Stratification of the Information Space in Web Based Information Systems

Marlene Villanova-Oliver
Laboratoire LSR-IMAG, marlene.villanova_oliver@imag.fr

Jerome Gensel
Laboratoire LSR-IMAG, jerome.gensel@imag.fr

Herve Martin
Laboratoire LSR-IMAG, herve.martin@imag.fr

Follow this and additional works at: http://aisel.aisnet.org/ecis2002

Recommended Citation
http://aisel.aisnet.org/ecis2002/85

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2002 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
STRATIFICATION OF THE INFORMATION SPACE
IN WEB-BASED INFORMATION SYSTEMS

Marlène Villanova-Oliver, Jérôme Gensel, Hervé Martin
Laboratoire LSR-IMAG, BP 72, 38402 Saint Martin d'Hères Cedex, France
tel: (+33) 476827280 fax: (+33) 476827287
e-mail: {Firstname.Name}@imag.fr

ABSTRACT

Web-based Information Systems (WIS) are used for diffusing and processing information over the Internet. Because of the large amounts of information they manage, it is crucial to adapt the delivered information to users and to give them a progressive access to information. For this purpose, we propose to stratify the Information Space of a WIS by decomposing it into personalized sub-Information Spaces. Multimedia data can also be stratified into multiple formats used for their presentation. These stratifications are described through a Progressive Access Model (PAM) written in UML. In order to personalize the progressive access, the PAM is linked to both the data model of the application domain and to a Generic User Model describing the rights and preferences of users. We also show that, based on these stratifications, navigation mechanisms allow users to access first some minimum and essential information, and then larger and/or smaller personalized Information Spaces.

1. INTRODUCTION

The Web is now considered a reliable medium for widely diffusing information and, unsurprisingly, more and more Information Systems, so-called Web-based Information Systems (WIS), are accessible on Internet. These WIS address various application domains such as e-business, education, geography, etc. As any other Information Systems, WIS are designed to manage very large sets of data and to offer specialized services. Moreover, in the context of the Web, they are supposed to offer navigation facilities (through hyperlinks and/or dynamic web pages) generally embedded into a multimedia presentation of information. Although this characteristic makes these systems attractive, the measurement of the quality of a WIS must not only take into account the graphical appearance or the quantity of available information but also its usefulness. The appropriateness of information a WIS delivers to its users turns out to be an acute problem when, designing such systems, one ignores that 1) all users do not need the same information, and 2) users do not need all the available information all the time.

The first point can be tackled by distinguishing users profiles and is often referred to as adaptability. Recent systems, developed by commercial companies (Interligo, 2001)(Broadvision, 2001), announce a new generation of web applications able to track down the behaviour of their users and to dynamically react, by adapting the presentation of information. More generally, a WIS is said to be adaptable when the user gets the impression that the system has been specially designed for her/him. Then, adaptability (and its dynamic version called adaptivity (Stephanidis & al, 1998)) can be defined as the ability a WIS has to provide its users with some relevant information with regard to their rights,
needs, individual characteristics and material configurations (WAP, browser, etc.), in terms of both content and presentation.

The second point concerns the delivery of information by WIS and the fact that users may occasionally only need some parts of information. Indeed, even when information is appropriate, users can be confronted to a cognitive overload, due to a too massive and/or difficult to understand quantity of information. In order to spare users the trouble of getting “lost in the hyperspace” (Theng, 1998), we propose to provide them with a gradual organization of information, which allows them to progressively access it.

In this paper, we describe how to organize, from the conceptual level, the data of a WIS in order to offer to its users a progressive and personal access to their own Information Space. An Information Space refers to a set of information related to the application domain generally described by a (Entity-Relation, Class-Association, XML, etc.) data model. The idea is to stratify the Information Space into different levels of detail, which will be exploited by progressive access mechanisms. Moreover, in order to improve adaptability, a stratification is attached to each identified (group of) user(s). Finally, the multiple representations of information resulting from this stratification can be used for defining the navigation schema of the WIS.

A Progressive Access Model (PAM) describes the way the information space is stratified. Through the PAM, each user or group of users can be given a personal information space, which is organized, in different levels through the notion of mask. When information is of multimedia type (text, image, sound, video), a multiple format can be associated with it in order to fit the user’s preferences at displaying time. The PAM is also a support for navigating in the WIS since the content of pages and hyperlinks between them can be obtained from the stratification of the Information Space.

