



# LIRICS

## Deliverable D6.2

### Lirics project presentation

|  |   |
|--|---|
| Project reference number                     | e-Content-22236-LIRICS  |
| Project acronym                              | LIRICS  |
| Project full title                           | Linguistic Infrastructure for Interoperable Resource and Systems  |
| Project contact point                        | Laurent Romary, INRIA-Loria<br>615, rue du jardin botanique BP101.<br>54602 Villers lès Nancy (France)<br>romary@loria.fr |
| Project web site                             | <a href="http://lirics.loria.fr">http://lirics.loria.fr</a>   |
| EC project officer                           | Erwin Valentini   |
| Document title                               | Lirics project presentation   |
| Deliverable ID                               | 6.2   |
| Document type                                | Report  |
| Dissemination level                          | Public  |
| Contractual date of delivery                 | M1  |
| Actual date of delivery                      | 31st January 2005   |
| Status & version                             | Draft   |
| Work package, task & deliverable responsible | INRIA-Loria   |
| Author(s) & affiliation(s)                   | Laurence Kbida (INRIA-Loria ) & Gil Francopoulo, (INRIA-Loria)  |
| Additional contributor(s)                    |   |
| Keywords                                     | Web site, mailing list, visibility  |

### Document evolution

| version | date              | version | date |
|---------|-------------------|---------|------|
| 1.0     | 31st January 2005 |         |      |
| 1.1     | 18th August 2005  |         |      |
|         |                   |         |      |
|         |                   |         |      |
|         |                   |         |      |

## **1 Introduction**

The aim of this document is to help dissemination activities and conferences amongst the Lirics partners and for EC purposes.

The document contains three sections:

- a power point presentation description
- a list of empty Lirics documents in order to get a common style sheet with a coherent look and feel
- the Lirics web site description

## **2 Lirics Power Point Presentation**

A Lirics power point presentation has been written and uploaded on the web server.

This presentation has been used during the kick off meeting in Luxembourg.

The presentation exists in two versions: a short presentation containing 8 pages and a longer one with 12 pages. Two formats are available: PDF and PPT. These files are available on the Lirics web site.

The long version can be found as an annex of the current document.

This presentation is not to be considered as mandatory for all Lirics members but is to be considered primarily as a help. So they can pick whatever slides they want.

## **3 List of empty Lirics documents**

In order to have a common style sheet two empty MS-Word files are provided:

- Empty Progress Report
- Empty word document

These two files are available on the Lirics web site.

## **4 Lirics web site description**

### **4.1 Introduction**

A Lirics web site has been set up and is maintained on one of the INRIA-Loria's servers. It can be visited at the address: <http://lirics.loria.fr>

### **4.2 Description**

The Lirics web site has two objectives:

