

Minutes of the LIRICS Industry Advisory Group
UPF 20-21 June 2005
Version 1 of the document

LIRICS Attendees:

Monica Monachini	CNR
Nuria Bel	UPF
Claudia Soria	CNR
Nicoletta Calzolari	CNR
Gerhard Budin	UW
Julien Nioche	USFD
Thierry Declerck	DFKI
Lee Gillam	Unis
Peter Wittenburg	MPI
Harry Bunt	UTiL
Gil Francopoulo	INRIA

IAG attendees

Elsa Sklavounou	Systran
Gudrun Magnusdottir	ESTeam AB
Steve Crowdy	Pearson
Theo van den Heuvel	Polderland
Frederik Cailliau	Sinequa
Maria Luisa Provitera	Thamus
Paolo Poto	Expert System
Gabor Proszeky	Morphologic
Pat Croke	HP

Author of the minutes: Gil Francopoulo with the help of Julien Nioche and Monica Monachini.

20th June 2005

The LIRICS members presented the files that are located on the LIRICS web site (see <http://lirics.loria.fr> + Event + LIRICS IAG meeting):

Gerhard Budin:

- LIRICS IAG meeting
- International Standards in Lexicography

Thierry Declerck

- LIRICS MAF API a quick overview
- LIRICS WP3
- Presentation of actual work on standardization for language resources at DIN
- MAF-SynAF

Monica Monachini

- LIRICS WP2 NLP lexica

Gil Francopoulo

- Data Category Registry, LMF, TMF etc. what is the situation?

21st June 2005

The LIRICS IAG members presented the files that are located on the LIRICS web site (see <http://lirics.loria.fr> + Event + LIRICS IAG meeting):

Paolo Poto

In the universe of knowledge with linguistic intelligence and semantic logic

Gabor Proszeky

Introducing morphologic to LIRICS

Theo van den Heuvel

Polderland Language & speech technology

Elsa Sklavounou

Company overview Systran applications customization for quality translations

Maria Luisa Provitera

Thamus final

Thamus presentation

Thamus products

Thamus schermate

Steve Crowdy (the file is not on the web site)

Lexicons at Pearson-Longman-Penguin

Pat Croke (the file is not on the web site)

Semantic content management at HP

Frederik Cailliau (strictly oral presentation)

Sinequa

Gudrun Magnusdottir (strictly oral presentation)

ESTeam AB

During the debate that follows, a certain number of points have been addressed. The IAG members said that:

The ISO documents must define the used terms in the standard documents because some technical terms like "collocation" have different meanings in the NLP field. It is important to avoid any misunderstanding among the readers.

Multiword expressions must be carefully taken into account. In certain languages, MWE are not completely frozen: some MWE allow a certain number of variations. And these variations usually respect the grammar rules of the given language.

The Hungarian member recalls that it is illusory to try to represent explicitly all the inflected forms of Hungarian in a lexicon (i.e. in extension). In Hungarian, the only possibility for describing morphology is to represent the forms in intension.

There is also another phenomenon that appears in Hungarian is that a word in Hungarian does not always match to a word in another language (say English). Due to the fact that the morphological system permits to mark a word with possessive or preposition marks, very often an Hungarian word does not match a word in another language but a fragment of text.