The paper is organized as follows. In section 2, we advocate for the stratification of the Information Space through the presentation of two notions for progressive access, called mask and multiple format. We present in section 3 the Progressive Access Model (PAM) in charge of describing the stratification of each personalized information space. Adaptability to users is established via the close relation between the PAM and a Generic User Model. In section 4, we propose some navigation mechanisms, which are based on the stratification described in the PAM, and upon which the personalized progressive access relies. Some related works are compared in section 5. Conclusions and perspectives end the paper.

2. STRATIFICATION OF THE INFORMATION SPACE

Our goal is to give a WIS designer the means to stratify the global Information Space according to the interest of each distinguished group of users. On the one hand, a stratification gives users a progressive access to information through the definition of different layers which correspond to different levels of detail. On the other hand, these layers can be used for security and confidentiality matters: one layer eventually corresponds to a confidential sub-information space some users are authorized to access, while some are not. The stratification of an Information Space is based on the two notions of mask and multiple formats we present below.

2.1 The notion of mask

The notion of mask allows representing more or less completely a set of information. Let us define a Maskable Entity (ME) as a set of information, which can be decomposed in different subsets, called Representations of Maskable Entity (RoME). Each RoME corresponds to a level of detail of the ME and defines a mask (cf. Fig. 1). Moreover, we define an ordered relation on these RoMEs which is based on the set inclusion.
Let $ME = \{e_1, e_2, \ldots, e_n\}$ be a Maskable Entity, where $e_i$ is an element of information and such as $n = \text{Card (} ME \text{)} \geq 2$. In order to set-up a progressive access to the whole set of information corresponding to a ME, we define four rules, to which each RoME of ME has to complied:

- RoME $\in$ Set of parts of ME.
- RoME$_i$ corresponds to the level of detail $i$ available on the ME, where $1 \leq i \leq m$, $m$ being the maximum detail level authorized.
- RoME$_1 \neq \emptyset$, even at the first level of detail, at least one element of information is visible (i.e. can be accessed).
- for each $i < j$, RoME$_i \subseteq$ RoME$_j$, so the elements of information visible at the level of detail $i$ are also visible at each greater level of detail $j$. Moreover, at the level of detail $i+1$, at least one more element of ME is visible (strict inclusion) than at level $i$.

It can be noticed that RoMEs are seen as masks since they offer a certain level of detail but they also hide some elements of information of the ME. This is the reason why they can then be used for confidentiality purpose.

![Figure 1. Representation of a ME using three masks](image)

### 2.2 The notion of multiple format

Multimedia information (text, video, audio, image, etc.) becomes more and more frequently managed by WIS. Such information generally corresponds to voluminous data. As a consequence, users may not want multimedia information to be displayed for various reasons: because they lack time or interest, or even because they have an unsuited material configuration. In most of existing WIS however, access to multimedia information is mainly ruled by an 'all or nothing' logic. We propose here the notion of *multiple format* for increasing the flexibility in the access to a multimedia datum. This notion introduces additional structuring features for representing a multimedia information. The idea is to associate such a datum with more or less detailed formats, in order to give users the opportunity to access, when needed, a less costly (in terms of time, resource, attention, etc.) representation of it. For instance, using a multiple format, a video can be substituted by an excerpt or an abstract of it. Then, even when users do not have enough time to entirely see the video, they can still access a partial representation of it. Basically, we propose to associate four formats to a multimedia datum. They are listed below in a decreasing order of detail:

- the *Most-Complete Authorized Format* (MCAF) corresponds to the version of the multimedia datum a user is authorized to access (which is not necessarily the integral version of the datum specially when some parts of it have to remain confidential for a group of users). Next formats are built on the basis of this format;
- the *abstract format* gives a global but concise view of the multimedia datum MCAF;
the excerpt format gives less details about the content of the multimedia datum MCAF. It corresponds to a fragment of the datum's content obtained by an extraction mechanism applied to the MCAF;

the meta-information format is the simplest representation of a multimedia datum in terms of content. Such a meta information may refer to its title, its author, its date of creation, its texture, etc.

Our objective here is not to propose techniques for building these different formats but rather to offer a descriptive way to handle them.

The next section presents the Progressive Access Model, which integrates masks and multiple formats.

3. THE PROGRESSIVE ACCESS MODEL

3.1 General overview

In order to perform a personalized stratification of the Information Space, one initially needs information about both the application domain and the users of the WIS. The former is embedded in a model, which represents the objects of the real world managed by the WIS. In our approach this model is called Data Model and is described using UML notations (Rumbaugh & al, 1999). Information about users necessitates an explicit modelling of users, we achieve through a model called the Generic User Model (GUM). The GUM, described by a UML class diagram in Fig. 2, is a minimal model, which can be extended according to the characteristics of users and to the requirements of the application domain. The main classes of the GUM are described below.

