



Национална конференция  
по информатика  
посветена на  
80-годишнината  
от рождението на  
проф. Петър Бърнев

12-13 ноември 2015  
София, България

## **СБОРНИК ДОКЛАДИ**

София, 2016  
Институт по математика и информатика – БАН

**Национална конференция по информатика, посветена на 80-годишнината от рождението на професор Петър Бърнев – сборник доклади**

Редактори: Красимира Иванова и Нели Манева

Технически редактор: Анна Салева

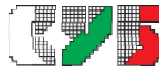
Издател: Институт по математика и информатика, БАН

Със съдействието на: Асоциация за развитие на информационното общество

Печат: „Ракурси“ ООД, гр. Пловдив

ISBN 978-954-8986-45-8

© 2016



## Предговор

През 2015 г. се навършиха 80 г. от рождението на проф. Петър Бърнев, един от основателите на направление „Информатика“ в Института по математика и информатика при БАН, основател и ръководител на секция „Математическо осигуряване“, дългогодишен заместник-директор на Единния център по математика и механика, дългогодишен заместник-председател на Съюза на математиците в България, основател и пръв председател на Асоциацията за развитие на информационното общество.

Проф. Бърнев беше един от пионерите в развитието на информатиката в България. Неговият принос има много аспекти – от изучаването и развиването на информатичната теория, през избора какво да е мястото на българската информатична общност на фона на общите процеси на развитие на информатиката и компютърната индустрия по света, до обучението на младите информатици не само в рамките на тясната специалност, но и в по-широк философски, обществен, а и чисто човешки план.

По повод на тази кръгла годишнина от рождението на проф. Бърнев, Институтът по математика и информатика при БАН (ИМИ-БАН), съвместно с Факултета по математика и информатика при Софийски университет „Св. Климент Охридски“, Пловдивския университет „Паисий Хилендарски“, Асоциацията за развитие на информационното общество, Българската секция на Асоциацията за информационни системи, Съюза на математиците в България и Съюза на учените в България организираха провеждането на юбилейна научна конференция по информатика в негова чест.

В приятелска атмосфера се събраха няколко поколения негови колеги и ученици, заразени малко или повече от нестандартния и многостранен поглед на проф. Бърнев към проблемите и събитията.

Казват, че един човек има голям научен принос, ако е успял да създаде школа от последователи на своите идеи в научен план. Хвърляйки поглед към миналото и настоящето, можем да заключим, че проф. Бърнев остави подълбока следа – по един или друг начин той повлия на много хора в оформянето на тяхното философско виждане за света, не капсулирането в конкретното научно поприще, на което човек се посвещава, а осмислянето на по-общите процеси на влияние на научния напредък върху човека и обществото.

Използвахме събитието, за да честитим и 80-годишнината на доц. Маргарита Бърнева – не само спътничка в живота на проф. Бърнев, но и на практика първата жена-информатик в България, дала своя съществен принос в развитието на програмирането у нас. Тя и сега продължава да участва активно в организирането на ежегодната Национална конференция „Образованието и изследванията в информационното общество“.

През 2015 г. се навършиха и други годишнини на колеги, които, за съжаление, вече не са между нас: проф. Томов – създател на направлението „Изкуствен интелект“ в България, проф. Додунеков – създател на групата по кодиране в България, доц. Бойчо Кокинов – основоположник на Когнитивната наука в България. Част от докладите в историческата секция на тома са посветени на тях или на резултатите от техния труд. Друга част проследяват развитието на основните информатични направления в ИМИ-БАН – като история, настояще и бъдещи планове.

Една наука не може да има развитие, ако няма последователи. Ето защо част от докладите в историческата част бяха посветени именно на обучението по информатика – щрихи от развитието на университетското образование по информатика в България, както и как информатиката стана олимпийска дисциплина за студентите.

Втората част на конференцията съдържа докладите, които представяха нови научни резултати във водещи направления на изследване в информатиката у нас.

Дело на проф. Бърнев е схемата, която дълги години продължава да се ползва като средството, с което за пет минути на всеки човек да може да се обясни какво представляват информатиката и информационните процеси. За нея по-подробно може да прочетете в ключовия доклад от тома, посветен на проф. Бърнев, с автори проф. Ескенази и проф. Станчев. Тази схема послужи за основа и на предложения от проф. Раденски научен форум „Събиране, съхранение, обработка и визуализация на големи данни“, който се състоя в рамките на конференцията.

Томът, макар и позакъснял, вече е пред вас. Надяваме се, поглеждайки историята и настоящето, да осмислим бъдещето си.

доц. д-р Красимира Иванова, проф. д-р Нели Манева  
25 март 2016 г.  
гр. София

## **Организатори:**

- Институт по математика и информатика при БАН
- ФМИ при Софийски университет „Св. Климент Охридски“
- Пловдивски университет „Паисий Хилендарски“
- Асоциация за развитие на информационното общество
- Българска секция на Асоциацията за информационни системи
- Съюз на математиците в България
- Съюз на учените в България

## **Програмен комитет**

- Петър Станчев (Институт по математика и информатика при БАН) – *председател*
- Павел Азълков (Pennsylvania State University)
- Красимир Атанасов (Институт по биофизика и биомедицинско инженерство при БАН)
- Ренета Бърнева (State University of New York at Fredonia)
- Александър Геров (Институт по математика и информатика при БАН)
- Дарина Дичева (Winston-Salem University)
- Милена Добрева (University of Malta)
- Аврам Ескенази (Институт по математика и информатика при БАН)
- Владимир Занев (Columbus State University)
- Фани Златарова (Elizabethtown College)
- Антон Илиев (Съюз на математиците в България, Пловдивски университет „Паисий Хилендарски“)
- Стефан Керпеджиев (General Dynamics Mission Systems)
- Иван Койчев (Асоциация за развитие на информационното общество, Софийски университет „Св. Климент Охридски“)
- Красимир Марков (Институт по математика и информатика при БАН)
- Мария Нишева (Софийски университет „Св. Климент Охридски“)
- Атанас Раденски (Charman University)
- Евгения Сендова (Съюз на математиците в България, Институт по математика и информатика при БАН)
- Георги Тотков (Пловдивски университет „Паисий Хилендарски“)

## **Организационен комитет**

- Стоян Порязов (Институт по математика и информатика при БАН) – *председател*
- Красимира Иванова (Институт по математика и информатика при БАН) – *секретар*
- Вася Атанасова (Институт по биофизика и биомедицинско инженерство при БАН)
- Тодор Брънзов (Институт по математика и информатика при БАН)
- Георги Врагов (Асоциация за развитие на информационното общество, Институт по математика и информатика при БАН)
- Детелина Димитрова (Асоциация за развитие на информационното общество, Институт по математика и информатика при БАН)
- Нели Манева (Институт по математика и информатика при БАН)
- Илия Митов (Институт по математика и информатика при БАН)
- Галя Стоянова (Асоциация за развитие на информационното общество, Институт по математика и информатика при БАН)



## Съдържание

<b>Доклади, посветени на историята на развитието на информатиката в България .....</b>	<b>9</b>
Професор Петър Бърнев – 80 години от рождението му <i>Аврам Ескенази, Петър Станчев .....</i>	11
Софтуерни технологии в ИМИ <i>Румен Радев, Аврам Ескенази, Нели Манева, Весела Ангелова .....</i>	21
Развитие и приложение на системи за компютърна алгебра в ИМИ-БАН <i>Маргарита Спиридонова, Александър Геров, Мария Нишева .....</i>	30
Компютърната лингвистика в ИМИ – история, проекти и резултати <i>Людмила Димитрова .....</i>	39
История и развитие на направлението Дигитална хуманитаристика в секция ПИТХН на ИМИ <i>Калина Сотирова, Анна Семева .....</i>	58
Секция Математически основи на информатиката в ИМИ-БАН: Теория на кодирането и криптография <i>Петър Бойваленков .....</i>	64
Бойчо Кокинов и развитието на когнитивната наука в България <i>Красимира Иванова .....</i>	68
История на обучението по информатика в Шуменския университет <i>Теодоси Теодосиев .....</i>	83
Републиканска студентска олимпиада по програмиране <i>Николай Киров .....</i>	87
<b>Научни доклади .....</b>	<b>93</b>
Приложение на итеративни принципи и обектно-ориентиран подход при разработката на образователен софтуер <i>Калинка Калоянова, Андрей Антонов .....</i>	95
Methodological Approach for Modelling, Simulation & Assessment of Complex Discrete Systems <i>Zlatogor Minchev .....</i>	102
Методологии за проектиране и разработване на сигурни уеб услуги: съвременен състояние <i>Малинка Иванова .....</i>	111
Password Manager with 3-Step Authentication System <i>Zhelyazko Petrov, Razvan Ragazan .....</i>	121

Компютърно творчество – вчера, днес и утре <i>Евгения Сендова, Христо Стайков</i> .....	129
Challenges to the Support of Bulgarian Open Institutional Repositories <i>Juliana Peneva, Stanislav Ivanov, Krassimira Ivanova, Peter Stanchev</i> .....	143
Онтологичен модел за музейна документация: пример от българската практика <i>Калина Сотирова</i> .....	153
Цифровизирани колекции на пловдивски културни институции <i>Владимир Балчев, Димитър Минев, Георги Врагов</i> .....	165
Моделиране на бизнес-обекти чрез СУБД с външна логика на данните <i>Анани Ризов, Илия Митов, Красимира Иванова</i> .....	173
Variable Ordering Based on Distinguishing Sets <i>Vesela Angelova</i> .....	185
SDBC-Based Technology-Independent Functionality Models (Applications in Emergency Medical Response) <i>Boris Shishkov</i> .....	194
A Decision Making Support Method in Agile Software Development <i>Neli Maneva</i> .....	205
Оптимизиране на процесите по интеграция между информационните системи и изграждане на единен потребителски профил <i>Младен Георгиев</i> .....	215
Параметрични вектори за разпознаване на емоциите в говора <i>Александър Илиев</i> .....	225
Реклами, реклами, ... – намиране без виждане <i>Мирослав Иванов, Красимира Иванова, Илия Митов, Евгения Великова</i> .....	239
Индекс на авторите .....	249

**Доклади, посветени на историята  
на развитието на информатиката  
в България**



# **Професор Петър Бърнев – 80 години от рождението му**

**Аврам Ескенази, Петър Станчев**

Институт по математика и информатика при БАН, София  
[eskenazi@math.bas.bg](mailto:eskenazi@math.bas.bg), [stanchev@math.bas.bg](mailto:stanchev@math.bas.bg)

## **1. Някои биографични данни**

Петър Бърнев е роден на 28 февруари 1935 в София. Женен, с две дъщери. Съпругата му и двете му дъщери са завършили математика в СУ. Съпругата му Маргарита е учен и преподавател по информатика, добре позната на колегията, днес пенсионер. Голямата му дъщеря, Ренета, е професор и ръководител на катедра в Щатския университет на Ню Йорк, САЩ, а малката дъщеря – Светлана, живее и работи в Монреал, Канада.

Завършил е през 1958 г. специалността математика – научно-производствен профил, в Софийския университет, специализирал две години (1961-63) в Обединения институт за ядрени изследвания, гр. Дубна, и 6 месеца в Института „Блез Паскал“, Париж, както и краткосрочно в редица чуждестранни научни организации.

През 1968 г. е защитил дисертация за научната степен „Кандидат на физико-математическите науки“.

След като една година е учител по математика, през 1960 постъпва с конкурс в Института по математика и информатика при БАН, където е последователно научен сътрудник, старши научен сътрудник и професор (от 1984). Основател (1964) и ръководител е на секцията „Информационни изследвания“ (бивша „Автоматизация на програмирането“, а по-късно – „Математическо осигуряване“) от основаването ѝ до пенсионирането му през 2005 г., както и на Лабораторията по приложение на математиката – Пловдив през целия период на съществуването ѝ.

Починал на 29 април 2010 година.

## **2. Приноси в науката**

Петър Бърнев остави името си в науката преди всичко като информатик. Научните му постижения отразяват изключително широките му от тематична гледна точка интереси. Добре известно е, че през целия период на развитието си (примерно през последните 6-7 десетилетия) информатиката е била сред най-динамичните научни направления. Обективните данни показват, че П. Бърнев е публикувал свои и съвместни резултати през много голяма част от

---

Тази работа е представена за първи път на конференцията, посветена на 80 г. от рождението на проф. Бърнев, на 12.11.2015. По-късно нейни варианти са публикувани в сп. Наука бр.1/2016 и в Сборника "Математика и математическо образование" на СМБ, 2016.

този период – от 1961 практически до смъртта си през 2010. През това време, след кратък начален интервал, когато той навлиза в направлението, от средата на 1960-те години неговите резултати са свързани в повечето случаи с най-активно решаваните информатични проблеми.

В областта на *приложението на компютрите в лингвистиката* Бърнев изследва заедно с Д. М. Добрев и Р. Киркова (1967, 1972) за пръв път с компютър честоти на конструкции в българския език, като авторите създават необходимите за целта алгоритми и програми. Добре известно е как тази пионерска работа беше последвана от дългогодишни изследвания на български учени, които от години се ползват със световна известност и признание.

Още през 1964 г., съвместно с М. Апостолова, публикува статия с обзорен характер по въпросите на автоматизацията на програмирането, а самостоятелно изнася доклад по подобна тематика. Скоро след това започва публикуване на научните си приноси в областта на *езиците за програмиране и методите за трансляция*. Независимо от МакКарти предлага да се провежда статистически анализ на фактическото използване на конструкциите в езиците за програмиране и прилага резултатите от изследване на програми на Фортран при създаване на оригинален транслятор. Ръководи създаването на първите системи за програмиране в България. Естествено, започва се с Минск 2 – първият реално внедрен и експлоатиран компютър в България, в Изчислителния център на тогавашния Математически институт на БАН: Диспечерска система МИД за машината Минск-2 (1966), Библиотека стандартни програми за машината „Минск-2“ (1966), Система за символично програмиране на машината Минск-2 (1967). Някои от публикацииите по тази тема са направени у нас, има обаче и други – в полско научно списание. Ясно е, че по това време разработването на такъв комплексен инструментариум е изисквало усилия и творчески резултати, а не само прилагане на отработени и проверени методи и процедури – такива просто е нямало или са били нови и недостъпни. Струва си да отбележим и още един важен факт. В информатиката, в болшинството от случаите, както и в конкретния разглеждан, резултатите се получават с колективни усилия. Не случайно казахме по-горе за Петър Бърнев „ръководи“ – негови съавтори са П. Петров, М. Бърнева, Д. Петрова, Д. Тошков, Р. Калтинска и др. От този момент нататък, практически във всички подобни случаи на колективни резултати, той е ръководителят. По тези причини навсякъде по-долу, когато цитираме съавтори на Петър Бърнев, ще имаме предвид, че във всяка публикация или приложна разработка той е ръководителят или водещият в нея. Ако има изключение, то ще бъде изрично отбелязвано. По-долу в друг раздел ще бъде отбелязана всестранната дейност на П. Бърнев по разработването на софтуера на следващия, доставен вече в по-широки мащаби, компютър „Минск“ – модел 32, част от която дейност се основава отново на разработки на Бърнев и под негово ръководство в областта на езиците за програмиране.

Петър Бърнев реагира на развитата от Заде и много модерна в края на 1960-те години теория на *размитите множества*, като творчески я прилага в социологическите проучвания заедно с Вл. Димитров и П. Станчев (1974, 1975) и към изучаването на изкуствените езици (1975).

В областта на *кооперативните (съвместните) информационни дейности* е сред първите изследователи в тази област и организира международна конференция по тези въпроси още през 1978 г., където изнася основния доклад. Още през 1976 обаче той е автор на доклад по темата на международния конгрес на УВМ по приложна математика в Солун, а през 1977 публикува обширна статия в сп. Сердика. Десетина години по-късно П. Бърнев се връща към тематиката и заедно с А. Григорова и М. Шишкова изследва информационни дейности при съвместната работа на група автономни изпълнители или при противодействие на две групи (1990, 1991, 1994).

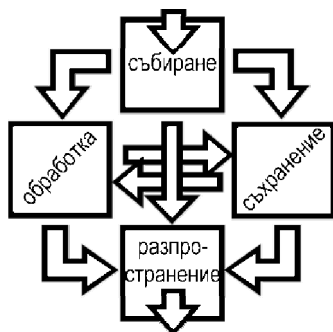
Впрочем, дълги години, в съответствие с европейските тенденции, тези дейности се развиват под названието *бюротика* – въведено у нас от П. Бърнев. Самият той определя за тяхно начало 1977, по-точно разработената от П. Азълов, тогава аспирант на П. Бърнев, система „Рекор“ за автоматизиране на административната дейност, свързана с регистрация на решения и кореспонденция и контрол за изпълнението им. Следват реализация на автоматизиран личен бележник (1983), на система за автоматизиране на организационните дейности на съвет (1984), разгледани са възможностите за автоматизиране на дейностите на Българска енциклопедия (1985), изследвано е автоматизирането на дейността на ръководител с участието на М. Бърнева и нейния аспирант Д. Обретенков. По-късно изследванията продължават със статистически анализ на поставяните резолюции върху документите (1989), на деловите разговори по телефон (1990), създаването на автоматизирано информационно табло като средство за общуване (1990) и др. Струва си да си дадем сметка, че това са всъщност предшествениците на същите инструменти, използвани днес почти от всекиго благодарение на новите технологии. От чисто научна гледна точка вероятно най-значими са резултатите по автоматизирано съставяне (синтез) на текстови документи и свързания с това метод за извличане на данни от текстове. В тези работи участват аспирантите на Бърнев от Виетнам До Виет Нга и Занг Конг Тхе и особено задълбочено, продължително и резултатно – друг негов ученик и сътрудник – Ст. Керпеджиев (1983-1988). Бърнев развива широка популяризаторска и учебна дейност по бюротика, особено след 1985, а през 1994 издава книгата „Бюротика“, предназначена за учебни цели, но съдържаща и някои оригинални елементи.

В областта на *информационните системи* постиженията на П. Бърнев са разнообразни по характер. От една страна е специфична разработка, в която той заедно с Д. Дичева, независимо от Лирски, развива методология и система за работа с непълна и ненапълно достоверна информация (1979). От друга страна са работи с по-традиционен характер, посветени на разработката на конкретни информационни системи, например свързани с кадрови и подобни справки (1975).

В областта на *когнитивната наука* П. Бърнев работи от 1983, предимно с друг свой ученик – Б. Кокинов, по въпроси на представяния на информация в човешката памет. Организира национален семинар по когнитивна наука. В резултат през 1990 се създава дружество по когнитивна наука. Тук ще цитирам рано напусналия ни Бойчо Кокинов, несъмнено най-авторитетната българска и

с международен авторитет фигура в тази област: „...той [Бърнев] спомогна за създаването на още едно научно направление в България – когнитивната наука. Той никога не е споменавал това и не е имал такива претенции, но е факт, че когнитивната наука се роди от семинара „Информатика и психология“, който той организира в Института по математика през 1983 година. Тогава аз бях студент 3 курс, но той активно ме подкрепяше и избутваше напред. Без него когнитивната наука нямаше да се роди в България, поне не и през 20. век“...

С важно методологическо значение са изследванията на Бърнев в областта на **информацията и информационните процеси**, особено тези върху общите информационни структури и основните информационни дейности. Прочута е предложената от него графика, която в синтезиран вид представя модел на тези структури и процеси. Дълги години тази графика беше използвана като емблема на многобройни национални и международни събития.



Освен в информатиката, Петър Бърнев постига сериозни научни резултати и в други области.

В областта на **изчислителната математика** съвместно с К. Дочев и П. Русев (1964, 1969) създава числен метод за едновременно намиране на всички корени на алгебрични уравнения, който по-късно се развива от други автори (Кернер и др.) по света.

В областта на **математическата физика** съвместно с Р. Денчев и Кми Зе Пхен (1962, 1963) работи по решаване на уравненията на Чу-Лоу, по изследване на една хипотеза за кълбовидната мълния и по някои други въпроси.

В областта на **геофизиката** в периода 1963-1965 самостоятелно или съвместно с В. Томов и М. Бърнева определя понятието гравитационна плоча (ядро) на тяло и определя ядрата на различни тела<sup>1</sup>. Впрочем, кандидатската му дисертация, защитена през 1968, е свързана с тази тематика. Съвместно с Ив. Недялков през същото десетилетие работи по създаване на методи за пресмятане на гравитационни и други полета в посока на техните особености.

---

<sup>1</sup> Гравитационното ядро е образуване с по-ниска размерност, в което, ако се съсредоточи по подходящ начин масата на тялото, то гравитационните потенциали на ядрото и тялото са едни и същи извън тялото.

Тук следва да отбележим, че съвместната научна дейност на П. Бърнев с проф. Иван Недялков продължава от 1961 в продължение на няколко години. П. Бърнев с уважение го счита за един от своите учители и му посвещава очерк (съавтор М. Бърнева) с описание на научната, педагогическата и организационната му дейност, както и на чисто човешките му качества.

### **3. Приноси в образованието**

Интересът на Бърнев към въпросите на образованието по информатика и математика е дълготраен, на всички нива, намира израз в разнообразни форми (учебни планове, програми, учебници, учебни помагала, публикации, доклади, участие и организиране на различни събития). Дейностите му в това направление изиграват много важна роля. За дълъг период от време той е фактически водещата фигура у нас по въпросите на обучението по информатика.

Вероятно първият му принос е прочутата книга „Електронни сметачни машини“ (1966), написана заедно с Д. Богданов, В. Василев и Бл. Сендов. Макар да е била предназначена за по-широк кръг читатели, заедно с това тя е и изиграла роля на първия университетски учебник по изчислителна техника и програмиране у нас<sup>1</sup>.

