Implementing an Engineering Data Management System for the LHC Accelerator and Experiments: The CEDAR Project

B. Rousseau, N. Høimyr, C. Hauviller
on behalf of the EDMS Task Force

Bertrand.Rousseau@cern.ch *CERN/ECP/IPT*
Nils.Hoimyr@cern.ch *CERN/IT/CE*
Claude.Hauviller@cern.ch *CERN/PPE/TA*

The huge quantities of data required for the design, building and operation of the LHC and its experiments require consistent management and storage. The CEDAR project aims at implementing an Engineering Data Management System (EDMS) at CERN. After having defined what is an EDMS and what are the expected benefits for the LHC, we present the activities held in CEDAR, during the selection and pilot project phases. Lessons learned and conclusions reached through the pilot projects are then summarised.

1 Introduction

The design, construction and operation of the LHC and its experiments will generate a huge, complex and multi-disciplinary collection of data. This data cover documents and pieces of information such as CAD drawings, technical reports, product configurations, catalogues of components, simulation and analysis results, calibration data, software source code, meeting minutes and slides, detectors images, project plans, etc. Data production will be performed in many different institutes, located all over the world using a whole variety of methods and tools, ranging from word-processors to CAD systems. The need for ease of access, for consistency, maintenance, for 25 years availability and for rational organisation of the work with such a large and complex amount of data dictate the use of an Engineering Data Management System, an EDMS [1] [2]. Examples from industry shows that an EDMS can drastically reduce the time spent on search for information and lead to improved quality by ensuring use of the correct information [3].

2 What is an EDMS

An EDMS has numerous functions to manage data and documents, process and workflows, product structures and configurations, classifications, programmes, etc [4] [5]. We detail below 3 important facets of these functions.

2.1 An EDMS is a data vault

An EDMS is built around a database (often called a data vault), in which all kinds of data used to define, manufacture and support products are stored, managed and controlled. The basic idea is that files produced by users with various applications (e.g. Euclid, FrameMaker, MS Project, Geant, etc) are made known and referenced by the EDMS, to become “Managed Files”. The EDMS may be accessed by users providing and/or consulting documents. Users may be located at CERN, in external institutes or in companies. Once a document has been checked-in to the database, it is under the control of the EDMS: further access to the document must be achieved through the EDMS, that checks user privileges for reading, updating or releasing the document. *This approach guarantees that the documents are centrally available in an updated form*.

2.2 An EDMS records data structure and semantics

An EDMS stores not only document files, but also the so called metadata, objects that give structure and semantics to the set of controlled documents (see Figure 1). For instance, the various components of a detector can be modelled as a tree shaped *Product Breakdown Structure* (PBS). CAD drawings, specifications, project plans, minutes of meetings, etc.are attached to the corresponding nodes of the PBS. This kind of structure, made of links between objects, constitutes the basis for navigating the data vault. The user may either browse the data vault by following links between objects of interest, or perform requests based on names of objects, values of attributes, etc. To do so, it can use the native EDMS user interface, or a WWW interface that guarantees world-wide access. *Metadata helps the task of navigating and locating data*.

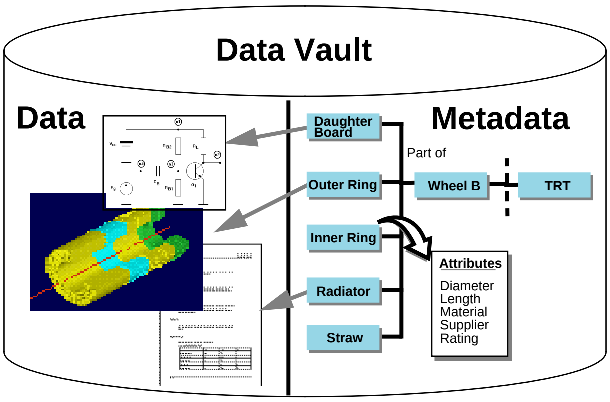

Figure 1: A Product description is stored in the *Data Vault* by means of *data files* and of *metadata* (objects, attributes and links between objects) that superimposes structure and semantics to the data. In this example, metadata represents a part of the Product Breakdown Structure of the ATLAS Transition Radiation Tracker. This linked structure allows users to navigate in the part hierarchy and find all information related to each part.

