" The World according to Arcane " : an operating instrumental paradigm for schoolarly edition

Éric-Olivier Lochard, Dominique Taurisson-Mouret

"The World according to Arcane": an operating instrumental paradigm for schoolarly edition

Éric-Olivier LOCHARD - University of Montpellier - France
Dominique TAURISSON - CNRS – France

Version non corrigée par l'éditeur

Abstract. " The World according to Arcane" is an operating instrumental   paradigm for the electronic edition of scientifically established texts and knowledge, currently being used in several projects of scholarly edition. The world of   knowledge is edited in a   data base the architecture of which is both generic, so as to be applicable to numerous domains, and simple: any information is a subject of interest, a multimedia document, a relation between subjects or an enrichment. Internal or external documents are enriched with the editing module, independently of the media and the final publication.   The reading module offers very powerful procedures to investigate and browse electronic work: typified links inferred by the architecture,   sophisticated indexation,   dynamic composition of   virtual documents,   naturally formulated requests, formal treatments, and reading itineraries. The publishing module allows to export information in various formats (HTML, XML, TEX), to compose   paper books, and to produce electronic books in the form of autonomous applications distributed on CD-ROM, DVD-ROM,   web site, or data base system.

"The World according to Arcane" is an instrumental paradigm for the electronic edition of scientifically established knowledge, or more exactly, for the edition of electronic books dedicated to the diffusion of this knowledge.
By instrumental paradigm, I mean a coherent set of notions and concepts, methods and tools.

This paradigm is the outcome of research work originally carried out in a CNRS research group belonging under my direction, "computer models for   electronic text analysis and edition".
It was first presented in 1994 to the Human Sciences House in Paris, after being tested with the edition of *Denis Fonvizine's Lettres de France* published in the Archives de

l'Est series (dir. G. Dulac, CNRS-Edition). It is currently being used in several projects: the edition of *Pierre Bayle's complete Correspondence* (dir. In. McKenna, Voltaire Foundation, Oxford), of which I have just corrected the first proof of the second volume, a model for a CDRom's of the *Port-Royal Dictionary* (dir. In. McKenna, Champion, Paris), the inventory of *Samuel Formey's Correspondence* (dir. J. Häseler, Champion, Paris), and the electronic edition of *Le journal de Corberon*; I myself use it for the web edition of my university lectures.

In this conference, I shall first mention the objectives and principles which guided the conception of Arcane.
Then I shall briefly explain what we mean by electronic book, work and electronic edition in the paradigm. I shall then electronically leaf through "a world according to Arcane" in an edition in progress, *Corberon's world described by his journal* , by choosing examples which seem to me very interesting within the framework of this colloquium. I shall do this with a software I have developed according to the paradigm. In this selective reading I shall say almost nothing of the work of establishment and enrichment of the information, leaving this for questions.
Finally I shall show some published examples of electronic books implementing a world according to Arcane.

Objectives

Several objectives were aimed at in the conception of the instrumentation, which appeared to us a priori as considerable progress with regard to paper edition:

• first not to limit oneself to the computerization of the usual editing processes, in particular the digitalization of works already published, but to envisage all the new forms of edition, on the one hand those that the state of the art in electronics and computer science makes conceivable, and on the other hand those which can enrich the activity of traditional scholarly edition with more powerful, more sophisticated means.
• obviously, to exploit all the resources given us by the electronic and computer industries, particularly the use of multimedia materials, hypergraphic architecture and the computer 's power of computation;
• to base the instrument on an operating model of the un-linear reading and handwriting processes, which would exploit the processing semantic capacities of the human brain, as well as the technological mechanisms;
• to favor cooperative work, by bringing together on an editorial project specialists of different domains, and editors of varied levels of competence;

• to bring about the accumulation of knowledge, the extension of the paratext in time and in space, by exchanges with other editions following standardized procedures and protocols (XML, SQL, TEX for example);

• to give an epistemological status to edited information, in order both to make use of scientific methods and to support our editions in the looming competition, particularly on the Internet;

• to protect the rights of authors;

• to facilitate the realization of the editorial economic investment, by radically singling out the editing and publishing activities, to allow the production and publication, from one single piece of work, of various electronic books, at various points in time, on different media, and aimed at different categories of readers.

The foundations of the paradigm

The paradigm is based on two assumptions:

The first one deals with the human activity of reading: we assume that the meaning of a text is the product of a complex process carried out by the reader at the time of reading; a text has no other meaning than the one that the reader gives to it; it behaves as an operative interface, in an exchange between editor and reader.

