A Cooperative Environment for Editing Documents with Workflow Technology

José Valdeni de Lima[1], Nina Edelweiss[1], Manuele Kirsch-Pinheiro[1]

Vincent Quint[2], Nabil Layaïda[2], Tayeb Lemlouma[2]

[1] Instituto de Informática - Universidade Federal do Rio Grande do Sul, Caixa Postal 15064, CEP 91591-970, Porto Alegre – RS - Brazil

{valdeni, nina, manuele}@inf.ufrgs.br

[2] Institut National de Recherche en Informatique et Automatique (INRIA), ZIRST – 655, Avenue de l'Europe – 38330 Montbonnot Saint Martin – France

{Vincent.Quint, Nabil.Layaida, Tayeb.Lemlouma}@inrialpes.fr

Abstract. *In the last years the Web became a great communication medium for publishing and delivering online multimedia documents. Therefore, the production (authoring, generation, transformation) of such documents became one of the most active research fields. This paper presents a web-based environment for cooperatively editing documents with workflow technology. The workflow technology provides a group of authors scattered across the network the means of guiding the authoring process by controlling the cooperative work. Second, this environment uses the most modern and promising Web technologies, such as XML, SVG, and SMIL to obtain a solution deployable world wide. Finally, we present another aspect concerning the content and user characteristics using CC/PP technology. Following such model offers not only a good description of the environment and its capabilities but defines also strategies about how some efficient interactions can be achieved.*

1. Introduction

Document production is not an easy task. When documents are intended for the Web, this task is even more complicated. If we add the use of multimedia and other advanced technologies, the complexity of the system increases substantially. In order to diminish this complexity, the production process is often achieved according to a cooperative process. A group of people may work together to define, design and deploy the content on the underlying computer infrastructure. One of the most important problems encountered when working with a group of editors is the increase in management effort. To overcome this problem, the workflow technology may be used to provide the group management.

Besides, the presence of Web technologies deployed at a world wide level should be highly considered. The use of the Web for multiple purposes is unquestionable, and is growing even more with the latest Web technologies, like XML and SMIL. The Web is not only a medium to present information. When the construction of documents is concerned, the Web may also be used as a medium to connect a group of authors, which may be geographically distributed.

This paper presents a cooperative authoring environment for Web multimedia documents. To build this environment, recent technologies, such as XML, SMIL and WebDAV, and open source software for the Web, such as Amaya and Annotea, are used. Moreover, on top of these components, we purpose the use of a workflow engine to control the access and creation of documents. The workflow guides the document production process, helping the group to produce their material.
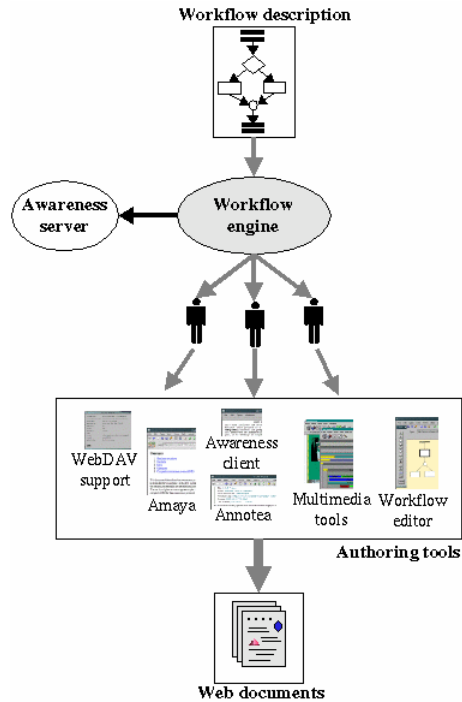
Figure 1: Environment overview.

The result is an environment divided in layers, presented in Figure 1: (1) a workflow engine that controls the document production process, based on (2) a workflow description; (3) Web authoring and multimedia tools, used by the authors to create the documents, and a set of functions supporting the cooperative work, as awareness and annotations. Those elements are discussed separately in the next sections: Section 2 presents the main Web authoring tool used, Amaya. Section 3 shows the first cooperative aspect, the WebDAV protocol, used to share documents among the authors. Section 4 presents the communication system used, Annotea, and Section 5 introduces the awareness support mechanism. Section 6 presents the workflow support provided in the environment, integrating the previously presented tools. Finally, some perspectives for this work are presented in Section 7, and the conclusions are drawn in Section 8.

