Diploma Thesis

Designing an ETL Process with BPEL

Beginn: immediately

Background
Distributed business processes in industry have to be compliant to certain regulations. To measure the degree of compliance they can be instrumented in a way that they generate events during their execution that contain information about the process status. Based on these events a data warehouse (DW) can be built that facilitates a comprehensive analysis of the business process by applying data mining techniques. By utilizing these techniques it can for instance be checked to which extent predefined goals (e.g. compliance goals) and key performance indicators were met during the execution of the process. Moreover, the reason of violations can be determined and also predictions can be made about future process behavior.

Tasks
In order to populate the DW with data the generated events have to be extracted from heterogeneous data sources, transformed to fit to the schema of the DW and loaded into the DW. This process is called ETL (Extraction, Transformation, Load). In a first step, it has to be analyzed which information (e.g. events) are required by the DW, how the data sources have to be instrumented to obtain the required information and how this information has to be transformed. The ETL process has to be modeled in BPEL. As the data sources usually contain huge amounts of information special attention has to be turned to the efficient transportation of the data from the data sources to the DW. The data sources, data format and even the DW schema may evolve over time, thus the ETL process must be highly configurable. In order to achieve that a configuration mechanism has to be elaborated and implemented that enables a non-technical user to adapt the ETL process to the new requirements without being bothered with the technical details of BPEL.

Required knowledge and experience:
- Good SOA and BPEL skills
- Database knowledge (SQL, Data Warehouse basics)
- Good Java skills

Literature:

Contact:
Sebastian Wagner, sebastian.wagner@iaas.uni-stuttgart.de
Daniel Schleicher, daniel.schleicher@iaas.uni-stuttgart.de