The second assumption has to do with the economy of edition: rather than to invest in computer technologies in order to find or generate meaning in documents published at low cost, we prefer to invest in editorial work and to offer documents having strong added value to a low cost reading.

Electronic book

In Arcane a book, paper book, digital or electronic book is the material or virtual object distributed to the reader: it is a coherent set of information established according to scientific methods, intended to answer an imagined need of readers (not of consumers), and labelised as such by a socio-economic authority (any set of information does not constitute a book).

What fundamentally distinguishes a paper book from a digital or electronic book is the reading system: the paper book is its own reading system, while the digital or electronic book can be read only with an electronic device.

We can then distinguish the digital book from the electronic book, which is Arcane's real concern.

The former is in a sense a digitized paper book, a "collection of a rather large number of pages" according to Le Petit Robert, recorded on digital media, through which the

reader can navigate as he pleases by using an ad-hoc electronic tool, possibly carrying out previously planned operations.

The electronic book is a more complex device, conceived to allow calculation, production, composition of informations contained in the book and their dynamic posting in views, in answer to requests formulated by the reader on the system of reading; the reading of an electronic book is by nature a productive, deliberate activity, which a "good" electronic book has to favor, due to the architecture of the information it contains, and to the flexibility given to the reader to manipulate (process) this information and to program views of it; furthermore, because it is read with a computer, an electronic book can lend itself to statistical, logical, formal, linguistic, philological, graphic treatments, to dynamic composition of virtual documents, to reader's annotation, and obviously to communication with other electronic books.

Knowing what is a "good" electronic book is still an open question.

Edited work versus published book

The very nature of the electronic book allows to consider the process of its edition in a radically different way from that of a paper or digital book.

In the last two cases, the object on which the editor works becomes in fact identified, in some way, with the published book; a first proof is that in French the term "éditer" can refer, the editing activity as well as to the publishing one; a second proof is that the " imprimatur " for publishing marks the end of the edition.

In the process of editing an electronic book on the contrary, nothing obliges this identification. The object on which one or several editors work can give birth to several publications, to several electronic books, on different media, for different publics, and even to successive or continuous (on the Web) publication: it is no longer **one** book but a richer object with strong added value, that we call an electronic piece of work; of course, this opportunity requires, in order to be actual, as many elaborated editorial strategies as appropraite methods of edition.

The editing process in Arcane

The founding principle of the Arcane paradigm is to articulate the exchange among editors and readers around a world of information semantically identified in a metalanguage.

Electronic edition then appears as a cooperative, unlimited, process consisting in:
•   define and prescribe the architecture of the edited world as an instance of Arcane's meta-architecture;
•   scientifically building, in the data base, a world of the subjects of interest of the editorial project;

•   enriching this world with added informations: multimedia documents   and illustrations, possibly contradictory interpretations, and relations   increasing its combinatorial semantic;
•   preparing documents for publication;
•   extracting from it, at properly chosen moments, coherent subsets of information to produce and publish electronic books.

