

# An Integrated Approach to Rating and Filtering Web Content

Elisa Bertino<sup>1</sup>, Elena Ferrari<sup>2</sup>, Andrea Perego<sup>3</sup>, and Gian Piero Zarri<sup>4</sup>

<sup>1</sup> CERIAS, Purdue University, IN, USA  
`bertino@cerias.purdue.edu`

<sup>2</sup> DSCPI, Università degli Studi dell'Insubria, Como, Italy  
`elena.ferrari@uninsubria.it`

<sup>3</sup> DICO, Università degli Studi di Milano, Italy  
`perego@dico.unimi.it`

<sup>4</sup> LaLICC, Université Paris IV/Sorbonne, France  
`gpzarri@paris4.sorbonne.fr`

**Abstract.** In this poster, we will illustrate an integrated approach to Web filtering, whose main features are flexible filtering policies taking into account both users' characteristics and resource content, the specification of an ontology for the filtering domain, and the support for the main filtering strategies currently available. Our approach has been implemented in two prototypes, which address the needs of both home and institutional users, and which enforce filtering strategies more sophisticated and flexible than the ones currently available.

Web content filtering concerns the evaluation of Web resources in order to verify whether they satisfy given parameters. Although such definition is quite general, and it applies to diverse applications, Web filtering has been enforced so far mainly in order to protect users (e.g., minors) from possible 'harmful' content (e.g., pornography, violence, racism).

The filtering systems currently available can be grouped into two main classes. The former adopts a *list*-based approach, according to which Web sites are classified either as 'appropriate' (*white lists*) or 'inappropriate' (*black lists*). In the latter, Web resources are described by metadata associated with them, which are used for evaluating whether they can be accessed or not, depending on the preferences specified by the end user or a supervisor. Such approach is adopted mainly by the rating systems based on the PICS (Platform for Internet Content Selection) W3C standard [1], which defines a general format for *content labels* to be associated with Web sites.

Both such strategies have been criticized for enforcing a restrictive and rather ineffective filtering. In fact, their classification of Web resources is semantically poor, which does not allow to distinguish between categories concerning similar contents (e.g., pornography and gynecology). For the same reason, they often *under-* and/or *over-*block the access to the Web—i.e., respectively, they allow users to access inappropriate resources, or they prevent users from accessing appropriate resources. The metadata-based approach should overcome such drawbacks, since it would allow one to specify a precise and unambiguous description

of resources, but this is not true for the available metadata-based rating and filtering systems.

In order to address the issues of Web content rating and filtering, we developed an integrated approach which, besides supporting both the list- and metadata-based strategies, defines content labels providing an accurate description of Web resources and takes into account users' characteristics in order to enforce flexible filtering policies. The outcome of our work, formerly carried out in the framework of the EU project EUFORBIA<sup>1</sup>, has been two prototypes, addressing the needs of institutional and home users, and an ontology (namely, the EUFORBIA ontology) for the specification of content labels.

The EUFORBIA ontology is an extension concerning the pornography, violence, and racism domains, of the general NKRL (Narrative Knowledge Representation Language) ontology [2]. NKRL is used to specify the EUFORBIA content labels, which consist of three sections: the first concerns the aims of the Web site, the second describes its relevant characteristics and content, whereas the third outlines the Web site's main sections. It is important to note that, differently from the currently available PICS-based content labels, a EUFORBIA label does not rate a Web site only with respect to the contents liable to be filtered, but, since the NKRL ontology is general purpose, it provides a precise and objective description of its content and characteristics. As a result, we can specify policies more sophisticated than, e.g., "user  $u$  cannot access pornographic Web sites", and it is possible to distinguish more precisely between, e.g., an actually pornographic Web site and a Web site addressing sexual topics and contents from a non-pornographic (e.g., medical) point of view.

The EUFORBIA ontology and the corresponding labels are used by two filtering prototypes which enforce complementary strategies for addressing end users' needs.