2. Editing Web documents

The first element of the environment is a Web editor to create the Web pages. We have chosen Amaya, a Web client developed by W3C with some contributions from some of the authors of this paper. Amaya is a complete Web browsing and authoring environment with a WYSIWYG style of interface. It integrates seamlessly the browsing and editing functionalities and accesses (read/write) documents directly on the Web [15]. As it is used by W3C to demonstrate and test new developments in Web technologies, it includes very modern features, like XML. Thus, by using Amaya, users can edit pages in HTML and XHTML, with MathML expressions, and SVG vector graphics. They can also create and update CSS style sheets associated to these Web pages.

In complement to Amaya, a well configured HTTP server is needed. This server permits the authors to publish their pages and make these pages available for co-authors, allowing the construction of the cooperative environment needed for the documents authoring process. However, just this combination Amaya/HTTP does not cover the question of "lost update", when

a user's update is overwritten by another user's update. Figure 2 shows how this can happen: (1) first, two users take the original version of a document from the Web server; (2) then, both users make different changes in the document separately, but (3) user 1 submits his/her own version before the other, so (4) when user 2 submits his/her version, he/she overwrites his/her colleague's version.
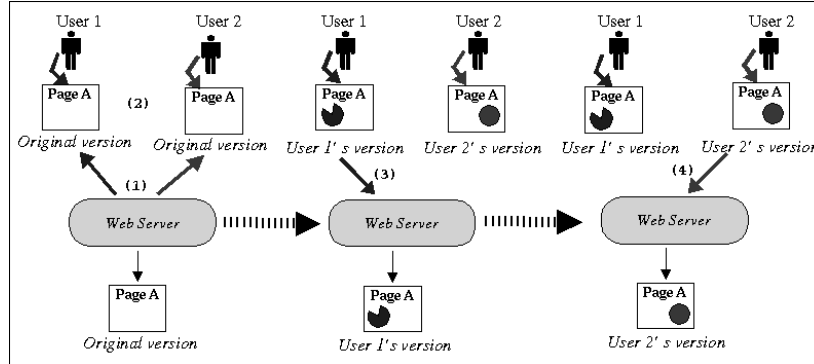

FIGURE 2: The lost update problem.

3. Editing cooperatively: sharing documents

The idea of a group work is to gain efficiency in time and quality to solve complex problems, in order to reach a better final product. We expect to achieve higher quality and productivity through a cycle of harmonic enhancements and cooperative interaction among the group members [3]. So, in order to integrate this aspect in the document construction process, group support for the design and creation of Web documents should be provided. Thus, it is possible to divide the document construction process in smaller parts, to enhance productivity, to take advantage of the authors' expertise, and also to discuss, exchange points of view, etc.

One of the main problems to cooperatively write Web documents is the lost update problem, showed in Figure 2. The WebDAV protocol has been designed to solve this problem. WebDAV (WWW Distributed Authoring and Versioning) is a set of extensions to the HTTP protocol, which allows users to collaboratively edit and manage files on remote Web servers [18].

We introduced the WebDAV protocol into Amaya to allow a cooperative edition of Web documents. WebDAV defines new methods and properties for the manipulation of shared objects (Web resources). Among the manipulation functions defined by WebDAV, there are methods to lock and unlock resources. By using these methods, an author can reserve to himself/herself the write access to a document, avoiding the lost update problem. Besides, considering that users not engaged in a group can also use Amaya to edit their Web pages, it should work individually with some resources as well as cooperatively with others, during a cooperative document production process.
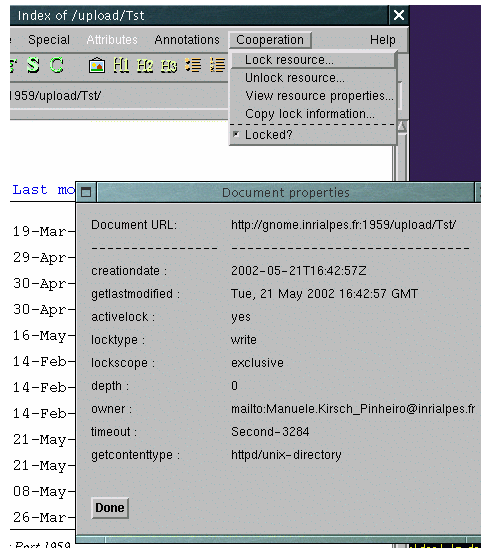
Figure 3: A fragment of WebDAV support under Amaya.