Figure 1 - Editing process in Arcane

Readings

The meta-architecture of the world suggests natural readings, investigations from the elements of this meta-architecture: subjects, enriched multimedia documents, relations, and a vocabulary used by the reader to formulate requests and interrogations.
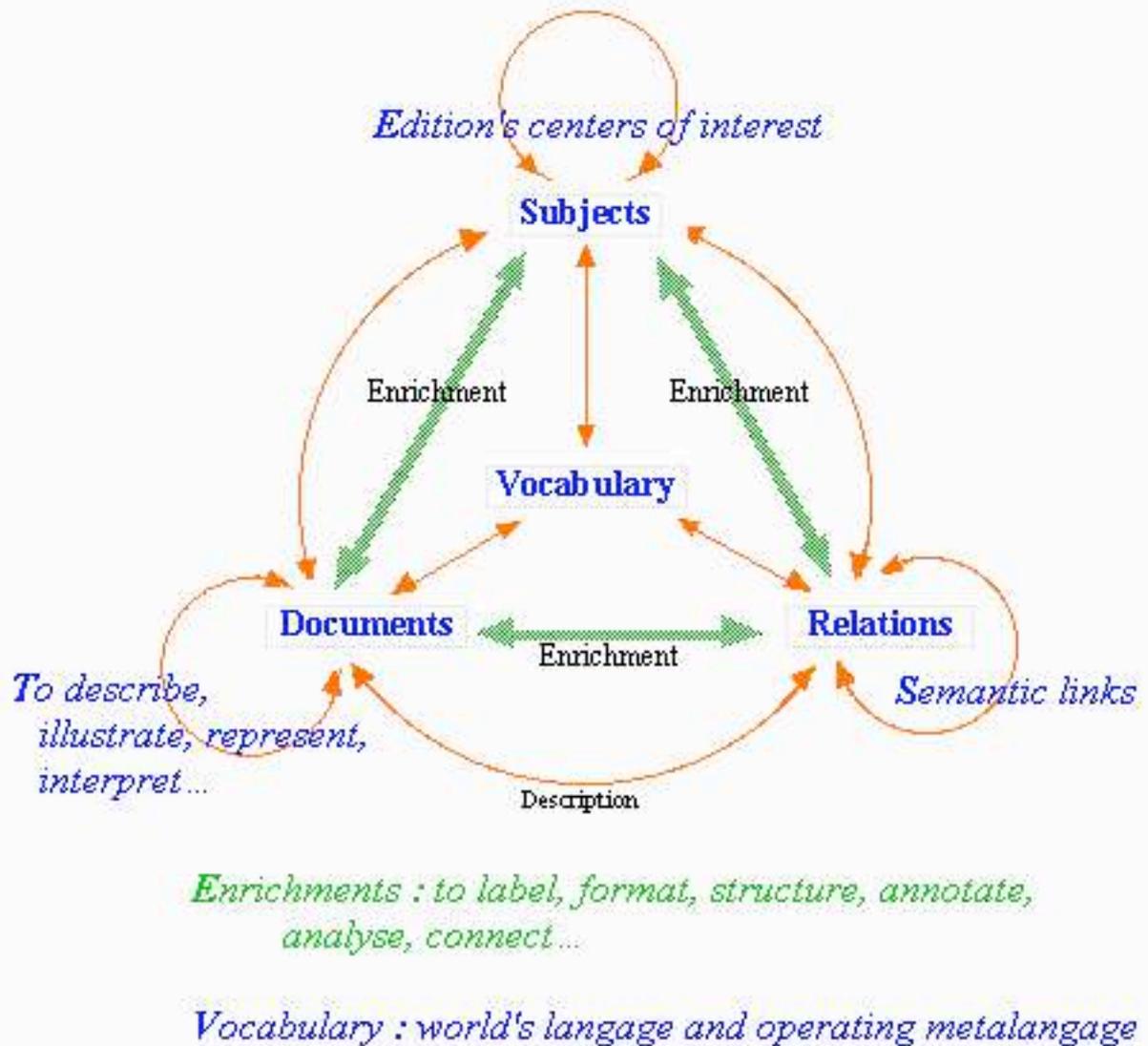
Figure 2 - The meta-architecture of the world.

The subjects of the world

The subjects are the centres of interest of the editorial project, organised according to various types: people, correspondence, works, events, places, subjects of study, notions, etc; they are described with Arcane multimedia documents (figure 3) and fields which depend on their type.