Тази книга е последвана през годините (започвайки от 1971) от над 30 учебника, сборника, помагала и др., част от тях в рамките на международни колективи, част – с експериментален характер.

Тъй като няма как да представя пълния набор от писмените изяви на Бърнев в образователното направление, ще илюстрираме интереса му към различните образователни равнища и разновидности като посочим само четири от тях:

- съвместната му с Д. М. Добрев статия в сборник (1976) по осъвременяване на обучението по математика в средното училище във връзка с кибернетиката и изчислителната техника,
- самостоятелната му публикация и доклад (1976) за образованието по информатика във ВТУЗ на IV национална конференция „Обучение на студентите от нематематическите ВУЗ по числени методи и програмиране“,
- съвместната му с П. Азълов, Д. Добрев и Ц. Бистеров публикация за проект за програма по информатика за ЕСПУ в сп. „Обучението по математика“ (1986),
- съвместната му с М. Бърнева публикация върху образованието по информатика и информационното общество в сборник от международна конференция през 2001.

Що се отнася до университетската лекционна дейност на Петър Бърнев, вероятно това е преподавателят с най-дългата и най-обемната такава в

---

<sup>1</sup> Самият П. Бърнев, който беше добре известен с безупречната си научна етика, в своя статия за Д. М. Добрев пише: “Съавтор е [Д. М. Добрев] на първия университетски учебник по информатика заедно със свои ученици“

България, започнала през далечната 1964 година с курса „ЕСМ и програмиране“ и продължила повече от четири десетилетия. Не е възможно тук да се изброят всички прочетени от проф. Бърнев курсове, никак не са малко и тези, прочетени за първи път в университет в България. Интересна (но и твърде трудна) би била задачата да се пресметне колко са студентите, които са слушали поне един курс на Бърнев, като вземем предвид, че по наши данни това се е случвало поне в СУ, НБУ, в университетите в Пловдив, Бургас, Шумен, в Академията на МВР. Все пак да споменем някои от тези курсове в опит да покажем тяхното разнообразие и широта на спектъра им: Програмиране за АСМ, Методи за трансляция, Алгоритмични езици, Операционни системи за работа в режим на разпределение на ресурсите, Паралелно програмиране, Бюротика, Изчислителна математика, Проектиране на АСУ, Размити множества, Информатика и бизнес. Разбира се, заслугата на Петър Бърнев е не само в количеството и новостта на курсовете, а и в тяхното качество. Ще цитираме по този повод В. Вълев (днес професор): „Проф. Бърнев четеше лекциите на другия поток, като в програмата си беше включил и изучаване на езика PL/I. Този език се изучаваше тогава само теоретично, тъй като единствената изчислителна машина по онова време в Института по математика и механика беше Минск 32, която имаше транслятор само за езика Фортран. Започнах да посещавам неговите лекции. Проф. Бърнев имаше невероятни педагогически умения. Владееше аудиторията. Преподаваше с невероятна вещина не само езика за програмиране, но и основите на програмирането в детайли. Мога да кажа, че от него получих тези основи. Още пазя записките си оттогава. Години по-късно, когато беше внесена машината ЕС 1040 и имаше транслятор за PL/I, благодарение на знанията, които получих от проф. Бърнев, написах доста програми на този език. Преподавах този език и като спецкурс на студентите по математика във Висшия педагогически институт в Шумен.“

Заедно с тази богата лекционна дейност е налице и активна работа с дипломанти и аспиранти. Успешно защитилите дипломанти са около 60, аспирантите – над 20, някои от тях – чужденци. От тези възпитаници на проф. Бърнев израснаха видни учени и университетски преподаватели у нас и в чужбина, известни софтуеристи, ръководители в областта на производството и използването на софтуер, и на изчислителна техника.

#### **4. Научно-приложна и проектна дейност**

Както вече се видя, в много случаи в областта на информатиката е твърде трудно да се постави границата – къде свършват чисто научните резултати и къде започват научно-приложните. Тъй като не е мястото тук да решаваме този въпрос и след като по-горе в раздела за научната дейност споменахме „Минск 32“, отделяме тук малко място и на тази страна от дейността на Бърнев.

В началото на 1970-те в България са внесени около 15 компютъра „Минск-32“, като един от първите е за Математическия институт. За да се координира работата по създаване и използване на софтуера на тези компютри е решено да се създаде Асоциация на организациите, притежаващи тези компютри,

ръководена от Математическия институт. Асоциацията е утвърдена от тогавашния ДКНТП и е с ръководител Петър Бърнев. Работата на Асоциацията се контролира от няколко нива на ДКНТП до зам.-председател. По-късно (поради смешни днес идеологически причини) приемникът на ДКНТП – КНТПВО – преимува асоциацията в клуб „Минск-32“.

Асоциацията организира редица национални съвещания, две национални конференции, семинари. Активно сътрудничи със съветските специалисти, изпратени дългосрочно със задачи по внедряването и експлоатацията на Минск 32. Под чадъра и от името на Математическия институт, П. Бърнев организира няколко курса по различни аспекти на „Минск-32“ за сътрудници на членове на Асоциацията. Освен това, чрез националната асоциация България членува активно в международната асоциация на клубовете „Минск-32“.

Заедно с това, под ръководството на П. Бърнев са разработени разнообразни компоненти на операционната система (например ФОР 32 за работа с Фортран), развити и усъвършенствани други. Самият П. Бърнев самостоятелно се занимава с теоретически интересни и особено важен за практиката в този момент въпрос за програмната съвместимост (1973). Да не забравяме, че в страната се експлоатират внесените 15 броя „Минск-32“, в ИЦ на МИ още функционира „Минск-2“, а компютрите от серията ЕС (аналози на революционните IBM/360) започват да завладяват СИБ, включително и България.

Впрочем от казаното става ясно, че цялата тази дейност би следвало да бъде отбелязана и в раздела за организационната такава.

Вероятно първата мащабна научно-приложна дейност, в която е участвал Бърнев, е създаването на първия български компютър – „Витоша“ под ръководството на акад. Любомир Илиев (1961-1963) и в рамките на колектив от десетина млади специалисти, имената на които остават в историята на българската изчислителна техника и информатика.

Не е възможно в рамките на този доклад дори да се изброят многобройните научно-приложни разработки (а в някои случаи и такива с фундаментален характер), изпълнени от колективи под ръководството на Петър Бърнев в рамките на различни договори. Сред тях са пет с МОН – ФНИ (или техни предшественици), посветени на бюротика, когнитивно моделиране, системи с изкуствен интелект за специални цели, както и внедреното „Автоматизирано работно място на програмиста“. В рамките на научното сътрудничество на академиите на страните от СИБ по изчислителна техника (КНВВТ) Бърнев е ръководил българските колективи по разработка на езици за програмиране (1964-1973), целия международен колектив по проекта КАПГ за бази данни в геофизиката (1981-1985-1990). Ръководител е на значителен брой двустранни теми с академични институции от Италия, СССР, Румъния, Монголия, Виетнам, бившите ГДР и Чехословакия с широк спектър от теми.

## **5. Организационна, обществена и популяризаторска дейност**

Почти през целия си професионален живот П. Бърнев заема ръководни позиции в институциите, в които работи, и в други организации и институции.

След като основава секция „Автоматизация на програмирането“ в рамките на Математическия институт на БАН през 1964, той я ръководи до пенсионирането си през 2005, когато тя е вече под името „Информационни изследвания“. През тази секция преминават над 140 сътрудници на различни академични, експертни и технически позиции. Някои от тях завършват кариерата си в секцията, други преминават в други организации, учебни заведения и институти, където малко или повече пренасят творческите импулси и стил на работа от секцията на проф. Бърнев.

Бил е заместник директор на Математическия институт на БАН и на Единния център по математика и механика, непрекъснато е бил член на Научния съвет на Института. Бил е член на редица други научни съвети по линия на ВАК, председател на СНС по информатика и приложна математика (1998-2004) и на СНС по информатика и математическо моделиране (2004-2007).

Един от най-важните приноси на П. Бърнев към българската информатика е организираната от него международна школа по програмиране. Първото ѝ издание е през 1976 г. с участието на около 30 слушатели и лектори от България. Още от следващата година тя става международна с 10 лектори и докладчици от чужбина. В течение на годините школата започва да се фокусира всяка година върху определена важна тема, като едновременно с това се обогатява със съпътстващи семинари и други събития. От 1991 програмният комитет става международен, а от 1992 школата става част от международната конференция ICT&P – Information and Communication Technology and Programming. От началото на 21. век организацията се поема от Красимир Марков, който постепенно развива събитието до комплекс от свързани международни конференции. П. Бърнев обаче продължава да играе важна роля в оформянето на програмата и изданията на събитията. Огромна заслуга на Бърнев е, че благодарение на международния си авторитет, той успява да покани и осигури участието на световни светила – Ф. Л. Бауер, В. Брауер, С. Лавров, А. П. Ершов, И. Поттосин, Вр. Курочкин, Т. Вамош, Л. Лукашевич, Н. Спиратос, С. Спакапиев и др.

Специално внимание следва да се отдели на дългогодишната му активна дейност в Съюза на математиците в България. Учредител е на Българското математическо дружество (БМД) (1971) и избран от тази дата за един от зам.-председателите на дружеството. По-късно (1977) е учредител на Съюза на математиците в България и е избран за един от зам.-председателите, длъжност на която е преизбиран и заема до 1992.

Редовен участник е (обикновено с доклади и като организатор) в първите 21 (1972-1992) Пролетни конференции на БМД/СМБ и председател на програмните комитети на 4-тата ПК (Перник, 1975), 11 ПК (1982, с домакин В. Търново) и 21 ПК (1992, София).

Не по-малко активна, дългосрочна и разнообразна е дейността на Петър Бърнев по линия на Съюза на научните работници. Участник е във II (1976) и III (1982) конгрес на СНРБ. Бил член на Президиума на СНРБ, а на V-я извънреден конгрес на Съюза на учените в България (СУБ – новото название на

СНРБ), както и на VI-я конгрес е избран за член на УС и на изпълнителното бюро на СУБ. През 1990-те години в продължение на няколко години е главен редактор на списанието „Наука“ на СУБ, което, по общо мнение, се отразява много благоприятно на списанието.

Председател е на Физико-математическата секция на СНРБ (1979 – 1990). В това си качество организира ежегодни колоквиуми на секцията в периода 1978-1983.

През 1997 учредява, оглавява и до смъртта си е моторът на една от първите неправителствени организации в България – Асоциацията „Развитие на информационното общество“ – известна като АРИО. Семинарите на АРИО в продължение на дълги години бяха най-важният център на интелектуалния и професионален живот в областта на информационното общество, често поставяйки на свободно обсъждане много важни по-общи въпроси за развитието на обществото и страната. АРИО и днес продължава да е организатор на традиционната национална научна конференция „Образованието в информационното общество“ и издава в сборник рецензираните и изнесени на нея научни доклади. Със сигурност можем да твърдим, че материали, разработени от АРИО във връзка с държавната стратегия за информационното общество през 1998 г., са били взети под внимание от колектива, разработил тази стратегия. Особено трябва да отбележим и месечния бюлетин на АРИО, чийто главен редактор до смъртта си беше П. Бърнев. Благодарение на него бюлетинът беше не само източник на актуална професионална информация, отразяваше подробно споменатите семинари, но и беше свободна трибуна за изказване на становища, мнения и идеи, свързани с информационното общество, но често и в много по-широк аспект. Самият Бърнев беше не само редактор, но и много активен автор на материали с широк тематичен спектър.

От средата на 1970-те до 1986 вкл. П. Бърнев беше представител на България в Технически комитет 2 (Софтуер) на Международната федерация по обработка на информацията – ИФИП.

По много научни и обществени въпроси П. Бърнев имаше собствено мнение, понякога несъвпадащо с общото. Той не го криеше, дори във времена, когато това можеше да бъде твърде опасно. Бърнев например смяташе копирането на компютрите от серията IBM/360 и преди всичко на софтуера (по който всъщност беше специалист) за голяма грешка. Нямаło е как да знае (защото това стана публично известно едва преди няколко години), че Дайкстра, когато по време на една конференция по софтуерни технологии на НАТО в Рим през 1969 научава за въпросните планове на СИБ за копиране, казва, че това е най-голямата победа на американците през Студената война. Като се абстрахираме от хиперболата на Дайкстра – като че ли и двамата излязоха прави. Ние създадохме у нас много добри софтуеристи, дълго време обаче ограничени в не особено творческите методи на IBM за този конкретен тип софтуер. Едва в последните петнайсетина години българските софтуеристи намират радушен прием по света в софтуерни фирми от ранга на Майкрософт и Гугъл и, разбира се, всякакви други. Българските софтуерни фирми работят по

поръчки от чужбина, софтуерни гиганти откриха сериозни клонове тук, а вече имаме и първата българска софтуерна фирма, продадена за стотици милиони в чужбина.

Що се отнася до популяризаторската дейност на П. Бърнев, тя е била винаги неотменна и важна част от професионалния му живот. Документално доказателство за началото ѝ е статията „Същност и проблеми на кибернетиката“, публикувана във вестник „Авангард“, издание на тогавашния ВМЕИ, в броя му от 16.03.1960 г. (Да отбележим, че въпреки известно „размразяване“, по това време Тодор Павлов е председател на БАН, има позиции във висшите ешелони на партийната номенклатура и властта и продължава да счита кибернетиката за „буржоазна лъженаука“). В следващите години, но с особено засилващ се интензитет от 1970-те нататък, проф. Бърнев под различни форми довежда до разнообразни кръгове от обществото (ученици, студенти, ръководители и пр.) важни въпроси и теми, пряко или косвено свързани с информатиката. Не е възможно системно изброяване в рамките на този доклад, затова отново ще припомним някои представителни примери:

- цикъл лекции „Върху понятието алгоритъм“ – пред ученици от МГ-Бургас в периода 10-20.06.1974;
- член на комисия на Седми национален преглед на техническото и научно творчество на младежта. София, 1-5.10.1974;
- „Ще ги наричаме просто магистрите“ – Бърнев в интервю, взето от Еми Барух относно ЕЦНПКММ, в. „Студентска трибуна“, 2.04.1974;
- „Автоматично пресмятане на изрази“, сп. Математика, № 3, 1975, с. 5-8;
- „Математиката в помощ на политиката“, сп. Наука, издание на СУБ, № 1, 1991;
- „Мултимедия и древна европейска култура“, доклад на семинара „Компютърна археология, мултимедия и Интернет представяне“ на Форум на древната култура, в памет на Петър Детев. Пловдив, Дом на учените, 9-11.10.2000.

## **Заклучение**

За своята всестранна и активна дейност проф. Петър Бърнев получава много награди – 6 държавни ордена и медала, сред които два са „Кирил и Методий“ 1. степен, множество почетни знаци и грамоти от министерства, държавни комитети и от БАН, както и медал за значителен принос в съвременната наука от Института по математика и информатика на БАН през 2008 година.

Ценността на тези награди е несъмнена.

Сигурни сме обаче, че проф. Бърнев винаги е държал повече на високата оценка, признание и дълбоко уважение на цялата информатична колегия през целия му професионален живот, които днес отново му засвидетелстваме. Нямаме никакви съмнения, че огромното му дело не може да бъде забравено, а резултатите от него са и ще бъдат пред всички.

## Софтуерни технологии в ИМИ

**Румен Радев, Аврам Ескенази,  
Нели Манева, Весела Ангелова**

Институт по математика и информатика при БАН, София  
[rtradev@math.bas.bg](mailto:rtradev@math.bas.bg), [eskenazi@math.bas.bg](mailto:eskenazi@math.bas.bg), [neman@math.bas.bg](mailto:neman@math.bas.bg), [vaa@math.bas.bg](mailto:vaa@math.bas.bg)

**Резюме:** Настоящият доклад разглежда дейността и постигнатите резултати в обширната област на софтуерните технологии и близки до нея в продължение на няколко десетилетия. Важно внимание е отделено на научно-изследователската дейност като база, от която са получени интересни научни резултати, разнообразни практически и експериментални разработки. Основните направления на научни изследвания включват софтуерните технологии, езиките за програмиране и специфициране, базите данни и информационните системи, прилагането на информационните технологии в обучението. Научно-изследователската дейност намира своето естествено продължение в сферата на преподавателската и образователна дейност, изготвянето на учебници, учебни помагала и друга литература. Отразено е и участието в редица български и международни проекти.

### 1. История

През 1961 г. е създаден Изчислителният център към Математическия институт на БАН. Малко по-късно, през 1963 г. е създадена група „**Теория на крайните автомати**“ с ръководител доц. Димитър М. Добрев. Това се счита и за рождена дата на изследванията в областта на софтуера, тъй като оттогава, въпреки значителната динамика и в състава, и в разработваните теми, личи ясна приемственост и устойчивост на интересите, разработките, резултатите и сътрудниците. Тъй като става ясно, че тематиката, която тази група развива, се разширява твърде бързо, през 1966 г. е създадена секция „**Теоретични въпроси на кибернетиката**“ с ръководител доц. Д. Добрев. От самото начало сътрудниците на секцията се занимават не само с основната си задача – научно-изследователска работа, но и с приложение на научните резултати в практиката, както и с педагогическа дейност (четене на лекции, водене на упражнения, писане и издаване на учебници и учебни помагала и пр.). Още тогава доц. Добрев ръководи и т.нар. „Група 10“, която е съставена от сътрудници на секции на Математическия институт (МИ) и има за задача обслужването, поддържането и развитието на изчислителната техника в МИ – преди всичко МИНСК-2 – първата внесена електронно-изчислителна машина в България.

По-нататък през годините секцията следва развитието на Института, към който принадлежи. През 1970 г. е създаден ЕЦНПКММ (Единен център за наука и подготовка на кадри по математика и механика). Като негово звено, на

базата на секция „Теоретични въпроси на кибернетиката“, е създаден сектор „Основи на кибернетиката и теория на управлението“ (ОКТУ) отново с ръководител доц. Д. Добрев.

През 1984 г. приемник на сектор ОКТУ, който се разделя на 3 секции, става секция „Софтуерни технологии“, тогава създадена под името „Автоматизация на софтуерното производство“. За ръководител на секцията е избран ст.н.с. Аврам Ескенази, който я ръководи до 2012. След това ръководител е проф. Нели Манева до вливането на секцията през 2014 г. в новата „Софтуерни технологии и информационни системи“ (СофТИС).

## **2. Научно-изследователска дейност**

Научно-изследователската дейност на ИМИ–БАН в направление „Софтуерни технологии“ е неразривно свързана както с разработването на експериментален софтуер, така и с активна образователна и популяризаторска дейност в следните области:

- софтуерни технологии;
- езици за програмиране и специфициране;
- бази от данни и информационни системи;
- информационни технологии в обучението.

През различните периоди от време тези области са се развивали с различна интензивност.

Като пример ще посочим конкретните изследвания, проведени през последните няколко години и съответно постигнатите резултати.

**Ползваемост и ергономика на софтуера.** Основните изследвания и резултати са по:

- създаване на базов модел на ергономичното качество за обект „потребителски интерфейс“;
- сравнителен анализ за прилагане на ергономичните принципи в различни ситуации (при разработване, поръчване, закупуване и съпровождане на софтуер);
- анализ на целесъобразността при прилагане на ергономичния принцип, че ефикасността е възможна само ако изключва принуда, злоупотреба и пресищане.

Специално внимание е отделено на важния аспект как да бъде стимулиран потребителят, за да използва активно интерактивни средства.

**Нов подход към оценката на различни аспекти на реализация на софтуерни проекти от конкретен тип (ERP).** Изследва се проблемът за оценка на разходите за разработване на ERP системи. Анализирани са нови аспекти, свързани преди всичко с повторната използваемост, както и с функционалната мощност на продуктите. Теоретичните резултати се експериментират върху безспорния лидер за този тип приложения – SAP/3.

**Разработване на учебно съдържание съгласно СС 2005 и реализацията на представителна част чрез съвременни комуникационни и софтуерни**

**средства.** Поради бързото развитие на ИКТ (информационните и комуникационни технологии) преподаването на информатика във висшите училища е динамичен процес, свързан както със структурата на учебните планове, така и с актуализирането на учебното съдържание. В тази връзка, след проведен анализ на изискванията на Computing Curricula 2005, са разработени паспорти на различни курсове по приложна информатика. Реализирани са съответни курсове като част от учебния план на студентите от специалност „Информатика“ в Нов български университет (НБУ). Разработени са учебни съдържания и електронни учебни ресурси за курсове по бази от данни, информационни системи, информационно моделиране и др. Сътрудници на секцията са дали основен принос по оформяне на учебни планове и програми по софтуерни технологии за докторанти в рамките на европейски проект.

Едновременно с това е предложен и нов апарат за един аспект на реализацията на това учебно съдържание, когато това става компютризирано. Разработен е подход за моделиране на адаптивен сценарий за учене в система за електронно обучение. Определени са факторите, влияещи върху избора на технологични средства и сценарии за учене – стил на учене, постигнати цели на обучение. Разработени са концептуален и физически модел на пътя за учене и алгоритми за проектирането и реализацията им. Разработен е формален метод за верификация на моделите и са формулирани основни изисквания към системата за електронно обучение, в която могат да се реализират адаптивните възможности на модела по отношение на стила на учене на обучаемите и техните постижения по отношение на поставените учебни цели.

