Web Services for Effective NLP Application Development and Evaluation: Using and Contributing to the Language Application (LAPPS) Grid

May 26, 2014

Presenters:

Nancy Ide, James Pustejovsky, Chris Cieri, Denise DiPersio, Keith Suderman, Marc Verhagen, Di Wang, Chunqi Shi

Tutorial Programme/Overview

Part I. Introduction

1. Overview of web service technology
   a. What are web services?
   b. Advantages of web service use for NLP application development
   c. How web services work: communication
      i. Protocols and wrappers
      ii. Data interchange: JSON and JSON-LD for syntactic interoperability
      iii. Data interchange: issues and options for semantic interoperability (ISO Web Service Exchange Vocabulary)
2. The LAPPS Grid
   a. Project overview
   b. Architecture
      i. Core nodes and distributed services
      ii. The composer platform
      III. Licensing
   c. Available services
      i. Data sources
      ii. NLP components
      iii. Open Advancement evaluation services

Part II. Contributing to the LAPPS Grid: Hands-on exercise

1. Wrapping tools as web services
2. Mapping input and output to the exchange vocabulary
3. Registering tools in the LAPPS Grid

Coffee Break

Part III. Using the LAPPS Grid: Hands-on exercise

1. Using the composer
   a. Selecting a data source
   b. Creating a processing pipeline: Sample task: Co-reference annotation
   c. Deploying the pipeline
   d. Evaluating performance: Using the Open Advancement Evaluation Environment
   e. Iterative improvement
      i. Modifying the pipeline on the basis of OA analysis
      ii. Assessing results

Tutorial Description/Outline/Contents

This tutorial will provide an overview of and hands-on experience with the LAPPS Grid, an open, web-based infrastructure through which massive and distributed resources can be easily accessed, and within which tailored language services can be efficiently composed, evaluated, disseminated and consumed by researchers, developers, and students across a wide variety of disciplines. The tutorial will consist of two components:

(1) Contributing to the LAPPS Grid: participants will be shown how to wrap their own tools as web services to be suitable for inclusion in the LAPPS Grid, including how to map input and output to exchange protocols in oder to enable full interoperability. They will then be shown how to register their tools in the Grid. At the end of the session the tools will be accessible to all LAPPS Grid users.

(2) Using the LAPPS Grid: participants will be introduced to the LAPPS Grid and shown how to select from available resources and select and deploy processing pipelines to perform various NLP tasks. Participants will be urged to utilize their own tools registered in the earlier part of the tutorial in the constructed pipelines. We will demonstrate the use of the Open Advancement Evaluation Environment, which flag incorrect outputs and provide a statistical analysis of the errors on a component-by-component basis. Participants will iterate the process over subsequent runs, during which problematic modules may be tuned or replaced in order to achieve the best results.