The types of subjects, their fields and the kind of documents are freely introduced into the world by means of prescriptions.
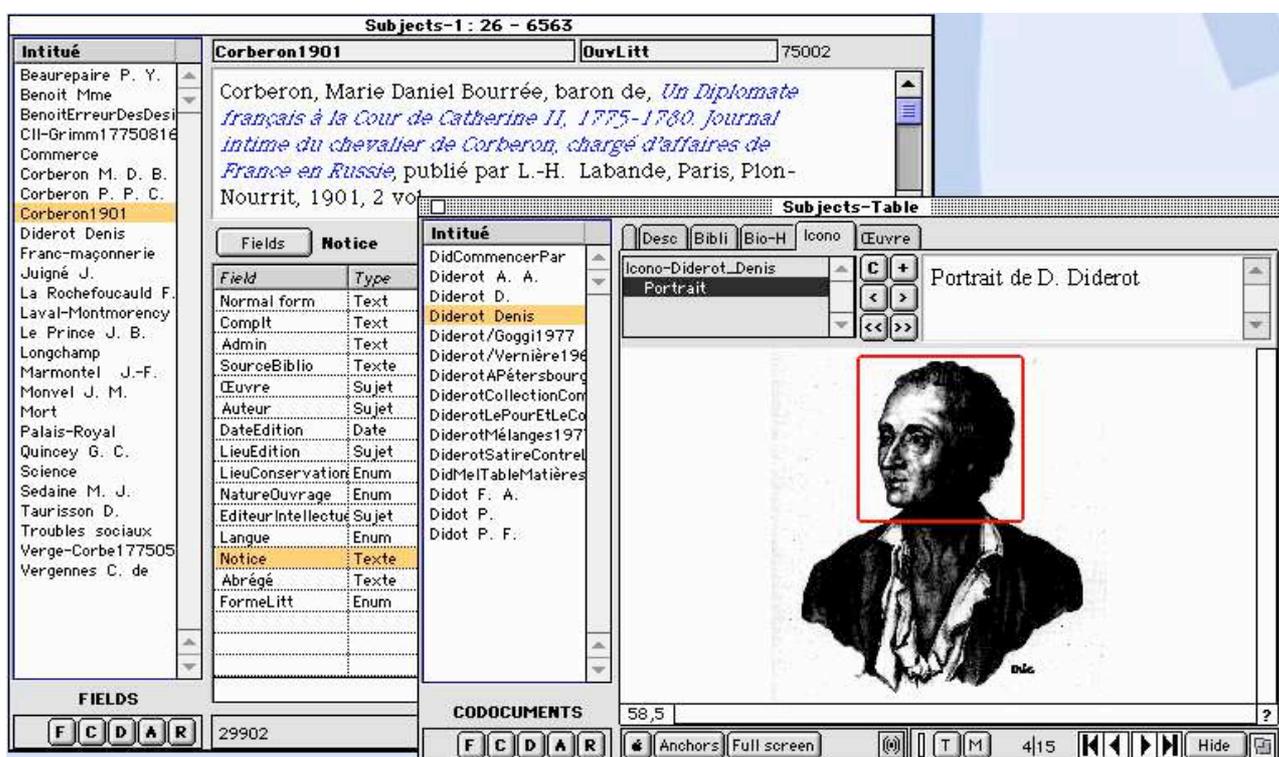
Figure 3 - Subjects in the fields and codocument views.

Subjects are used by the editors for the indexing of documents and the composition of Arcane relations; from a methodological point of view, they facilitate rigorous and exhaustive plans of work.

For the reader, subjects are entries to investigate and read world, natural access itineraries to a large part of the information which it contains, essentially through dynamic links recorded in indexing, relations and enrichments.

Figure 4 - A subject in the index view.

Enriched documents

Documents are multimedia information intended for explanation, illustration, structuring, comment, interpretation and annotation.

In Arcane the documents are distributed in three classes: autonomous documents, codocuments of document, subject or relation and footnote generalisation, and dynamic documents. The latter are defined by a script to be applied to elements of the world: they are effectively calculated and produced with the reading module in answer to an

instruction from the reader (for example the textual document composed with all the indexing anchors by a given subject).

Every document belongs to a kind (article, illustration, sound commentary, etc.) which comes with style sheets for representation, and possibly a type with its fields.

Arcane distinguishes the multimedia source document and its enrichments created explicitly to format it, structure, comment on or annotate it: a document to be   read is then an enriched document, defined as one source document interpreted by a set of enrichments. This choice allows the same source document to give rise to several structural representations (e.g. a working representation, printing on paper or display on a Web browser), as well as semantic interpretations giving a reading of different points of view.

Enrichments

These allow the editor to add information to subjects, relations and documents so as to format, structure, annotate, index, analyse, connect, or increase the combinatorial of the world.