**Изследвания по езици за програмиране.** Извършени са сравнително-аналитични изследвания на множество езици за програмиране. Предложени са оригинални програмни конструкции – обобщения на широко използваните – за пресмятане на изрази, за описване на циклични действия и др. Разработени са основи на общ подход за изразяване на структурни отношения в програми, универсален по отношение на различните модели (парадигми) на програмиране. Разработен е език за аотиране на програми. Разработени са и са реализирани интерпретатори на два приложни езика: за създаване на планиметрични чертежи и за онагледяване на геометрични данни. Част от резултатите са представени под формата на уникален по съдържанието си уебсайт, представящ множество езици за програмиране.

**Изследване на нагласите за прилагането на съвременни технологии при разработването на софтуер от малки и средни предприятия.** Въпреки развитието на разнообразни технологии за разработване на софтуер, не може да се твърди, че съществува стандартен метод за изграждане на информационни системи. Затова интересна задача е да се изследва приложимостта на някои от технологиите за разработването на информационни системи, които поддържат основни организационни дейности. В тази връзка е проведено проучване на необходимостта от подходящо професионално обучение в софтуерните фирми за придобиване на конкретни умения по UML. Изследвани са нагласите за прилагането на съвременни технологии за разработването на софтуер от малки и средни предприятия. Разработени са съответни стратегии за професионално

обучение. Освен специфични резултати, свързани с определени продукти, са уточнени предпочитанията на служителите към методите и формите на продължаващото им обучение.

**Приложение на класификационни методи към задача, свързана с ползването на е-правителството.** В рамките на европейския проект ЕЛОСТ, в който секцията представя България, са направени приложения на разработени по-рано в секцията софтуерни средства за класификация и окачествяване на програмни продукти на основата на класификационни методи – разработени или усъвършенствани в секцията. В резултат са направени съдържателни заключения относно разнообразни характеристики, вкл. прилики и разлики на ползването на е-правителството от групи с нисък социален или икономически статус в шест различни страни.

**Различни аспекти от приложението на гъвкавите методологии за производство на софтуер** – въпроси на управлението на качеството, съчетаване с други методологии (PMBOK, Lean, Kanban, Scrum), проблеми на приложението им в практиката.

### 3. Педагогическа дейност

Освен научна дейност, членовете на секцията ежегодно са се занимавали и продължават да се занимават с педагогическа дейност (четене на лекции, водене на упражнения, четене на специализирани курсове за нуждите на други организации, ръководство на млади научни кадри).

За периода на съществуването на секцията в различните ѝ форми своите дипломни работи и дисертации са защитили над 400 **дипломанти** и около 45 **докторанти**. Всяка година чрез разнообразни курсове са били обучени много специалисти, които по-късно са се реализирали успешно, както в България, така и в чужбина.

Членове на секцията са чели **лекции** или водили упражнения в различни градове и учебни заведения. Сред градовете са София, Бургас, Пловдив, Варна, Благоевград, Габрово, Велико Търново, Шумен, а сред университетите – Софийски (различни факултети – математика и информатика, биологически, журналистика, философски), Пловдивски, Шуменски, Бургаски свободен, Великотърновски, Югозападен (Благоевград), Американски (Благоевград), ВСУ (Варна), ИУ (Варна), НБУ (София), УНСС (София), УАСГ (София), УНИБИТ (София).

Сред четените **курсове** са: Теоретични основи на информатиката, Дискретна математика, Теория на крайните автомати, Основи на програмирането, Понятия и структури в езиките за програмиране, Съвременен функционално програмиране, Изчислителна геометрия, Разпознаване на образи, Структури от данни и организация на данните, Бази от данни, Информационни системи, Разпределени бази данни и системи, Системи за управление и обработка на данни, Автоматизация на обучението, Софтуерни технологии, Оценка и осигуряване на качеството на софтуера, Компресиране на данни, Електронна търговия, Софтуерен бизнес, Информационно моделиране.

Немалка част от тези курсове са били разработени и четени за първи път в страната от сътрудници на секцията.

Четени са курсове в областта на информатиката и пред специалисти или учени от други области.

В продължение на години сътрудници на секцията активно работят с изявени ученици в областта на състезателното програмиране: ръководят школа към Софийската математическа гимназия, участват в подготовката на националния отбор по информатика, изнасят лекции в Националната лагер-школа по програмиране и в Лятната изследователска школа към Ученическия институт при ИМИ и др. Освен пряко с ученици от страната, работи се и с техни преподаватели. Създадена е, регулярно се обновява и се разпространява сред ученици и учители сбирка с помощни материали: книги, статии, програми, задачи и др.

#### **Учебници, учебни помагала и друга литература**

През 1979 г. колектив в състав Д. М. Добрев, Р. Киркова, Й. Денев, А. Ескенази и Й. Швертнер подготви и издаде първия университетски учебник по информатика. Членове на секцията са участвали в създаването и на първите сборници със задачи и решени примери по програмиране (1976 и 1981 – разширен и допълнен).

Основен съавтор на първия учебник по дискретна математика е Й. Денев – той е издаден през 1984 г. в съавторство с Р. Павлов и Я. Деметрович.

Сътрудниците на секцията А. Ескенази и Н. Манева са написали и издали и първия у нас учебник по софтуерни технологии (2001 и 2006 – разширен и допълнен), а А. Ескенази – и едни от първите учебници по електронна търговия (2002, 2005, 2006).

Сътрудници на секцията са превели и издали няколко книги в областта на информатиката и изчислителната техника.

#### **4. Връзка с практиката**

Паралелно с провеждане на научните изследвания винаги се е търсела връзка между теорията и практиката от позициите на най-актуалната и най-необходимата за дадения момент област. Като най-значителни примери могат да се споменат:

- 1962-1969: разработване на основни принципи по архитектура и организация на съвременните изчислителни машини, включително и чрез моделиране на една машина върху друга;
- 1968-1985: създаване на множество специализирани информационни системи за различни потребители, както в рамките на БАН, така и за други организации (като примери могат да се споменат системите СИС за информационно обслужване на обучението в едно учебно заведение, системите ФИКС за автоматизиране на дейностите по заплащане на труда, серията от системи КИСЕС по автоматизиране на дейностите за подготовка и провеждане на научни мероприятия, серията от системи

АСУЧПИ за автоматизация на библиотечните дейности, серията ТЕСТ за нуждите на автоматизираната проверка на знанията, вкл. за нуждите на КАТ);

- 1971-1980: създаване на базово програмно осигуряване за ефективно използване на най-разпространената в дадения момент изчислителна техника в България (тук спадат машината МИНСК-32, серията СМ-4, машините от серията ЕС);
- 1968-1978: разработване, експериментирание и внедряване на разнотипни изчислителни устройства (като пример може да се посочат периферното устройство контролер ИЗОТ 5501.Е за включване на магнитни дискове към машината МИНСК-32, както и специализираното устройство за обучение и изпитване ЕМИ-25);
- 1971-1983: създаване на пакети с базови програмни средства за автоматизиране на процесите по създаване на информационни системи с общо предназначение на системи за управление на бази от данни (като примери могат да се споменат пакетите УГИС и БИС32 за машината МИНСК-32, БИСЕС и ДАКОМС за машините от серията ЕС);
- 1977-1983: обединяване на изчислителните мощности в големи изчислителни комплекси с цел тяхната ефективна експлоатация;
- 1984-1987: създаване по поръчка на Министерство на образованието на пилотен софтуер за автоматизация на специализирани дейности по обучение и управление на обучението;
- 1993-1994: създаване на специализиран софтуер за ТРЗ по поръчка на VallInformatique (Франция);
- 1995-1996: създаване по поръчка на Dornier (Германия) на софтуерна система за проектиране и разработка на софтуер (системата AGRIS);
- 1993-1995 и 1996-2001: създаване по поръчка на EADS (Германия) на софтуерна система за проектиране и разработка на учебен софтуер (система STORYIST);
- 2009-2011: проектиране, разработване и внедряване на вътрешна информационна система на ВАК (закрита); създаване и поддържане на интернет-базиран Регистър на научните степени и звания в Република България.

Почти всички от изброените програмни пакети имат редица внедрявания в различни научни или стопански организации, както в България, така и в чужбина, като някои са достигнали до повече от 800 внедрявания.

От създаването на Фонда за научни изследвания (имената са се сменяли) през 1987 г. секцията почти непрекъснато е изпълнявала договори с него, като през 1998 изпълняваният от секцията проект „Изследване на софтуерни обекти и процеси и приложение към софтуерния маркетинг“ (И-406) 1994-1997 получи грамотата на МОН за постигнати съществени резултати.

## **5. Международна дейност**

### **Двустранно сътрудничество**

Както е известно, всички звена на Института по математика и информатика са изпълнявали и изпълняват широка програма от теми и задачи по линия на двустранното научно сътрудничество. Сред по-значителните, в които активно са работили членове на секцията, са: Създаване на център за колективно ползване (с ИЦ при Сибирското отделение на АН СССР); Изчислителни мрежи (с Института по електроника и изчислителна техника при АН на Латвийската ССР); Математическа кибернетика и обработка на данни (с Изследователския институт по изчислителна техника и автоматизация при Унгарската АН); Автоматизиране на дейностите по информационно обслужване и управление на академични организации (с Централен институт по изчислителна техника и с Изчислителен център при АН ГДР); Архитектурни принципи на СУБД (със Секцията по обработка на информацията при ТУ Дрезден); Теоретична информатика и информационни системи (с Института по информатика при Университета в Хамбург).

Съществено и полезно участие секцията е имала при разработването на тематиката „Нови поколения изчислителни машини“ (в частта „софтуер“) през периода 1985-1990.

От 1997 година е установено трайно и ползотворно сътрудничество с Университета в Тел Авив, с който в момента се разработва пети съвместен проект. Разработваните досега проекти са следните:

- Оценяване на информационните системи и на технологиите за разработването им;
- Дистанционно разработване на информационни системи и изследване на тяхната използваемост;
- Разработване на информационни системи с използване на авангардни технологии;
- Оценка и избор на сложни софтуерни системи;
- Разработване и усъвършенстване на управленческите аспекти на гъвкавите методологии в софтуерното производство.

### **Многостранно сътрудничество**

По линия на многостранното научно сътрудничество трябва да се отбележи участието на членове на секцията и особено на доц. Д. М. Добрев в дейностите по линия на Проблемната комисия „Научни въпроси на изчислителната техника“ (КНВВТ) между Академиите на науките на бившите социалистически страни от нейното създаване през 1962 год. до 1991 год. Секцията активно е работила както в научната и организационна работа на КНВВТ като цяло, така и в много от нейните работни групи (РГ): РГ-2 по „Разпознаване на образи“; РГ-3 „Изчислителни мрежи“; РГ-5 „Системи за автоматизация на обучението“; РГ-11 „Системи за управление на бази данни и информационни системи“ (ръководител доц. Д. Добрев); РГ-19 „Системи за управление на разпределени бази данни“ (ръководители доц. Д. Добрев и ст.н.с. Р. Киркова); РГ-25

„Интеграция на информационни системи“ (ръководител ст.н.с. Р. Киркова); РГ-26 „Системи за вземане на решения“.

### **Международни организации**

Друго основно направление е участието на сътрудници на секцията в Международната федерация по обработка на информацията (ИФИП). В рамките на ИФИП през годините се създаваха техническите комитети (ТС), в някои от които работеха и работят активно членове на секцията: ТС-2 „Софтуер“ (А. Ескенази); ТС-6 „Комуникации“ (К. Боянов); ТС-8 „Информационни системи“ (Д. Добрев). Следва да се отбележи, че с наше съдействие и в България беше учреден ТС-9 „Общество и компютри“.

С активното участие на секцията бяха организирани през годините и няколко научни и организационни мероприятия на ИФИП и неговите комитети, последното от които през 2007 година – редовното годишно заседание на ТС-2 по теория и практика на софтуера.

Сътрудници на секцията участват в българския клон на АСМ от основаването му през 1990, както и в самата АСМ. Те са участвали както в научните мероприятия под егидата на АСМ, така и в тяхната организация, вкл. и във високо оценената от АСМ Международна лятна школа по мултимедия за млади специалисти, Албена, 1997.

### **Международни мероприятия**

Една от основните дейности на Института по математика и информатика е представянето на научната продукция в различни научни мероприятия у нас и в чужбина. В тази насока членове на секцията са участвали както в организирането (включително и автоматизирането) на научни конференции, симпозиуми, конгреси и пр., така и като членове на програмните им комитети, като рецензенти на представяни за докладване материали, а също и в тяхната научна програма с научни съобщения или пленарни доклади, които са били пряко свързани с развиваната научна тематика. Списъкът на тези мероприятия е прекалено дълъг, за да бъде даден тук, още по-голям е списъкът от участията на сътрудниците на секцията в тях.

### **Международни проекти**

Особено активно в последните години е участието на секцията в международни проекти, финансирани предимно от ЕС. Това участие обаче, е започнало още от 1986. Ето списък на международните проекти с участието на секцията като главен изпълнител и координатор или като изпълнител:

- Контракт с ИПИ АН СССР 38-80/82578-10/1.1.9: 1986-1988 по разработка на иновативен софтуер;
- Подпроект в рамките на проекта BUL/003 UNESCO, информационни технологии в образованието: 1986-1990;
- Международен европейски проект Multiversum „Мултимедията като епистемологичен ресурс: към многопосочно, вместо към линейно обучение“ в рамките на програмата Socrates: 2000-2003;

- Международен проект ECIST-2001-34488 eXPERT по екстремно програмиране: 2002-2003;
- ЕС проект Erasmus 114046-CP-1-2004-1-BGETNDEC по хармонизиране на обучението на докторанти: 2005-2007;
- ЕС проект по 6. рамкова програма Elost 027287 – е-правление за групи с ниско социално и икономическо положение: 2006-2008;
- ЕС проект TRICE 142399-LLP-1-2008-1-BG-ERASMUS-ENW – Teaching, Research, Innovation in Computing Education: 2008-2010.

## **6. Експертни дейности**

Секцията като цяло, а също и отделни нейни сътрудници в качеството си на експерти, са участвали в разнообразни дейности с висока обществена значимост. Следват някои примери за това:

Разработване на Националната стратегия за обучението по ИКТ в средното училище – по поръчка на Министеръа на МОН: 1997-1998.

Разработване на Национална стратегия за информационното общество – след конкурс и по поръчка на Правителството на Република България: 1998.

Разработване на Национална стратегия за високите технологии – по поръчка на Правителството на Република България: 1998-1999.

## **Списък на сътрудници на секцията в различни периоди от време**

В този списък са включени имената на сътрудници, които са били членове на секцията или на някое от предшестващите я звена (по азбучен ред):

Аврам Ескенази, Александър Ахегукян, Александър Людсканов, Анри Овадиев, Антон Петков, Атанас Терзиев, Бойко Банчев, Боянка Главчева, Валентин Събев, Валя Петрова, Весела Ангелова, Венцислав Вълев, Владимир Гюлев, Владимир Тончев, Галя Младенова, Георги Тупаров, Данаил Дочев, Димитър Димитров, Димитър М. Добрев, Елена Паскалева, Емилия Живкова, Заре Зарев, Иван Мицев, Йордан Денев, Йосиф Швертнер, Кирил Боянов, Красимир Манев, Людмил Манасиев, Мая Данева, Мая Ингилийска, Моника Филипова, Начо Василев, Нели Манева, Николай Тодоров, Петър Първанов, Пламен Думков, Радостина Терзиева, Росен Рашев, Румен Радев, Румяна Киркова, Румяна Лесева, Снежана Костова, Снежана Молерова, Снежина Карагъзова, Соня Лазарова, Станчо Станков, Стоян Капралов, Татяна Шамрай, Тодор Брънзов, Тодор Панделиев, Христина Зашева, Юлиана Василева, Юлиа Джамбазка, Юлиана Пенева.

## Развитие и приложение на системи за компютърна алгебра в ИМИ-БАН

Маргарита Спиридонова<sup>1</sup>, Александър Геров<sup>1</sup>,  
Мария Нишева<sup>1&2</sup>

1. Институт по математика и информатика при БАН, София  
2. ФМИ на СУ „Св. Кл. Охридски“, София  
[mspirid@math.bas.bg](mailto:mspirid@math.bas.bg), [gerov@math.bas.bg](mailto:gerov@math.bas.bg), [marian@fmi.uni-sofia.bg](mailto:marian@fmi.uni-sofia.bg)

**Резюме.** Представени са основни насоки на научните изследвания в Института по математика и информатика (ИМИ) на БАН, свързани с развитието и приложението на т. нар. системи за компютърна алгебра (СКА). Разгледани са накратко някои общи характеристики на тези системи и множество конкретни специализирани СКА, разработени в ИМИ.

Началото на споменатите изследвания бе поставено в секция „Математическо осигуряване“ на ИМИ, ръководена от проф. Петър Бърнев. След създаването на секция „Изкуствен интелект“ с ръководител проф. Валентин Томов, разглежданата тематика стана предмет на изследвания в секцията. Тук следва да подчертаем специалната роля на проф. Валентин Томов за развитието на това направление в ИМИ. Нека отбележим, че през 2015 г. той би навършил 80 години.

Изследванията и приложните разработки, представени в текста, са осъществени предимно в секциите „Математическо осигуряване“ и „Изкуствен интелект“ на ИМИ при БАН. Споменати са и други секции на ИМИ, аспекти на изследванията в които са свързани със СКА.

**Ключови думи:** символни и алгебрични преобразования, системи за компютърна алгебра, приложения.

**ACM 1998 Classification Keywords:** I.1 Symbolic and Algebraic Manipulation, I.1.3: Languages and Systems: Special-purpose algebraic systems, I.1.4 Applications.

### 1. Увод

В първите години на нашите изследвания по разглежданата тематика използвахме термини като „аналитични преобразования“ и „символни пресмятания“, за да характеризираме възможностите на системите REDUCE 2, SAC-2 и други подобни. Едва по-късно възприехме утвърдили се вече термин „системи за компютърна алгебра“. В настоящия текст се използва предимно този термин (не обсъждаме терминологичните нюанси и периодите на тяхното използване, включително в наши публикации).

За разлика от системите за числени пресмятания, в СКА се използва символно представяне както на входната информация, така и на извеждания резултат, дори да съдържа само числа. Има специфика и в алгоритмите за обработка на изразите, подадени на входа от потребителя (с цел преобразуване, опростяване или извършване на определени математически операции). Тези алгоритми също са предмет на компютърната алгебра. Разглеждана като област от сечението на информатиката и математиката, компютърната алгебра включва следните основни направления: алгоритми, системи и приложения. Нашите изследвания, които са представени накратко в този текст, касаят предимно системите за компютърна алгебра, тяхното развитие и приложение.

Съвременните системи за компютърна алгебра са интегрирани програмни среди за символни и числени пресмятания, за визуализация на функции и данни, за създаване на програмни продукти, за решаване на задачи с различна сложност в почти всички области на науката и техниката. Тяхното използване в процеса на обучение в университетите и в средните училища е в основата на ефективното приложение на информационните технологии в образованието.

СКА се създават от около 6 десетилетия и за този период са разработени няколко десетки такива системи – с общо предназначение (с богати възможности за приложение в различни предметни области) и специализирани (ориентирани към приложение в определена област). У нас са добре познати например СКА Maple, Matlab, Mathematica, Reduce, Derive, Macsyma и др. С някои от тези системи са свързани изследванията, разгледани по-долу.

## **2. Началото**

Началото на нашата изследователска дейност, касаеща СКА, беше поставено през 1979-80 г. в секция „Математическо осигуряване“ (МО) на ИМИ с проучване и експериментиране на системите REDUCE 2 и SAC-2, като особено значение за това имаха контактите ни с ОИЯИ в Дубна, и Италианския център за научни изследвания, тъй като чрез тях получихме достъп до тези системи и се запознахме с опита на колегите, които са работили с тях. Много полезни научни контакти имахме и с колеги от Техническия университет в Дрезден и Унгарската Академия на науките.

През 1980 г. в секция МО беше създадена изследователска група с име „Искусствен интелект в математическото осигуряване“ (ИИМО) и ръководител проф. Валентин Томов. Работата в групата беше свързана предимно със системите REDUCE 2 и SAC-2 и формиралите се вече в групата идеи за разработки с използване на тези системи.

През 1985 г. групата ИИМО прерасна в секция „Искусствен интелект“ с ръководител проф. Валентин Томов. След преждевременната кончина на проф. Томов през 1995 г., ръководството на секцията беше поето от доц. Александър Геров. През 2010 г. няколко сътрудници на секцията се пенсионираха и в края на годината секцията прекрати дейността си. Някои изследвания, започнати в секцията обаче, бяха продължени и след това (част от тях – в рамките на секция „Информационно моделиране“).

През цялото съществуване на секция „Изкуствен интелект“ на ИМИ в нея се извършваше активна научно-изследователска и научно-приложна дейност. В тематиката ѝ неизменно присъстваше компютърната алгебра.

### **3. Изследвания и разработки**

По време на формирането на групата ИИМО развитието на изчислителната техника и информатиката беше създало предпоставки за широко използване на компютрите в научните изследвания, при решаване на научно-технически задачи и в управлението. Проучвайки възможностите на СКА, ние популяризирахме тези възможности [1–4] и търсехме подходящи области за тяхното приложение. Една такава възможност беше използване на СКА REDUCE за решаване на задача от областта на сеизмичната механика с колеги от Лабораторията по сеизмична механика при БАН [5]. Тази задача породидеята за реализация на Лапласово преобразование в REDUCE. Беше разработен съответен програмен пакет, включен по-късно в библиотеката на системата [6, 7].

