

РЕЦЕНЗИЯ

от проф. д-р Иванка Милошева Цаковска,
Институт по биофизика и биомедицинско инженерство – БАН

на дисертационен труд за присъждане на образователната и научна степен „доктор“

по: област на висше образование **4. Природни науки, математика и информатика**

професионално направление **4.2. Химически науки**

докторска програма **"Теоретична химия"**

Автор: **Гергана Илиева Танчева**

Тема: **ПРИЛОЖЕНИЕ НА МЕТОДИТЕ НА ХИМИЧНАТА ИНФОРМАТИКА ПРИ МУЛТИКОМПОНЕНТНИ СУБСТАНЦИИ И НАНОМАТЕРИАЛИ**

Научен ръководител: **доц. д-р Николай Тодоров Кочев, Пловдивски университет „Паисий Хилендарски“**

1. Общо описание на представените материали

Със заповед № РД-21-1840 от 23.10.2024 г. на Ректора на Пловдивския университет „Паисий Хилендарски“ (ПУ) съм определена за член на научното жури за провеждане на процедура за защита на дисертационен труд на тема **„Приложение на методите на химичната информатика при мултикомпонентни субстанции и наноматериали“** за придобиване на образователната и научна степен „доктор“ в област на висше образование 4. Природни науки, математика и информатика; професионално направление 4.2. Химически науки; докторска програма "Теоретична химия". Автор на дисертационния труд е Гергана Танчева – докторант в задочна форма на обучение към катедра "Аналитична химия и компютърна химия" на Химическия факултет, с научен ръководител доц. д-р Николай Кочев от Пловдивския университет „Паисий Хилендарски“.

Представеният от Гергана Танчева комплект материали е в съответствие с Чл.36 (1) от Правилника за развитие на академичния състав на ПУ и включва следните документи: (1) молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационен труд; (2) автобиография в европейски формат; (3) протокол от катедрения съвет, свързан с докладване на готовността за откриване на процедурата и с предварително обсъждане на дисертационния труд; (4) дисертационен труд; (5) автореферат; (6) списък на научните публикации по темата на дисертацията; (7) копия на научните публикации (три броя); (8) декларация за ори-

гиналност и достоверност на приложените документи; (8) справка за минималните изисквания; (9) становище от научния ръководител.

Прегледът на материалите показва, че всички законови изисквания по провеждането на процедурата са спазени.

2. Кратки биографични данни за докторанта

Гергана Танчева е завършила бакалавърска програма по медицинска химия и магистратура по спектрохимичен анализ в ПУ „Паисий Хилендарски“, Химически факултет, катедра “Аналитична химия и компютърна химия”. През 2019 г. е зачислена като задочен докторант в същата катедра и е отчислена с право на защита през 2024 г. Дисертационният ѝ труд е разгледан и насочен към защита от разширения катедрен съвет на катедра “Аналитична химия и компютърна химия” на заседание, проведено на 14.10.2024 г. По време на докторантурата си Гергана Танчева е завършила редица обучителни курсове, които са подбрани много адекватно, така че да я подпомогнат за осъществяване на дисертационните изследвания на високо ниво. Впечатляващо е и активното ѝ участие в международни научни проекти, в рамките на които са осъществени научните изследвания по дисертацията. Успоредно с това, от края на 2022 г. докторант Танчева работи във фирмата ИДЕЯКОНСУЛТ ООД, която е ключов партньор в тези проекти.

3. Актуалност на тематиката и целесъобразност на поставените цели и задачи

Химичната информатика има ключово значение и утвърдено приложение в областта на лекарствения дизайн и изчислителната токсикология по отношение на моделирането на добре дефинирани химични структури. Когато става въпрос обаче за химични субстанции, които се състоят от повече химични структури, както и за наноматериали, методите на химичната информатика са в процес на адаптация, за да отразят коректно и пълно новия тип структурно пространство и да позволят структурирано съхранение на всички необходими мета данни за моделиране и вземане на решения, особено в областта на токсикологията. **В този смисъл дисертационното изследване е особено актуално и перспективно, като резултати от него са приложени в редица международни научни проекти, целящи да подпомогнат вземането на регулаторни решения по отношение на безопасност на наноматериалите.**

Целта е ясно формулирана и за нейното ефективно изпълнение са идентифицирани необходимите изследователски задачи. Тук имам критичен коментар по отношение на формулирането на задачи, свързани с изучаване на съществуващи софтуерни системи, семантични модели, онтологии и др. Разбирането ми е, че това е част от литературния обзор, а изследова-

телските задачи на едно дисертационно изследване трябва да отразят конкретните изследвания, които ще се осъществят, за постигане на дисертационната цел.