Every enrichment is attached to an anchor, a sequence selected by the editor; it can be associated with typified attributes, keyboarded at the time of the enrichment, typographic styles for the representation of the anchor or automatic mark-up used for the export of the enriched document in various formats (TeX, XML, HTML).
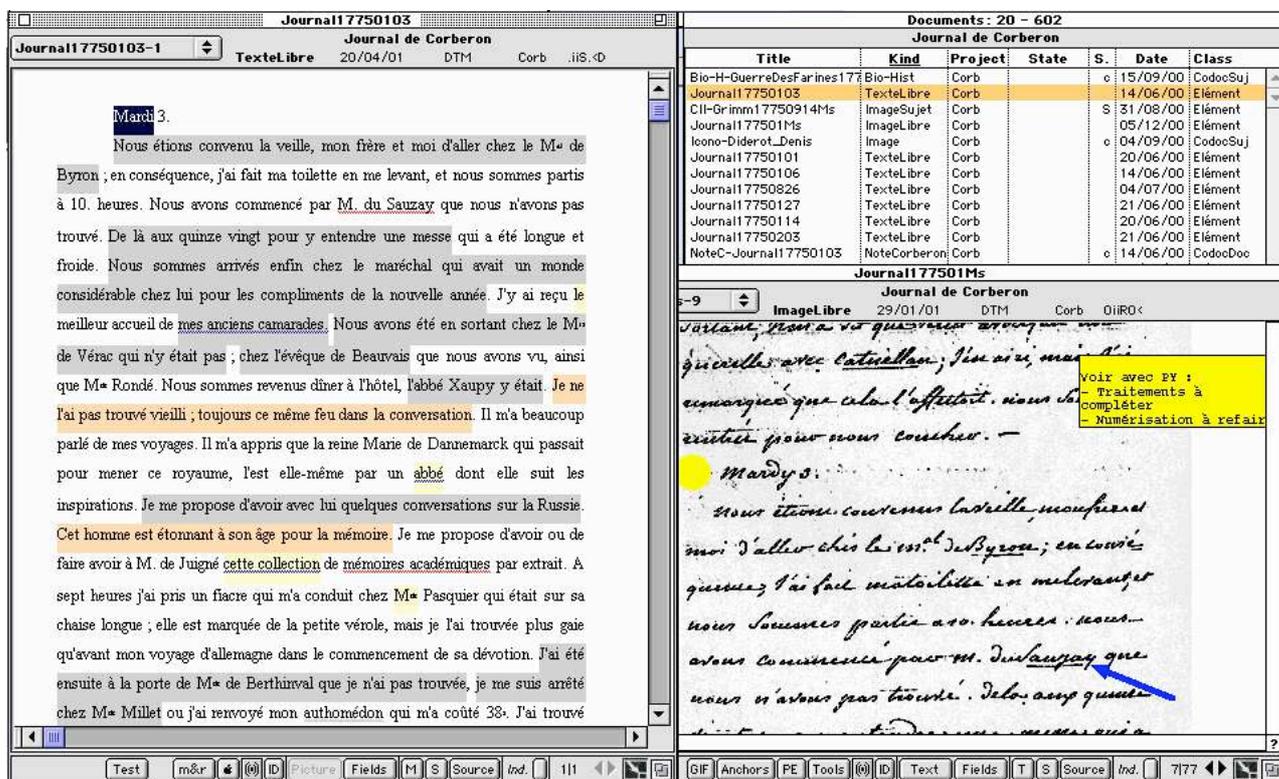
Figure 5 - Enriched documents.

The production tools of enrichments, proposed to the editors in menus, are freely prescribed with their possible attributes, their styles, and the objects to which they can be applied (types of subject, kinds of document); the editors can thus have a contextual work environment, which is coherent and convenient, particularly by the use of graphic charters.

When it is natural, the enrichments automatically generate dynamic links for investigation and navigation.

Relations

Defined as a labelled list of typified arguments, relations serve for establishing semantic and dynamic links between subjects or for building formal representations of the content of documents.

In *Le journal de Corberon* the relations are used to index activities connected with sociability and networks: meetings, journeys, readings, etc.

Relations are very effective reading tools; for example, they allow an increase in combinatorial semantics of the world by operating definition of new terms in the vocabulary to compose dynamic views of the subject; they lend themselves to all sorts of dynamic representation and treatment (statistical, graphic, logical, etc), to the reader's taste, or that suggested by the editor.

Figure 6 - A graphical representation of relations: Corberon's meetings in january 1775.

Reading procedures and itineraries

Among procedures which we have already presented, some can be naturally considered as reading procedures: for example the views of subjects give main information to be read about these subjects.

As in any database, the reader can also process requests to his taste, compose dynamic lists of the world's objects, dynamic documents, etc, which are so many reading itineraries.

Another interesting reading procedure is the "deliberate reading itinerary": by drafting a scenario, a synchronized succession of multimedia views in windows, the editor can propose to the reader passive (the scenario takes place automatically) or active itineraries (the reader activates the views himself) for some presentations of the edited world. The reader himself can build his own, as an educational exercise for example.

