

РЕЦЕНЗИЯ

от доц. д-р Красимира Иванова,
Институт по математика и информатика, БАН

на дисертационен труд за присъждане на образователната и научна степен "доктор"
в област на висше образование: 4. Природни науки, математика и информатика
професионално направление: 4.6 Информатика и компютърни науки

автор: **Александър Насков Трайков**

тема: **Стандартизиране и интегриране на разнотипни информационни ресурси**

научен ръководител: **проф. д.м.н. Георги Атанасов Тотков**

1 **Общо описание на представените материали и документи**

Със заповед № Р33-947 от 10.03.2017г. на Ректора на Пловдивския университет "Паисий Хилендарски" съм определена за член на научното жури за защитата на дисертационен труд на тема **"Стандартизиране и интегриране на разнотипни информационни ресурси"** за придобиване на образователната и научна степен "доктор" в област на висше образование 4. Природни науки, математика и информатика; професионално направление 4.6. Информатика и компютърни науки; докторска програма Информатика. Автор на дисертацията е **Александър Насков Трайков**, отчислен с право на защита докторант редовна форма на обучение към катедра Компютърна информатика на Факултета по математика и информатика при Пловдивския университет "Паисий Хилендарски".

Представеният от Александър Трайков комплект материали на електронен носител включва следните документи:

- молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационния труд;
- автобиография;
- копие от диплома за висше образование (ОКС "магистър");
- заповед за зачисляване в докторантура;
- заповед за провеждане на изпит от индивидуалния план и съответен протокол за издържан изпит по специалността с успех Отличен (5.50);
- заповед за отчисляване с право на защита;
- извлечение от протокол от катедрен съвет, свързан с докладване на готовност за предварително обсъждане на дисертационния труд;
- извлечение от протокол от катедрен съвет за проведено предварително обсъждане на дисертационния труд, предложение за откриване на процедура по защита, състав на научното жури и дата на защита;
- дисертационен труд;
- автореферат;
- декларация за оригиналност и достоверност на приложените документи;

- справка за спазване на специфичните изисквания на ФМИ при ПУ за придобиване на ОКС "доктор";
- списък на научни публикации по темата на дисертацията, както и копия на самите публикации (4 статии в научни сборници и 3 глави от книга);
- списък на забелязани цитирания и копия на цитиранията (4 цитирани публикации).

Представените материали съответстват на изискванията на Чл.36(1) от Правилника за развитие на академичния състав на ПУ. На основа на тези материали установявам, че са изпълнени всички нормативни изисквания. Докторантът има образователно-квалификационната степен "магистър". От приложените заповеди и протоколи е видно, че са спазени всички процедури по провеждане на докторантурата, предварителното обсъждане на дисертационния труд и насочването му към защита.

2 Кратки биографични данни

Александър Насков Трайков е роден през 1988г., има бакалавърска степен по информатика (2011г.) и магистърска степен по Бизнес информатика (2012г.), придобити във ФМИ на ПУ. Зачислен е в редовна докторантура през 2013г. и е отчислен с право на защита на 01.03.2016г. Участието му в 2 европейски проекта на Пловдивския университет (BG051-PO001-4.3.04-0064 "Пловдивски електронен университет (ПеУ): национален еталон за провеждане на качествено е-обучение в системата на висшето образование" /2012-2014/ и BG051-PO001-3.1.08-0041 "Стандартизиране и интегриране на разнотипни информационни и управленски университетски системи" /2013-2014/) способства за натрупване на знания и опит за изграждане на цялостно виждане на проблемите и намиране на пътища за тяхното разрешаване.

3 Актуалност на тематиката и целесъобразност на поставените цели и задачи

Информационните технологии навлязоха в почти всички сфери на дейност по събиране, съхраняване, обработка и разпространение на информационни ресурси. Различните цели, задачи, форми, както и различните конкретни системи, които ги пораждаат и обработват, пораждаат многообразие от формати на съществуване на тези ресурси. По естествен път възниква необходимостта от изследване и предлагане на методи и средства за тяхното автоматизирано събиране, категоризиране и съхранение, както и за създаване на предпоставки за обмен на данни между различни системи и подсистеми в една организация или виртуална общност.

