to focus on the roles of components, the constraints upon their interaction with other components, and their interpretation of significant data elements. http://en.wikipedia.org/wiki/Representational_state_transfer
RSS	Family of standard web feed formats used to publish frequently updated information. http://www.rss.com/
Semantic Web	Collaborative movement led by W3C that promotes common data formats on the World Wide Web. It aims at converting the current web, dominated by unstructured and semi-structured documents into a "web of data". It builds on the W3C's RDF. http://en.wikipedia.org/wiki/Semantic_Web
Shibboleth	Software package for web single sign-on across or within organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner. http://shibboleth.net/
Single Sign-On	Property of access control of multiple related, but independent software systems. With this property a user logs in once and gains access to all

	systems without being prompted to log in again at each of them. http://en.wikipedia.org/wiki/Single_sign-on
SPARQL	Query language for databases, able to retrieve and manipulate data stored in RDF format. http://www.w3.org/TR/rdf-sparql-query/
stSPARQL	The query language of Strabon for data represented in an extension of RDF called stRDF. http://www.w3.org/2001/sw/wiki/Strabon
Strabon database	Fully implemented semantic geospatial database system that can be used to store linked geospatial data expressed in RDF and query them using an extension of SPARQL. http://www.strabon.di.uoa.gr/
Turtle	Format for expressing data in the RDF data model, similar to SPARQL. http://www.w3.org/TR/turtle/
URI	Uniform Resource Identifier is a string of characters used to identify a name of a web resource. Such identification enables interaction with representations of the web resource over a network (typically the World Wide Web) using specific protocols. http://en.wikipedia.org/wiki/Uniform_resource_identifier
Virtual Organisation	Dynamic set of individuals or institutions defined around a set of resource-sharing rules and conditions. http://en.wikipedia.org/wiki/Virtual_organization_%28grid_computing%29