**UserCategory** is an abstract class of the GUM dedicated to the description of users and groups of users. Its sub-classes **Group** and **IndividualUser** maintain information concerning respectively a group of users and an individual user. The membership of an individual user to one or more groups is modelled by the relation **is_member_of**. The Boolean attribute **isMainGroup** in the class association **Main_Group** links a user to one and only one group by default.

![Figure 2. Description of the Generic User Model in UML](image)

In the remainder of this section, we refer to the GUM via the class **UserCategory** and we call **user(s)** object(s) of this class in order to simplify the PAM description.

3.2 Minimal description of the PAM

The PAM associates one stratification of the Data Model with one user identified in the GUM as shown in the UML class diagram of Fig. 3.

The class **Information** refers to the information to be stratified. The class **Representation** denotes the different representations (RoMEs) available for information as they are described in section 2.1. The association **structures** between these classes is so that:
– 0 to \( n \) instance(s) of *Representation* can be linked to one couple of instances (*UserCategory*, *Information*). No (0) instance means that the user is given no progressive access to information. On the contrary, \( n \) instances of *Representation* offer to the user \( n \) levels of representation for this information.

– 1 to \( n \) instance(s) of *UserCategory* can be linked to one couple of instances (*Information*, *Representation*), allowing several users to share the same stratification.

– one and only one instance of *Information* is linked to one couple of instances (*UserCategory*, *Representation*), meaning that the representation is proper to one user information space.

![Figure 3. Minimal Description of the PAM in UML](image)

The previous description is minimal. Classes *Information* and *Representation* have to be extended in order to describe the stratification according to the formalism adopted in the data model. In the next section, these classes are extended assuming that the Data Model is described using a UML class diagram.

### 3.3 Description of the PAM for a Data Model in UML

As presented in section 2, we propose multiple representations for Maskable Entities and Attributes. In the context of a Data Model expressed by an UML class diagram, a Maskable Entity is either the Data Model itself or one of the structured constructs involved in the class diagram, namely a class, a class association or an association. In the PAM, we refer to these different UML entities through the meta-model of UML. These entities are stereotyped “meta-class” in Fig. 4.

We describe below the parts of the PAM dedicated to both the mask and the multiple format descriptions, starting with the specialization of the class *Information* into the sub-classes *Maskable_Entity* and *Attribute* respectively.

**Mask Description.** The class *Maskable_Entity* describes a ME as defined in section 2. A *Maskable_Entity* is either a data model (class *DataModel*) – expressed through a class diagram in our approach – or a *Construct*\(^1\) (i.e. a *Class*, a *Class_Association* or an *Association*) of this diagram.

The class *Representation* specializes into the class *RoME* in order to describe the different representations of a *Maskable_Entity\(^2\)*. Instances of the class *RoME*, linked to a couple of instances (*Maskable_Entity*, *UserCategory*) have all a distinct and sequential level number (attribute *levelNum*) corresponding to the detail level of this RoME according to the rule 4 (cf. section 2.1).

Two classes specialize the class *RoME*. *RoME_Schema* and *RoME_Construct* are used to define respectively how a *RoME* is built for a *DataModel* considered as a *Maskable_Entity* and for a *Construct* seen as a *Maskable_Entity*.

The class *RoME_Schema* is linked to the class *Construct* through the association *has_construct*. Linking a set of instances of *Construct* to one instance of *RoME_Schema* reflects that these constructs

\(^1\) This abstraction level simplifies the definition of Representations of Maskable Entity.

\(^2\) The specialization of the association *structures* actually expresses this in the PAM, but this is not shown in Fig. 4. for a question of visibility.
of the data model (whether they are classes, classes association or associations) are accessible at the level of detail to which corresponds this instance of RoME_Schema.

The class RoME_Construct is linked to the metaclass Attribute through the association has_attribute. Linking a set of instances of Attribute to one instance of RoME_Construct signifies that these attributes are accessible at the level of detail to which corresponds this instance of RoME_Construct.

**Multiple Format Description.** The class Attribute, as a sub-class of Information, is linked to the sub-class of Representation called Multiple Format. This class specializes into four sub-classes, each corresponding to a format as described in section 2.

### 3.4 Strategy of Stratification

When designing a WIS, progressive access is specified once the data model and the GUM have been described. The PAM design task consists for a WIS designer in defining for each group\(^3\), the appropriate stratification in terms of masks and multiple formats. The different levels of detail can be arbitrarily determined. One possible strategy is to estimate *a priori* degrees of relevance of information for a given user and to correlate them with the levels of detail: the more relevant information is, the lower the level at which this information is accessible.