Publication and communication

In Arcane, publication is the publisherâs initiative. It consists of selecting enriched documents and/or subjects to be published and in applying a treatment described by   a script to each of these elements, essentially to code basic information and add prescribed mark-up. The result is a file or set of files in a particular format: TeX file for a paper book or pdf documents, HTML files for a Web site, or XML files for exchanges.
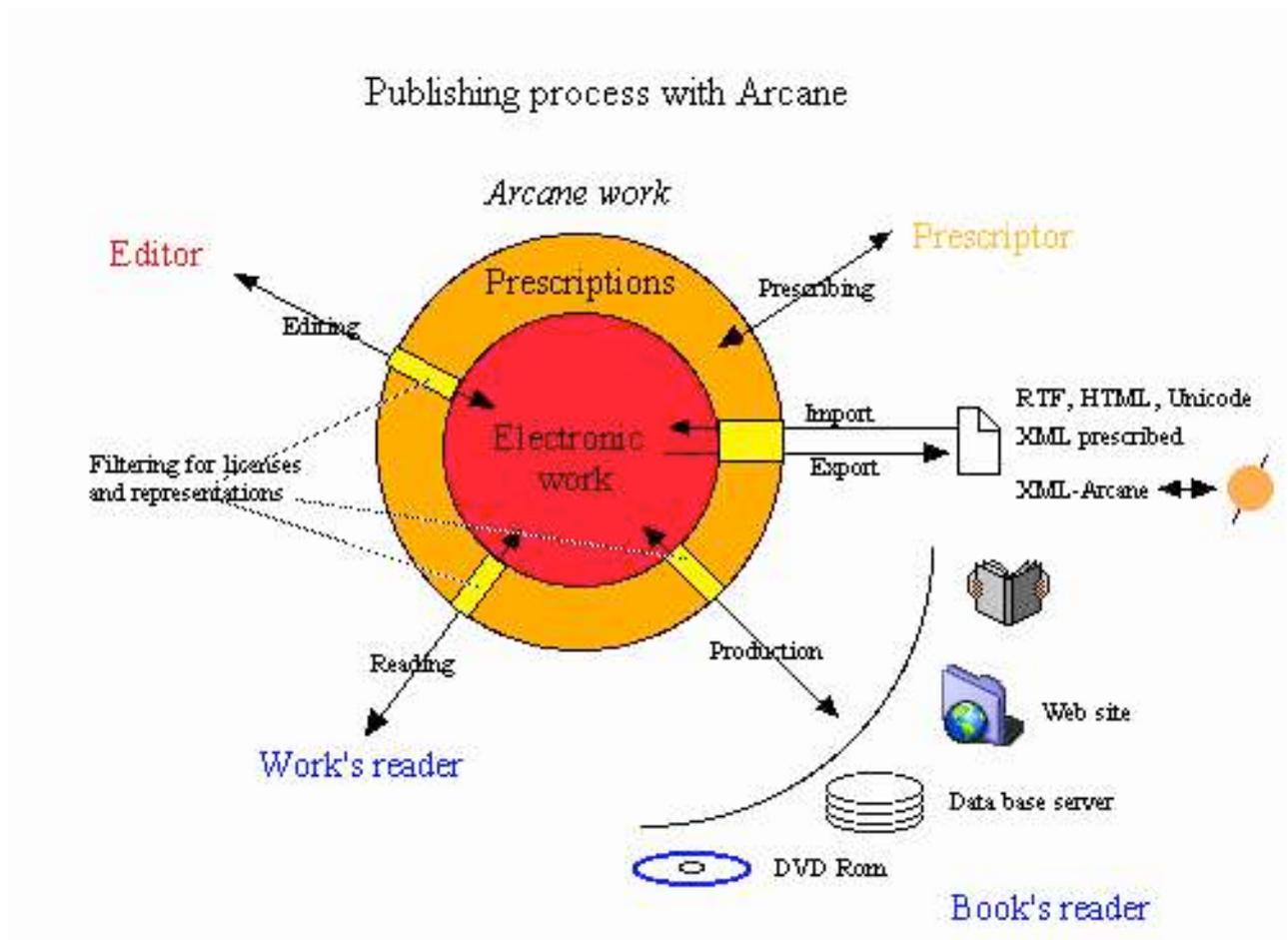
Figure 7 - Publishing process

Arcane was conceived from the very beginning as an open instrument to allow communication with other edited worlds or other editions; it can naturally import documents in different formats, the only constraint on interpreting XML documents being that the architecture of the imported information should be compatible with the Arcane architecture, in other words that its DTD (document type definition) is an Arcane DTD.

eric-olivier.lochard@univ-montp3.fr
dominique.taurisson@univmed.fr