Thus, we implemented the WebDAV protocol in two levels. First, we implemented all WebDAV functions in a new module for the W3C network library, Libwww. This library is the base library that supports all network operations performed by Amaya. By using the new WebDAV functions in this library, we could implement the second level, the lock functions under Amaya (Figure 3), as well as functions to consult the properties associated to the document. These new functionalities introduced in Amaya allow users in a group to control their access to the shared objects, and then prevent the lost update problem.

However, to work cooperatively, a group needs more than shared documents. To reach the desired quality of the cooperation, a group needs also communication, awareness and coordination support.

4. Communicating through annotations

Users working cooperatively need a communication support. In fact, people should communicate in order to coordinate their efforts and cooperate on an objective [4]. A significant part of the work that takes place in an office is based on informal communication, and even the hierarchical, strict formal bureaucracy is entwined with officially invisible, informal cooperative work [13]. Indeed, through a communication support, users can improve their knowledge about their colleagues, about their qualities, preferences, etc. As a consequence, the interactions between those users become more informal and the work more fluid.

An effective way for authors to communicate and exchange ideas is to use annotations. Annotations act as a communication resource, which gives to the group the possibility to asynchronously discuss about the work they are doing together. Using annotations, the co-authors of a document may exchange experiences, opinions, comments, questions, etc. As a result, the cooperation process will be enriched, leading the group to better results.

In this work, we proposed the use of a new annotation system, Annotea, based only on Web standards. Annotea is a shared annotation system based on the general-purpose RDF infrastructure, where annotations are modelled as a class of metadata [5]. It treats annotations as

external statements made by an author about a Web document, by using XLink and XPointer technologies, and it stores these annotations in RDF bases, in one or more annotation servers.

Using Annotea, authors involved in the authoring process can associate annotations to their documents. Amaya supports an Annotea client. Thus, using Amaya with this client, authors have a cooperative function and a communication channel, which allow them to exchange comments, ideas, experiences, etc. They can easily use this resource without changing the authoring tool, doing theirs annotations directly on the documents they are constructing.

5. Awareness

In a cooperative work, it is important to have information about the evolution of shared data, users' actions, etc., to improve individual contributions. These contributions may be improved by understanding the activities of the whole group, and that is why awareness is important in the development of collaborative activities [12]. Awareness is the knowledge of group activities, including past, present and future activities. It is also the knowledge about the group itself and its overall status [7]. Awareness information forms a context for individual work by comparing it with the activities of the rest of the group. It helps users to synchronize their work, coordinating themselves around their context. Without this context, the individuals are unable to measure their own work with respect to the group's goals and progress [4].

Comparing awareness support and communication support, we can perceive that awareness support should be provided by the system and shall be obtained and presented automatically, in order not to interrupt the work. On the other hand, communication support, like annotations should be controlled by the user. An user needs to explicitly create an annotation, and his/her colleagues need to explicitly ask for the annotations.

Thus, to give users some support for awareness, we propose the use of an awareness module composed by an awareness protocol and an awareness server. The awareness protocol is a connection-oriented protocol based on the request/response paradigm: a client sends requests in XML format to the server, which also answers in XML format. It allows the exchange of awareness information between many client applications used by the group members and an awareness server, allowing the diffusion of the awareness information in the complete authoring environment. Awareness information includes information about the users (team members), their preferences and roles, in addition to the activities produced by them.

As a result, by exchanging this information, authors involved in document production can inform their colleagues about their own activities. They can also be informed about their colleagues' activities. Figure 4 shows an example of a request of the awareness protocol, where a client asks for awareness information, and the server answers with the events produced in the environment in accordance with the user's profiles.

Figure 4: An example of awareness protocol request.

The user produces the awareness information on the clients' applications, as Amaya, which format and pass this information to the server using the awareness protocol. The server is responsible of the control of this information, storing it in a reliable database. To build this server we use a framework, called BW Framework, designed to provide awareness support, especially in asynchronous environments [8]. Combining this framework with the authoring tools, authors can receive the notification of many events and even perform queries to filter them. Such notifications are related to document changes or new annotations associated with a document. The goal here is to optimise the knowledge about the group activities among the group members.

Besides, the BW framework filters the awareness information available by using profiles, where users may indicate their global preferences and their preferences when playing a specific role. For example, a team's coordinator may define, by setting a profile, what information should be presented to him/her only when he/she is playing the coordinator role. By considering those profiles, the awareness server can filter the information available in its database, showing only the useful information.