4. Познаване на проблема

Докторантката показва много добро познаване на полето на химичната информатика и изключително разностранни знания - от представяне и моделиране на ефектите на химични съединения до подходи за решаване на сложни проблеми в областта на наноинформатиката. Всичко това се е опитала да представи в литературния обзор на дисертацията. От една страна е похвално с колко подходи и методологии се е запознала, от друга страна литературният обзор се възприема сравнително трудно предвид голямото количество информация, извън конкретните цели и задачи, както и известната липса на прецизност в структурирането на съдържанието (напр. в т.1.2 Класификация на химичните субстанции са включени "Свойства на наноматериалите", както и "Подходи за синтез на наноматериали"; Подходът на разработване на Adverse Outcome Pathways е разгледан в т. Валидиране на модели и това, както и много оскъдното му представяне го прави неясен за разбиране и др.).

Също така докторантката се спира обстойно на универсалните принципи на OECD за валидиране на QSAR модели като се базира на общоприет през последните 15 години в регулаторните среди референтен документ (забележка: необходима е корекция на цитирания литературен източник 180). Биха препоръчала да се има предвид обаче и последният, актуализиран рамков документ, разработен от OECD в края на 2023 г. за оценка на QSAR модели, а именно „(Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions” [<https://doi.org/10.1787/d96118f6-en>].

5. Методика на изследването

Методологията на изследването е адекватна и позволява да бъде постигната дисертационната цел, както и да бъдат решени конкретните изследователски задачи.

6. Характеристика и оценка на дисертационния труд

Дисертационният труд е оформен на 196 стр., които включват 70 страници литературен обзор; цел и задачи; 78 страници собствени изследвания и обобщена дискусия; заключения и приноси. Библиографията се състои от 246 литературни източници, които дават много добра картина на изследователското поле. Включени са и четири приложения, които допринасят за пълната прозрачност и отвореност на дисертационните данни.

В собствените изследвания на докторантката, за описание на данни за мултикомпонентни субстанции и наноматериали е избран моделът за данни на *Ambit/eNanoMapper*, като FAIR описанието на мултикомпонентни субстанции и наноматериали е представено чрез три основни слоя с метаданни (i) идентификация на веществото, (ii) описание на композицията и (iii) записи за измервания с богато множество метаданни. На тази основа са направени различни подобрения в изчислителната инфраструктура за управление на токсикологични данни за наноматериали *eNanoMapper*. Например е разработен конфигурируем инструмент *NMDataParser*, който улеснява подготовката на данни и качването им в базата данни *eNanoMapper*. Заслужава да се отбележи работата по обработка на голямо количество EXCEL файлове с експериментални данни, събирани в рамките на множество международни научни проекти с цел FAIR-ификация на данните. Конвертираните данни са обогатили базата данни на *eNanoMapper* с голямо разнообразие от физикохимични и биологични анализи. Разработен е също така уеб базиран софтуерен инструмент - *eNanoMapper Template Wizard*, чрез който да се описват хармонизирано данни с необходимите метаданни, което е много полезно за специалистите при необходимост от хармонизация и стандартизация на техните експериментални данни. Допълнително, в рамките на дисертационния труд е обогатена разработената в рамките на *eNanoMapper* онтология с въвеждането на 15 допълнителни научни термина.

Съществен принос на дисертационното изследване е разработването на идентификатори за структурно представяне на мултикомпонентни субстанции като разширение на подходите, прилагани за идентификаторите на химични съединения. В тази връзка е разработена софтуерната библиотека с отворен код *Ambit-SLN*, за представяне на химични субстанции чрез линейната нотация *SYBYL Line Notation*. Библиотеката е интегрирана в софтуерната платформа с отворен код *AMBIT*.

Във връзка с необходимостта от обработване на експерименталните данни, свързани с охарактеризиране и безопасност на наноматериалите са разработени електронни бележници съответно за извличане, обработка и визуализация на данни от базата данни на *eNanoMapper* и за проверка на пълнотата на данните съгласно хармонизираните шаблони. Съществени резултати са постигнати по автоматизиране на обработката на данни от високопроизводителен биологичен скрининг (*highthroughput screening, HTS*). В частност е разработена софтуерна библиотека *ToxFAIRy* за прочитане и анотиране на HTS данни, предварителната обработка, FAIR-ификация и оценяване на токсичност. В допълнение е разработен потребителски интерфейс към библиотеката за платформата *Orange*. Библиотеката е успешно използвана за автоматична обработка и оценка на данни в рамките на два международни научни проекта.

7. Приноси и значимост на разработката за науката и практиката

Резултатите от дисертацията са вече приложени по отношение на структуриране и обработване на информация за наноматериали в рамките на редица европейски проекти и недвусмислено са демонстрирали принос по отношение на разширяване на подходите на химичната информатика за многокомпонентни химични системи и особено по отношение на наноматериали. Реализирани са редица софтуерни компоненти основно в рамките на проекта eNanoMapper. Приемам формулираните от докторантката съществени приноси като такива с научен и научно-приложен характер. Първите включват: (i) разработване на софтуерен инструмент NMDataParser за подготовката и качването на данни в базата данни eNanoMapper в унисон с концепцията за FAIR-ификация на данни за химични субстанции и наноматериали; (ii) разработване на прототип за идентификатор за наноматериали на базата на линейната нотация SLN, който кодира както структурни данни, така и множество експериментални данни за всеки конкретен наноматериал; (iii) разработване на концепция за аотиране на данни от високопроизводителен биологичен скрининг, включваща метаданни, предварителна обработка, както и изчисляване на токсикологичен приоритизиращ индекс.