Затова и целта на дисертационния труд е особено актуална – изследване, предлагане и апробиране на стандарти, методи и средства за автоматизирано генериране и архивиране на разнотипни информационни ресурси, които са резултат от работата на информационни системи, работещи в рамките на една организация.

За постигане на тази цел са формулирани следните основни задачи: проучване на съвременното състояние на проблема, създаване на модели на данни и архитектура на система за депозиране и архивиране на информационни ресурси с възможност за автоматично генериране на метаданни в рамките на процесите, протичащи в институционална информационна структура; изграждане на софтуерен прототип на система за депозиране и архивиране на съдържание и метаданни за информационни

ресурси (наречена СИРИР) и тестване на работоспособността на предложените модели и архитектура, реализирани в прототипа.

Целта е ясно поставена и за реализирането ѝ коректно са определени задачите, които следва да се извършат, което е основна предпоставка за успешното извършване на изследването.

4 Познание на проблема

Считам, че докторантът познава добре съвременното състояние на проблема. Щателно е разгледал съществуващите стандарти и методи за интегриране на разнотипни информационни ресурси, както и е проучил основните цифрови хранилища и платформи за съхранение и генериране на информационни ресурси, използвани в академичната общност (в повечето случаи – тези с отворен достъп). Списъкът на използваните литературни източници включва 161 заглавия, повечето са препратки към стандарти или системна документация, което е нормално пред вид спецификата на работата. От друга страна има изключително добър професионален (информатичен) опит, което му позволява да реализира идеите си на много високо ниво.

5 Методика на изследването

Избраната методика на изследване е стандартна за дисертация в областта на информатиката. Въз основа на резултатите от проучване на съвременното състояние на проблема са предложени модели на данни и архитектура на система за архивиране на информационни ресурси с възможност за автоматично извличане на метаданни, реализиран е прототип на базата на тази архитектура, който е интегриран в общата среда на университетската електронна среда и са проведени и анализирани експерименти за валидиране на предложените идеи.

6 Характеристика и оценка на дисертационния труд

Дисертацията е от 172 стр. (има грешка в номерирането на страниците) и се състои от увод, 4 глави, заключение, списък на публикации по темата на дисертационния труд, списък на забелязаните цитирания, 3 приложения, литература и декларация за оригиналност и достоверност. Има и списъци на използваните съкращения, фигури и таблици. Текстът е форматиран в страници с около 30 реда и 90 символа на ред, така че общият обем на дисертацията е в рамките на обичайния (имайки пред вид, че част от конкретиката е изнесена в приложенията).

Изложението е добре структурирано, ясно, точно и богато илюстрирано. Текстът е добре премислен и свързан. Приятно впечатление прави наличието на много малко правописни грешки (бел. – на места във фигурите следва да се поправят текстове, отрязани от форматирането на полетата).

В *Увода* е обоснована необходимостта от решаването на проблема по интегриране на разнотипни ресурси. Формулирана е основната цел на дисертационното изследване и правилно са поставени 4 задачи, чрез които тя да се постигне. Накратко е описана структурата на дисертацията.

Глава 1 въвежда основните понятия – ресурс, информационен ресурс (ИР), информационна система, структурирано описание на данни чрез метаданни, и други

понятия, свързани с типизацията, организацията и съхранението на метаданните. Изложението е логически последователно, дефинициите са точни и ясни. После докторантът се спира на основните организации, занимаващи се със стандартизация на международно и национално ниво и основните стандарти за метаданни за работа с различни типове информационни ресурси. В тази част лично аз бих предпочела организациите да са отделени от стандартите. Третата част разглежда процесите по интеграция на разнотипни информационни ресурси, отделяйки различните нива, на които може да се извърши тази интеграция – на ниво схема, запис, хранилище и детайлно описвайки различните случаи, които могат да възникнат и начините за решаването им. Текстът по естествен начин преминава към разглеждане на основните използвани в момента цифрови хранилища и техните платформи. Докторантът прави съпоставка на цифровите хранилища по няколко критерия и най-вече кои протоколи за оперативна съвместимост поддържат. На базата на проведеното изследване той изтъква основните наблюдавани недостатъци и формулира идеята за платформа, реализираща работна рамка, която поддържа нормализацията (обработка, аташиране на метаданни и запазване на съдържанието) на разнотипните ресурси (СИРИР). В този смисъл поставената **Задача 1** е изпълнена и изводите от проведения анализ дават пълното основание да се продължи в набелязаната насока.