6. Coordinating with Workflow

A workflow can be used to connect and organize all the features presented in the sections above, providing the coordination needed for the group. The workflow defines the tasks to be executed, the agents responsible for their execution, and the synchronisation among them [20].

The convergence between workflow and cooperative authoring occurs in order to make a specification of cooperation at the conceptual level, where we can formalise the cooperative editing model, and also at the execution level, where the workflow engine controls the execution of the model. At the conceptual level, a sequence of tasks, needed for the construction of the document, is defined. This sequence shall be followed by the group in the authoring of documents. At the execution level, the application controls the activities that involve coordinated execution of multiple tasks.

To specify this coordination workflow, an XML DTD was defined. Furthermore, a graphic workflow editor was implemented, producing workflow descriptions according to this DTD.

6.1 Workflow DTD

An XML DTD was defined to represent the same elements specified by Casati et al. in their workflow model [2], [19]. The following elements are represented: begin; end; task, multi-task, and super-task; connector; total, partial, and interactive join; total, conditional and non-deterministic fork. Each of these workflow elements is represented by one element in the DTD. All these elements are direct descendents of the root element called "workflow". As an example, Figure 5 presents the elements defined for begin, task and connector.

```
<!ATTLIST     begin id ID #REQUIRED
              status CDATA #IMPLIED >

<!ELEMENT     task EMPTY>
<!ATTLIST     task id ID #REQUIRED
              task_name CDATA #REQUIRED
              role CDATA #IMPLIED
              pre_condition CDATA #IMPLIED
              post_condition CDATA #IMPLIED
              description CDATA #IMPLIED
              status CDATA #IMPLIED
              agent CDATA #IMPLIED
              supertask CDATA #IMPLIED
              type CDATA #IMPLIED >

<!ELEMENT     connector EMPTY>
<!ATTLIST     connector id ID #REQUIRED
              status CDATA #IMPLIED
              answer CDATA #IMPLIED
              preID CDATA #REQUIRED
              postID CDATA #REQUIRED >
```

FIGURE 5: A fragment of the workflow definition language DTD.

6.2 Workflow Editor

To help the workflow specification, a graphical editor was built. This editor uses Amaya and its SVG [14] support to manipulate graphical symbols, which represent the workflow tasks, super-tasks, etc. The modelling technique is based on the WIDE methodology (Workflow Development Methodology) [19], using SVG drawings to represent the symbols defined in this methodology.

With this editor, the workflow author is able to define the workflow using graphical shapes and their attributes. At the end of a workflow definition, the editor generates a XML description of the workflow according to the DTD presented above. The resulting XML code contains the commands, instructions and parameters for the workflow engine to control and coordinate the whole authoring process. Figure 6 shows the palette used by the editor to produce the workflow definition and the XML code generated by it. Thus, the user, through a visual interface, may construct a workflow to control the entire cooperation process.
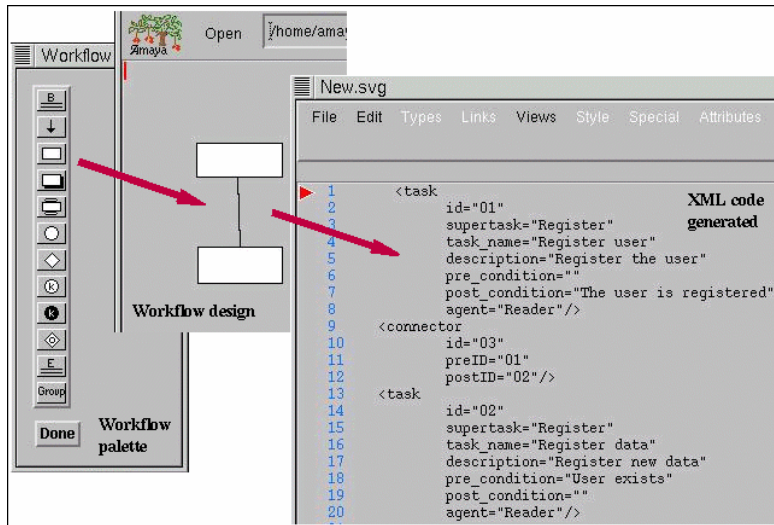
FIGURE 6: Workflow editor interface and an example of XML code generated by it.