Научно-приложните приноси включват: (i) разширение на базата данни на eNanoMapper с FAIR данни за физико-химични и биологични параметри на наноматериали, изследвани в рамките на европейски проекти, в които научната група активно участва; (ii) разширение на онтологията на eNanoMapper в областта на екотоксичност и опазване на околната среда; (iii) разработване на софтуерна библиотека ToxFAIRy за аотиране, обработка, токсикологична приоритизация и FAIR-ификация на данни от високопроизводителен биологичен скрининг; (iv) разработване на модул Orange3-ToxFAIR към платформата за машинно обучение и анализ на данни Orange, който може да се използва като потребителски интерфейс за софтуерната библиотека ToxFAIRy; (v) Разработване на автоматизиран процес за обработка, групиране и FAIR-ификация на данни от високопроизводителен биологичен скрининг, позволяващ ефективна обработка на голям обем от данни.

Докторантката е представила амбициозни насоки за бъдещо развитие, които логично стъпват върху достиженията до този момент по отношение на модели за представяне на структурна и експериментална информация за мултикомпонентни субстанции, с оглед моделиране на свойства/ефекти на химични субстанции/наноматериали.

8. Преценка на публикациите по дисертационния труд

Публикувани са три научни статии във връзка с дисертацията – две в тематично свързани с полето на дисертацията научни списания с импакт фактор в категории Q1 и Q2, както

и една глава от книга. Върху публикациите са забелязани 17 цитата, което е атестат за разпознаваемост в научната общност. Съдържанието на публикациите е тясно свързано с дисертационните изследвания. Резултатите са активно популяризирани чрез множество участия в научни форуми, както и са приложени в няколко големи научни проекта по програмата Horizon 2020. **Въз основа на това оценявам, че минималните национални изисквания за придобиване на образователната и научна степен „доктор“ са убедително изпълнени.**

9. Лично участие на докторантката

Въпреки че представените изследвания са изцяло екипни, в унисон със съвременните научни тенденции, нямам съмнение, че **дисертационният труд е лично дело на докторантката**, както и че тя е взела активно участие в докладваните изследвания. За това свидетелства и изцяло положителното становище от страна на научния ѝ ръководител.

10. Автореферат

Авторефератът съумява в синтезиран вид да предаде съществените достижения на **многостранныте дисертационни изследвания** и коректно отразява основните резултати, постигнати в дисертацията.

11. Критични забележки, препоръки, въпроси

Критичните бележки и препоръки са представени в по-горните секции на рецензията.

Към докторантката имам следните въпроси:

1. Как се разграничават понятията „многокомпонентна субстанция“ и „химична смес“, например защо показаната композиция 2 на фиг. 17, състояща се от два главни компонента (бензен и толуен) се класифицира като субстанция, а не като химична смес?

2. Може ли докторантката да дискутира възможности за моделиране на токсичност на многокомпонентни химични субстанции, предвид информацията в дисертацията, че предизвиканата от тях токсичност (визират се материали с нови или подобрени функционалности) "може да не бъде точно оценена въз основа на измерената токсичност, предизвикана от химията на отделните им компоненти, а на база на уникалната комбинация от всички компоненти"?

3. В т. "Химични бази данни" на литературния обзор са дадени примери за бази данни предимно на химични съединения, по-нататък се дискутира базата данни на eNA노Mapper. Съществуват ли други такива структурирани бази данни на химични субстанции/наноматериали??

12. Лични впечатления

Нямам лични впечатления от докторантката. Познавам добре, обаче, работата на колектива, който е сред водещите научни групи в областта на приложение на химичната информатика за целите на регулаторната токсикология, прокарващи трасето на наноинформатиката в редица Европейски научни проекти, за оценка на безопасност на наноматериали. Без съмнение тази среда е изградила отлична основа за по-нататъшното ѝ развитие като утвърден, самостоятелен изследовател.

ЗАКЛЮЧЕНИЕ

Дисертационният труд **съдържа научни и научно-приложни резултати, които представляват оригинален принос в науката и отговарят на всички** изисквания на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за прилагане на ЗРАСРБ и съответния Правилник на ПУ „Паисий Хилендарски“.

Дисертационният труд показва, че докторантката **притежава** задълбочени теоретични знания и професионални умения в областта на химичната информатика като **демонстрира** качества и умения за самостоятелно провеждане на научно изследване.

Поради гореизложеното, убедено давам своята **положителна оценка** за проведеното изследване, представено от рецензираните по-горе дисертационен труд, автореферат, публикации, постигнати резултати и приноси, и **предлагам на почитаемото научно жури да гласува положително за присъждане на образователната и научна степен „доктор“ на Гергана Илиева Танчева** в област на висше образование: 4. Природни науки, математика и информатика, професионално направление 4.2. Химически науки, докторска програма "Теоретична химия".

18.12.2024 г.

Рецензент:

/проф. д-р Иванка Цаковска/