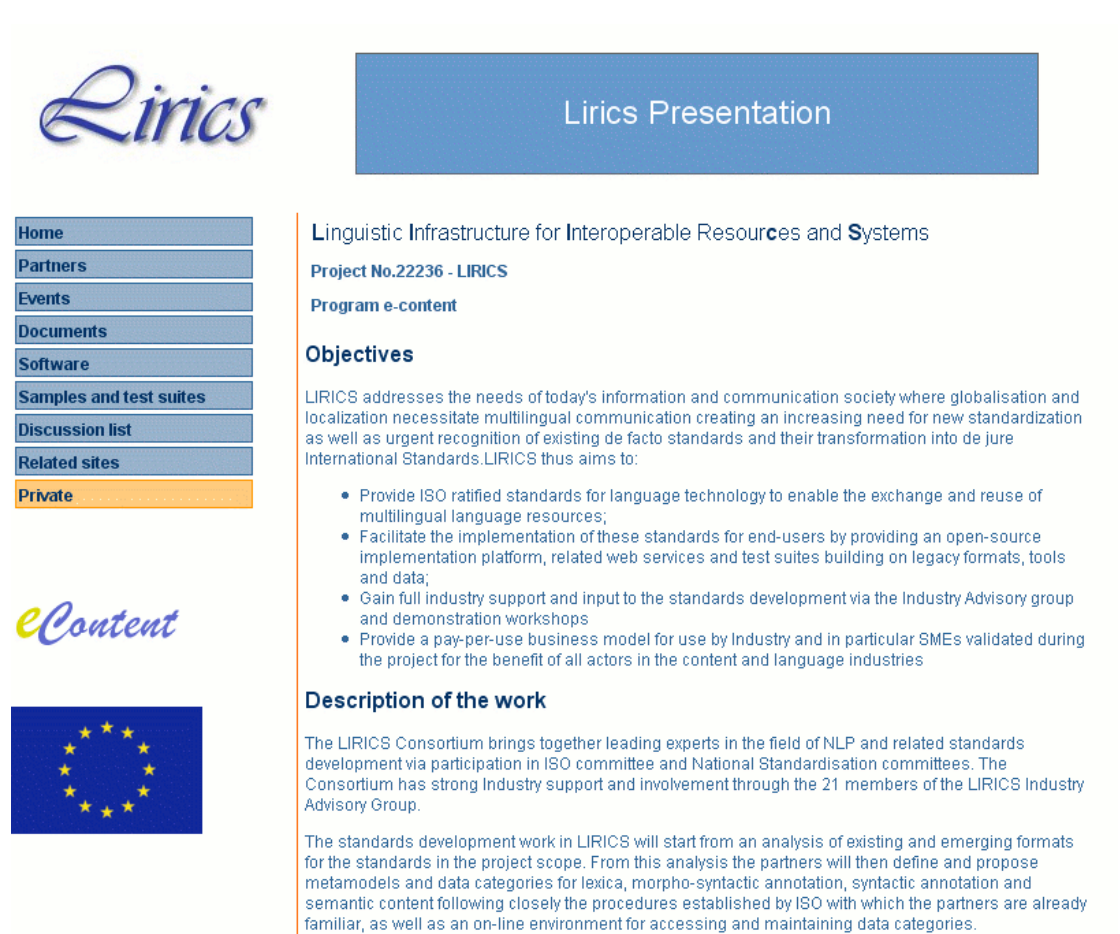
- provide information to visitors
- provide an efficient tool for the Lirics partners in order to share intermediate results

The web site provides visibility for the project for both visitors and partners. It presents an overview of the project: objectives, partners, progress and events.

It's also an area to archive various Lirics related documents:

- ISO documents
- Lirics deliverables
- Presentations
- Auxiliary technical reports

Each Lirics member can ask the web master to upload a file. Each member can download these files from the server to their own computing facilities.



The screenshot shows a presentation slide for the Lirics project. At the top left is the 'Lirics' logo in a blue, cursive font. To its right is a blue rectangular box with the text 'Lirics Presentation' in white. Below the logo is a vertical navigation menu with buttons for 'Home', 'Partners', 'Events', 'Documents', 'Software', 'Samples and test suites', 'Discussion list', 'Related sites', and 'Private' (which is highlighted in orange). To the right of the menu, the main content area is titled 'Linguistic Infrastructure for Interoperable Resources and Systems' and 'Project No.22236 - LIRICS'. It includes a section for 'Program e-content' and 'Objectives', which lists four bullet points: providing ISO ratified standards, facilitating implementation, gaining industry support, and providing a pay-per-use business model. Below this is a 'Description of the work' section, which explains that the LIRICS Consortium brings together leading experts in NLP and related standards development via participation in ISO and National Standardisation committees. It also mentions that standards development work will start from an analysis of existing and emerging formats for the standards in the project scope.

|                         |
|-------------------------|
| Home                    |
| Partners                |
| Events                  |
| Documents               |
| Software                |
| Samples and test suites |
| Discussion list         |
| Related sites           |
| Private                 |
| <b>Workpackages</b>     |
| Meta-data               |
| Morpho-syntax           |
| Syntax                  |
| Lexicon                 |
| Semantics               |