Приложенията на системата REDUCE в задачи от механиката по-късно бяха разширени [8–10]. Интересни приложения бяха направени и в задачи от областта на висшата геодезия [11–13], беше разработен и специализиран програмен пакет с някои нестандартни реализационни решения [14]. Участваха специалисти от Лабораторията по висша геодезия при БАН и катедра Геодезия на Университета по архитектура, строителство и геодезия (УАСГ) в София.

Полезни програмни пакети, реализирани със системата REDUCE, бяха свързани с операции над степенни редове и с изследване на функции [15–17]. Заслужава отбелязване и подходът за организация на работата с формули, реализиран програмно на базата на REDUCE [18–20].

Паралелно със споменатите разработки в групата, под ръководството на В. Томов бяха изградени концепции [21–24] за:

- аналитично-изчислителни системи (АИС) – с акцент върху ефективното реализиране както на символните и числените пресмятания, така и на „прехода“ от единия към другия тип пресмятания;
- интелектуализирани аналитично-изчислителни системи (ИАИС) – в които т. нар. интелектуализация включва възможност за потребителя да формулира задачата си на език, близък до естествения, системата да анализира тази формулировка и да планира необходимите действия за решаване на задачата.

На базата на проучените възможности на системата SAC-2 и на различни изследователски подходи при използването на СКА започна програмната реализация на две специализирани СКА – за операции от линейната алгебра [25–28] и за работа с непрекъснати дроби [30, 31]. В тези две системи бяха реализирани елементи на споменатите концепции. Особено следва да се отбележи фактът, че и двете реализации бяха оформени като докторски дисертации [29, 32].

Изследванията, свързани с концепциите за АИС и ИАИС продължиха. След преминаването на М. Нишева на работа във ФМИ на СУ „Св. Климент Охридски“, под нейно ръководство беше създаден прототип на интелигентна (основана на знания) система за компютърна алгебра [33] и бяха изследвани възможностите за приложение на такъв тип системи в обучението по алгебра в средното училище [34].

Още две специализирани СКА бяха разработени с използване на SAC-2 – за пресмятане на неопределени интегрални от дробно-рационални функции [35] и за факторизация на полиноми [36].

По-късно, съвместно с колеги от INRIA Sophia Antipolis, Франция, беше реализирано символно извеждане на така наречените модифицирани (частни диференциални) уравнения [37, 38] с използване на СКА MAPLE, с цел автоматизирано прилагане на т.нар. подход на модифицираните уравнения при числено решаване на някои класове частни диференциални уравнения. Паралелно с това бяха проучени възможностите на СКА MAPLE. С използване на тази система беше реализирано и едно преобразование от изпъкнали анализ (по идея на акад. П. Кендеров) – преобразованието на Лъожандър [39] (преди това то беше реализирано с помощта на REDUCE [40]).

С използване на едно от копията на системата *Mathematica 2*, получена в ИМИ в рамките на международен проект, направихме някои експериментални приложения, в това число с теоремата на Понселе от проективната геометрия (пак по идея на акад. П. Кендеров).

По-късно с помощта на следващи версии на системата бяха реализирани програмни средства, позволяващи приложение на подхода на операционното смятане за решаване на задачи с начални и гранични условия за някои типове линейни обикновени и частни диференциални уравнения [41–45]. Идеята за тези изследвания и основните теоретични постановки принадлежат на чл.-кор. Иван Димовски. Разработените алгоритми и програмни средства станаха основа за докторска дисертация [46]. Изследванията, свързани с различни приложения на подхода на операционното смятане, бяха продължени [47–49]. Някои резултати са отразени в публикации от последните години [50–52].

В този период беше експериментирано и приложение на СКА *Mathematica* в задачи, свързани с изследване на телекомуникационни системи [56].

Накрая, макар и не на последно място по важност, нека кажем и това, че с някои наши доклади на семинари и конференции (напр. [53–55]) и осъществени контакти с преподаватели се постаряхме да покажем поне част от възможностите на СКА, подходящи за приложение при преподаване и изучаване на различни математически дисциплини.

Част от представените тук изследвания бяха осъществени в рамките на следните **проекти, финансирани от Фонд „Научни изследвания“**:

- И-220 „Програмни системи за компютърна алгебра, ориентирани към научно-изследователски и инженерни задачи“ (1992-1994);
- И-520 „Разработване и приложение на програмни средства за компютърна алгебра“ (1995-1998);

- V Рп-И-5 „Представяне, обработка и обмен на знания с използване на системи за компютърна алгебра“ (1999-2000);
- И-1002 „Програмни средства за компютърна алгебра, ориентирани към обучението по математика“ (2001-2004).

Важна особеност на всички наши изследвания, свързани с приложни разработки, беше участието (на определени етапи от работата) на специалисти в областите, от които са разглежданите задачи.

В голяма част от споменатите по-горе разработки са участвали аспиранти на съответните секции и дипломанти от ФМИ на СУ „Св. Климент Охридски“.

#### **4. Други дейности, съпътстващи представените изследвания**

По инициатива на проф. В. Томов през 1992 г. беше организиран **Национален семинар по СКА**, в който участваха учени от академични институции в цялата страна. Той се провеждаше един път месечно (без летните месеци) в периода 1992-96 г. и беше много полезен с възможността за получаване на актуална информация за развитието на СКА и за изследвания в тази област у нас и в чужбина.

**Участието ни в наши и международни научни конференции и гостуването у нас на учени със световна известност** съдействаше както за представяне на нашите резултати, така и за разширяване на научните ни контакти.

Подобно беше значението и на **участието ни в международни работни групи** (по изкуствен интелект и компютърна алгебра) и **в съвместни научно-изследователски теми с чуждестранни институти** (с институти на АН на Русия и Унгария, с ОИЯИ (Дубна), с национални научни центрове на Италия, Франция и др.).

Секция „Изкуствен интелект“ беше **организатор и съорганизатор на редица международни съвещания и конференции** по тематика, свързана с нашите изследвания. Особено внимание заслужава организирането на международните конференции по приложение на компютърната алгебра АСА 2006 [<http://www.math.bas.bg/artint/mspirid/ACA2006/>] и АСА 2012 [<http://www.math.bas.bg/ACA2012/>].

Важно е да отбележим, че изследвания, свързани с компютърната алгебра, са провеждани и в други секции на ИМИ – секция „Алгебра“ (сега „Алгебра и логика“), „Биоматематика“ (сега част от секция „Математическо моделиране и числен анализ“), „Математически основи на информатиката“ и др. СКА се използват и сега в редица секции на ИМИ, също така и в Софийския университет „Св. Климент Охридски“, Техническите университети в София и Габрово, УАСГ, Великотърновския университет „Св.св. Кирил и Методий“, Пловдивския университет „Паисий Хилендарски“, Нов български университет, Бургаския свободен университет и още в редица висши и средни училища у нас. С някои от тях имаме регулярни контакти.

*Лекции по системи за компютърна алгебра* са четени в СУ „Св. Климент Охридски“ и други измежду изброените университети от сътрудници на секция „Искусствен интелект“ и преподаватели от тези университети.

Представените изследвания и разработки, свързани със СКА, отразяват в значителна степен (без да изчерпват) резултатите от нашата научно-изследователска и приложна дейност в тази област. Те са отразени в над 100 **публикации**. Голяма част от тях са включени в списъка на литературните източници, цитирани в настоящия доклад.

## **5. Заключение**

СКА все повече се използват като инструмент в изследователската работа на много учени, инженери, преподаватели и студенти. В същото време тези системи се развиват и обогатяват. Разработват се нови алгоритми, развитието на информационните технологии неминуемо оказва своето влияние. В редица случаи приложенията им са „нестандартни“, както се вижда например в програмите на някои конференции. Именно по докладите, съдържащи се в програмите на различни международни форуми, по много публикации от последните години, по новите версии на най-известните големи СКА и т.н. се вижда както тяхното алгоритмично обогатяване, така и развитието им като програмни системи и все по-широкото им приложение. С радост можем да кажем, че сме участвали и продължаваме да участваме в този процес.

## **Литература**

1. **Томов В., М. Спиридонова, А. Геров**, Възможности на съвременните системи за аналитични преобразования, Сп. Автоматизирани системи за управление, № 1, **1983**, 31-38.
2. **Томов В., М. Спиридонова, А. Геров, С. Върбанов**, Аналитично-изчислителни системи, Сб. Доклади на III Международна школа „Автоматизация и научно приборостроене“, Варна, 24-31 октомври, **1984**, 245-252.
3. **Томов В., М. Спиридонова, А. Геров**, Развитие и приложение систем аналитических преобразований, Сб. докл. Междунар. конференции „Автоматизация научных исследований“, Пловдив, 15-20 октябрия, **1984**, 58-64.
4. **Томов В., А. Геров, М. Спиридонова**, Системи за аналитични преобразования (доклад по покана), Математика и математическо образование, Сб. доклади на 14-та Пролетна конференция на СМБ, Сл. бряг, 6-9 април, 1985, София, **1985**, 92-103.
5. **Brankov G., M. Spiridonova, K. Ishtev, Ph. Philipov**, Modelling of Building Structure Behaviour under Seismic Loading with Use of an Algebraic Manipulation system, Proc. of the 7-th European Conference on Earthquake Engineering, Athens, **1982**, 585-592.
6. **Спиридонова М., Хр. Казасов**, Аналитично извършване на Лапласово преобразование с помощта на ЕИМ, Сп. Автоматизирани системи за управление, № 3, **1983**, 90-94.
7. **Kazasov C.** Laplace Transformations in REDUCE 3, Proc. of Eurocal '87, Lecture Notes in Comp. Sci. 378, Springer-Verlag, **1987**, 132-133.
8. **Спиридонова М., Л. Хаджиков, П. Динева**, Исследование реакции динамической модели со сосредоточенными массами при помощи системы аналитических преобразований REDUCE 2, Сб. докладов конференции по системам для автоматизированного информационного обслуживания, Варна, 3-8 октябрия, **1983**.

9. **Спиридонова М., К. Ишев, Ф. Филипов**, Применение структурного метода при аналитическом решении волновых задач, Сб. докладов конференции по системам для автоматизированного информационного обслуживания, Варна, 3-8 октября, **1983**.
10. **Цветкова Г.В., М. Спиридонова**, Цифрово моделиране на нелинейни съставящи в динамиката на работи. 8-ма Национална научно-техническа конференция с международно участие „Автоматизация на дискретното производство“, Научни известия на НТС по машиностроене, 4-та година, № 8, **1997**, 144-149.
11. **Тренков И., Спиридонова М.**, Символьное решение математических задач для применения в геодезии, Сб. трудов Междунар. семинара „Проблемы и применения искусственного интеллекта“, Варна, 21-25 септември, **1987**, 206-210.
12. **Trenkov I., M. Spiridonova, M. Daskalova**, An Application of REDUCE System for Solving a Mathematical Geodesy Problem, Proc. of the Intern. Symposium on Symbolic and Algebraic Computation ISSAC'91, July 15-17, 1991, Bonn, Germany, ACM Press, **1991**, 448-449.
13. **Trenkov I., M. Spiridonova, M. Daskalova**, Bases of symbolic transformations for mathematical geodesy problems, Proc. of the Intern. conf. on Computer Algebra Systems in Phys. Res., Dubna, 1990, New Scientist, Singapore, **1991**, 372-376.
14. **Daskalova M., M. Tomev, M. Spiridonova**, Symbolic and Numerical Computations for Mathematical Geodesy Problem Solving by MGCOMP, Proc. of the First National Conference Informatics'94, Sofia, **1994**, 75-82.
15. **Спиридонова М., Ч. Докев, И. Диков**, Система за аналитични преобразования на степенни редове, Мат. и мат. образование, Сб. докл. на 18-та Прол. конф. на СМБ, Албена, 3-6 април, **1989**, 494-497.
16. **Dikov I.**, Symbolic Computation of the General Term of Taylor Series, Proc. of the Intern. conf. on Computer Algebra Systems in Phys. Res., Dubna, 1990, New Scientist, Singapore, **1991**, 159-163.
17. **Спиридонова М., И. Диков, Л. Цонкова**, Пресмятане на граници на функции със системата REDUCE. Приложение за изследване на рационални функции и работа със степенни редове. Мат. и мат. образование, Сб. докл. на 21-та Прол. конф. на СМБ, София, **1992**, 235-241.
18. **Спиридонова М., М. Джамбазова**, Создание и использование баз математических знаний при работе с системой REDUCE 2, Сб. докл. Междунар. конференции „Автоматизация научных исследований“, Пловдив, 15-20 октября, **1984**, 463-466.
19. **Спиридонова М.**, Организация и применение баз символьных преобразований на основе системы REDUCE 2, Сб. трудов Международного совещания по аналитическим вычислениям на ЭВМ и их применению в теоретической физике, Дубна, 17-20 сентября, **1985**, 28-32.
20. **Spiridonova M.**, Some Extensions and Applications of REDUCE System, Proc. of EUROCAL '87, Leiptzig, 1987, Lecture Notes in Computer Science 378, Springer-Verlag, **1989**, 136-137.
21. **Геров А., Ю. Капитонова, М. Спиридонова, В. Томов**, Интеллектуальные системы аналитических преобразований, Представление знаний в человеко-машинных и робототехнических системах, Том С: Прикладные человеко-машинные системы, ориентированные на знания, ВЦ АН СССР, ВИНТИ, Москва, **1984**, 112-136.
22. **Спиридонова М., М. Нишева**, Аналитико-вычислительные системы: состояние и перспективы, Сб. трудов Междунар. семинара „Проблемы и применения искусственного интеллекта“, Варна, 21-25 сентября, **1987**, 184-188.
23. **Томов В., А. Геров**, Интеллектуальные прикладные программные системы, Сб. трудов Междунар. семинара „Проблемы и применения искусственного интеллекта“, Варна, 21-25 сентября, **1987**, 58-66.
24. **Спиридонова М., М. Нишева**, Интеллектуализированные системы аналитических преобразований и их применение в научных исследованиях и проектировании, Сб. трудов

- Междунар. семинара „Теория и приложение на искусственного интеллекта”, Созопол, 29 мая – 2 июня, **1989**, 313-318.
25. **Геров А., А. Григоров**, Аналитично-вычислительная система для работы с матрицами, Сб. докладов конференции по системам для автоматизированного информационного обслуживания, Варна, 3-8 октября, **1983**.
  26. **Геров А.**, Интегрированные численные и символьные вычисления на основе языка ALDES. Сб. трудов Международного совещания по аналитическим вычислениям на ЭВМ и их применению в теоретической физике, Дубна, 17-20 сентября, **1985**, 28-32.
  27. **Геров А., Е. Дечева**, Аналитическое вычисление собственных значений и собственных векторов матриц, Сб. докл. Междунар. конференции „Автоматизация научных исследований“, Пловдив, 15-20 октября, **1984**, 232-239.
  28. **Геров А., Н. Куманов**, Планиране на действията в АИС, Математика и математическо образование, Сб. Доклады на 14-та Пролетна конференция на СМБ, Сл. бряг, 6-9 април, 1985, София, **1985**, 362-367.
  29. **Геров А.**, Специализирана аналитично-изчислителна система за работа с матрици, Автореферат на дисертация за получаване на научната степен „кандидат на математическите науки“, София, **1988**.
  30. **Nisheva M., T. Tonev**, A System for Continued Fractions Manipulation. Complex Analysis and Applications'85, Sofia, Publishing House of the Bulgarian Academy of Sciences, **1986**, 482-488.
  31. **Tomov V., M. Nisheva, T. Tonev**, Computer Algebra System for Continued Fractions Manipulation. Lecture Notes in Computer Science 378, Springer-Verlag, **1989**, 52-53.
  32. **Нишева М.**, Аналитично-изчислителна система за работа с непрекъснати дроби. Автореферат на дисертация за получаване на научната степен „кандидат на математическите науки“, София, **1988**.
  33. **Nisheva-Pavlova M.**, Knowledge Representation and Problem Solving in the Intelligent Computer Algebra System STRAMS. Annuaire de l'Université de Sofia "St. Kliment Ohridski", Faculté de Mathématiques et Informatique, Tome 91 (**1997**), Livre 2, 193-202.
  34. **Nisheva-Pavlova M.**, An Intelligent Computer Algebra System and its Applicability in Mathematics Education. International Journal of Computer Algebra in Mathematics Education, ISSN 1362-7368, Vol. 6 (**1999**), No. 1, 3-16.
  35. **Томов В., Й. Харизанов**, Система аналитического вычисления неопределенных интегралов рациональных функций, Сб. докладов конференции по системам для автоматизированного информационного обслуживания, Варна, 3-8 октября, **1983**.
  36. **С.И. Вырбанав**. Система факторизации полиномов Сб. трудов Международного совещания по аналитическим вычислениям на ЭВМ и их применению в теоретической физике, Дубна, 17-20 сентября, **1985**, 59-64.
  37. **Spiridonova M., J.-A. Desideri**, Symbolic Computations for the Analysis of Finite-Difference Schemes by the Modified Equation Approach, Intern. Workshop "Computer Algebra Applications", Kiev, July 9, 1993, Extended Abstracts of Contributions, St.Petersburg, **1993**, 46-49.
  38. **Spiridonova M., J.-A. Desideri**. Symbolic Derivation of Linear and Nonlinear Modified (Partial Differential) Equations, In: Mathematics and Education in Mathematics, Proc. of the 27th Spring Conf. of the Union of Bulgarian Mathematicians, Pleven, **1998**, 234-238.
  39. **Spiridonova M., L. Alamanov, P. Kenderov**, Implementation of Convex Analysis Operations Using Maple, W. Kuechlin (Ed.), ISSAC'96 Posters, ETH, Zurich, July 24-26, 1996, ACM Press, **1996**, 79-82.
  40. **Спиридонова М., Б. Байчева**, Символьное вычисление преобразования Лежандра, Ж. Программирование, № 4, Наука, Москва, **1994**, 77-85.

41. **Dimovski I., M. Spiridonova**, Numerical solution of boundary value problems for the heat and related equations. In: Computer Algebra and its Application to Physics CAAP 2001, Dubna, **2001**, 32-42.
42. **Dimovski I., M. Spiridonova**, The Fourier Method and the problem of its computer implementation. In: Mathematics and Education in Mathematics. Proc of the 32<sup>nd</sup> Spring Conf. of the Union of Bulgarian Mathematicians, Sl. bryag, **2003**, 117-127.
43. **Dimovski I., M. Spiridonova**, Computer Implementation of Solutions of BVP for Finite Vibrating Systems, J. Mathematica Balkanica, Vol 18, **2004**, Fasc. 3-4, 277-285.
44. **Spiridonova M.**, Computation of Periodic Solutions of Differential Equations by the Operational Calculus Approach, In: Mathematics and Education in Mathematics, Proc. of the 34<sup>th</sup> Spring Conf. of the Union of Bulgarian Mathematicians, Borovets, **2005**, 191-195.
45. **Dimovski I., M. Spiridonova**, Computational Approach to Nonlocal Boundary Value Problems by Multivariate Operational Calculus, Math. Sci. Res. J., Vol 9, No 12, **2005**, 315-329.
46. **Спиридонова М.**, Директни операционни методи в програмна среда за компютърна алгебра, Автореферат на дисертация за получаване на научната степен „доктор по информатика“, София, **2008**.
47. **Spiridonova M.**, Operational Methods in the Environment of a Computer Algebra System, Serdica J. Computing (Inst. of Math. and Informatics, Bulgarian Acad. of Sci.), Vol 3, No 4, **2009**, 381-424.
48. **Dimovski I.H., M.N. Spiridonova**, Construction of Nonlocal Linear Vibration Models Using a Computer Algebra System. J. Programming and Computer Software, Vol 37, No 2, **2011**, 71-77.
49. **Dimovski I., M. Spiridonova**, Operational Calculus Approach to Nonlocal Cauchy Problems, J. Mathematics in Computer Science, Vol 4, No 2-3, **2011**, 243-258.
50. **Dimovski I., M. Spiridonova**, Extension of Duhamel Principle for Linear Nonlocal Initial-Boundary Value Problems, "Complex Analysis and Applications '13" (Proc. Intern. Conf., Sofia, 31 Oct.-2 Nov., 2013) – Electronic Book, Institute of Mathematics and Informatics, Bulg. Acad. Sci., Sofia, **2013**; ISBN 978-954-8986-37-3 and ISBN 978-954-8986-38-0, 85-92.
51. **Dimovski I., M. Spiridonova**, Operational Calculi for Nonlocal Cauchy Problems in Resonance Cases, Lecture Notes in Computer Science 8372, Springer-Verlag Berlin Heidelberg **2014**, 83-95.
52. **Dimovski I., M. Spiridonova**, Operational Calculus Approach to Explicit Solving of Initial and Boundary Value Problems, Particles and Nuclei, Letters, Vol 12, No 3, **2015**, 619-623.
53. **Спиридонова М.**, Системите за компютърна алгебра – естествен помощник в обучението по математика (плен. доклад), Мат. и мат. образование, Сб. докл. на 29-та Прол. конф. на СМБ, Ловеч, 3-6 април, **2000**, 67-78.
54. **Spiridonova M., B. Kuyumdjieva**, The Computer Algebra Systems as an Element of the New Technologies in Education, Proc. of the Second National Conference “The New Technologies in Education and Professional Learning”, Plovdiv, Centre for European Education and Multicultural Communications, Sofia, **2004**, 186-194.
55. **Спиридонова М., М. Нишева, Д. Николова, Б. Куюмджиева**, Системи за компютърна алгебра в обучението по математика и информатика, Математика и математическо образование, Доклади на 32-та прол. конф. на СМБ, Слънчев бряг, 5-8 април, **2003**, 416-420.
56. **Poryazov S., E. Saranova, M. Spiridonova**, Modeling of Telecommunication Processes in an Overall Complex System, Particles and Nuclei, Letters, Vol 12, No 3, **2015**, 657-662.