В **Глава 2** докторантът определя основни изисквания към моделите на данни на информационни ресурси. Данните са разделени на два основни вида – статични и динамични. Статичните са: документи; медийни файлове; скриптове. При тях СИРИР идентифицира само файловете атрибути. Динамичните данни се създават в СИРИР и представляват комбинация от съдържание и метаданни, резултат от изпълнението на процес в СИРИР (справка). Справките използват йерархичността на структурата на метаданни и чрез обратното наследяване на класове се постига адекватно и пълно извличане на обектите. Основен принос тук е предложеният процес за класифициране на метаданни чрез задаване на акумулиращи функции, които могат да се променят динамично.

Във втората част докторантът предлага архитектура на СИРИР, ясно подчертавайки, че задачата на системата е да бъде посредник при генерирането и автоматизирането на динамични данни. Основната идея е данните да се асимилират в момента на преминаване през системата, така че да се интегрират с други типове документи необременени от стандартите, на които са били подчинени дотогава. За тази цел докторантът определя основните слоеве на архитектурата, позволяващо капсулиране на специфичните функционалности, така че локалните изменения да не засягат цялостната система. Отделните слоеве са описани с отчитане на най-новите технологични изисквания и възможности.

На следващата стъпка докторантът описва подходите, използвани при процесите – генериране на метаданни за ИР, търсене на метаданни, оформяне на резултатите от справките като ИР (бележката ми е само към термина "метод", който според мене се е появил в описателния текст под влиянието на ООП). В тази част са обхванати прецизно възможните случаи на пораждаване и използване на метаданните и връзката им с ИР.

В този смисъл и поставената **Задача 2** е изпълнена, задавайки теоретичната основа за по-нататъшно реализиране на СИРИР.

Глава 3 описва реализацията на предложения прототип. Тук особено ясно личи как използването на средствата на ООП по естествен начин разгъват възможностите за реализиране на наследяването и капсулирането на процесите и резултатите. За оптимизиране на процеса на обратно наследяване се въвеждат допълнителни шаблони "връзка" (чрез който могат се проследяват пътищата към родителя) и "контекст" (обединяващ общите обекти, с които работят метаданните). Дефиниран е общ модел на структура на метаданни, която описва разнотипни ресурси – дървовидна структура, адаптираща се към различни изисквания. Въведена е гъвкава рамка за съхранение на справки, базирана на общ интерфейс, което е предпоставката за успешната интеграция на разнотипните данни. Детайлно е описан процесът на генериране на метаданни.

Докторантът е успял да представи създадения прототип изключително ясно и точно, съчетавайки умело описанието на обектите и процесите с технически детайли от програмната имплементация. В този смисъл, считам, че поставената **Задача 3** е изпълнена на професионално ниво.

Глава 4 описва имплементирането на СИРИР като ядро на цифровото хранилище, използвано от подсистемите на ПеУ и изследване на резултатите от работата ѝ. При избора на база от данни правилно е насочен погледа върху нерелационните бази, а от разгледаните MongoDB и Elasticsearch е избрана втората, поради по-добрия набор от функционалности за търсене и значително по-добро бързодействие (за съжаление референцията [DB_PERF] е пропусната в Литературата). Подобно внимателно изследване е направено и при избора на файлов склад, програмен език и работна рамка за реализация на ядрото, както и другите компоненти, необходими за реализиране на зададените в Глава 2 архитектурни слоеве и в Глава 3 модели за метаданните, автоматизирани процеси и справки.

Проведени са 3 експеримента с реални студентски данни в ПеУ – (1) автоматично генериране на справки и оценка на близост между ресурси, на примера на подаване на заявления и генериране на ведомости при кандидатстване за стипендии в ПУ; (2) депозиране и разглеждане на статични документи чрез сканиране и запис на първичния информационен ресурс и разпознатия текст, на примера на научни статии и документи за прием; (3) търсене на статични и динамични данни по съдържание и метаданни.