2.3 An EDMS helps managing the life cycle of data

Data, as well as metadata, are not static objects. Along their life cycle, they can take several successive *states*. For instance, a document may be *in preparation*, *being reviewed*, *released* or *obsolete*. Each one of these states corresponds to different access rights: once in the *released* state, a document shall not be updated. When moving from one state to another, a document can automatically trigger actions. As an example, when a *Drawing* moves from *in preparation* to *being reviewed*, managers are asked by email to approve, reject or comment. Once approved the *Drawing* can move from *being reviewed* to *released*, and a notification is sent by email to all affected users. In parallel, another action is launched automatically to convert the *Drawing* into the appropriate archive format (e.g. HPGL). *By allowing the capture and the formalisation of such life cycle processes, an EDMS ensures the integrity of data and contributes to a better, more consistent organisation of the work.*

3 The CEDAR Project

The CEDAR project [6] is managed by the EDMS Task Force, whose goal is to organise and manage the selection, and to propose a plan to the CERN management for purchasing, customisation and installation, of a unique EDMS for the whole CERN. The Task Force started off in January 1995, and is composed of representatives of 3 CERN sectors: Research (LHC experiments), Accelerator (LHC machine) and Technical (information technology and engineering support).

3.1 The Selection Phase

A *Market Survey*, followed by a *Call for Tender* (December 95) were used to filter the available product and vendors from 70 firms to a shortlist of 3 candidates. Part of the Call for Tender was the User Requirements Document (URD) describing the capabilities of the EDMS expected by the users, and the constraints for its integration in the CERN environment. The Software Engineering Standards of the European Space Agency (ESA PSS-05) [7] was used as a framework on which to build a URD. The suppliers were asked to describe how their product matches CERN's User Requirements. The Task Force organised visits and benchmarks to see the EDMS in action with real LHC data and scenario. In June, the decision of using Matrix from ADRA Systems Inc. [8] for the pilot projects was taken.

3.2 The Pilot Projects

Three concurrent pilot projects started in September 1996. They involved physicists, mechanical and electronics engineers/designers, as well as software engineers to evaluate both the usefulness of an EDMS at CERN and the suitability of the chosen product.

- The LHC pilot project targeted the management of the design and integration of the *Short Straight Section* of the LHC machine.
- The ATLAS pilot project was oriented towards the design of the *Transition Radiation Tracker.*
- The CMS pilot project was oriented towards the integration effort of the detector, and the assembly of the ECAL detector.

In addition to these projects, two parallel activities were launched in order to implement a Matrix/WWW interface, and to use OMT [9] and CASE tools as a way to define and maintain data models of the EDMS.

4 Lesson Learned

The conclusions from the three pilot projects are the following:

- An EDMS can be used in the CERN environment and will be needed to manage the engineering data for the LHC project (machine and experiments).
- The WWW is the required user interface for most users. It removes all platform dependencies, and has now become familiar to everybody. The WWW allows the development of user interfaces that may be simpler and better targeted than the EDMS native interfaces.
- Starting from a “naked” EDMS, substantial work is needed to adapt it to the needs of a given project. Customisation includes the definition and implementation of the data models for metadata and processes, the development of scripts to automate complex operations on data and metadata and the coupling with external software.
- A CASE Tool supporting an object-oriented notation such as OMT is of great help when designing the data models. It ensures a clear and common understanding of the EDMS data base schema, facilitates maintenance and preserves against a too heavy dependency on a given EDMS product.
- Successful use of an EDMS depends more on organisation and project management than on the software tool itself. Prior to its introduction into an EDMS, data has to be structured and engineering processes formalised (see for instance [10]). The EDMS itself can act as a catalyst for the definition of rules and processes, but cannot solve any organisational problems on its own.
- Members of the pilot projects have not reached agreement on the suitability of Matrix as LHC’s EDMS. In order to obtain a better foundation for the EDMS choice, a new project started in January 97 to re-implement the ATLAS/TRT pilot with another EDMS: CADIM/EDB from Eigner + Partner [11]. The final decision for the product choice will take place in April 97.

5 References

[1] J.L.Chevalley and al., *Progress Report of the EDMS Task Force*, January 96, CN/96/2 http://cadd.cern.ch/cedar/documents/progrep95/progress.html

[2] G. Bachy, A.P.Hameri, M.Mottier, *Engineering data management - a tool for technical coordination*, CERN MT-95-07(DI), LHC Note 345

[3] A.P.Hameri, *EDMS - Concepts, Motivations and Basic Requirements*, Proceedings of CERN School of Computing 96, CERN 96-08, Nov. 96

[4] CIMData Home Page http://www.CIMdata.com/

[5] Product Data Management Information Center http://www.pdmic.com/

[6] CEDAR Project Home Page http://cadd.cern.ch/cedar/

[7] C. Mazza and al., *Software Engineering Guides*, Prentice Hall, 1996.

[8] ADRA Systems, Inc. Home Page http://www.adra.com/

[9] J.Rumbaugh and al., *Object-Oriented Modeling and Design*, ISBN 0-13-629841-9, Prentice Hall, 1991

[10] ATLAS Quality Assurance Plan http://atlasinfo.cern.ch/atlas-bin/list?dir=QA

[11] Eigner + Partner Home Page http://www.ep-ka.de/