The former prototype, referred to as NKRL-EUFORBIA [3], allows end users to generate and associate EUFORBIA labels with Web resources, and to build a user profile by specifying NKRL-encoded filtering policies. NKRL-EUFORBIA can run either server- or client-side, and it consists of three main modules: the *Annotation Manager*, which allows the creation of well-formed NKRL 'conceptual annotations' to be used for encoding EUFORBIA labels, the *Web Browser*, which allows the definition of a user profile and a safe navigation over the Internet, and finally the *Web Filter*, which is used by the Web Browser module in order to determine whether the access to a requested resource must be granted or not.

The latter EUFORBIA prototype [3, 4], whose current version is referred to as MFILTER [5], is a proxy filter specifically designed for institutional users, who must manage the access to Web content for a high number of heterogeneous users. MFILTER implements a model according to which filtering policies can be specified on either users'/resource identity or characteristics. Users are characterized by associating with them *ratings*, organized into a hierarchy and denoted

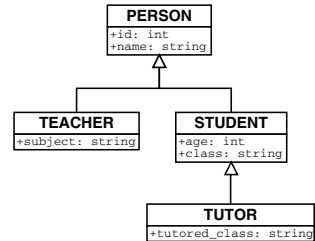
---

<sup>1</sup> For detailed information concerning EUFORBIA, we refer the reader to the project Web site: <http://semioweb.msh-paris.fr/euforbias>

by a set, possibly empty, of attributes. An example of user rating system is depicted in Figure 1. Thus, differently from the available filtering systems, which make use of predefined and static profiles, MFILTER allows one to specify policies which take into account both user ratings and attribute values (e.g., “all the students whose age is less than 16”).

Resources are rated according to the metadata-based strategy, with the difference that MFILTER makes use of multiple rating systems, for which a uniform hierarchical representation is adopted. Currently, MFILTER supports the EUFORBIA ontology and any PICS-based rating systems. Thanks to the hierarchical organization of both user and resource ratings, we can exploit a policy propagation principle according to which a policy concerning a given rating applies also to its children. As a result, we can dramatically reduce the number of policies that need to be specified. Moreover, MFILTER supports both positive and negative policies in order to allow for exceptions to the propagation principle mentioned above. Finally, since the decisions about what users can or cannot access may be shared among several persons (e.g., parents and teachers, in a school context), MFILTER enforces *supervised* filtering techniques, according to which the policies specified by the service administrator must be validated by *supervisors*.

The integrated approach we present in this poster is an attempt to enhance metadata-based rating and filtering in order to address the heterogeneous requirements of the Web users’ community by supporting accurate content labels and flexible filtering policies. Moreover, our approach is fully compliant with the core Semantic Web technologies—namely, RDF and OWL—and it can be easily integrated into the W3C Web Service architecture, of which Web filtering is one of the main components.



**Fig. 1.** A user rating system

## References

1. Resnick, P., Miller, J.: PICS: Internet access controls without censorship. *Communications of the ACM* **39** (1996) 87–93
2. Zarri, G.P.: A conceptual model for representing narratives. In Jain, R., Abraham, A., Faucher, C., van der Zwaag, B., eds.: *Innovations in Knowledge Engineering*. Advanced Knowledge International, Adelaide (2003)
3. Zarri, G.P., Wu, P., Nallet, B., Pires, F., Abreu, P., Allen, D., Bertino, E., Ferrari, E., Perego, A., Mantegazza, D.: Report on the Final Enhancements to the EUFORBIA Demonstrator. EUFORBIA Deliverable D12, CNRS, Paris (2003) <http://semioweb.msh-paris.fr/euforbia/download/D12pdf.zip>.
4. Bertino, E., Ferrari, E., Perego, A.: Content-based filtering of Web documents: The MaX system and the EUFORBIA project. *International Journal of Information Security* **2** (2003) 45–58
5. Bertino, E., Ferrari, E., Perego, A.: Web content filtering. In Ferrari, E., Thuraisingham, B., eds.: *Web and Information Security*. IDEA Group (2005) To appear.