Описанието на постановъчната среда, проведените експерименти, получените резултати и тяхната оценка е направено ясно, последователно и на много добро ниво.

Изнасянето на реализациите на шаблони за описание на метаданни, за генериране на еднотипови ИР и за описание на процес за генериране на метаданни е позволило от една страна текстът в тази глава да не бъде натоварен с твърде много техническа информация, а от друга страна да е наличен в случай на конкретен интерес от читателя.

В **ЗаклЮчението** се посочват решените задачи, чрез които е постигната поставената цел; основните приноси, които има работата според докторанта; начините на популяризация на резултатите; практическото им приложение; перспективите за бъдещо развитие. Съгласна съм с така дефинираните приноси на дисертационния труд. От идеите за бъдещо развитие тази за "създаване на модул за дигитализация и архивиране на отчети и документи от системата за управление на качеството СиДОК в цифровото хранилище на ПУ" е практическа задача, чиято реализация ще обогати

цялостната информационна инфраструктура на ПеУ, докато задачите за "откриване на релации между документи и ресурси, генерирани от процеси на различни информационни системи, вследствие на повторяемост на стойности и близост на свойствата" и "използване на метода за извличане на знания от данни като опция за генериране на метаданни от статични данни и откриване на невидими релации между справки" е научно предизвикателство, чието решение бих следила с интерес по-нататък (според мене първата може да се разглежда като частен случай на втората доколкото става въпрос за търсене на асоциации).

7 Автореферат

Авторефератът е в обем от 30 стр. и следва структурата на дисертацията. Докторантът е успял да представи кратко и ясно най-важното от всяка част на дисертационния труд, така че читателят да получи пълна представа за изследвания проблем и постигнатите резултати.

8 Преценка на публикациите по дисертационния труд

Условията на специфичните изисквания на ФМИ на ПУ за поне 4 публикации в рецензирани издания, от които една в списание (или глава от книга на български език), са удовлетворени. Представени са 4 публикации в сборници на конференции (2 международни и 2 национални) и 3 глави от книга. Една от работите е самостоятелна, останалите са съвместни с научния ръководител. Публикациите са на български език. Удовлетворено е изискването най-съществените резултати от дисертационното изследване да бъдат публикувани.

Четири от публикациите са цитирани в две други дисертации.

Препоръката ми е да се публикува и на английски език с цел по-голяма популяризация на идеите и резултатите.

9 Лично участие на докторанта

Аз нямам съмнение в същественния личен принос на докторанта в публикациите, в които е съавтор, доколкото те отразяват основни части от неговия дисертационен труд. Самата дисертация е добре структурирана и едностилна. Резултатите от работата на докторанта са дали съществено отражение по успешното изпълнение на двата европейски проекта, в които той участва.

10 Забележки и препоръки

Бележките и препоръките, които имах, посочих по-горе.

Като цяло общата ми оценка за дисертацията е положителна.

11 Лични впечатления

Не познавам докторанта и нямам преки впечатления от работата му. От прочетеното в предоставените материали си изградих представа за млад човек, който усърдно и последователно работи върху своето развитие, добър професионалист, с възпитани умения за работа в екип, с изградени възможности за провеждане на цялостно самостоятелно научно изследване.

12 ЗАКЛЮЧЕНИЕ

Дисертационният труд съдържа оригинални приноси 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", както и на специфичните изисквания на ФМИ, приети във връзка с Правилника на ПУ за приложение на ЗРАСРБ.

Представеният дисертационен труд и съпътстващи материали показват, че докторантът Александър Трайков притежава задълбочени теоретични знания и професионални умения за извършване на научни изследвания. Поради това убедено давам своята положителна оценка и предлагам на почитаемото научно жури да присъди образователната и научна степен "доктор" на Александър Насков Трайков в област на висше образование 4. Природни науки, математика и информатика, професионално направление 4.6. Информатика и компютърни науки, докторска програма Информатика.

20.04.2017

Подпис: