

Author	Gian Piero Zarri	Organisation	CNRS		
Title	Report on the 2 nd six months meeting of the EUFORBIA project				
Publication Date	24/1/2002	Document Reference	CNRS/Note/3	Task Reference	WP3, WP4
Circulation (C,R,P)	R	Document Type	Note	Document Version	1.1
Deliverable Reference	D5, D9				

EUFORBIA SECOND SIX MONTHS MEETING

(AXON, Hotel Le Meridien, Oporto, Portugal)

MINUTES

January 10th – 12th, 2002

Delegates:

Centre National de la Recherche Scientifique, France: Gian Piero ZARRI;

AXON, Portugal: Gilberto MACHADO, Fátima PIRES, Pedro ABREU;

PIRA International, United Kingdom: Don ALLEN, Anthony CONSTABLE-MAXWELL;

Department of Computer Science of the University of Milan, Italy: Elena FERRARI.
Andrea PEREGO, Andrea MERCANTE.

Apologies for absence have been received by Peter STOCKINGER (MSH, France).

Thursday, January 10th, 2001

The meeting opens at 2 pm at Le Meridien Hotel in Oporto. Gilberto Machado, C.E.O. of AXON, welcomes the delegates.

Gian Piero Zarri introduces shortly some administrative topics. The time has come to prepare the new Six Months Report. To do this, all the EUFORBIA teams must send to CNRS i) the (signed) cost statements for the period between July 1st, 2001, and December 31st, 2001; ii) a short presentation of the scientific work they have accomplished in the last semester. With respect to Deliverable D10 – which, in principle, should be devoted to the description of the final EUFORBIA architecture – all the EUFORBIA partners have already agreed to make use of this Deliverable to present a “state of the art” of the project. The architecture did not undergo, in fact, any major changes from the first, “provisional” version presented in Deliverable D3. Mr. Zarri hopes to send to the Commission both the Sixth Months Report and the Deliverable D10 by the beginning of February. The EUFORBIA Official Review is still expected for the mid-April, 2002; the importance of this event for the EUFORBIA consortium must be continuously emphasised.

Fátima Pires presents the demo concerning the software module (the “Annotation Manager”) that allows to associate “EUFORBIA labels” with the potentially harmful Internet sites: the sites can be labelled either i) during their first construction, or ii) at the moment they are submitted to some major revamping. This module represents both the Deliverable D5 and the conclusion of the task T3.2, “Implementation of the EUFORBIA conceptual annotation module”. Fátima emphasises that this work has implied a complete rebuilding of the analogous module inherited from the CONCERTO project, which was especially “unfriendly” to use. In particular, a new, important feature has been added that should allow to speed up considerably the coding of the labels. When an NKRL template has been selected, the operator can choice to fill its slots either by using the keyboard – and in this case the system will check the coherence of the new entry

with respect to the H_CLASS ontology – or s/he can choose to open H_CLASS and, by clicking on the concept/individual s/he wants to insert, transfer directly the right term in the slot. Even if, of course, many improvements could be added to this module, the audience estimate unanimously that it has already attained a remarkable level of flexibility and utility, and do not estimate that further work on it is really necessary. Some additions/improvements can, in case, be envisaged for the last phase of the project, after the Official Revue.

Andrea Perego presents the new Java implementation of the “Milan model”, task T4.2 and Deliverable D9. The demo cannot be presented entirely because of the lack of an Internet connection in the meeting room: in particular, the proper filtering functions cannot be demonstrated. Andrea adds that, anyway, these functions must still be perfected. The demo deals then mainly with a presentation of the functionalities of the Web interface that has been built up for the Java implementation of the model. The interface is organised into three main components: i) a “Configuration interface”, by which the system administrator can build an EUFORBIA “user profile” according to a credential based mode of functioning; ii) an “administrative interface”, for the management of the filtering system; iii) a “User Interface”, which deals with user authentication. A “General Online Help” and a “Contextual Online help” are also provided. This work is favourably received by the assistance. Some remarks concerns the layout of the interface, which is considered as somewhat too “crowded”. The Milan team will try to introduce some modifications in this interface before the Review.

The two demos will be, of course, both presented at the Official Review in Luxembourg. After a short general discussion, the meeting is adjourned at 5h50 pm.

Friday, January 11th, 2002

The meeting resumes at 9h30 am. Don Allen presents the work accomplished at Pira to collect and organise a first set of terms (dealing mainly with pornography) to be used to specialise the NKRL ontology, H_CLASS, towards the specific EUFORBIA tasks. Don illustrates both the methodology followed (an application of Vickery’s methodology of “facets”) and the sources utilised. The final result is quite impressive, given that about 600 terms have been collected only for the specific “pornography” domain. The final version of this work will be included in Deliverable D10; Pira will now start an analogous work in the “violence”, “racism” and “hate speech” domains.

G.P. Zarri remarks that, to pass from the “lexical” (linguistic) level represented by the terms collected by Pira to the “conceptual” level, the main work to be accomplished has concerned grouping together under a unique conceptual label terms that, even if they were completely different from the lexical point of view, they referred in reality the same “concept” – sometimes, of course, some details may be lost in the process. A trivial example concerns, e.g., “prostitute” and “whore”. The concepts obtained after this “grouping” operation have been structured hierarchically and added to the CONCERTO

version of H_CLASS. This solution has been preferred to that of building up a specific EUFORBIA ontology mainly because other concepts, in addition to those specifically associated with the pornography/racism/violence/hate speech domains, will have to be used for the filtering operations.

The last part of the meeting has concerned a preliminary discussion about a possible organisation of the Official Review. Assuming that this review will last a full day, the morning could be devoted to a general presentation of the project and the results obtained, reserving the afternoon for the demos. Note that (at least) a demo on the filtering of harmful sites making a direct use of the EUFORBIA labels will be added to the two presented during the Oporto meeting. More specifically, the morning presentation should include:

- A reminder about the aims of the project and its basic assumptions: very in short, the assumption that an *in-depth description* of the “semantic content” of the Internet sites – which is impossible to obtain with the traditional approaches – should allow the implementation of more sophisticated filtering strategies.
- A description – very plain and without any formal development – of the basic principles underpinning the two main tools used in EUFORBIA, NKRL and the Milan model.
- A description of the work accomplished in the ontological domain – as already noticed in other occasions, we think that this work will constitute one of the main results of the project.
- A description of the architecture of the general EUFORBIA prototype, by illustrating its functionalities and by emphasising the existence of a “bridge” between the two components (NKRL and Milan model) of the prototype represented by the existence of a unique ontology (H_CLASS).

If the provisional date (mid-April) for the Official Review is confirmed, the consortium agrees to have a technical meeting at the University of Milan on Monday, March 18th, 2002.

The meeting ends at 12h30 p.m. of Friday.

In the afternoon of Friday, an internal meeting between CNRS and AXON has concerned the set up of the filtering module that will operate directly on the EUFORBIA labels. In the morning of Saturday, January 12th, the delegates have been invited by M. Machado to assist to a demo of the ALICE system developed by AXON. This system concerns the overall automation of the real estate registration office in Oporto; the EUFORBIA delegates have particularly appreciated this demo.