

## РЕЦЕНЗИЯ

от

Д-р Георги Георгиев Марков, Професор в Институт по биоразнообразие и екосистемни  
изследвания на Българска академия на науките

на дисертационен труд за присъждане на образователната и научна степен 'доктор'

в област на висше образование *Природни науки, математика и информатика*

професионално направление 4.6 *Информатика и компютърни науки*

докторска програма *Информатика 01.01.12*

**Автор:** Виктор Ернестов Сендеров

**Тема:** "OpenBiodiv: отворена система за управление на данни за биологичното разнообразие"

**Научен ръководител:** проф. Любомир Пенев (ИБЕИ-БАН)

**Научен ръководител:** доц. Кирил Симов (ИИКТ-БАН)

### 1. Общо описание на представените материали

Със заповед № 86 от 30. 04. 2019г. на Директора на ИИКТ-БАН съм определен за член на научното жури за осигуряване на процедура за защита на дисертационен труд на тема "OpenBiodiv: отворена система за управление на данни за биологичното разнообразие" за придобиване на образователната и научна степен 'доктор' в област на висше образование „Природни науки, математика и информатика“, професионално направление „4.6 Информатика и компютърни науки“, докторска програма „Информатика 01.01.12“. Автор на дисертационния труд е Виктор Сендеров – докторант на самостоятелна подготовка към ИИКТ-БАН с научен ръководител проф. Любомир Пенев и научен консултант доц. Кирил Симов.

Представеният от Виктор Сендеров комплект материали включва

1. Дисертационен труд (book.pdf)
2. Автореферат (summary.pdf)
3. Публикации по темата на дисертационния труд
4. CV

**Бележки и коментар по документите:** в приложения пакет са пропуснати две от публикациите на дисертанта (RIO\_article\_10445.pdf; BDJ\_article\_10356.pdf), които допълнително бяха предоставени.

### **Кратки биографични данни за докторанта**

Виктор Сендеров е докторант на ЕС по програма (ITN) BIG4 на Мария-Склодовска Кюри, бенефициент, на която е „Пенсофт“, като обучителното му звено е ИИКТ-БАН, където докторантът е на самостоятелна подготовка. Научен ръководител е проф. Любомир Пенев („Пенсофт“/ИБЕИ-БАН) и научен консултант доц. Кирил Симов (ИИКТ-БАН). Докторантът притежава магистърска степен по биостатистика от Университет Мюнхен (LMU) и бакалавърска степен по компютърна математика от Университет Магдебург (OvGU). Докторската дисертация е изготвена в индустрията (фирма „Пенсофт“) със съдействието на научните институти ИИКТ и ИБЕИ от БАН.

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

Конференцията от 1992 г. на ООН за околната среда и развитието (Earth Summit) определя биологичното разнообразие като „многообразието от живи организми от всички форми, включително, сухоземни, морски и други водни екосистеми, и екологичните комплекси, към които те принадлежат: това включва разнообразие в рамките на отделния вид, между видовете и в екосистемите“. Изучаването му предоставя данни които подпомагат определянето на политиката за биологичното разнообразие, както в световен така и в национален мащаб, която да бъде основана на знанието в сферата на опазването на природата и развитието на „зелената стратегия“. Разнообразието на флората и фауната на всяка страна има значима стойност като биологичен ресурс, важен за поминъка на населението и икономиката ѝ. Тяхното устойчиво управление е важно за развитието и повишаването на общественото благосъстояние.

Но биоразнообразието е застрашено навсякъде по света. Редица растения и животни са на изчезване, главно поради човешка дейност. Необратимостта на загубата на биологично разнообразие предопределя необходимостта от неговото научно изследване и съхранение. Разглеждана в този аспект темата на дисертацията, която включва създаването на теоретичен модел-онтология за представяне на знание за биологичното разнообразие в машинен вид и също така създаването на Свързани Отворени Данни (LOD) за биоразнообразието, стъпващи на модела е с висока актуалност. Нейната актуалност се предопределя и от реалността, че към дадения момент, данните за биологичното разнообразие са разпръснати в печатни монографии и статии, цифрови монографии и статии, както и в несвързани помежду си бази