## Workpackage list

|     | Title  | Lead Contractor | Person-Months | Start Month | End Month | Deliverables   |
|-----|--|-----------------|---------------|-------------|-----------|--|
| WP1 | Infrastructure for Standards Development and Quality Assurance | INRIA           | 34            | M1          | M30       | <b>D1.1, D1.2, D1.3</b>  |
| WP2 | NLP Lexica   | CNR-ILC         | 45.1          | M3          | M30       | <b>D2.1, D2.2A, D2.2B, D2.3</b>  |
| WP3 | Morpho-syntactic and syntactic annotations                     | DFKI            | 43            | M3          | M30       | <b>D3.1, D3.2A, D3.2B, D3.3A, D3.3B, D3.4</b>  |
| WP4 | Semantic content   | UTIL            | 33.4          | M3          | M30       | <b>D4.1, D4.2, D4.3, D4.4</b>  |
| WP5 | LIRICS Reference Implementation platform                       | USFD            | 87.2          | M3          | M30       | <b>D5.1A, D5.1B, D5.1C, D5.1D, D5.1E, D5.2A, D5.2B, D5.2C, D5.2D, D5.3A, D5.3B, D5.4A, D5.4B</b> |
| WP6 | Dissemination and Exploitation                                 | UW              | 39.9          | M1          | M30       | <b>D6.1, D6.2, D6.3A, D6.3B, D6.4</b>  |
| WP7 | Management   | INRIA           | 24.5          | M1          | M30       | <b>D7.1, D7.2, D7.3, D7.4, D7.5</b>  |
|     | <b>TOTAL</b>   |                 | <b>307.1</b>  |             |           |  |

## Deliverables List

| Deliverable | Title   | Date           | Nature | Dissemination level |
|-------------|---|----------------|--------|---------------------|
| D6.2        | <b>LIRICS Project presentation (pdf) (ppt)</b>                            | M1             | Report | PU                  |
| D6.1        | LIRICS Dissemination plan   | M2             | Report | PU                  |
| D7.1        | Risk Management plan  | M2             | Report | PP                  |
| D1.1        | <b>Guidelines and tools for producing standards, test-suites and APIs</b> | M3             | Report | PU                  |
| D7.2        | <b>LIRICS Quality Plan</b>  | M3             | Report | PP                  |
| D5.1.A      | API for Data Category Registry  | M6 v1, M1 2 v2 | Report | PU                  |
| D7.3        | Periodic Progress Report (M6)   | M6             | Report | PP                  |

The site is divided in two parts corresponding to two different access rights:

- a public area for any visitor who wants to obtain information about the project,
- a private area for the participants of the project with the complete list of the deliverables, access to the internal documents or products, information.

Depending on specific needs, more sections will be implemented as the project gradually progresses. At the end of the project, the site will bring together all the information and documents of the project.

The web site is associated with the Lirics mailing list: [lirics@loria.fr](mailto:lirics@loria.fr) which allows the participants to communicate and exchange about work documents and organization.

## Annex A Lirics Presentation

# LIRICS

## Linguistic Infrastructure for Interoperable Resources and Systems

- ▶ Kick off meeting presentation
- ▶ Proposal N° 22.236
- ▶ Presented by Laurent Romary  
(INRIA, France, chair of ISO-TC37/SC4)

24 Jan 2005

Kick off meeting (Luxembourg)

1

## Scope

- Europe being a mosaic of languages, the processing of multilingual linguistic data concerns a lot of people in Europe
- And the recent expansion to 10 new EU members will intensify this task
- Of course, various linguistic data already exist all over Europe
- **But today there exists no established standard to enable interoperability and re-use of multilingual data**

24 Jan 2005

Kick off meeting (Luxembourg)

2

## Scope (cont.)

- And these data need to be improved, extended, processed, merged, used and re-used
- Of course, translation is directly concerned
- And to address the whole European population, localised tools regarding to various markets and languages are also concerned

But at present, these tasks form a timely and costly part of daily work of Europe's industry

24 Jan 2005

Kick off meeting (Luxembourg)

3

## Objectives

To lower this cost, LIRICS will:

- Provide Europe with a set of industry validated standards for language resource management ratified within the project lifetime
- Facilitate the acceptance of these standards by providing an open-source reference implementation platform, related web services and test suites
- Gain full industry support and input to the standards development via the Industry Advisory group and demonstration workshops
- Provide a pay-per-use business model for use by industry validated during the project

24 Jan 2005

Kick off meeting (Luxembourg)

4

# Consortium

The LIRICS consortium bring together leading experts in the field of Natural Language Processing via participation in ISO committees