**Figure 4. Extension of the PAM to fit a Data Model described by a UML class diagram**

### 4. NAVIGATION FEATURES BASED ON THE PAM

We show here how set up the navigational features of a WIS using the PAM. From the specifications of this model, we automatically provide user with a personalized information space through which she/he can navigate and progressively access information.

The PAM addresses the stratification considering three granularities of information corresponding to the two kinds of Maskable Entities (i.e. data schema and constructs) plus the attributes. For each of these granularities of information, several representations are offered (RoME and Multiple Format).

---

\(^3\) Individual users are initially provided with the same progressive access as their group. They will define their own structuring using the WIS, if needed and if they are allowed to.
The navigation mechanisms we propose exploit both granularity and representations of information. We distinguish between two types of navigation which are described in the next sections.

4.1 Intra-Granular Navigation

The Intra-granular navigation gives access to the different representations inside a Maskable Entity or an Attribute. The mechanism we propose for this navigation is based on a masking and an unmasking functions. Concerning a Maskable Entity (cf. Fig. 5a), from a given RoME_i, the masking function gives the immediately less detailed RoME_{i-1}, if it exists, while the unmasking function gives the immediately more detailed RoME_{i+1}, if it exists:

- \( \text{masking} (\text{RoME}_i) \rightarrow \text{RoME}_{i-1} \) with \( 2 \leq i \leq m \), \( m \) being the maximum detail level,
- \( \text{unmasking} (\text{RoME}_i) \rightarrow \text{RoME}_{i+1} \) with \( 1 \leq i \leq m-1 \).

For an Attribute, the two functions are defined as follows (cf. Fig. 5b):

- \( \text{masking} (\text{MCAF}) \rightarrow \text{Abstract} \)
- \( \text{unmasking} (\text{Meta-Info}) \rightarrow \text{Excerpt} \)
- \( \text{masking} (\text{Abstract}) \rightarrow \text{Excerpt} \)
- \( \text{unmasking} (\text{Excerpt}) \rightarrow \text{Abstract} \)
- \( \text{masking} (\text{Excerpt}) \rightarrow \text{Meta-Info} \)
- \( \text{unmasking} (\text{Abstract}) \rightarrow \text{MCAF} \)

![Figure 5. Intra-granular navigation using the masking and unmasking functions, applied a) to the RoMEs of a Makable Entity and b) to the Multiple Format of an Attribute.](image)

4.2 Inter-Granular Navigation

The Inter-granular navigation allows to navigate between the different granularities of information. From one representation at a given granularity, one can access a representation at a coarser or finer granularity, if it exists, and provided that these representations are linked by a structure/element relation. For instance (cf. Fig 6), from the RoME_i of a data schema, a representation defined for a class \( C \) can be reached, provided that \( C \) belongs to RoME_i. We call zooming in the action of accessing information at a finer granularity, and zooming out, the opposite (e.g. accessing RoME_i of the data schema from a representation of \( C \)). The zooming in function applies on a target (i.e. one of the elements of the current representation of information) and gives access to the less detailed representation of this element. The zooming out function applies on the current representation of information gives access to the representation of the structure to which this current representation belongs. We explain below the possible cases of the inter-granular navigation:

- from S.RoME_i, a representation of a data schema S, the zooming in function can be applied to any class or class association \( C \) of S.RoME_i:
  \( \text{zooming in} (C) \rightarrow C.\text{RoME}_i \)
- from C.RoME_i, a representation of a class \( C \), the zooming in function can be applied to any attribute \( A \) of C.RoME_i, multiple formats being available for \( A \):
  \( \text{zooming in} (A) \rightarrow A.\text{MetaInformation} \)
– from A.MFi a multiple format of an attribute A, applying the \textit{zooming out} function gives access to the class C to which belongs A. C is represented by its last visited\(^4\) RoME\(_j\):

\[ \text{zooming out} \rightarrow \text{C. RoME}\(_j\) \]

– from C.RoME\(_i\) a representation of a class C, the application of the \textit{zooming out} function allows to access the data schema S to which belongs C. S is represented by its last visited\(^4\) RoME\(_j\):

\[ \text{zooming out} \rightarrow \text{S.RoME}\(_j\) \]

The implementation of the different functions we propose for the intra- and inter- granularity allows personalized progressive access to information, since the stratifications described by the PAM – and exploited by the functions – are themselves associated with groups of users.

\[\text{Figure 6. Inter-granular navigation using zooming in and zooming out functions}\]

### 5. RELATED WORK

We compare here our work with proposals integrating notions close to the stratification. \textit{Views} (Tsichritzis \& \textit{al}, 1977) which were first defined in Data Bases (but since broadly applied to OODBMS (Abiteboul \& \textit{al}, 1991), to XML (Abiteboul \& \textit{al}, 1999), etc.) are also used to provide the user with a convenient representation of information, according to her/his needs. Then RoMEs can be considered as ordered series of views. But, to our knowledge, there is no work dealing with a stratification relying on several views as the one presented in this paper and favouring a progressive access to information.

