

# WP5

**Adam Funk**

a.funk@dcs.shef.ac.uk

**University of Sheffield**

8th October 2007

# WP5: Objectives

- ✦ Define APIs following the LIRICS standards defined in WP1-4.
- ✦ Provide open-source implementation of these APIs.
- ✦ Demonstrate improved NLP module integration from different partners into an end-to-end NLP system

# WP5: Outstanding results

- ✦ APIs defined for DCR, LMF, MAF & SynAF as SOAP web-services.
- ✦ Servers and clients provided for DCR, LMF, MAF and SynAF; much of this software re-uses existing tools and platforms, as agreed.

# WP5: Outstanding results

- ✦ LMF client implemented as a GATE Language Resource.
- ✦ MAF and SynAF clients for English, French and Bulgarian implemented as GATE Processing Resources.

# WP5: Outstanding results

- ✦ MAF and SynAF web-services for English, French and Bulgarian implemented with GATE back-ends.
- ✦ MAF for English uses standard ANNIE and GATE components
- ✦ MAF for Bulgarian uses TreeTagger trained on BTB

# WP5: Outstanding results

- ✦ MAF and SynAF for French use Tagmatica's tagger and parser.
- ✦ SynAF for English uses the Stanford Parser with the supplied data (trained on the Penn Treebank).
- ✦ SynAF for Bulgarian uses the Stanford Parser trained on the BTB.

# WP5: Outstanding results

- ✦ The state of the art in automatic semantic annotation is not adequate to demonstrate the SemAF standard, so we have provided a corpus instead.

# WP5: Critical overview

- ✦ As mentioned above, it is not possible to provide automatic semantic annotation tools.
- ✦ Because the standards were being developed during the life of the same project as the reference software, it was difficult to aim at the “moving target”.



# WP5: Dissemination

- ✦ USFD's software is open-source and available to the public.

# WP5: Future activities

- ✦ USFD will recycle the parsing code developed for English and Bulgarian into a plug-in to be distributed with GATE.
- ✦ We will keep the source code for our LIRICS software available.

# WP5: Future activities

- ✦ Experiences with DCR and API implementation have resulted in the decision to develop the DCR into a next version (Isocat).
- ✦ Possibly modifications and extensions to the DCR API will result from this.
- ✦ Usage of DCR API will be extended to include other applications developed at the MPI (Elan, Annex).
- ✦ API for lexica will be further developed to support interaction between annotation tool and lexicon tool.