данни. Значимостта на темата се свързва и с факта, че въпреки високите темпове на загуба на биологично разнообразие и потенциално катастрофалните последици от този факт за екологичното състояние на Планетата и за потенциалните възможности за изхранване на населението в световен мащаб, е налице недостиг на учени и средства, както и липса на информираност у обществеността за науката биологична таксономия и систематика. От гледна точка на информатиката, въпреки че вече съществуват речници за основните понятия от биологичната таксономия, не е решен проблемът за изграждане на връзки между базите данни, съхраняващи таксономично знание, както и проблемът за извличане на знанието, поместено в научни статии за биологичното разнообразие.

#### **4. Познаване на проблема**

Във Въведението към дисертацията, Виктор Сендеров въвежда научния проблем, като достатъчно място е отделено за историческа справка. Прави впечатление, че въпреки дългогодишни опити да се стандартизира областта на информатиката на биоразнообразието, тази стандартизация все още е в своя начален стадий. След историческата справка се преминава към преглед на литературата, в който са разгледани две направления: (1) Свързани Отворени Данни и (2) Публикуване на знания и данни за биоразнообразието. Интересно е да се отбележи разнородната литература (от сферата на таксономията, базите данни и изкуствения интелект), с която дисертантът се е запознал при проектирането на системата и творческата им оценка. Изведени са три констатации: (1) информатиката на биоразнообразието работи с изключително хетерогенни данни, (2) универсална система за идентификация на понятията не е налична, и (3) налице са много първични източници под формата на бази данни и цифровизирани монографии и научни статии, които могат да послужат за захранване на система от знания.

На базата на тези три констатации са формулирани целите и задачи на докторантурата, като целта е определена като създаване на система за управление на знанието за биологичното разнообразие на базата на научни статии и публични бази данни. Тази цел е разбита на няколко задачи, като отделна глава е посветена на всяка една от тях.

Смятам, че докторантът се е запознал в дълбочина с литературата по проблема и удачно е идентифицирал неизследвани проблеми, към които се е насочил.

#### **5. Методика на изследването**

Във Въведението се представя методиката на изследването. В нея, за конкретните задачи са използвани подходящи и модерни методи.

## 6. Характеристика и оценка на дисертационния труд

Глава 1 представя архитектурата на системата, като се разяснява смисълът, който дисертантът влага в понятието система от знания: а именно база данни от структурирано и свързано знание и логически правила, които позволяват машинен извод от фактите в базата. По-нататък се представят основните източници на информация, които хранят системата. Това са списанията на академично издателство „Пенсофт“ и базите данни на „Плаци“ и GBIF. Избраните източници на информация покриват значителна част от публично-достъпната информация относно биоразнообразието. Освен това тези източници предлагат полу-структурирано знание, което е подходящо за изграждането на ядро на система от знания, която да се развива и по-нататък.

В Глава 2 е представен основният теоретичен резултат от работата на докторанта: система OpenBiodiv-O. Текстът на главата е публикуван като публикация 8 в Journal of Biomedical Semantics. Иновативна е трактовката на таксономичната концепция като отделно понятие от таксономично име: а именно таксономична концепция представлява идеята, формирана в научна публикация за биологичен вид и може да се символизира от таксономично име, последвано от цитат на научната публикация, която третира въпросното таксономично име.

В Глава 3 са представени Свързани Отворени Данни, разработени по схемата OpenBiodiv-O. Главата завършва с дискусия относно изпълнението на Принципите за Свързани Отворени Данни и относно изчислителната ефикасност на процеса по преобразуване.

В Глава 4 са изложени техническите детайли и имплементацията на алгоритмите, използвани за създаване на Свързани Отворени Данни. Реализацията се базира на софтуерна библиотека с отворен код.

В Глава 5 са разгледани работни потоци, засягащи попълването на системата с нови данни. Тези потоци могат да бъдат от полза за авторите на научни статии, така щото данните, които публикуват да се озоват по-бързо в семантичната мрежа.

