

## RESUME

Of Master of Science thesis on the subject: "Classification of administrative documents for the medical workers"

Angel Veselinov Velikov, FN M21318

Bio- and medical informatics

The thesis is directed to the module of the software solution "Hippocrates" that is producing the reports and financial documents needed in the work of the general practitioners and specialists.

The global tasks of the thesis are:

- To improve the flexibility of the software;
- To give more freedom to the users;

In the thesis is suggested an optimization of the report generation by single execution of the classifying rules over the whole set of documents and storing the results for further use.

The report generation is reduced to simple counting of the documents that have been classified to a given type.

An acceptable solution is found for the problems of the synchronizing of the stored classifications with the document data over which the rules were executed. This problem has two aspects. The first aspect is that when the document related data is changed there should be an appropriate change in the classification of the document. The second aspect is that we need information about the relation between a given document data and the results depending on it.

An architectural reorganization is implemented in order to separate the logic of the classifying rules from the source code of the software.

The thesis concentrates on a suggested approach for better formalization of the rules for classification of the documents. Its main idea is that all of the queries should be dismantled into separate expressions and that the strongly coupled parts are removed. It is shown that after the list of the elementary expressions are prepared all the rules can be described as a list of elementary expressions which are concatenated together by using logical operators.

A general solution has been proposed for a better flexibility of the rules for document classification. A structure of the rules has been built that follows the normal forms of the logical expressions from the Boolean algebra.

Using the new rules definition structure that were implemented an algorithm was proposed to generate an human readable answer to the question why a given document is not classified to a given type.

In this way is established the base for future opportunities for additional solutions, that can analyze the corrections of the classification and extract new rules from them, that are effective only in this administrative region, and application of other methods and algorithms from the fields of the logic and artificial intelligence.