**Abstract**

We propose the PlayCoref game, whose purpose is to obtain substantial amount of text data with the coreference annotation. We provide a description of the game design that covers the strategy, the instructions for the players, the input texts selection and preparation, and the score evaluation.

**1 Coreference**

A: Jak býste definovali národ?
B: Jsou to lidé, kterí se odišli akcentem, gramatikou a slovníkem, ale... CZECH

A: How would you define a nation?
B: It is a body of people who have different accents, grammars and vocabularies but... ENGLISH

**2 Motivation**

Issue: Collecting high quality data is resource-demanding.

PDT 2.0 [1] coreference annotation (annotation scheme design, annotators training, technical and linguistic support, and annotation correction) spanned from summer 2002 till autumn 2004. Each annotator worked with a half of 3,165 documents.

Idea: Let us exploit the capacity of Internet users who love to play games.

Let us assume a two-player game for marking coreference links in the document, at least one document is present in each session, the session lasts up to 5 minutes and the players play half an hour a day. Then 3,165 documents will be annotated by... players in... days.

**3 Related Projects**

Game With a Purpose (GWAP) [2] methodology:
- images [3]
- tunes [4]
- video+text [5]
- text [6, 7]

Reading text is more difficult, burdensome and time-consuming.

**4 LGame Portal**

http://www.lgame.cz

**Key Properties of the Games**

1. Collecting data for the NLP tasks.
2. Only a basic knowledge of grammar is required. No extra linguistic knowledge is needed.
3. Games rules are language independent.
4. By default, the game is designed for Czech and English.
5. At least a general idea of what the opponents do.
6. Game is for two players at least.
7. If possible, the game offers several levels of difficulty.

**Toy games:**
- Place the Space - word segmentation
- The Shannon Game - hidden words

Ongoing work:
- PlayCoref - coreference
- PlayNE - named entities
- PlayDoc - contents of documents

**5 PlayCoref**

It is a body of people who have different accents, grammars and vocabularies but... speak the same language.

**REFERENCES**