В Глава 6 е представен потребителският интерфейс на системата, представляващ уеб сайт с три целеви групи потребители: обикновена, специализирана, и програмисти. Сайтът е публично достъпен на адрес <http://openbiodiv.net>. Сайтът не претоварва потребителя с графични елементи като показва обикновен search bar, под който са поместени икони на различните приложения (apps). При търсене по ключова дума, напр. за автор „Lyubomir Penev“ се зарежда информационна страница показваща информация за автора и семантични връзки към статиите, които е публикувал и неговите съавтори. Удачно е, че е помислено за различ-



ните нива на компютърна компетентност на потребителите за системата, но за съжаление предвидените app-ове все още не са включени.

Останалите глави представляват листинги на изходен код, социологически детайли за потребителите на системата, авторска справка и литературна справка.

### **7. Приноси и значимост на разработката за науката и практиката**

В Заключението дисертантът очертава, как работата изпълнява целите и задачите и дискутира проблеми, които все още не са решени от системата. Очертават се три основни приноса:

1. Създаване на система OpenBiodiv-O на публикуване на знание за биоразнообразието.
2. Създаване на Свързани Отворени Данни, OpenBiodiv-LOD, за биоразнообразието.
3. Разработване на софтуерно обезпечение, позволяващо генерирането на поддръжането на Свързани Отворени Данни

Прави впечатление сравнително широкият спектър от проблеми, който е разгледан, и това че OpenBiodiv-O прекарва мост между академичната издателска дейност и моделирането на биологични данни. Разработката на дисертацията в издателска среда, (под ръководството на ръководител и консултант от БАН), определя и по-приложния характер на някои от главите. Фактът, че теоретичните разработки са превърнати в действителен продукт (Beta версия) и вече внедрени в работата на „Пенсофт“ - <http://openbiodiv.net>, и <http://graph.openbiodiv.net/>, поддръжани от „Пенсофт“ е допълнителен принос, позволяващ нейното ползване и от научните среди.

### **8. Преценка на публикациите по дисертационния труд**

Преценката е извършена съгласно Специфичните изисквания на ИИКТ-БАН, за изчисляване на точки за „Научни публикации в издания, които са реферирани и индексирани в световноизвестни бази данни с научна информация (Web of Science и Scopus, Zentralblatt, MathSciNet, ACM Digital Library, IEEE Xplore и AIS eLibrary) 50 за публикувани в Q1, 40 за публ. в Q2, 30 за публикувани в Q3, 24 за публ. в Q4, 20 за публикувани в издание със SJR без IF, 12 за други“ показатели в професионално направление 4.6. Информатика и компютърни науки.

№	бр. автори	Издание/ издателство	Индексиране	Точки
1	2	RIO/ Pensoft	други	0
2	21	RIO/ Pensoft	други	12
3	5	BDJ/ Pensoft	Q3	30
4	3	RIO/ Pensoft	други	12
5	13	RIO/ Pensoft	други	12
6	9	RIO/ Pensoft	други	12
7	5	ZooKeys/ Pensoft	Q2	40
8	9	J. Biomed. Semantics Springer Nature	Q1	50
9	1	Cybernetics and Inform. Tech./ BAS		0

- 3 публикации (№ 3, 7, 8), публикувани в реферирани научни списания в SCOPUS, с импакт-ранк SJR, освен това всичките три публикации са реферирани и в Web of Science, като две от тях (№ 7 и 8) имат и импакт фактор и попадат в квантил Q3. Публикации № 3 и 8 са индексирани и в PubMed.

- 5 публикации (№ 1, 2, 4, 5, 6) в RIO Journal, електронно списание, публикуващо както традиционни научни изследвания, така и трудове относно аспекти от научния процес, които не са традиционно обхванати от академичната издателска дейност. Напр. публикация № 1 не е научно изследване, а работен план на докторантурата и не е включена в точкуването.

- 1 публикация (№ 9), която в момента се рецензира и не е все още публикувана.

В публикация № 9 дисертантът е единствен автор, а останалите публикации са в съавторство, както с научните му ръководители, така и с български и чуждестранни учени. В четири от публикациите дисертантът е първи автор (№ 1, 4, 8, 9).