# Компютърната лингвистика в ИМИ – история, проекти и резултати

Людмила Димитрова

Институт по математика и информатика при БАН, София  
[ludmila@math.bas.bg](mailto:ludmila@math.bas.bg)

**Резюме:** Статията представя накратко развитието на компютърната лингвистика в ИМИ-БАН. Секция „Математическа лингвистика“ на ИМИ е пионер в тази област за България. Описани са изследвания в областта на автоматичната обработка на български език, разработката на алгоритми и системи за приложенията им. Проектите на секцията в областта на езиковото инженерство и резултатите от тяхното изпълнение – представяне и обработка на лингвистични знания, разработка на цифрови многоезични ресурси, средства и системи за представянето и управлението им, както и приложенията им – също са представени.

**Ключови думи:** Езикови ресурси, компютърна лингвистика, лексически бази данни, цифрови корпуси.

## 1. Кратки исторически бележки

Почти 10 години след първия опит за автоматичен руско-английски превод, реализиран в САЩ през 1954 г., към тогавашния Математически институт при БАН (МИ-БАН) е създадена Група по математическа лингвистика и машинен превод. В 1964 г. по идея на професор Александър Людсканов (1926-1976), известен специалист по автоматична обработка на езика и машинен превод, в МИ-БАН започват статистически изследвания на български текстове и първи опити за автоматизиран превод между родствени езици (руски и български) [Людсканов 1968], [Ludskanov 1971].

Секция „Математическа лингвистика“ на ИМИ-БАН е пионер в развитието на компютърната лингвистика в България. Основана е в 1977 г. като Лаборатория за математическа лингвистика и е наследник на Групата по математическа лингвистика и машинен превод към МИ-БАН. Лабораторията за математическа лингвистика е преименувана в секция „Математическа лингвистика“ в 1985 г.

## 2. Развитие

**Главните тематични направления,** свързани с изследванията в областта на автоматичната обработка на естествен език и развивани през годините в секцията, са:

1. компютърна лингвистика – средства за автоматична обработка на български език,

2. математическа лингвистика – логико-граматически формализми и компютърни системи за моделиране на резултати от теория на формалните езици,
3. компютърна лингвистика – представяне и обработка на лингвистични знания, разработка на цифрови многоезични ресурси, средства и системи за представянето и управлението им,
4. изследователски е-инфраструктури – създаване на концептуален модел на е-инфраструктура за цифрова лексикография.

Секцията е участник в две международни изследователски е-инфраструктури в областта на автоматична обработка на езика, финансирани от Европейската комисия:

- **ELSNET: European Network of Excellence in Language and Speech, 1993–2004**, (национален координатор – проф. Радослав Павлов),
- **CLARIN: Common Language Resources and Technology Infrastructure, 2006–2012**, проект, (координатор за ИМИ – проф. Людмила Димитрова).

### **3. Компютърна лингвистика – автоматична обработка на естествен език**

#### **✓ Алгоритми и системи за автоматична обработка на български език**

Това направление включва пионерните разработки върху автоматичната обработка на български език в периода 1980-1985 г.

**Първата система за автоматично сегментиране на българските словоформи без използване на лексемен речник** (речник на българските основи) и без каквито и да е ограничения, наложени върху входния текст, е представена в [Dimitrova, Paskaleva, Nenova, 1981]. Системата реализира описания в [Paskaleva, Dimitrova, Nenova, 1981] **първи алгоритъм за автоматичното сегментиране** (определяне на основа и окончание) на словоформите в текста. Алгоритъмът е комбинация от евристични съображения и точно определени множества от езикови елементи (множествата на българските окончания), той е разработен въз основа на анализа на 700 000 български словоформи, които като обем представят съдържанието на няколко тълковни речника и могат да се образуват от 70-те хиляди единици на Обратния речник на българския език. Системата съдържа и съответните обслужващи програми за създаване, поддържане и обновяване на базовата лингвистична информация.

**Програмната система за автоматичен анализ на български текст** реализира за първи път за български език **алгоритъм за автоматично снемане омонимията на окончанията** (автоматично причисляване на дадена словоформа към определен граматичен клас) [Dimitrova, Isusova, 1985]. Системата, работеща в режим на диалог на български език, дава възможност за изследване вида на лингвистичната информация, която може да се извлече от текста в резултат на този вид автоматичен анализ (присъединяване на дадена словоформа към определен граматически клас). Системата реализира

автоматичен преход „текст → речник“, включващ сегментиране на български текст; автоматично построяване на различни видове речници и речникови архиви: конкорданси на даден текст или даден автор с добавяне на натрупаната в процеса на обработка статистическа информация, честотни речници, комбинирани речници. По зададена от потребителя заявка, системата реализира търсене и автоматично извличане на словоформи и словосъчетания от даден български текст, с контекст или не, с цел изучаване на поведението на дадената лексическа единица, синтактична конструкция или определен граматичен клас в текста. Системата е разработена в две версии (зависещи от наличните по това време технологични платформи): едната версия е реализирана на APPLE-PASCAL за персонален компютър APPLE II, а другата е реализирана на PL/1 в DOS за EC1040.

#### **4. Математическа лингвистика – приложения**

##### **✓ Приложения на логико-граматически формализми и компютърни системи за моделиране на резултати от теория на формалните езици (1986-1995)**

Изследванията в това направление (с ясна теоретическа и практическата важност) са свързани с опитите (които продължават след Ноам Чомски) за формално представяне на естествените езици, в случая – на българския.

**Програмната система за моделиране на основни резултати от теорията на формалните езици и граматика** [Dimitrova, Todorova, 1993], [Dimitrova 1993], е реализирана за нуждите на обучението. Теоретичната база на системата включва ред теореми, доказани в [Dimitrova, Todorova, 1991], които свързват определени типове формални езици с магазинни автомати. Програмната система е диалогова и е предназначена за демонстриране на различни еквивалентности при описание и генериране на формални езици.

**Логическите граматика**, създадени в периода 1978-1989 г., използват предикатната логика от първи ред за описание на естествено-езикови структури и явления, като осигуряват ефективен синтактичен и семантичен анализ и генериране на фрази при частична обработка или за ограничен фрагмент. Най-напред са използвани за моделиране на свободен словоред, координация, лява и дясна екстрапозиция и други езикови явления в английски език, по-късно – за испански и френски.

В [Dimitrova, Pavlov, 1993] задълбочено и последователно са представени четири вида логически граматика – *metamorphosis grammars*, *definite clause grammars*, *extraposition grammars* и *discontinuous grammars* – и са разгледани възможностите им за описване на специфични особености на българския език.

За първи път в [Dimitrova, Pavlov, 1995] е показано приложение на прекъснатите логически граматика за формалното описване на свободния словоред в българския език. Предложен е пример за описание на „управлението“ на относителността в граматичен смисъл, който илюстрира тясната връзка между лингвистичната теория и теорията на формалните езици.

Приносът е в използването за първи път на развити логико-граматически формализми за описание и моделиране на особено тежки за формално описание явления в българския език (аспекти на синтаксиса), които имат пряко приложение към автоматизираната обработка на езика.

## **5. Компютърна лингвистика – проекти по Европейски програми**

### **✓ Проект на ЕК: COP 106 MULTEXT-East *Multilingual Text Tools and Corpora for Central and Eastern European Languages, 1995-1997***

**Координатор:** Jean Véronis – CNRS-France, ръководител за България – проф. Радослав Павлов.

**Участници:** CNRS (Laboratoire Parole et Langage) & Universite de Provence, Aix-en-Provence; Bulgarian Academy of Science, Sofia, Bulgaria; Charles University, Prague, Czech Republic; University of Tartu, Tartu, Estonia; Hungarian Academy of Science, Budapest, Hungary; Research Institute for Informatics, Bucharest, Romania; Jožef Stefan Institute, Ljubljana, Slovenia.

За първи път българският език участва в многоезичните ресурси на финансиран от ЕК проект – **MULTEXT-East**. Проектът включва шест централни и източно-европейски езици: *български, естонски, румънски, словенски, унгарски и чешки* (т.нар. **МТЕ-езици**), и *английски* като „опорен език“.

**MULTEXT-East** е продължение на проекта **MULTEXT *Multilingual Text Tools and Corpora (1994-1996)***, един от най-известните проекти на ЕК в областта на езиковото инженерство за създаване на цифрови ресурси и софтуер за обработка на английски, френски, испански, немски, италиански, финландски и холандски. Ресурсите, разработени в рамките на **MULTEXT-East (МТЕ)** [Dimitrova et al., 1998], са свързани с научната и технологична готовност за интегриране на българския език в многоезични приложни системи на ЕК, системи за машинен превод (МП) и др.

**Резултати:** Разработка на МТЕ-многоезичните ресурси:

- **средства за представяне и аотиране на лингвистични знания;**
- **специфични (езиково-зависими) ресурси;**
- **цифрови корпуси** (обемни хранилища на естественоезикови текстове).

### **Представяне на лингвистични знания**

В тази група попадат разработките на стандартизирани средства за представяне на лингвистични знания за изследвания и приложения в областта на езиковото инженерство: многоезични цифрови ресурси за МТЕ-езиците и английски език. Ресурсите включват средства за представяне и аотиране на лингвистични знания в съответствие с международните стандарти за автоматична обработка на естествени езици. Тези ресурси са специфични (езиково-зависими) ресурси. Средствата за представяне и аотиране на лингвистични знания обединяват хармонизираните формални лексически

описания на морфосинтактичните характеристики за всеки от шестте МТЕ-езика (вкл. български).

### **Разработка на стандартни лексически описания**

Формалните лексически описания (**MorphoSyntactical Descriptions – MSD**) представят морфосинтактичните характеристики на всеки естествен език. Характеристиките са дефинирани като стойности на атрибути за всеки клас – определена част на речта – съгласно граматиката на съответния език. Специфичните за всеки език явления, които нямат аналог в другите пет езика на проекта, са описани със същия формализъм, но са представени като отделен „специфичен“ езиков ресурс. По този начин се осигурява необходимата за автоматичния анализ – морфологичен, синтактичен или семантичен – информация за различните езици. Разработените за български език формални описатели (MSDs) са 326. Този формализъм дава възможност да се използват програмните средства, създадени в рамките на други европейски проекти (например MULTEXT) за обработка и на български текстове, маркирани с SGML (Standard General Mark-up Language) и снабдени със стандартните лексически описания (MSD). В [Dimitrova 2008] е описано приложението на българските специфични МТЕ-езикови ресурси за автоматично снемане на морфосинтактична многозначност в аотирани български текстове. В [Dimitrova, Rashkov, 2009] са разгледани МТЕ морфосинтактичните спецификации на някои български глаголни форми (причастията) и са предложени нови, съответстващи на съвременната българска граматика.

### **Разработка на стандартизирани специфични езикови ресурси за сегментиране на текстове**

Специфичните езикови ресурси за сегментиране на аотирани текстове са разработени като стандартни файлове, съдържащи за всеки език таблици с пунктуационните правила, най-употребяваните съкращения, съставните съкращения, клитиките и сложните съставни думи. Описаните ресурси се използват в системи за автоматично сегментиране на текст (т.нар. „сегментори“). Анализът на текст при автоматично сегментиране преминава през няколко етапа: от разпознаване на препинателните знаци до морфологичен анализ и преход „текст → реч“. Специфичните езикови ресурси са входни данни не само за „самообучаващи се“ софтуерни пакети, наречени „тагери“, но и в системи за „предсказване“ на възможното множество от тагове (определен тип характеристики) за непознати, предварително несрещани в текст думи, които липсват и в лексиконите. В [Dimitrova et al. 1996] са описани разработените за проекта MULTEXT-East лексикони (специализирани речници) и специфични (езиково-зависими) ресурси, необходими за сегментиране на аотирани езикови текстове.

### **Лексикони и системи от правила, необходими за сегментиране и аотиране на текстове**

Лексиконите имат стандартна структура, съставени са от речникови статии, които съдържат заглавна дума (словоформа), основната форма на заглавната

дума и съответното ѝ морфосинтактично описание (от множеството на хармонизираните MULTEXT-East-описания). Лексиконите за български език са 4: *MTE-Orwell-1984.bg*, *MTE-Fiction.bg*, *MTE-News.bg* и *MTE.Bg*. Лексикон ***MTE-Orwell.Bg*** съдържа 55182 речникови статии (от тях 50810 са словоформи, 4205 – съкращения и имена, 1044 числови данни), ***MTE-Fiction.Bg*** съдържа 24837 речникови статии (от тях 23837 словоформи, 772 съкращения и имена, 226 числови данни), лексикон ***MTE-News.Bg*** съдържа 26200 речникови статии (от тях 23164 словоформи, 2178 съкращения и имена, 858 числа). Лексиконът ***MTE.Bg*** е обединение на първите три с изключване на повтарящите се единици.

### **Цифрови многоезични ресурси**

Цифровите корпуси са обемни хранилища на естественоезикови текстове. В зависимост от броя на езиците те биват **едно-** и **многоезични**. Многоезичните корпуси, в зависимост от вида на текстовете, които съдържат, са **сравними** (текстовете на съответните езици не са преводни съответствия, а са сравними по обем – броя на словоформите в тях), **паралелни** (текстовете са преводни съответствия) и **подравнени** (паралелни текстове с връзки между съответните изречения, сегменти или параграфи). **Анотирани** са текстове, чиито единици са придружени от допълнителни данни. Анотацията на текст(-ове) или корпуси е структурна (маркирани са структурни единици в текста, като части, глави, параграфи, изречения) или лингвистична (към всяка словоформа в текста са присъединени данни, кодиращи принадлежността ѝ към определен клас, напр. част на речта).

Корпусите, разработени в проекта MULTEXT-East, са **анотирани паралелни** и **сравними** (за всеки език); **подравнени** (6 двойки „МТЕ-език – английски език“).

### **Паралелен МТЕ-корпус**

Паралелният многоезичен МТЕ-корпус е съставен от **три версии**, всяка съдържаща **7 текста**:

- **1-ва версия „Оригинални текстове“:**

Тази версия съдържа текстовете на оригинала на романа „1984“ на Джордж Оруел и 6-те превода на романа на МТЕ-езиците на проекта. Текстът на българския превод на „1984“ на Джордж Оруел съдържа **87235** словоформи.

- **2-ра версия „1984“ CesDOC encoding:**

7-те паралелни текста на „1984“ са **структурирани автоматично** съгласно международния стандарт за кодиране на корпуси – CES (Corpus Encoding Standard) на TEI (Text Encoding Initiative);

- маркирани с `<div>-</div>`, `<p>-</p>`, `<s>-</s>` tag-ове;
- анотирани ръчно с маркери за **имена**, **названия**, **дати**, **цитати**, **пряка/непряка реч**, **поезия**.

Пример от *CesDOC encoding* (структурна анотация):

```
<teiHeader> ..... </teiHeader>
<text id=„mteo-bg.“ lang=„bg“>
<body id=„bg1984“ lang=„bg“>
<div id=„Obj.1“ n=„1“ type=„part“>
  <head>I.</head>
  <div id=„Obj.1.1“ n=„1“ type=„chapter“>
    <p id=„Obj.1.1.1“>
      <s id=„Obj.1.1.1.1“>Априлският ден бе ясен и студен, часовниците биеха
      тринайсет часа.</s>
      <s id=„Obj.1.1.4.5“><q id=„Obj.1.1.4.5.1“ rend=„caps“ type=„slogan“>
        <name type=„person“>Големия брат</name> те наблюдава</q>
        , гласеше надписът, а тъмните очи пронизваха
        <name type=„person“>Уинстън</name>.</s>
```

### - 3-та версия „1984“ *CesANA encoding*:

Паралелните текстове на „1984“ на МТЕ-езиците в корпуса *MTE CesANA encoding* са автоматично ТОК-низирани: разделени са на *TOKEN* – думи и препинателни знаци, като *TOKEN* са маркирани с маркери за <tok>-</tok>, <orth>-</orth>.

Корпусът е анотиран автоматично на ниво *словоформа*: текстовете са *лематизирани* автоматично, т.е. добавена е лингвистична информация за основната форма (лема) и морфосинтактичното описание (MSD) на всяка словоформа. Словоформите в текста са маркирани с маркери за <disamb>-</disamb>, <lex>-</lex>, <base>-</base>, <msd>-</msd>, <ctag>-</ctag>.

3-та версия на корпуса е получена след *автоматично снемане на синтактична многозначност* и човешка пост-редакция за отстраняване на грешки: паралелните текстове са *ръчно* проверени за достоверност на препинателни знаци и граници на изречение, принадлежност на словоформите към класовете на частите на речта (т. нар. *POS-annotation*) и за снемане на морфосинтактична многозначност (присъединяване на основата с най-голяма вероятност към всяка словоформа в текста).

Пример на *CesANA encoding* за ден бе в изречението „Априлският ден бе ясен и студен, часовниците биеха тринайсет часа.“

```
<tok type=WORD from='Obj.1.1.1.1\12'>
  <orth>ден</orth>
  <disamb><base>ден</base><ctag>NCMS-N</ctag></disamb>
  <lex><base>ден</base><msd>Ncms-n</msd><ctag>NCMS-N</ctag></lex>
</tok>
<tok type=WORD from='Obj.1.1.1.1\16'>
  <orth>бе</orth>
  <disamb><base>съм</base><ctag>VAIA3S</ctag></disamb>
  <lex><base>бе</base><msd>Qgs</msd><ctag>QG</ctag></lex>
```

```
<lex><base>съм</base><msd>Vaia2s</msd><ctag>VAIA2S</ctag></lex>  
<lex><base>съм</base><msd>Vaia3s</msd><ctag>VAIA3S</ctag></lex>  
</tok>
```

### Сравним МТЕ-корпус

Сравнимият многоезичен МТЕ-корпус е съставен от текстове на МТЕ-езиците (за всеки език в обем около 200000 словоформи). Сравнимият корпус за всеки език е съставен от два корпуса *Fiction.L* и *News.L* (в *CesDOC encoding*). За български език това са корпусите *MTE.Fiction.Bg* и *MTE.News.Bg*:

*MTE.Fiction.Bg* – съдържа съвременна българска литература, романът *PASSION* или *смъртта на Алиса* на Емилия Дворянова; главите 1-4 от романа *Искам, Вярвам, Мога* на Юлия Берберян. Обемът му е **97251** словоформи. Корпусът е маркиран на ниво „параграф“ (<p> </p>), за пряка реч (<q rend=„PRE=mdash“> </q>) и за имена (<name> </name>).

Пример:

```
<p><q rend=„PRE=mdash“>Какво е това, дето го чете момичето,  
госпожо, неясно ми се вижда, да не й повлияе зле.</q></p>  
<p>А тя ми отвърна като знаеща:</p>  
<p><q rend=„PRE=mdash“> Философия, <name  
type=„person“>Йо</name>.</q></p>
```

*MTE.News.Bg* – съдържа откъси от вестници с обем **96538** словоформи, текстовете са структурирани ръчно на ниво „параграф“ и анотирани ръчно с маркери за ниво „параграф“, на ниво „sub-paragraph“ са анотирани ръчно за съкращения, дати, имена.

### Паралелен аудио МТЕ-корпус

МТЕ-аудио-корпусът е съставен от четири аудио файла на естонски, унгарски, румънски и словенски, и шест текстови файла за 6-те езика на проекта, с по 40 записа (кратки откъси, всеки от 5 тематично свързани изречения).

### MULTEXT-East подравнен корпус

Подравняването (*alignment*) на текстове е процес на свързване на двойки от думи (словоформи), фрази, термини или изречения в текстове от различни езици, които са преводни еквиваленти, то е и форма на анотиране, извършвана над паралелни корпуси, за да се улесни построяването и оценката на

преводните модели, запазени в паметта и използвани в „подпомогнат от компютър“ превод.

**Подравненият МТЕ корпус** съдържа 6 двойки автоматично подравнени текстове: „превод на МТЕ език – английски оригинал“.

Пример:

1-1 подравнени изречения	< <b>Obj.1.1.7.4</b> >Още три сгради, подобни по външен вид и размери, бяха посети из <b>Лондон</b> . < <b>Oen.1.1.9.2</b> >Scattered about <b>London</b> there were just three other buildings of similar appearance and size.
1-1 подравнени изречения	< <b>Obj.1.1.7.5</b> >И дотолкова се извисяваха над околните здания, че от покрива на жилищен дом <b>Победа</b> можеха да се видят и четирите едновременно. < <b>Oen.1.1.9.3</b> >So completely did they dwarf the surrounding architecture that from the roof of <b>Victory Mansions</b> you could see all four of them simultaneously.
3-1 подравнени изречения	< <b>Obj.1.1.29.6</b> >Тъмнокосото момиче зад <b>Уинстън</b> крещеше: „Свиня!“< <b>Obj.1.1.29.7</b> >„Свиня!“< <b>Obj.1.1.29.8</b> >„Свиня!“, после изведнъж грабна тежък речник по новговор и го запокити по телекрана. < <b>Oen.1.1.30.6</b> >The dark-haired girl behind <b>Winston</b> had begun crying out „Swine! Swine! Swine!“ and suddenly she picked up a heavy Newspeak dictionary and flung it at the screen.

