EntimICE search – how to index the repository and how to define queries with domain specific languages?

While the traditional approach of searching directly in the database works well for small repositories, it has performance limitations with a growing repository size. Storing repository information in an index can make searching for keywords or phrases in content very fast. In order to provide this functionality, a separate entimICE server component has been created – an indexing server that will be demonstrated in the session. The indexing server is based on cutting-edge search technologies that are integrated in the solution (e.g. Lucene, TIKA, Akka) and is fully integrated with the entimICE access rights mechanisms. A speciality of the entimICE indexing server is that the information which is to be indexed represents various content types supported by the entimICE application – documents (e.g. PDF, Word), SAS datasets, data warehouse tables, CSV files and even files which reside outside of the entimICE repository on remote servers. The index also includes all metadata from the entimICE metadata repository and thus allows combining content and metadata searches. Search queries can be written in the Lucene native query language or based on a domain specific search language (DSL). entimICE allows creating custom search grammars and thus simplifies the creation of search queries.