prof. dr hab. inż. Dariusz Dereniowski
Politechnika Gdańska
ul. G.Narutowicza 11/12, 80-233 Gdańsk
dariusz.dereniowski@pg.edu.pl

16.11.2026

Reviewer's opinion
on Ph.D. dissertation authored by
Adam Aron Rynkiewicz
entitled:
From Literature to Models: Semi-Automatic Construction of Biomedical Petri Net-Based Models

1. Problem and its impact

The main problem discussed in the dissertation, very broadly speaking, is that of processing scattered data into an organized representation. Here, the data covers biomedial research publications giving partial information on selected chemical/biological processes. For the representation of this data Petri nets have been chosen.

The problem consists both scientific and practical elements. On one hand, it has a potential for development of data processing tools, statistical analysis and algorithmic investigations, as well as touching computational complexity issues. On the other hand, the main goal is specifically set for targeting specific applications in biomedical data representation. In the latter lies a potential for practical applicability. In particular, growth rate of new scientific data makes "manual" analysis not feasible and semi-automated or fully-automated methods need to be developed.

2. Contribution

The contribution of the thesis is in chapters 4-7. Chapter 4 presents an example of Petri net construction where HDL metabolic network is presented. This chapter is, in my opinion, not directly linked to the main claims of the thesis, because it focuses more on deriving specific claims regarding metabolic flows, rather than the central points of the thesis like entity recognition or linking. The latter are covered in subsequent chapters. Chapter 5 discusses and presents approaches for named entity recognition. This part, in my view, is important because low quality solutions at this step basically make the whole Petri net construction infeasible. Author presents in chapter 5 an evidence that selected methods give good precision results. Chapter 6 describes the architecture of FELA system (designed and developed by the Author). The goal of the system is to solve the entity linking problem. What is valuable is that the Author identified weak points of EL (like issues with normalization) as well as strong ones (like overcoming the necessity of training on relatively small data sets). Chapter 7 is the final and main component of the dissertation and includes comparative analysis of semi-automatic Petri net construction methods restricted to biomedical literature. This part also describes tuning of several components of the system to improve the overall accuracy.

The dissertation has clear practical and experimental components, while I have found no theoretical contributions to the state of the art knowledge. Regarding the practical component, the main issue is the fact that the work on the thesis started at time when LLM models had very low capabilities, while their current development makes an approach taken in this thesis partially outdated. For example, one could conduct new set of experiments using the current strong models as a benchmark for obtained solutions. It is hence difficult to place this thesis in relation to the current state of knowledge. Adam Rynkiewicz addresses this issue in the work directly at the beginning of chapter 7: it is noted that the classical pipeline (NER → EL → RE) needs to be altered in the context of the new tools. The author discusses why earlier components, including NER and FELA, are no longer essential to the final Petri-net construction pipeline, while identifying independent uses for them. The FELA system has been designed to be modular to address this issue. While pointing this issue out, my following evaluation does not treat the thesis as incomplete or set on wrong assumption due to recent general LLMs' development.

The experimental component is therefore the main strong point of the thesis. By this I mean detailed analysis and multiple computer experiments that have been performed. The quality of the solution that has been obtained can be mainly measured by the numerical results. In my opinion it has been demonstrated that the Author managed to construct a working system that performs well against selected benchmarks.

Adam Rynkiewicz co-authored several publications, but for unknown to the reviewer reasons, only one of them, namely the one published in *International Journal of Molecular Sciences* is explicitly listed in the dissertation. I should mention that this publication appears in a highly valued journal. It is quite surprising why other works have been omited, like "Universal entity linking" in *Engineering Applications of Artificial Intelligence*, or "Daily Reports on Phage-Host Interactions" in *Frontiers in Microbiology*, or "Bridging Global Language Models and Local Spatial Data: The JackDaw Approach to Context-Aware Agriculture and Rural Planning" in *FedCSIS* 2025. The reason this is surprising is that they correlate with the subject of the dissertation.

From the practical standpoint, contributions include availability of developed tool for external use. The strongest point here is the publicly available code of *Watson*. Having a front-end interface, this software is close to black-box use tool. The publicly available repository (at the time of preparing this review) contains an instruction that allows for quick startup and use of the platform.

3. Correctness

Generally, to the point that it can be checked by evaluating the manuscript, the methodology does not show indications of errors or important missing components. Below I list a number of comments that give potentially weaker points.

The dissertation lacks of mathematical precision, for example "where $P = p_1, p_2, \ldots, p_n$ is a finite set of places" (one should write $P = \{ \ldots \}$ for a set; in some places of the work this notation is actually used correctly), "$W \rightarrow Z+$ is a weight function" (this omits the domain of the function, so it should state W: $F \rightarrow Z+$). The Author says that a t-invariant "corresponds to a set of transitions whose collective execution do not change the state" and interprets it as a balanced subprocess or pathways.

However, $Ax = 0$ does not seem to imply that there is a firing sequence realizing x from a given marking.

The theory regarding the Petri nets used changes of token markings induced by transition firings. This has been used directly e.g. in chapter 4. The Petri nets have been described in a sufficiently precise way including the software used for that. However, the simulations of token dynamics seem to be only partially described with the main focus on outlining certain cases found during experiments. It is not fully clear how many experiments, and which ones were conducted. Also, certain conclusions seem too strong: it is said that certain transitions do not belong to a non-trivial MCT set, but how this allows for concluding their independence?

Chapter 4 is very detailed regarding experiments conducted on the basis of the Petri net constructed for HDL metabolism. However, due to the subject of the thesis, I would expect more details of the network generation process. This would include how much was created automatically and what was edited manually as part of "expert knowledge".

Due to a rapid development of the discipline, it would be more informative for a reader if statements like "However, existing LLM-based approaches frequently impose constraints that limit their applicability in realistic information-extraction pipelines.". More precisely, I would refer here to specific models and their versions, at least some examples of those.

4. Knowledge of the candidate

The first three chapters of the dissertation give an introduction, present the problem and the basic terminology and overview the state of the art. In my opinion this part is a sufficient guide to the topic and shows good knowledge of the Candidate regarding the subject studied. In particular, allows to conclude good general knowledge and understanding of Petri nets' foundations and methods of their construction and analysis, which can be categorized as part of Information and Communication Technology discipline. Chapter 4 can be also partially seen as a demonstration of general knowledge of Petri net construction and partially as research task with specific results.

In my opinion, chapter 7 stands out. It shows a strong technical ability of Adam Rynkiewicz. The amount of technical work behind it is substantial and includes fine tuning of several components, construction of domain specific training data, performing improvements once the first experiments provide evidence of shortcoming of a taken approach. Another indication of well-done research is the broadness of the dissertation, i.e., the fact that it covers (chapters 4-6) broad spectrum of Petri net development.

5. Conclusion

Taking into account what I have presented above and the requirements imposed by Article 187 of the *Act of 20 July 2018 - The Law on Higher Education and Science* (with amendments)[1], my evaluation of the dissertation according to the three basic criteria is the following:

[1] http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000276

A. Does the dissertation present an original solution to a scientific problem? (the selected option is marked with X)

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

B. After reading the dissertation, would you agree that the candidate has general theoretical knowledge and understanding of the discipline of **Information and Communication Technology**?

[ ]	[X]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

C. Does the dissertation support the claim that the candidate is able to conduct scientific work?

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

Signature