✓ **Проект PL96-1142 CONCEDE Consortium for Central European Dictionary Encoding на програмата на ЕК INCO-COPERNICUS, 1998-2000**

**Координатори:** Roger Evans и Adam Kilgarriff – ITRI, Brighton, UK, ръководител за България – проф. Радослав Павлов.

**Участници в Консорциума:** University of Brighton, Brighton, UK; XRCE, Grenoble, France; Vassar College, New York, USA; Bulgarian Academy of Science, Sofia, Bulgaria; Charles University, Prague, Czech Republic; University of Tartu, Tartu, Estonia; Hungarian Academy of Science, Budapest, Hungary; Research Institute for Informatics, Bucharest, Romania; Josef Stefan Institute, Ljubljana, Slovenia.

Проектът има за цел да хармонизира методологията, средствата и ресурсите за създаване на обща лексическа база данни (LDB) за шестте европейски езика на проекта МТЕ – български, естонски, румънски, словенски, унгарски и чешки. Разработените ресурси са преносими и многократно използваеми, тъй като използват стандартите на Text Encoding Initiative Dictionary Working Group

(TEI-DWG) и затова са съвместими други TEI-ресурси. Ръководствата за кодиране на лексикографски ресурси, разработени най-напред за западноевропейски езици, са валидирани за МТЕ-езиците чрез прилагане на ISO8879SGML. CONCEDE LDB е интегрирана с паралелния подравнен МТЕ-корпус и е предназначена за интегрални изследвания и приложения в системи за езиково инженерство, машинен превод и пр.

**Резултати:** Разработка на лексическата бази данни CONCEDE.

Ресурсите, разработени в рамките на проекта CONCEDE, включват както средства за кодиране на речници и за представяне и аотиране на лингвистични знания в съответствие с международните стандарти, така и лексически бази данни (едно- и многоезични). Схемата от лексикографски спецификации, в съответствие с международните стандарти за кодиране на електронни речници, е разработена в проекта CONCEDE с цел приложения в многоезични приложни системи на ЕС и системи за машинен превод. Моделът CONCEDE за кодиране на речници е използван за създаване на стандартни (съгласно международните стандарти за кодиране на текстове на Text Encoding Initiative) лексически бази данни на шестте МТЕ-езика и английски като „опорен“ език.

#### **Лексическа база данни за български език**

Първата лексическа база данни на български език **Bulgarian CONCEDE LDB** е създадена за проекта CONCEDE. Резултатите от изследванията за българския език, изпълнени в рамките на проекта, са отразени в [Dimitrova, Pavlov, Simov, 2002]. Те включват разработка на лексикографските спецификации за българския език съгласно международните стандарти за кодиране на електронни речници, подбора (според определената в проекта методика) на заглавните думи за едноезичната лексическа база данни за български език като част от многоезичната база данни CONCEDE, комплектоването и кодирането на речниковите статии за тези думи. Bulgarian CONCEDE LDB съдържа 2700 речникови единици, базирана е на българския тълковен речник и използва лексикографските спецификации, разработени в рамките на проекта.

*Разработката на лексическата база данни на български език Bulgarian CONCEDE LDB е определена като едно от най-важните научно-приложни постижения на ИМИ-БАН през 2000 г.*

Пример на речникова статия:

```
<entry>
<hw>автом|ат</hw>
  <gen>м.</gen>
  <struc type=„1“>
    <def> Апарат, който сам извършва определени действия.</def>
  </struc>
  <struc type=„2“>
    <register>прен.</register>
    <def>човек, който действа механически, без обмисляне.</def>
```

```
</struc>
<struc type="3">
  <register>воен.</register>
  <def>къса автоматична пушка.</def>
</struc>
<etym>[от гр.]</etym>
</entry>
```

✓ **Проект 211938 MONDILEX Conceptual Modelling of Networking of Centres for High-Quality Research in Slavic Lexicography and Their Digital Resources (2008-2010), проект по 7ПП на ЕК**

**Координатор:** ИМИ на БАН, ръководител на международния екип е проф. Людмила Димитрова.

**Участници в Консорциума:** Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia (Coordinator); Institute of Slavic Studies, Polish Academy of Sciences, Warsaw; Ludovit Štúr Institute of Linguistics, Slovak Academy of Sciences, Bratislava; Jožef Stefan Institute, Ljubljana, Slovenia; Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow; Ukrainian Lingua-Information Fund, National Academy of Sciences of Ukraine, Kiev.

**Цел на проекта** – създаване на **концептуален модел на изследователска инфраструктура** поддържаща мрежа от центрове за висококачествени изследвания в областта на цифровата славянска лексикография (**славянски езици** – езикова група от групата на индоевропейските езици; около 400 милиона души в Европа говорят славянски езици: хора от европейската популация в Централна и Източна Европа, Балканите и някои части на Азия).

**Необходимост от такава инфраструктура** – несъответствието между важноста на славянските езици и обема и качеството на цифровите ресурси на тези езици [Dimitrova et al. 2010a].

**Резултат:** Изследователска Е-инфраструктура MONDILEX.

**Създаване на концептуален модел на Е-инфраструктура за цифрова лексикография.**

Разработен е Концептуален модел на грид-базирана европейска изследователска инфраструктура от центрове в областта на цифровата лексикография и техните езикови ресурси. Моделът обхваща всички съвременни езикови технологии и видове цифрови езикови ресурси: корпуси, речници, бази данни, лексикони, тезауруси, онтологии, които трябва да бъдат включени в изследователска инфраструктура за славянска лексикография, както и технологичната платформа за нейната поддръжка [Dimitrova et al. 2011]. В рамките на проекта са публикувани 124 научни статии и 1 монография [Dimitrova et al. 2010b].

*Разработката на модела на грид-базирана европейска изследователска инфраструктура от центрове в областта на цифровата лексикография е определена като едно от най-важните научно-приложни постижения на ИМИ-БАН през 2010 г.*

## 6. Междуакадемични проекти

- ✓ Проект *Generic Task of Knowledge Discovery (1997–1998)* между USA – National Science Foundation, Carnegie Mellon University и ИМИ-БАН (координатор проф. Владимир Периклиев)

**Резултати:** Машинно откритие в лингвистиката.

В областта на машинното откритие в лингвистиката са разработени системи за откритие в различни лингвистични дисциплини (семантика, типология, историческа граматика и др.) [Pericliev, 1995]. В областта на типологията е създадена програма, която описва откритията си във формата на цялостен текст (на английски език) и два такива текста са публикувани в списание без допълнителна редакция от човек. Това е първата програма, генерира научна статия. В областта на историческата граматика машинно е генерирана хипотезата за генетична връзка между езиково семейство, говорещо се в Южна Америка на Атлантическия океан, и езиково семейство, говорещо се в повечето острови в Тихия океан [Pericliev, Valdés-Pérez, 1998], [Pericliev, 2013]. Получените резултати са важни не само за лингвистиката и компютърната лингвистика, но и за други научни области като генетика, археология, антропология и предистория.

*Разработката на система за машинно откритие е определена като едно от най-важните научно-приложни постижения на ИМИ-БАН през 2011 г., в периода 2010–2013 са публикувани 3 монографии.*

- ✓ Проект *Електронни корпуси – съпоставително изследване с цел проектиране на българо-словашки електронни езикови ресурси (2008–2010, 2012–2014)* между ИМИ-БАН и Института по лингвистика „Людовик Щур“ на Словашката академия на науките, координатори проф. Людмила Димитрова (ИМИ-БАН) и д-р Радован Гарабик (ИЛ-САН)

**Резултати:** Българо-словашки електронни езикови ресурси:

- българо-словашки паралелен корпус;
- българо-словашки речник на фрази, автоматично извлечени от подравнения корпус.

## Българо-словашки паралелен и подравнен корпус

Паралелният и подравнен българо-словашки корпус съдържа около 163 милиона словоформи в текстове от Official Journal of the EU и литературни текстове:

- текстове от Official Journal of the EU, достъпни в Internet:
  - на български език: 5 140 296 изречения, 85 млн. словоформи,
  - на словашки език: 8 650 610 изречения, 78 млн. словоформи,
- литературни текстове (романи, повести, разкази): над 1 200 000 словоформи.

Текстовете са подравнени автоматично на ниво „изречение“ със софтуерния пакет *Hunalign* без използване на двуезичен речник [Dimitrova, Garabík 2011]. За автоматичната морфологична анотация на текстовете в корпуса са използвани: тагер Морџе (обучен автоматично с tagset, разработен в Словашкия национален корпус) за словашките текстове, а за българските текстове – тагер TreeTagger.

Българо-словашкият корпус е представен в Интернет:

- тест-версия (свободен достъп на par-skbkg-free-0.1) реализирана през 2014,
- търсене в текстовете – чрез NoSketch Engine, на [http://korpus.sk/skbkg\\_en.html](http://korpus.sk/skbkg_en.html).

## Българо-словашки речник на фрази, автоматично извлечени от подравнения корпус

Създаден е двуезичен речник на фрази, автоматично извлечени от българо-словашкия подравнен корпус. Експерименталната версия на двуезичния речник съдържа около 290 000 речникови единици, в които българските и словашките словоформи са около 793000 [Dimitrova, Garabík 2014].

Пример Search interface, queried word **rieka**, the highlighted example of word alignment, **plavebných komôr** and **шлюзовете**, съответствие 2:1:

rieka Find Strategy: Match words exactly

Results for: rieka

**Frázy z paralelného slovensko-bulharského korpusu.** [More information.](#)

cesta, železnica a rieka = пътна, железопътна и речна  
jazerá, rieky alebo kanály = езера, реки или канали  
korce, rieky a jazerá = хълмове, реки и езера  
na sever od rieky = на север от река  
podniku z jazier, riek = стопанството от езера, реки  
priehrad na rieke Mekong = язовири по река Меконг  
riek, kanálov, **plavebných komôr** = на реките, каналите, **шлюзовете**  
tiahne severne od rieky = простира северно от река  
...

Input: UTF-8 Output: UTF-8 In English · Po slovensky

Пример NoSketch interface to the parallel corpus, query word **вода**:

New query		Hits: 11,709 ( 137.84 i.p.m.; related to the whole corpus)   ARF: 2,617   Result is shuffled		
Word List		Page 1 of 586 <a href="#">Go</a> <a href="#">Next</a>   <a href="#">Last</a>		
<b>Save</b> <b>View</b> <a href="#">Custom</a> <a href="#">KWIC</a> <a href="#">Sentence</a> <a href="#">Alignment</a> <b>Sort</b> <a href="#">Custom</a> <a href="#">Left   Right</a> <a href="#">Node</a> <a href="#">References</a> <a href="#">Shuffle</a> <a href="#">Filter</a> <b>Frequency</b> <a href="#">Custom</a> <a href="#">Node tags</a> <a href="#">Node forms</a> <a href="#">Doc IDs</a> <a href="#">Collocations</a> <a href="#">ConcDesc</a>	<a href="#">skbg-free-0.1-bg</a>	<a href="#">skbg-free-0.1-sk</a>		
	<a href="#">oj23385</a> 41120000 - 6 Непитейна <b>вода</b>	<a href="#">oj23385</a> 41120000 - 6 Nepritejná voda		
	<a href="#">oj27341</a> Препаратите за лъскане на стъкло , които са съставени обикновено от <b>вода</b> , от алкохол , от малко количество амонячна вода или от киселини ( оксалова , винена и други ) и от мек абразив ;	<a href="#">oj27341</a> 1 . leštidlá na sklo obsahujúce spravidla vodu , alkohol a malé množstvo amoniaku alebo kyseliny ( napr . šťavelovej alebo vinnej ) a jemné brúsiivo ;		
	<a href="#">oj78333</a> Doxivex 10 % Доксциклин хиклат 100 mg / ml Концентрат за перорален разтвор пилета ( бройлери ) 0 , 5 – 1 , 0 ml продукт / 1 литър вода за пиене . 50 – 100 mg доксциклин хиклат / 1 литър <b>вода</b> за пиене / ден за 3 – 5 дни	<a href="#">oj78333</a> Doxivex 10 % Doxycyklin hyklát 100 mg / ml Koncentrát na perorálny roztok Kurčatá ( brojleri ) 0 , 5 - 1 , 0 ml pitnej vody 50 - 100 mg doxycyklin hyklátu / 1 l pitnej vody / deň , počas 3 - 5 dní		
	<a href="#">oj5979</a> Посредством стандартния процес на обработка на питейната <b>вода</b> ( озониране ) този метаболит се превръща в нитрозамин ( NDMA ) , който е вреден за здравето .	<a href="#">oj5979</a> Štandardným procesom na úpravu pitnej vody ( ozonizovanie ) sa tento metabolit premení na nitrosamin ( NDMA ) , ktorý je škodlivý pre zdravie .		
	<a href="#">oj28784</a> Най - високото допустимо водно съдържание на разфасовките от домашни птици се определя въз основа на съдържанието на белтъчини на разфасовките , което може да бъде свързано с естественото съдържание на <b>вода</b> .	<a href="#">oj28784</a> Najvyšší povolený celkový obsah vody v dieloch hydiny sa určí na základe obsahu bielkovín v kusoch , ktorý sa môže vzťahovať na obsah fyziologickej vody .		
<a href="#">oj53788</a> Директива 77 / 649 / ЕИО на Съвета от 27 септември 1977 г . за сближаване на законодателствата на държавите - членки относно зрителното поле на <b>водачите</b> на моторни превозни средства	<a href="#">oj53788</a> smernica Rady 77 / 649 / EHS z 27 . septembra 1977 o aproximácii právnych predpisov členských štátov týkajúcich sa výhľadu vodičov motorových vozidiel			

✓ Проект *Семантика и съпоставителна лингвистика, ориентирани към разработване на двуезичен електронен речник (2006-2017) между ИМИ-БАН и Института по славистика на Полската академия на науките*, координатори проф. Людмила Димитрова (ИМИ-БАН) и проф. Виолетта Косеска (ИС-ПАН).

В рамките на проекта са разработени първите българо-полски цифрови ресурси – двуезичен българо-полски корпуси, триезичен българо-полско-литовски корпус, следващи МТЕ-модела за електронни корпуси, LDB за поддръжка на българо-полски онлайн речник. Публикувани са 24 научни статии и 1 монография [Dimitrova, Koseska 2014]

**Резултати:** Първи българо-полски цифрови ресурси

- Първи българо-полски корпус;
- Първи българо-полско-литовски корпус;
- LDB на двуезичен онлайн речник.

**Първи българо-полски корпус**

Корпусът съдържа около 5 милиона словоформи. Съставен от 2 корпуса:

- паралелен (над 3 милиона словоформи), вкл. **подравнен** корпус (около 1 милион словоформи),
- **сравним**.

**Българо-полският подравнен корпус** съдържа текстове на български литературни произведения и преводите им на полски, между които *Далеч от Толедо* на Анжел Вагенщйн, *Приказки и разкази* за деца от Ем. Станев, Св. Минков, А. Каралийчев и др., и текстове на полски литературни произведения и преводите им на български, между които части от *Solaris* и *Powrót z gwiazd* на Stanisław Lem, *Jeszcze dzień życia*, *Heban*, *Wojna futbolowa* на Ryszard Kapuściński, *Papioły* на Stefan Żeromski [Dimitrova, Koseska 2009], [Dimitrova, Dutsova 2013], [Dutsova 2014].

Подравняването на паралелните текстове на ниво „изречение“ е изпълнено с езиково независимите софтуерни пакети: Memory Translation 2007, (TextAlign: <http://mt2007-cat.ru/index.html>) или с Bitext Aligner/Converter (<http://bitext2tmx.sourceforge.net/>).

Пример на 1-1 подравнени изречения – откъс от *Далеч от Толедо*:

```
<tu tuid=„0000000001“>
  <tuv xml:lang=„Bulgarian“>
    <seg>Пъстри и различни били в стари времена названията на тези
изоставени от евреите места.</seg>
  </tuv>
  <tuv xml:lang=„Polish“>
    <seg>Barwne i różne były w owych dawnych czasach nazwy tych porzuconych
przez Żydów miejsc.</seg>
  </tuv>
</tu>
```

### Българо-полско-литовски корпус

Първият българо-полско-литовски корпус е с обем около 3 млн. словоформи. Подравненият триезичен корпус дава възможност за сравняване на езици от две различни езикови групи: славянската (български и полски) и балтийската (литовски). Триезичният корпус има широко приложение в съпоставителните изследвания [Dimitrova, Koseska, Roszko, D., Roszko, R. 2010].

Корпусът съдържа две части:

- **Подравнен корпус** – около 1 милион словоформи, който съдържа:
  - Литературни текстове на един от трите езика и преводите им на останалите два, подравнени автоматично на ниво **изречение**: части от *Solaris* на Stanisław Lem,
  - Преводите на български, полски и литовски на литература на друг език, подравнени на ниво **изречение**, вкл.: *Le Petit Prince* на Antoine de Saint-Exupéry, *Master and Margarita* на Michael Bulgakov,
  - Специализирани текстове от документи на ЕК достъпни в Internet, подравнени на ниво **параграф**.
- **Сравнимият корпус** съдържа:

- Текстове от електронни медии, които са сравними не по обем, а по **съдържание** – описания на 3-те езика на едно и също събитие (английският текст също е включен), такива множества от текстове сме нарекли „**parallel descriptions of content**“,
- Литературни творби, представящи съвременната българска, полска и литовска литература (от втората половина на 20. век), с размер на текстовете сравним за трите езика.

Пример от *Solaris* на Stanisław Lem:

```
<tu tuid=„000000011“>  
  <tuv xml:lang=„polish“>  
    <seg>Widziałem już seledynowy kontur jedynego wskaźnika.</seg>  
  </tuv>  
  <tuv xml:lang=„bulgarian“>  
    <seg>Вече различавах светлозелените контури на универсалния  
указател.</seg>  
  </tuv>  
  <tuv xml:lang=„lithuanian“>  
    <seg>Jau išskyriau žalsvus universalaus indikatoriaus kontūrus.</seg>  
  </tuv>  
</tu>
```

### LDB на двуезичен онлайн речник

Моделът CONCEDE (който е разработен за едноезична LDB) е разширен с цел разработване на двуезична LBD за поддържане на двуезични онлайн речници [Dimitrova, Panova, Dutsova 2009]. Добавени са нови tag-ове за:

- представяне на спрежението – <**conjugation**> tag и <**type**> tag (за 3-те спрежения в българския език),
- представяне на семантична информация: <**semantic**> tag и <**type**> tag (1 – за глаголи, изразяващи „състояние“ и 2 – за изразяващите „събитие“),
- вида на глаголите (aspect of verbs) в tag <**gram**> (за свършен вид и за несвършен вид),
- представяне на специфична информация за преходност/непреходност (transitivity/intransitivity): в tag <**subc**>.

Разработен е и синтактичен анализатор за трансфер на данните от LBD в релационния модел на речника. За реализирането на синтактичния анализатор е избрана технологията DOM и езикът за програмиране Java (JDK *Java Development Kit версия 1.6*). Той трябва да бъде инсталиран на компютъра, на който е инсталирана СУБД, в случая MySQL (с отворен код, предоставя интерфейси за програмиране за езиците C, C++, Eiffel, Java, Perl, PHP и Python). MySQL се разпространява безплатно, използва се за web-базирани приложения и е една от най-добрите алтернативи за изграждане на система от бази данни (висока скорост и гъвкавост).

Резултатите от проектирането и разработването на LDB за поддържане на онлайн речници са отразени в [Dimitrova, Dutsova 2012]. Предвидено е Web-

приложение за представяне на първия българо-полски онлайн речник (с два основни програмни модула: „администратор“ и „за потребителя“), базирано на LDB, чието проектиране и разработване са описани детайлно в [Dutsova 2015].

Експерименталният вариант на речника работи оптимално на Internet Explorer 6.0+ (Windows), Firefox 2.0.1+ (Windows, Linux). За среда за разработка са използвани IntelliJ IDEA и TextPad. Избраната технология предоставя възможности за първоначално инициране на релационния модел с данни от LDB, които да служат за последващо модифициране, разширяване и визуализиране от web-базираното приложение на речника. Ще напомним, че LDB намират приложение не само за поддържане на онлайн речници в Интернет, но и в образованието [Kovacheva 2014, 2015].

Пример:

повя'рва|м, -ш *вр.* uwierzyć; не мо'га да ~м на очи'те си *разг.* nie mogę uwierzyć swoim oczom, nie wierzę swoim oczom

Представяне в базата данни:

```
<entry>
  <hw>повя'рва|м</hw>
  <pos>гл.</pos>
  <gram>св.</gram>
  <conjugation><orth>-ш</orth><type>3</type></conjugation>
  <semantic><orth>състояние</orth><type>1</type></semantic>
  <subc>преходен</subc>
  <struc type=„Sense“ n=„1“>
    <trans>uwierzyć</trans>
  </struc>
  <eg>
    <q>не мо'га да ~м на очи'те си</q>
    <usg type=„register“>pot</usg>
    <trans>nie mogę uwierzyć swoim oczom, nie wierzę swoim oczom</trans>
  </eg>
</entry>
```

## Заклучение

Авторът (водещ изпълнител или ръководител на проекти в секция „Математическа лингвистика“) представя изследванията и разработките в областта на компютърната лингвистика в ИМИ-БАН. Описани са резултатите от шест международни проекта – езикови ресурси и системи за разработването и управлението им. Посочени са и области на приложението им, освен всекидневния живот, това са и системи на езиковото инженерство, интегриране в среди за Е-обучение, за образователни цели – чуждоезиково обучение, за подготовка/тренинг на преводачи – хора или програми за машинен превод, в съпоставителните езикови изследвания и др.