7. Future perspectives

To achieve a complete Web document authoring environment, some features are still needed: They are currently under development.

7.1 Workflow Engine

The effectiveness of the team coordination can only be achieved through a workflow engine, able to execute the workflow specification. This workflow engine is currently under construction. The specification of this engine is described below.

In the document construction environment, the workflow engine shall be introduced in the server used to publish the documents under construction. Communication between this workflow engine and the authors can be done through a Web interface. The engine will read and execute the XML workflow description. By interpreting this description, the engine will understand the workflow tasks, resources and constraints, and will be able to control the workflow execution. Besides, the workflow engine may use the awareness information to better control the authoring process. For example, it can see which documents are complete through such awareness events as "final document save". This awareness information is maintained in the awareness server by the BW framework. This base can be periodically consulted by the workflow engine to make decisions during workflow progress. The workflow engine may use all the features presented here to make decisions. For example, when the author decides to work on a Web page, the workflow makes the edition tools available and, if necessary, locks this page on the server.

On the other hand, using a Web interface to reach the workflow engine, through a tool such as Amaya, authors may request a task, following the workflow sequence and constraints. In addition, the workflow may indicate which tool should be used to execute the task. Another

feature extension planned on the client side is to give the authors the possibility of interactively supplying information about the realization of tasks manipulated by the workflow engine.

On the fly revision of the workflow is also to be studied in order to provided more flexibility to the overall system.

7.2. Adaptability consideration

In many situations, content is accessed through heterogeneous devices with different characteristics and capabilities. This fact makes the presentation of the same content different from a client to another and sometimes, due to the device limitations, the received content isn't well presented as in the original model to which it was edited for. In addition to this hardware heterogeneity, clients can use a variety of browsers and have different preferences concerning the use of the original content. In order to take into account the constraints caused by the heterogeneous environment, a particular attention and effort must be made during the authoring of the content. This point is of high importance if we aim to deliver adaptable documents that follow the device independence principles. Making the system able to offer adapted services requires some indispensable descriptions that concern all the elements involved in the content delivery from the original server, to the connection network, to the end user. These descriptions, which we call in general profiles, must follow an efficient model that defines the characteristics and capabilities of the environment in order to make easier the application of future adaptation processes.

In this context, the CC/PP model [16] represents a good tool to express the different profiles and also to define a way to exchange them between the server and clients [10][17]. CC/PP is based on RDF [9] which gives semantics to the different profiles and facilitates their processing. On the server side and during the authoring step, the author must provide a more complete description of the content. This will help the server to negotiate the appropriate information to be delivered in client accesses. The UPS schema [11] based on a CC/PP extension, can be used to ensure a description of the content characteristics and capabilities in terms of the used structure, included resources, formats, size, etc. The effort made by the author during document edition and profiles creation ensures the adaptability of the global system and increases delivery efficiency.

8. Conclusions

An editing environment based on workflow is presented in this paper. In this environment, the Amaya editor is used to create Web documents; the WebDAV protocol allows authors to share theirs documents, enabling the cooperation between them; Annotea provides a communication channel for the authoring group; awareness keeps a context for the group activities; and workflow controls the whole document production. Finally, one additional functionality introduced into Amaya is the graphical workflow editor, based on SVG.

The use of workflow for the integration of several people and tools in order to create a document production environment is one of the main focus of this paper. In a cooperative editorial environment several different features, like annotations, awareness, and lock mechanisms, are available. The effort needed to remember how to use these tools, together with their functions and parameters, is diminished with the use of workflow technology. When an author executes a

9

task programmed by the workflow, the whole context needed for this execution is automatically created by the workflow itself.

In addition to the use of a workflow to guide the document production process, the environment proposed here presents another innovation: the use of Web standards for the document authoring. Comparing this environment with other similar systems, such as BSCW, a shared workspace [1], POLITeam, a groupware based on the metaphors of electronic folders and shared workspaces [13], and Media [6], a cooperative environment for Web-based learning, the proposed environment is the only system based only on web standards. The use of standard systems is an important guarantee that the environment can be applied to many platforms and situations, and deployed world wide.

References

[1] Bentley, Richard; Horstmann, Thilo; Sikkel, Klass; Trevor, Jonathan. Suporting Cooperative Information Sharing with the World Wide Web: The BSCW Shared Worksapce System. Available at <http://www.w3.org/Conferences/WWW4/151/>. Access: July, 1998.