- INRIA (F) specialist in standardisation
- DFKI (D) sp. in morpho-syntax & syntax processing
- USFD (UK) provider of the GATE open source platform
- CNR-ILC (I) sp. in language resources & standardisation
- UW (A) sp. in terminology management & language codes
- Util (NL) sp. in computational semantics
- MPI (D) sp. in meta-data
- Unis (UK) sp. in language resources
- IULA-UPF (E) sp. in lexicons & grammars

24 Jan 2005

Kick off meeting (Luxembourg)

5

# Industry advisory group

- For the standards to have impacts, LIRICS will ensure their usability by consulting with a group of industrial users
- The Industry advisory group will be consulted to identify priorities and requirements
- 21 members:
  - ▶ NLP solution providers like Systran, Sinequa, Temis or Morphologic
  - ▶ Lexicon publishers like Longman-Pearson
  - ▶ End users like EADS-CCR, British Telecom, Telefonica Invest-Des. or HP
- Membership will be expanded

24 Jan 2005

Kick off meeting (Luxembourg)

6

## Description of the work

- The deliverables will be direct inputs to the ISO ballots
- **WP1: Infrastructure for standard development & quality assurance**
  - to guarantee that the documents produced within the project are designed in accordance with ISO
  - to guarantee that they reach maturity, soundness and adequacy with the market
  - attendance at ISO meeting & submission of LIRICS deliverables to ISO

24 Jan 2005

Kick off meeting (Luxembourg)

7

## Desc. of the work (cont.)

- **WP2: Lexicons (connected to ISO TC37/SC4/WG4)**
  - efforts to address standardization have been already undertaken in the past: GENELEX, EAGLES, PAROLE-SIMPLE & ISLE constitute a valuable point for LIRICS
  - LIRICS will rely on the experience accumulated at each centre and will capitalise on results of the above mentioned projects, together with European and non European national projects
  - high compatibility ensured by the formulation of data categories (ISO 12620)
  - following the ISO milestones, a Lexical Markup Framework ISO-TC37 committee draft will be submitted to ISO ballot at M18. And DIS ballot at M27

24 Jan 2005

Kick off meeting (Luxembourg)

8



## Desc. of the work (cont.)

- **WP3: morpho-syntactic & syntactic annotations (connected to ISO TC37/SC4/WG2)**
  - valuable recommendations, best practices and guidelines have been proposed, on which WP3 will base its work (e.g. Eagles, Multext-East)
  - LIRICS will benefit from ongoing work at the ISO level
  - check the consistency with legacy data from existing Treebanks (e.g. Penn treebank) and with existing grammars (e.g. Matrix framework from EU project Deepthought)

24 Jan 2005

Kick off meeting (Luxembourg)

9

## Desc. of the work (cont.)

- **WP4: semantic content (connected to ISO 12620 DCR)**
  - developing standards for all aspects of the semantic content is beyond the scope of LIRICS
  - but, analysis of recent and emerging systems for the representation and annotation of semantic content
  - a useful step is the identification of a range of data categories such as temporal-spatial information, verb subcat, reference annotation, word sense information and quantification
  - data category compilation to be endorsed by the ISO TC37/SC4 Thematic Domain Committee (semantic group)

24 Jan 2005

Kick off meeting (Luxembourg)

10

## Desc. of the work (cont.)

- **WP5: reference implementation platform**

- all LIRICS defined ISO standards will be defined on the basis of web services in order to support distributed NLP resources
- support « try before you buy » paradigm which enables NLP companies to give temporary access and charge on per-usage basis
- provide open-source reference implementation of wrappers for lexicons, morphological analysers, syntactic parsers and semantic annotators

24 Jan 2005

Kick off meeting (Luxembourg)

11

## Desc. of the work (cont.)

- **WP6: dissemination & exploitation**

- a requirement workshop (M6) to identify priorities and essential characteristics from the Ind. Adv. Gr.
- 3 demonstration workshops (M12-18-24) open to public for the presentation of working drafts before submission to ISO
- eContent workshop with the existing eContent projects to make language standards known in all relevant areas of industry and economy
- a web site and a mailing list will be set up and managed

24 Jan 2005

Kick off meeting (Luxembourg)

12