## **Литература**

- Dimitrova, L. (1993). Some Aspects of Computer Simulation of Formal Grammars and Automata Theory. In: (K. Chimev and Sl.Shtrakov Eds.) *Discrete Mathematics and Applications*, Blagoevgrad, 144-151.
- Dimitrova, L. (2008). Bulgarian Digital Resources as a Base for Automatic Disambiguation. In: *International Journal Études Cognitives*. Vol. 8. SOW, Warsaw, ISSN 1641-9758, 255-271.
- Dimitrova et al. (1996): Dimitrova, L., L. Sinapova, K. Simov, D. Popov, Sv. Manova-Vidinska, V. Petkevic, J. Klímová, V. Schmiedtová, H.-J. Kaalep, E. Toomsalu, C. Oravec, L. Tihanyi, S. Bruda, C. Diaconu, L. Diaconu, D. Tufis, T. Erjavec, P. Holozan, M. Romih (1996). *Multext-East: Language-Specific Resources*. MULTEXT-East Final Report D1.2F, Laboratoire Parole et Langage, CNRS, France, 1996, 28 p.
- Dimitrova et al. (1998): Dimitrova, L., T. Erjavec, N. Ide, H.-J. Kaalep, V. Petkevič, D. Tufiş. (1998). *Multext-East: Parallel and Comparable Corpora and Lexicons for Six Central and Eastern European Languages*. In: Proc. of the COLING-ACL '98, Montréal, Québec, Canada, 315-319.
- Dimitrova et al. (2010a): Dimitrova, L., V. Koseska–Toszewa, R. Garabík, T. Erjavec, L. Iomdin, V. Shyrokov (2010). *MONDILEX – Towards the Research Infrastructure for Digital Resources in Slavic Lexicography*. In: *International Journal Cognitive Studies/Études Cognitives*. Vol. 10, SOW, Warsaw, ISSN 2080-7147, 147-162.
- Dimitrova et al. (2010b): Dimitrova, L., V. Koseska, R. Garabík, T. Erjavec, L. Iomdin, V. Shyrokov (2010). *Conceptual Scheme for a Research Infrastructure Supporting Resources in Slavic Lexicography*. Sofia, Demetra Ltd. Publisher, ISBN 978-954-8986-33-5, 131 p.
- Dimitrova et al. (2011): Dimitrova, L., V. Koseska–Toszewa, R. Garabík, T. Erjavec, L. Iomdin, V. Shyrokov (2011). *Main Results of MONDILEX Project*. In: *International Journal Cognitive Studies/Études Cognitives*. Vol. 11, SOW, Warsaw, ISSN 2080-7147, 265-290.
- Dimitrova, L., R. Dutsova (2012). *Implementation of the Bulgarian-Polish Online Dictionary*. In: *International Journal Cognitive Studies/Études Cognitives*. Vol. 12, SOW, Warsaw. ISSN 2080-7147, 219-229, [IF ERIH]
- Dimitrova, L., R. Dutsova (2013). *Web-Application for the Presentation of Bilingual Corpora (Focusing on Bulgarian as One of the Paired Languages)*. In: *Journal Cognitive Studies|Études Cognitives*. Vol. 13, SOW, Warsaw, ISSN 2080-7147, 183-193.
- Dimitrova, L., R. Garabík (2011). *Bulgarian–Slovak Parallel Corpus*. In *Proceedings of the Sixth International Conference NLP, Multilinguality SLOVKO'2011*, 20-22 October 2011, Modra, Slovakia, ISBN 978-80-263-0049-6, 44-50.
- Dimitrova, L., R. Garabík (2014). *Translation Equivalence of Demonstrative Pronouns in Bulgarian-Slovak Parallel Texts*. In: *International Journal Cognitive Studies/Études Cognitives*. Vol. 14, SOW, Warsaw, ISSN 2080-7147, 77-86.
- Dimitrova L., N. Isusova (1985). *Some aspects of the computer processing of Bulgarian text*. *Proceedings of the First Conference of Program Designers*. Budapest, 11-12 July 1985, 39-44.
- Dimitrova, L., V. Koseska (2009). *Bulgarian-Polish Corpus*. In: *International Journal Cognitive Studies/Études Cognitives*. Vol. 9, SOW, Warsaw, 133-141.
- Dimitrova, L., V. Koseska–Toszewa (2014). *Semantics Properties of Selected Universal Language Categories in Digital Bilingual Resources*. Demetra Ltd Publishers, Sofia, ISBN 978-954-8986-40-3, 155 p.
- Dimitrova, L., V. Koseska, D. Roszko, R. Roszko (2010). *Application of Multilingual Corpus in Contrastive Studies (on the example of the Bulgarian-Polish-Lithuanian Parallel Corpus)*. In: *International Journal Cognitive Studies/Études Cognitives*. Vol. 10, SOW, Warsaw, ISSN 2080-7147, 217-240.

- Dimitrova, L., R. Panova, R. Dutsova (2009). Lexical Database of the Experimental Bulgarian-Polish online Dictionary. Garabik (Ed.), Metalanguage and Encoding scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, 15-16 April 2009, Bratislava, 36-47.
- Dimitrova, L., E. Paskaleva, I. Nenova (1981). Program Package for Automatic Segmentation of Bulgarian Text. Journal of Systems and Control, No. 3, Sofia, 19-25. (In Bulgarian)
- Dimitrova, L., R. Pavlov (1993). Natural Language and Logic Grammar's Formalisms. In: (K. Chimev and Sl. Shtrakov Eds.) Discrete Mathematics and Applications, Blagoevgrad, 73-94.
- Dimitrova, L., R. Pavlov (1995). Discontinuous Grammars and Some Constituents Movements in Bulgarian Language. In: (Sl. Shtrakov and Iv. Mirchev Eds.) Research in Mathematics, Vol. 5: Discrete Mathematics and Applications, Blagoevgrad, 141-148.
- Dimitrova, L., R. Pavlov, K. Simov (2002). The Bulgarian Dictionary in Multilingual Data Bases. In: Cybernetics and Information Technologies. Vol. 2, Num. 2. ISSN 1311-9702, 33-42.
- Dimitrova, L., P. Rashkov (2009). A New Version for Bulgarian MULTTEXT-East Morphosyntactic Specifications for Some Verbal Forms. In: Shyrokov, Dimitrova (Eds.), Organization and Development of Digital Lexical Resources. Kiev, ISBN 978-966-507-252-2, 30-37.
- Dimitrova, L., M. Todorova (1991). An Attempt for Computer Simulation of Push-Down Automata and Context-free Grammars. In Proc. of the Mathematics and Education in Mathematics (Math and Educ. Math), 210-215.
- Dimitrova, L., M. Todorova (1993). Experimental System for Computer Simulation of Push-Down Automata and Context-Free Grammars. In: International Journal Information Theories and Applications. 1, 8, 26-32.
- Dutsova, R. (2014). Web-based Software System for Processing Bilingual Digital Resources. Intern. J. Cognitive Studies/Études Cognitives. Vol. 14, SOW, Warsaw. ISSN 2392-2397 (online), 33-43, DOI: <http://dx.doi.org/10.11649/cs.2014.004>; <https://ispan.waw.pl> [IF ERIH]
- Dutsova, R. (2015). Web-based Digital Lexicographic Bilingual Resources. Intern. J. Cognitive Studies/Études Cognitives. Vol. 15, SOW, Warsaw, 2015, ISSN 2392-2397 (online), [IF ERIH]
- Kovacheva, St. (2014). How to Incorporate New Knowledge in the History Courses for 4th and 5th Grade Pupils. In: Proc. of the International Conference Digital Presentation and Preservation of Cultural and Scientific Heritage, Vol. 4, Demetra LTD, Sofia, ISSN 1314-4006, 224-234.
- Kovacheva, St. (2015). Electronic Dictionary as a Tool for Integration of Additional Learning Content. In: Journal Cognitive Studies/Études Cognitives, Vol. 15, SOW, Warsaw, ISSN 2392-2397 (online) (<https://ispan.waw.pl>)
- Ludskanov, A. (1971). Einige Fragen der Formalisierung in der Sprachwissenschaft. In: Linguistische Arbeitsberichte, No. 6, 1-51.
- Paskaleva, E., L. Dimitrova, L. Nenova (1981). Automatic Segmentation of Bulgarian Wordforms – Description of the Algorithm. Journal of Automated Systems and Control, No. 4, Sofia, 9-16. (In Bulgarian)
- Pericliev, V. (1995). Empirical discovery in linguistics. Spring Symposium of the American Association for Artificial Intelligence. Stanford University, California, March 1995, 68-73.
- Pericliev, V. (2013). Componential Analysis of Kinship Terminology: A Computational Perspective. Palgrave MacMillan: Basingstoke & New York, ISBN: 9781137031174, 178 p.
- Pericliev, V., R. Valdés-Pérez (1998). Automatic componential analysis of kinship semantics with a proposed structural solution to the problem of multiple models. Anthropological Linguistics, 40(2), Indiana, 272-317.
- Людсканов, А. (1968). Превеждат човекът и машината. Наука и изкуство, София, 157 стр. (Преводи: Paris, Dunod, 1969; Leipzig, Niemeyer, 1971)

# История и развитие на направлението Дигитална хуманитаристика в секция ПИТХН на ИМИ

Калина Сотирова, Анна Самева

Институт по математика и информатика при БАН, София  
[kalina@math.bas.bg](mailto:kalina@math.bas.bg), [sameva@math.bas.bg](mailto:sameva@math.bas.bg)

**Резюме:** Настоящият доклад представя накратко историята и развитието на направлението Дигитална хуманитаристика в секция Приложение на информационните технологии в хуманитарните науки (ПИТХН). Разгледан е контекстът, в който направлението се развива в други западноевропейски страни и са обобщени резултатите от дългогодишната работа на звеното в Института по математика и информатика при БАН.

Дигиталната хуманитаристика (Digital Humanities) навлиза в изследователското поле на Америка и Европа от 1949 г. насам. Широкият обхват на интердисциплинарното поле се покрива с няколко наименования: *Humanities Computing*, *Digital Humanities*, *Heritage Studies*. През 1946, когато компютрите били огромни машини, твърде бавни и сложни за съвременните стандарти, италианският йезуитски свещеник отец Роберто Бурса планира използването на машини при анализ на естествен език за организирането на текстовете на Тома Аквински. През 1949 г. отец Роберто завършва своя пионерски труд *Index Thomisticus*, благодарение на подкрепата на основателя на IBM, Томас Уотсън. Днес този текст е възприет като основополагащ труд, който за първи път свързва информатиката и компютърните науки с хуманитаристиката. Към 2015 година, дисциплината „дигитална хуманитаристика“ се изучава в американските и европейски университети, но у нас стъпките в тази посока са бавни. В контекста на информатичната наука, развивана в Института по математика и информатика, такива стъпки се правят за първи път в периода 2001 – 2003 г., когато директор на института е акад. Стефан Додунеков.

Инициатор и ръководител на направление *Дигитална хуманитаристика* в рамките на ИМИ, от 2001 г. до октомври 2009 г. е д-р Милена Добрева, член на Изпълнителния съвет на Националната комисия на Република България за UNESCO (2006). Мисията на сформирания от нея интердисциплинарен екип, назначен и подкрепен изцяло от акад. Стефан Додунеков, е да развива изследователска, образователна и приложна дейност в областта на цифровизацията на културно и научно наследство.

Д-р Добрева натрупва своя опит в индивидуални изследователски проекти, а през 1992–94 тя печели финансиране от Централноевропейския университет на тема *Computer Supported Reconstruction and Representation of Ancient Manuscripts*. През 1996–97 координира проекта *Text Variety in the Witnesses of*

Medieval Texts: Study from Co-operative Writing Perspective, финансиран от UNESCO. След успешна защита на докторска дисертация и постепенно привличане на млади специалисти, през октомври 2004 г. вече са назначени трима специалисти в различни хуманитарни области, работещи в рамките на секция „Телекомуникации“, под ръководството на д-р Пламен Матеев. Същата година в ИМИ се създава **временно звено** за дигитализация, чиито основни дейности са свързани с обучение на екипа и координация на европейския проект "Knowledge Transfer for Digitisation of Cultural and Scientific Heritage in Bulgaria" (KT-DigiCULT-BG, 6RP на ЕК). Това е първият проект в ИМИ, който поставя началото на интердисциплинарния подход в областта на цифровизацията на културно и научно наследство в България. Временното звено придобива статут на **самостоятелна секция** в ИМИ през юни 2007 г. с името “Приложение на информационните технологии в хуманитарните науки”.

Ръководителят на екипа е и един от инициаторите на създаването на асоциацията SEEDI (South Eastern Europe Digitization Initiative) – с основна цел сътрудничество на специалисти от Югоизточна Европа в областта на дигитализацията на научно и културно наследство. SEEDI организира редовни конференции и реализира общи проекти и до днес. В сътрудничество с Агенция РСИКТ (сега ДАИТС), ФМИ и НБКМ беше създаден XML-базиран каталог на българските ръкописи, съхранявани в България.



*Първият специализиран професионален скенер за книги, закупен за нуждите на  
Лабораторията по цифровизация в ИМИ*

Екипът на М. Добрева организира и участва в редица национални и международни семинари и конференции. Техниката, с която беше оборудвана Лабораторията, беше закупена от ИМИ в рамките на извънбюджетни проекти (скенери, сървър, работни станции, специализиран софтуер за сканиране, софтуер за обработка и съхранение на информацията).

Ръководството на екипа – д-р М. Добрева, а по-късно и д-р Н. Икономов – допринасят за участието на екипа в значими европейски проекти като:

- COMputational TOOLS for the librarian and philological work in Cultural Institutions (COMTOOCI, 2005-2007, партньор),
- Discovering our Past World with Digitised Maps (DIGMAP, 2006-2008, партньор);
- три проекта с национално финансиране (ДАИТС и МОНМ);
- двустранен проект за цифровизация на културно наследство с ИМ на САНУ (2004-2006);
- цифровизация на аудио архиви с финансиране от ЮНЕСКО;
- ИМИ, в лицето на секция ПИТХН, беше асоцииран член на европейските проекти MinervaPlus, MICHAEL, EUROPEANA и DARIAH.

След 2007 ръководители на екипа на секцията последователно са били инж. Н. Икономов, доц. Кр. Иванова, проф. П. Станчев.

През 2009 г. част от екипа на секция ПИТХН участва в международен работен семинар по проекта “PLANETS” (финансиран от Централна европейска инициатива, CEI) на тема „Дългосрочно съхранение на цифровизирани документи”. През същата година, съвместно с Центъра по архитектурознание – БАН и Столична община (програма „Култура”), беше извършена дигитализация на извадки от архитектурен архив на значими сградни проекти и беше подготвена изложба „Архитектурни конкурси в София 1878-1944”. Изложбата беше експонирана в Галерията за чуждестранно изкуство, беше открита от тогавашния министър на образованието и предизвика силен медиен интерес.

Част от дейностите на секция ПИТХН до 2011 г. са:

- 2005. Съвместно с Държавен архив с музей за история на София: дигитализация на архива на синагогата, джамията в София и част от архива на църквата „Св. Неделя“; издаване на компакт диск “София – религиозни пространства”.
- 2005. Асоцииран партньор в проекта MINERVA и съдействие за националното изследване на дигитализацията в България и подготовката на Minerva Global report.
- 2006. Експертно съдействие при осъществяването на изследване на ЮНЕСКО за дигитализацията в югоизточна Европа.
- Дигитализация на архивни фондове, съвместно с Главно управление на архивите; трансфер на знания и обучение на екипа им.
- Невмена нотопис – изследователска дейност и разработка на софтуер за разпознаване в ръкописни паметници (доц. Л. Ласков, понастоящем преподавател в НБУ).
- Дигитализация на математическа периодика, съхранявана в библиотеката на ИМИ и издавана от ИМИ; дигитализация на Физико-математическо списание 1904-1993, Известия на Математическия институт – 1954-1974, Списание „Сердика“ (достъпно онлайн), трудове

- на български математици, Протоколна книга на Физико-математическото дружество – ръкопис, Архив на ИМИ и др.
- Дигитализация на старопечатна периодика, притежание на ЦБ при БАН: Годишници от вестниците „Гайда“, „Мир“, „Лампион“, „Дъга“.
- Обучение на екип от Народна библиотека „Иван Вазов“, Пловдив, и дигитализация на техни документи и книги.
- Дигитализация на личния архив на основателя на БАН, проф. Марин Дринов, достъпен онлайн.
- Дигитализация и подготовка за онлайн публикация на архива на Висша атестационна комисия, достъпен онлайн.
- Дигитализация на музикална периодика – сканиране, обработка и разпознаване във връзка с включване в проект RIPM (1800 – 1950) – Ретроспективен индекс на музикалната периодика, Първото българско музикално списание „Гусла“– 1891 г., Списание „А.С.О.“ (Академически симфоничен оркестър), 1934 г.
- 2008. Участие в създаването на референтния модел за дигитални библиотеки от проекта DELOS.
- 2008-2011. Участие в работна група по дигитализация към Министерството на културата.
- 2011-2013. Участие в Координационното звено на Националния съвет по дигитализация към НДФ „13 века България“.
- 2010. Участие в Контролната комисия на ИКОМОС.
- 2010. Преподавателска дейност в специализирани курсове по дигитализация и дигитални библиотеки със съдействието на Столична община, Пловдивски университет „Паисий Хилендарски“, СУ „Св. Климент Охридски“.



*Дигитализиран личен архив  
на проф. М. Дринов, съхраняван в  
Научния архив на БАН.*



*Дигитализация на архива на синагогата,  
джамията в София и част от архива на  
църквата „Св. Неделя“.*

Във връзка с реструктурирането на БАН, от март 2011 г. секция ПИТХН се трансформира в **Лаборатория по цифровизация на научно и културно наследство** към секция „Информационни системи“ под ръководството на проф. Петър Станчев. Основно научно-приложно постижение на Лабораторията е създаването и прилагането на методологии за цифровизация, съобразена с добрите европейски практики. Основните дейности на Лаборатория ЦНКН включват цифровизация на архивни фондове, на българско математическо наследство, на стара българска музикална периодика, на аудиоархиви, на старопечатни книги и учебителни курсове по дигитализация. До 2013 г., когато лабораторията престава да съществува самостоятелно, са сканирани и обработени повече от 90 000 документи, между които: документи на Главно управление на архивите, стари вестници, фотографски колекции, архив на ВАК, архив на ИМИ.



*Примери на цифровизирани от екипа артефакти*

До днес ИМИ поддържа създадените със съдействието на екипа сътрудничества с Научен архив на БАН, ИБЕ, ИЛ, ИФИ, Етнографски институт с музей, ЦБ–БАН, ИПОИ, НАИМ, Столична община, Главно управление на архивите, НБКМ, Национална библиотека Пловдив, ВАК, ДАИТС, НХГ, НИМ Бояна, СУ, ПУ, множество музеи. Целта е с натрупания опит и знания експертите на екипа, които днес работят във и извън БАН, да се развива изследователската, приложната и преподавателската дейност на направление „Дигитална хуманитаристика“; да се създават мрежи от институции и експерти в областта у нас и в Европа; да се организира професионално обучение за специалисти от институциите за културна памет, специализирани училища и университети, да се съдейства с експертиза за формиране на национална културна политика и др. Екипът трупа преподавателски опит в национални и чуждестранни университети, по програми на ЮНЕСКО, Британски съвет, Софийския университет „Св. Кл. Охридски“, Пловдивски университет „Паисий Хилендарски“, Столична община, НАИМ-БАН, министерство на културата на Италия и програма „Учене през целия живот“. Понастоящем членове на някогашната *Лаборатория по дигитализация* продължава да работи в областта на дигиталните библиотеки, гражданската наука и дигитализация на музейни фондове.

## **Секция Математически основи на информатиката в ИМИ–БАН: Теория на кодирането и криптография**

**Петър Бойваленков**

Институт по математика и информатика при БАН, София  
[peter@math.bas.bg](mailto:peter@math.bas.bg)

Секция Математически основи на информатиката е основана през 1988 година от Стефан Додунков (1945-2012). Сътрудниците на секцията са публикували над 500 статии от 1975 насам в престижни специализирани международни списания в областта на теорията на кодирането и криптографията, защитили са 8 докторски дисертации по математически науки и над 30 дисертации за научната и образователна степен доктор. В момента на щат към секцията са 4-ма професори, 5-ма доценти, 3-ма асистенти и 5-ма докторанти.

Основните тематики, които се развиват в момента, са:

- Кодове и дизайн на полиномни метрични пространства
- Оптимални кодове
- Криптография
- Комбинаторни конфигурации

Като се започне от началото на 90-те години на 20. век, голям брой резултати са свързани с получаването и изследването на различни видове граници за кодове в полиномиални метрични пространства (включително евклидови сфери, реални, комплексни и кватернионни проективни пространства като безкрайни примери, и пространства на Хеминг и Джонсън като крайни примери). През последните години имаме важно развитие по въпроси, свързани с минимаксни енергийни проблеми в полиномиални пространства. Тези задачи са свързани с оптимални и близки до оптималните разпределения на точки в съответните пространства и намират приложения в теорията на кодирането, безжичните комуникации, интегрируемите системи, теорията на изчислителната сложност и др., както и приложения отвъд математиката – като нанотехнологиите например. Отново новаторските изчислителни подходи играят важна роля. Дизайните в полиномиалните метрични пространства привличат напоследък голям интерес поради тяхната важност в апроксимациите на големи системи, по-специално за проблемите с климатичните промени и др. Проблемите, които изучаваме, се отнасят до теоретичните и числовите аспекти на минималните и максималните енергийни проблеми и различните им приложения.