\textit{Splits objects} (Bardou \& \textit{al}, 1996) and RoMEs share the underlying common idea to split a concept in order to favour the access to different points of view of this concept. However, one main difference is that split objects, stemming from the theory of prototypes, are defined on objects only, while masks propose a stratified definition at a higher level of abstraction which concerns classes but also the whole schema of a data model.

The notion of \textit{context} (a higher order conceptual entity that describes a group of conceptual entities from a particular standpoint) (Motschnig-Pitrik, 1999) can also be considered as a basis for stratifying an information space into RoME, at a schema level. Masks can be seen as a specification of contexts having a variable size: the information space is stratified in order to be either enlarged or reduced. Our approach is close to the proposal of (Theodorakis \& \textit{al}, 1999) where the notion of context is enriched in order to partially mask information. However, this approach does \textit{not} put the emphasis on adaptability to user needs as ours does.

In WebML (Ceri \& \textit{al}, 2000), the concept of \textit{data unit} allows to define short or long, multimedia or textual, versions of an entity, what could be assimilated to our approach. But entities in WebML only

\[\text{As classically encountered in the context of a Web navigation.}\]
correspond to the notion of class in our work, which handles information at different granularities. Also, we define, at the conceptual level, navigation mechanisms based on the stratification and automatically adapted to users (thanks to the link between the PAM and the GUM). The WebML approach requires the explicit definition of this kind of links between the representations in the navigation schemas that are offered to users and does not supply any progressive access features.

6. CONCLUSIONS AND PERSPECTIVES

In this paper, we have presented a way to stratify the Information Space (data model) of a Web-based Information System (WIS) in order to provide users with a progressive access to information. This stratification applies on a set of information (called Maskable Entity) and consists in decomposing it into sub-sets (called RoME) ordered by a set inclusion relation. The stratification can also applies on multimedia data and consists in associating them with multiple formats to be used for their presentation. The description of this stratification is achieved through a Progressive Access Model (PAM) written in UML. In order to personalize the progressive access, the PAM is linked to both the data model of the application domain and a model, called Generic User Model (GUM) which maintains knowledge about users. We have described also how these models can be used to establish a navigation schema, by proposing two kinds of navigation. Based on the stratification, navigation mechanisms allow users of a WIS to access first essential information, through three levels of granularity, and then, depending on their interest, time or material configuration, etc., to access more or less information. To our opinion, the stratification and the navigation features proposed here through the PAM could be applied more generally on any kind of Information Systems, not only on Web-based ones, and not only on those adopting a class/association model.

Results of this study have been implemented into KIWIS (Villanova-Oliver & al, 2002), a platform for designing and deploying adaptable WIS. KIWIS integrates a Data Model, the GUM and the PAM. These three models have been implemented using AROM (Page & al, 2001), an Object-based Knowledge Representation System whose representation language is based on UML-like notations. For a designer, modelling a WIS using KIWIS consists in instantiating these different models, and to specify which graphical charters are to be used. These models are then translated in XML before KIWIS automatically generates the so-specified WIS. From the stratification described in the PAM, the associated navigation mechanisms and graphical charters for presentation, users of the generated WIS have a progressive and personalized access to information.

Using KIWIS, the stratification proposed here has been first experimented in the SPHERE project (Davoine & al, 2001) consisting of an Information System dedicated to geographic and historical data on river floods. Masks and multiple formats have shown to be of a rather intuitive use for different categories of users (experts, city hall employees, …) who consult data about the same theme (floods) but at different levels of detail.

Our research are now directed towards two directions. First, we address the security issue in WIS by integrating to our design approach a step for expressing groups and users rights on data at the conceptual level. The resulting security model will extend the GUM and will also offer a support for expressing of the security features of the actions performed on the system. Second, we investigate dynamic adaptability techniques in order to react more efficiently to end-users actions. The idea is to dynamically elaborate and modify both the stratification and the navigation schema of a WIS, learning from the user's behavior. The use of cookies technology to track information about users’ sessions in order to automatically adapt information, coupled with some meta-rules of navigation, is one way we have started to explore.
REFERENCES