[2] Casati, F., Ceri, S., Percini, B., Pozzi, G., Conceptual Modeling of Workflows, Proceedings of *The Object-Oriented and Entity-Relationship Conference*, Gold Coast/Australia: 1995. pp.341-354.

[3] Dias, M.S., Borges, M.R.S., Development of groupware systems with the COPSE infrastructure, Proceedings of *International Workshop on Groupware*, IEEE Computer Society, Cancun, Mexico, September 1999, pp.278-285.

[4] Fucks, H.; Assis, R. Facilitating perception on virtual learning based environments. *Journal of Systems and Information Technologies*, Vol. 5, No. 1. Edith Cowan University, 2001. pp.93-113.

[5] Kahan, José; Koivunen, Marja-Riitta; Prud'Hommeaux, Eric; R. Swick, Ralph. Annotea: An Open RDF Infrastructure for Shared Web Annotations. Proceedings of *WWW10 International Conference*, Hong Kong, 2001. Available at <http://www.w3.org/2001/ Annotea/Papers/www10/annotea-www10.html>. Access: Aug. 2001.

[6] Karacapilidis, N.; Khaled, O. A.; Pettenati, M.C.; Vainorbeck, C. Developing Higher-Order Skills with the MEDIT Web-based Learning Environment. *Educational Technology & Society*. Vol. 3, No. 1, 2000. International Forum of Educational Technology & Society (IFETS).

[7] Kirsch-Pinheiro, Manuele; Lima, José Valdeni; Borges, Marcos R.S. Awareness em Sistemas de Groupware. Proceedings of *IV Jornadas Iberoamericanas de Ingenieria de Requisitos y Ambientes de Software*. Santo Domingo. Costa Rica : CIT, 2001.

[8] Kirsch-Pinheiro, Manuele. BW : A Java Framework for Past Awareness Support. Available at <http://www.inf.ufrgs.br/~manuele/BW/>. Access: June, 2002.

[9] Lassila, Ora; Swick, Ralph. "Resource Description Framework (RDF) Model and Syntax Specification", W3C Recommendation: http://www.w3.org/TR/1999/REC-rdf-syntax

[10] Lemlouma, Tayeb; Layaïda, Nabil. "The Negotiation of Multimedia Content Services in Heterogeneous Environments", *The 8th International Conference on Multimedia Modeling*, CWI, Amsterdam, November 5-7, 2001.

[11] Lemlouma, Tayeb; Layaïda, Nabil. "Universal Profiling for Content Negotiation and Adaptation in Heterogeneous Environments", *W3C Workshop on Delivery Context*, W3C/INRIA Sophia-Antipolis, France, 4-5 March 2002.

[12] Preguiça, N.; Martins, J. L.; Domingues, H.; Duarte, S. Data management support for asunchronous groupware. Proceedings of *CSCW'00*. December 2-6, 2000. Philadelphia, USA. ACM Press. P. 69-78.

[13] Sohlenkamp, M.; Mambrey, P.; Prinz, W.; Fuchs, L.; Syri, A.; Pankoke-Babatz, U.; Klöckner, K.; Kolvenbach, S. Supporting the distributed German government with POLITeam. *Multimedia Tools and Applications*, Vol. 12, No. 2, 2000. Kluwer Academicc Publishers, Netherlands. P. 39-58.

[14] Lilley, Chris; Jackson, Dean. Scalable Vector Graphics (SVG). Available at <http://www.w3.org/Graphics/SVG/>. Access : June, 2002.

[15] Vatton, Irène. Welcome to Amaya. Available at <http://www.w3.org/Amaya/>. Access : August, 2001.

[16] W3C Working Draft, "Composite Capability/Preference Profiles(CC/PP): Structure and Vocabularies", 15 March 2001.

[17] W3C Note, "CC/PP Exchange Protocol Based on HTTP Extension Framework", 24 June 1999.

[18] WebDAV:org. Welcome to WebDAV Resources. Available in: <http://www.webdav.org/> Access : ago. 2001.

[19] WIDE. The WIDE Workflow Modeling Language. Available in: http://dis.sema.es/projects/WIDE/Documents/4080-2.pdf. Access: Jan. 2002.

[20] Workflow Management Coalition. The Workflow Reference Model. Brussels, Jan. 95. Available in: http://www.wfmc.org/standards/docs/tc003v11.pdf. Access: Jan. 2002.