Harmonization of Inventor and Applicant Names

Background
Patent documents contain bibliographical data such as name of the inventors and name of the applicants (company, or inventor if a private person applies), the inventor names can sometimes be reported differently in different countries or different times (e.g. surname first name or first name surname). Sometimes there is a wrong spelling, or the person might have changed their name, … Similar issues arise with company names.

The main problem with this arises when one wants to analyse a portfolio of (hundreds or thousands of patents) and this analysis will be difficult to understand if the same person or company is treated as several entities.

Project description
The goals of this project are, given a certain portfolio of patents to:

- Use an initial identification of patents with similar inventor names to analyse the content of the patents (text similarity, coherence of the patent classifications and time frames, co-inventors and name of employer (applicant), time frame of the invention etc.) and merge the names that are identified as belonging to the same person or company into one entity (one name).

- An additional user interaction phase might be required as larger corporations might have different business units with similar names but apply for patents separately and the user can be interested in only analysing one business unit and not the complete corporation.

Customer
IamIP Sverige AB, Stockholm

Contact at NLPLAB
Marco Kuhlmann, marco.kuhlmann@liu.se

Student profile
Knowledge about natural language processing (via courses such as TDDE09) and/or text mining (via courses such as TDDE16, 732A92).