Понеже публикация № 9, в която дисертантът е единствен автор, е в процес на рецензиране, тя не е включена в точкуването. Също така в точкуването не е включена и публикация 1, поради по-особения ѝ характер. В автореферата е приложен списък с намерени 23 цитата, което е похвално, имайки предвид краткия срок на разработката.



## **9. Лично участие на докторанта**

Докторантът прояви изключителна самостоятелност и инициатива при създаването на продуктивни контакти с учени от БАН, въпреки че е базиран в „Пенсофт“. По време на докторантурата той поравно се концентрира върху теоретичния модел, практическото приложение и публикуването на научните открития, за което свидетелстват множеството му научни статии. Дисертантът успешно работи в една интер-дисциплинарна и международна област, като си сътрудничи, както с научните си ръководител и консултант, с колеги от Пенсофт, така и с учени от два от институтите на БАН (ИБЕИ и ИИКТ), а също така и с колегите си от международната мрежа Мария Кюри, към която дисертантът принадлежи. Въпреки съвместния характер на неговата разработка, аз съм убеден, че той е допринесъл в много голяма степен за успешния ѝ завършек.

## **10. Автореферат**

Авторефератът на български отразява съдържанието на дисертацията добре, но търпи подобрене и малки съкращения. Автореферат е направен съответстващо на шаблона на ИИКТ.

## **11. Критични забележки и препоръки**

Интердисциплинарният характер на разработката, предопределя широкия ѝ профил и възможността тя да се използва от редица учени, работещи в областта на проучването на биоразнообразието. Препоръчвам на дисертанта да положи усилия „Система за управление на знанието за биологичното разнообразие OpenBiodiv“ да бъде широко популяризирана в средите на учените биолози, за нейното практическо приложение в тяхната научно изследователска работа.

## **12. Лични впечатления**

Личният ми контакт с дисертанта Виктор Сендеров остави положително впечатление. Познавам го от изпит по предмет и от някои негови презентации, на които съм присъствал.

## **13. Препоръки за бъдещо използване на дисертационните приноси и резултати**

От ключово значение е уеб-сайтът и данните да бъдат допълвани и да са достъпни за обществеността. В по-общ философски смисъл биологичното разнообразие може да бъде възприето като съвкупност от знания, изведени от развитието на биологичните видове в продължение на милиони години, за това как да оцелеем в изключително разнообразните условия на съществуване на Земята, осъзнавайки, че човечеството в момента изгаря „библиотеката на живота“. Таксономичните, функционалните, трофичните, генетичните и редица други

измерения на биологичното разнообразие все още са относително слабо количествено определени. Дори познаването на таксономичното разнообразие, най-известното измерение на биоразнообразието е непълно и силно предубедено към видовото ниво, мегафауна и компонентите, използвани от хората. Подпомагането на систематизирането на всички тези знания, с използването на предлаганата система OpenBiodiv-O ще позволи на учените по-детайлно да се запознаят с известните данни и ще подпомогне по-рационалното описание на биоразнообразието на Планетата.

## ЗАКЛЮЧЕНИЕ

Дисертационният труд съдържа научни, научно-приложни и приложни резултати, които представляват оригинален принос в науката и отговарят на изискванията на Закона за развитие на академичния състав в Република България (ЗРАСРБ) и Правилника за прилагане на ЗРАСРБ. Представените материали и дисертационни резултати съответстват и на Правилника за специфичните условия за придобиване на научни степени и за заемане на академични длъжности в ИИКТ-БАН. Дисертационният труд показва, че докторантът Виктор Сендеров притежава задълбочени теоретични знания и професионални умения по научна специалност „Информатика“ като демонстрира качества и умения за самостоятелно провеждане на научно изследване.

Поради гореизложеното, убедено давам своята положителна оценка за проведеното изследване, представено в рецензираните по-горе дисертационен труд, автореферат, постигнати резултати и приноси, и предлагам на почитаемото научно жури да присъди образователната и научна степен 'доктор' на Виктор Сендеров в област на висше образование: „Природни науки, математика и информатика“, професионално направление „4.6 Информатика и компютърни науки“ докторска програма „Информатика 01.01.12“

29.05. 2019 г.

Рецензент: Проф., Дбн Георги Марков

**NOT FOR  
PUBLIC RELEASE**