

STATEMENT

by

Acad. Atanas Ivanov Atanasov,
Bulgarian Academy of Sciences

on the defense of a dissertation on the topic:

Intelligent information systems in bioinformatics: semantic integration, analysis and classification of bio-medical data

author

Iliyan Nedkov Mihailov,
full-time doctoral student

at the Department of Information Technology, Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski "

for awarding the educational and scientific degree 'doctor'
in professional field 4.6 "Informatics and Computer Science", doctoral program "Information Technology - Bio and Medical Informatics"

Supervisor:

Assoc. Prof. Dr. Dimitar Ivanov Vassilev,
Department of Computer Informatics, Faculty of Mathematics and Informatics,
Sofia University "St. Kliment Ohridski "

As a result of his doctoral study and research work, doctoral student Iliyan Mihailov has presented a solidly written thesis both in volume and quality: 189 pages, with 49 figures, 21 tables, list of references (151 titles), glossary of terms, list of abbreviations. The dissertation is structured in 5 chapters: Chapter 1 - introduction to the problem domain, Chapter 2 - literature review and theoretical background, Chapter 3 - methodological aspects of the tasks in the dissertation, Chapter 4 - results, discussion and software implementation, Chapter 5 - contributions, perspectives, conclusions

Iliyan Mihailov has also presented 9 publications and 1 patent all in the scope of the topics in the thesis, which significantly exceeds the required minimum. All publications are in editions referred to by Scopus and Web of Science, half of them have an impact factor, there are 31 citations and h-index = 3.

The thesis is dedicated to various aspects related to the semantic integration of bio-medical data and their analysis. Problems of data integration are a very important area of data science *per se*, but they are also particularly important in the digitalization of various fields, and in the case of the dissertation the focus is on data related to bio-medical research and practice. In this line the presented dissertation has unconditional relevance. The significance of the work is expressed in the active citation of some of the publications dedicated to models for semantic integration of bio-medical data on the example of cancer data including: patient, clinical, laboratory and molecular analysis data.

Another important merit of the work is the development of a model and its software implementation for a decision support advisory system for generating dietary recommendations for patients with *diabetes mellitus* type 2. It is also important to emphasize the original author work on the integration of metagenomics studies data on microbial

resistance and the subsequent analysis and classification related to diversity and distribution. The thesis unconditionally has a very good level of complexity and methodological innovations in solving the announced goals and objectives. The created models for using ontologies for extracting rules, models for machine learning for the purposes of classification and analysis of different types of data. The proposed model of data compression is also of interest. A very good impression is made by the software implementation of all theoretically suggested and developed models on various topics for integration of bio-medical data in the dissertation. Naturally, this variety of developed topics in the dissertation is based on the ideas of the thesis related to creation of a model for a platform for intelligent data integration, which can be used as a common solution when working with data with different structures and formats, which need to be semantically linked and securely stored, and publicly accessible.

The parts of the thesis, dedicated to the review of the literature and introduction to the problem area and the subsequent comments on the theoretical background are written in detail, carrying information about systems with similar purpose and present the author's ideas for creating a universal (as far as possible) approach to semantic integration of bio-medical data. A possible critique of the first two chapters is that they are relatively long as a text (which could be regarded also as a critique of the entire dissertation).

The "Goals and Tasks" section is written briefly and clearly. Some stylistic inaccuracies could be addressed, such as the existence of a single goal, and the title targeting more than one goal. Here, a slightly better composition of tasks such as sense and implementation could be required.

The most important part of the dissertation is presented in Chapter 3, where the original methodological ideas of doctoral student Iliyan Mihailov are developed. This chapter shows in best manner the substantive merits of the author's work, as well as the choice of applications that have not only conceptual sense but are also definitely practical. This part of the thesis is the most valuable, but it could be written shorter and clearer. The main attribute related to this is the number of author's publications included in the dissertation.

The software implementation of the developed models and approaches used in the thesis is presented in Chapter 5, where the undoubted knowledge in the domain of informatics is clearly evident.

I would not like to miss the presentation of the contributions and the conclusion of the dissertation, which could have been presented more clearly.

My criticisms reflected in the text of the statement are mainly from the point of view of the organizational aspects of the presentation of the thesis. From a semantic point of view, I believe that the presented work has certain and obvious positive merits. Of course, in this thesis can be found syntactic errors and some stylistic ambiguities, foreign words, which are a matter of language culture and perhaps could be better controlled by the supervisor Assoc. Prof. Dr. Dimitar Vassilev, who nevertheless has done his job very well.

CONCLUSION

Regardless of the remarks, criticisms and recommendations, I firmly state that PhD student Iliyan Nedkov Mihailov has presented an excellent work based on very well formulated theses, filled with a lot of knowledge and diligence. The presented dissertation has an undoubted contributing character: interesting objects of research, new methods for practical implementation and analysis, many new and interesting results, the main part of the work has been published in extremely good journals. All this, and what I wrote above in the review of the work gives me reason to fully support the awarding of Iliyan Nedkov Mihailov with scientific and educational degree "Doctor (PhD)" in the professional field 4.6 "Informatics and

Computer Science", doctoral program "Information Technology – Bio- and Medical Informatics"

02.08.2021
Sofia

Prepared the statement:

Acad. Ātanas Iv. Atanasov