Доказани са необходими и достатъчни условия за оптималност на границите на Левенщайн и в случая, когато тези граници не са оптимални, е разработен

метод за получаване на нови граници. Това изследване е подкрепено от масивни изчисления посредством специално разработени програмни продукти. Понастоящем работим върху обобщения и разширения на границите на Левенщайн за максимални кодове и разработваме техники за получаване на нови граници за кодове със слабо ограничена структура, които имат потенциал да отворят нови направления в областта. Нови универсални граници на енергия на кодове в PMS, получени от Бойваленков, Драгнев, Хардин, Саф и Стоянова (<http://arxiv.org/abs/1503.07228>, приета за печат в *Constructive Approximations*), с използване на новоразработени методи, позволяват оценяването на важни характеристики.

Кодовете, коригиращи грешки, имат практически приложения в защитата на данни, във всички видове комуникационни технологии, системите за контрол и т.н. В поредица от статии, считано от края на 70-те години, са установени редица важни свойства на оптималните линейни кодове, които доведоха до класификационни резултати както в класическия двоичен случай, така и за  $q$ -ичните случаи,  $q > 2$ . Разработени са специализирани софтуерни продукти за подпомагане на научните изследвания в областта. Изследванията продължават, като се конструират нови кодове или се доказват резултати за несъществуване в различни случаи, осигуряващи по-подробна класификация.

Първият проблем, свързан с комбинаторна структура с определени параметри, е проблемът за съществуването. След като той бъде решен, възниква проблемът с класификацията, защото различните структури с еднакви параметри може да притежават различни групи от автоморфизми, подструктури и други свойства, които засягат тяхната пригодност за определени приложения. Накрая, може да се избере най-добрата структура с определена цел. Освен за пряко приложение, комбинаторните структури с малки параметри се използват като съставки в рекурсивни конструкции с по-големи параметри и това ги прави много полезни за конструирането и представянето на всички комбинаторни структури (с точност до еквивалентност) с определени параметри.

Методите, които използваме, са комбинация от математически техники на разсъждение и компютърно подпомагано търсене. Внимателното изучаване на свойствата на определена комбинаторна структура е използвано за разработване на бърз класификационен алгоритъм. Ние разчитаме на съвременните системи за компютърна алгебра (Maple, GAP, Mathematica), но също и на наш собствен софтуер, който обикновено разработваме на C++. Това позволява използване на паралелни внедрявания. Когато е възможно, получаваме резултатите чрез различни методи или различен софтуер, като проверка на верността им.

Наши публикации върху CRC-кодовете вече са намерили различни практически приложения. Кодове с малък радиус на покритие, а именно квази-съвършени кодове и кодове с възможно най-малък радиус на покритие, намират приложения в компресирането на данни, WOM-паметите, излъчването на сигнали във взаимосвързани мрежи и др. Оптичните ортогонални кодове (ООС) и кодовете за избягване на конфликти се прилагат в комуникационните

технологии за обмен с висока скорост, асинхронните трансмисии, мрежовия дизайн, осъществяване на разпръскване на данни, мобилни радиостанции, а напоследък и в т. нар. решетъчна криптография. Съществуването на ООС свързаните класификационни задачи се изследват интензивно от Байчева и Топалова, които получиха важна класификация.

Възможността за загуба на пакети при предаването на данни по интернет поднови интереса към кодовете за възстановяване на заличени символи. Кодовете, коригиращи вмъквания и заличавания, имат приложения и в записите на магнитни носители. Ние имаме значителен напредък в случая на едно изтриване, където т. нар. кодове на Варшамов-Тененголец осигуряват свършени кодове с голяма мощност.

Науката криптография се бори с всички предизвикателства, причинени от уязвимостта на ИКТ инфраструктурата към широк спектър от заплахи и атаки на сигурността. Тя се занимава с разработването на математически инструменти, които съдържат важни градивни елементи за посрещане на неотложни проблеми по сигурността на информацията и със създаването на нови възможности за решаването им. Криптографията се фокусира върху важните цели за осигуряване на поверителност, цялостност, автентикация и недопускане на намеса в електронно предавани данни.

Въз основа на опита ни в решаването на проблеми от Теория на кодирането и наскоро придобитата експертиза в криптографията, ние продължаваме изследвания по следните теми – да предложим сигурна архитектура за ОСДМА мрежи, базирани на криптография на решетки, като използваме класификационни резултати за ООС, различни видове схеми на разпределени механизми за доверие (тясно свързани с дейността на самодуалните кодове, пакети и дизайни), приложения на кодове в криптографски алгоритми за анализ на грешките (които използват кодове за контрол на грешки). Наскоро разработени методи за намиране на специален вид пермутации могат да бъдат използвани за получаване на подобрение в най-добрата известна към момента дифузия в блокови шифри, базирани на обобщени мрежи на Файстел (намирането на пермутациите, които водят до най-добрата дифузия, е отворен проблем).

В наши скорошни публикации в областта на криптоанализа (изучаването на шифротекст в опит да се възстанови оригиналното съобщение) ние изследваме възможностите за приложение на някои съвременни подходи, които могат да се прилагат за модификация (произтичаща от теоретични основания) на добре познатия Data Encryption Standard (DES) и неговите производни. Ние събираме и трупаме опит в използването на съвременни методи за криптоанализ (линейни, алгебрични и т.н.). Такива резултати предупреждават потребители в подходите им при избора и употребата на различни криптографски продукти. Ние възнамеряваме да развиваме нашите изследвания в тази насока в комбинация с техники от кодирането.

Комбинаторните конфигурации са тясно свързани с важни видове кодове, коригиращи грешки, като перфектни кодове, LDPC-кодове, самоортогонални кодове, ООС-кодове с постоянно тегло и др. Те също имат множество

криптографски приложения, като например в схеми за разпределение на тайни, използващи механизмите за разпределяне на секретна информация. Поспециално T-разпределенията в проективни пространства са свързани с линейни шифри, кодове с фиксирана размерност и разпределени кодове. T-паралелизмите на проективни пространства имат криптографски приложения в някои прагови схеми.

В секцията са получени многобройни класификационни резултати за комбинаторни дизайни с дадени автоморфизми, резолюции на дизайни, поддизайни, двойно разрешими дизайни и ортогонални резолюции. Ние класифицирахме разпределения с допълнителни свойства в някои проективни пространства и паралелизми на крайни проективни пространства с дадени групи от автоморфизми.

Секция Математически основи на информатиката поддържа активно международно сътрудничество. Традиционно провеждаме Международна конференция по Алгебрична и комбинаторна теория на кодирането – организиран съвместно с Института по проблеми на предаването на информация на Руската академия на науките, като досега са проведени 14 издания (от 1988 г. насам), а следващото е насрочено за юни 2016 г. в Албена, България. Секцията организира и Международна конференция по оптимални кодове, от която са проведени 7 издания (от 1995 г. насам) и следващото ще се проведе през юни 2017 г. Традиционен е и годишният ни национален семинар по теория на кодирането с над 30 издания досега (последно на 19-22 ноември, 2015 г.).

Секцията организира и школи, като например Приложения на техники от кодове, коригиращи грешки, в криптографски примитиви (Велико Търново, 6-9 октомври 2008 г.), Проектиране и сигурност на криптографски функции, алгоритми и устройства (Албена, 30 юни - 05 юли 2013 г.). Бяхме основни организатори на първата конференция BalkanCrypt (София, 7-8 ноември 2013 г.) и EUROCRYPT 2015 (26-30 април 2015 г.), като последната е сред най-големите международни математически конференции, провеждани някога в България.

## Бойчо Кокинов и развитието на когнитивната наука в България

Красимира Иванова

Институт по математика и информатика при БАН, София  
[kivanova@math.bas.bg](mailto:kivanova@math.bas.bg)

**Резюме:** Представен е приносът на Бойчо Кокинов за развитието на когнитивната наука в България, който може основно да се изрази в: задълбочени научни изследвания и постигнати резултати на световно ниво за разкриване на модела на човешкото мислене и влиянието на контекста и кратковременната памет; създаване на школа по когнитивна наука в България; създаване възможности за комуникации на българските учени и млади изследователи със световната научна общност от областта.

През 2015 г. се навършват 55 години от рождението му. Гордост за Института по математика и информатика е фактът, че той израсна като учен именно тук, в секцията, ръководена от проф. Бърнев.



*„Всяка мечта е постижима“*  
Бойчо Кокинов  
27.12.1960-10.05.2013

Началото на развитието на когнитивната наука в България е неразривно свързано с личността на нашия колега и приятел Бойчо Кокинов. По думите на самия Бойчо Кокинов: *„Когнитивната наука е една интердисциплинарна област, която обединява методите на множество традиционни науки, такива като психологията, философията, лингвистиката, антропологията, информатиката и невронауките. Всички те в един момент осъзнават, че си задават едни и същи въпроси – как е устроен човешкият ум; по какъв начин ние възприемаме света около нас; как разбираме това, за което говорим; по какви закони се извършва човешкото мислене. И, разбрали веднъж, че си задават сходни въпроси, в един момент решават да обединят усилията си. И създават тази интердисциплинарна област, в която идеята е, че, тъй като*

*тези въпроси са много сложни, единственият начин да си отговорим по-успешно на тях е да обединим методите на различните науки, така че фактите и изследванията, които се получават в една област, намират интерпретация в друга и обратно“.<sup>1</sup>*

### **Кратки биографични данни**

Бойчо Кокинов е роден на 27.12.1960 г. в гр. София. Баща му Никола Кокинов е главен счетоводител, а майка му Койна Кокинова е старши преподавател по математика в ПИС „Аврам Стоянов“.

От 6-то основно училище „Граф Игнатиев“, той продължава образованието си в Софийската математическа гимназия. Активно участва във всички математически олимпиади, достигайки до подборния кръг. През 1976 г. се класира на второ място на националната защита на научно-теоретичната конференция по математика. Завършва СМГ през 1978 г. с пълно отличие. Същата година е приет във Факултета по математика и информатика при СУ „Св. Климент Охридски“. По време на следването си е награден със златна значка „За отлично учение“. В характеристиката му, дадена от деканското ръководство, пише: „Отличава се с организаторски способности, безкомпромисност към себе си и колегите, и умение да довежда докрай започнатата работа“.

Научната си кариера започва още като студент, като на 6.11.1984 г. постъпва на длъжност програмист на 4 часа в Института по математика и информатика. И тогава имаше практика студенти, които имат интерес и показват способности да се занимават с наука, да бъдат привлечени за съвместна работа. По този начин работата по магистърската им теза е начална стъпка към по-нататъшното развитие. По думите на самия Бойчо, цялостното му развитие се дължи на уникалния шанс да се срещне с проф. Петър Бърнев на организирания от него семинар „Информатика и психология“ в Института по математика през 1983 година, когато Бойчо е студент трети курс. Както казва Бойчо в прощалното си слово към проф. Бърнев: „Ако не беше уникалното свойство на проф. Бърнев да оценява новото и перспективното и да му дава път, когнитивната наука нямаше да се роди в България, поне не и през 20. век“. Но все пак да не забравяме будистката поговорка „Когато ученикът е готов, учителят ще се появи“.

През 1985 г. Бойчо защитава блестящо своята дипломна работа на тема „Представяне на знанията в човешката памет“ и от 1.10.1985 г. по разпределение е назначен на пълен щат, а през 1986 г. е зачислен за редовен аспирант (сег. докторант) в тогавашния блок С. Научен ръководител му е проф. П. Бърнев. Блок С беше специална форма на аспирантура, в която хората се избираха по показан отличен успех и отличителни качества за провеждане на

---

<sup>1</sup> <http://www.btv.bg/video/shows/otpechataci/videos/dots-boicho-kokinov-za-kognitivnata-nauka.html>

научни изследвания. През 1989 г. е назначен за научен сътрудник III ст., а през 1999 г. е повишен на научен сътрудник I ст.

На 16.1.1991 г. става един от основателите на Българско дружество по когнитивна наука и негов пръв председател.

Самата защита на дисертацията му „AMBR1: Един хибриден модел на човешките разсъждения по аналогия“ се състоя сравнително късно – през 2004 г., като това беше един по-скоро формален акт на признаване на вече натрупания му опит и признание от световната общност в неговата област. Една от причините беше и новостта на направлението, на което той се яви пионер – на практика в България нямаше комисия, която да даде компетентна оценка на неговите постижения. А чисто патриотичните му идеали не му позволяваха да си помисли за евентуална защита другаде.

През 2005 г. в Нов български университет беше обявен конкурс за доцент, който той спечели и премина на работа там.

Бойчо Кокинов е работил в множество международни и национални проекти, свързани с когнитивните процеси и когнитивното моделиране. Публикувал е над 70 статии и студии.

## **Научни интереси и постижения**

Един от основните въпроси, които будеха неговия постоянен научен интерес, беше разкриването на механизмите на човешкото мислене, по-специално на начините на решаване на проблеми, разбирането, вземането на решения по аналогия. Информатичната му нагласа по естествен път го караше да търси решение за тези процесни проблеми в другата основна съставка: човешката памет. Считаше, че разбирането на механизмите на нейното активиране и в частност – конструктивният елемент, който внася епизодичната памет, ще доведе до търсеното обяснение на горните проблеми. В цялостния процес на търсене на решение той обръщаше особено внимание и на дълбоко философската част на феномена и намесата на теорията на отражението чрез разбирането на контекстно-зависимата природа на човешкото познание.

Екипът, с който той работеше, се състоеше главно от неговите ученици – будни младежи и девойки, които намериха по-нататъшна реализация в различни престижни институции в Европа и САЩ.

Основата, на която Бойчо изгради по-нататъшните си изследвания, е създадената **когнитивна архитектура с общо предназначение DUAL**, която интегрира символните и конекционистките подходи на микроиво. Всички когнитивни процеси, моделирани чрез DUAL, се базират на симулиране на колективно поведение чрез взаимодействия на много хибридни микро-агенти. [7,8,9,12,18,21].

Хибридноста в DUAL се проявява по два начина. От една страна тя се базира на разграничението и интеграцията между символния и конекционисткия подход, от друга страна – на декларативния и процесуалния подход. Освен това характерни черти са динамичността и контекстната зависимост.

Цялата обработка и представяне на знания в архитектурата се извършва от т.нар. DUAL агенти. Няма централен изпълнител, който да контролира системата, да отпуска средствата, да решава конфликтите и т.н. Решенията възникват на базата на местните взаимодействия между локалните DUAL агенти. Важна особеност е и влиянието на околната среда върху решението на конкретните DUAL агенти.

Всеки DUAL агент е хибридно образувание, изпълняващо и представителните функции, и обработващите процеси. Всеки агент е сравнително прост, има достъп само до локална информация, и си взаимодейства със съседните агенти. Неговият символен процесор може да изпълнява проста манипулация на символи (дискретни композиционни структури) и да ги прехвърля на други агенти. Другият аспект на процесора е разпространението на активност (непрекъснато добавяно количество) между агентите, с което те се припознават като възли в мрежата.

Скоростта на символната преработка, извършвана от даден DUAL агент, зависи от нивото на неговата активност. Активните агенти работят бързо, по-слабо активните агенти работят бавно, а неактивните агенти не работят изобщо. По този начин всеки агент допринася за общото изчисляване в системата в различна степен. Едновременно с това нивата им на активност се променят непрекъснато, което води и до постоянна промяна на скоростта на символната обработка.

Дългосрочната памет на архитектурата се състои от общия брой на населението на всички постоянни DUAL агенти. Активното подмножество от тях плюс някои временни агенти представляват работната памет на системата. Съдържанието на работната памет се променя динамично, отразявайки промените в околната среда и вътрешния ход на изчисление. Това е още един фактор за гъвкавост и контекстна чувствителност.

На базата на DUAL Бойчо Кокинов разработи системата **AMBR** като **модел на човешкото мислене**, опитвайки се да обедини аналогията, дедукцията и обобщението. Моделът интегрира подпроцеси на достъп до паметта, осъществяване на съответствие (mapping) и трансфер, които работят паралелно и взаимодействат помежду си. [19,25,28]

Следващото развитие включва разработване на **модел на човешката епизодична памет**. Този модел е интегриран с AMBR и демонстрира взаимодействието между памет и разсъждения. Моделът прогнозира, че намесването на един епизод в друг, който няма нищо общо с него, може да се получи поради проявата на две отделни аналогии с трети епизод. Прогнозата на модела намира и експериментално потвърждение, и по този начин разкрива възможни механизми за възстановяване на паметта. [23,37,38,55]

Архитектурата DUAL се използва и при изграждането на **модел на човешката преценка – JUDGEMAP**. Моделът прогнозира, че оценката на дадена ситуация силно може да бъде изкривена поради наличието на нерелевантни факти в околната среда, които въздействат върху решението. Тези прогнози са потвърдени експериментално. [27,29,31-34,36,39-41]

По-късно екипът разработва **модел на човешкото възприятие PEAN**, който разширява AMBR със способности за възприятия и на тази база са проучени взаимодействията между аналогия, памет и възприятие. [42-45]

Бойчо Кокинов работи върху изграждането на **теорията за контекста**, стъпвайки върху разбирането на ролята на контекста при човешкото мислене и влиянието на динамиката на контекста върху човешките познавателни процеси. Теорията включва разкриването на същността на понятието контекст и инициращия му ефект върху разсъжденията, решаването на проблеми, вземането на решения. [26,46-51..73]

В памет на Бойчо, през 2013 г. Александър Петров, който е един от основните участници в реализацията на DUAL, издава книгата „Associative Memory-Based Reasoning: A Computational Model of Analogy-Making in a Decentralized Multi-Agent Cognitive Architecture“ (Lambert Academic Publishing. ISBN 978-3-659-26248-7)<sup>1</sup>, съдържаща текста на дисертационния труд на Александър, както и глава от Георги Петков (последният докторант на Бойчо Кокинов), описваща резултатите от JUDGE MAP.



### **Бойчо и младите**

През 1992 г. Бойчо създава магистърска програма по когнитивна наука в Нов български университет, която е първата университетска програма по когнитивна наука в Централна и Източна Европа.

В Нов български университет преподавателската му дейност е свързана с много курсове, сред които „Въведение в когнитивната наука“, „Мислене, разсъждения и решаване на проблеми“, „Аналогия“, „Познание в контекст“, „Конструктивна памет“. Бил е лектор в различни европейски университети като: Ейотвош-Лоранд университет в Будапеща, Ягелонския университет в Краков, Университета в Атина, Университета в Сиена, Университета в Бремен, Университета във Виена.

В периода от 1999 до 2007 г. е заместник-ректор на НБУ по научноизследователската работа и системата за качество, а между 2010 и 2013 г. е ръководител на департамент „Когнитивна наука и психология“.

Под негово ръководство през 1997 г. се дипломират първите магистри по специалност „Когнитивна наука“ за България. През годините 12 негови дипломанти успешно защитават своите тези, а някои от тях по-късно

---

<sup>1</sup> <http://alexpetrov.com/pub/AMBRbook/>

продължават научната си кариера като докторанти на престижни университети в Европа и САЩ.

През 1998 г. когнитивната наука в България има своя първи доктор. Пред международна комисия със състав Наум Якимов (председател), Здравко Марков, Kenneth Forbus, Dedre Gentner, Keith Holyoak, John Hummel и Pentti Kanerva докторантът на Бойчо Александър Петров представя своя труд „A Dynamic Emergent Computational Model of Analogy-Making Based on Decentralized Representations“. След него през годините следват докторските защиты на още 6-ма ученици на Бойчо, като всяка от дисертациите развива и обогатява общата платформа AMBR с нови възможности.

Той беше блестящ учен не само защото имаше задълбочени научни постижения, но и защото изгради цяла армия от последователи, запалени от неговите идеи.

Докато си кореспондирахме с Александър Петров по повод написването на тази статия, видях неговото мото в края на подписа на писмата му „It is better to light one candle than to curse the darkness“. Бойчо беше човек, който „запали много свещи“. В творческата му автобиография особено място заема списъкът на неговите ученици – защитили магистри и докторанти, показвайки, че той се гордее повече с техните постижения, отколкото със своите. Факт е и обратното – в паметното платно с магистри и докторанти, защитили под неговото ръководство, изработено от самите тях за аудиторията в НБУ, личи особената гордост, с която те, представяйки своите постижения, засвидетелстват неговия талант като учен и учител.

*Магистри по когнитивна наука в НБУ, завършили под негово ръководство:*

- Калина Хаджилиева (Калина Кристоф): магистър (1997); докторат в Станфордския университет; постдокторска специализация в Кеймбридж; доцент в Университета на Британска Колумбия, Ванкувър, Канада
- Андреа Кулаков: магистър (1998); професор в Университета в Скопие
- Володомир Иванченко: магистър (2000); докторат в Университета в Рочестър; софтуерен инженер в Самсунг и Айфлуенс
- Ивайло Влаев: магистър (2000); докторат в Университета на Оксфорд; постдокторски специализации в Университета Уоруик и Лондонски университетски колеж; професор в Университета Уоруик, Великобритания
- Найджъл Мойс: магистър (2001); психолог-консултант в клиника в Оксфорд
- Бойка Братанова: магистър (2003); докторат в Университета на Мелбърн, Австралия; постдокторски специализации в Университета Съри, Великобритания и Университета на Мелбърн, Австралия
- Даниела Раева-Бери: магистър (2003); докторат в Екол Нормал Супериор, Пиза; преподавател в Университета в Бирмингам, Великобритания

- Адриан Нестор: магистър (2003); докторат в Университета Браун, САЩ; постдокторска специализация в Университета Карнеги-Мелън, САЩ; асистент в Университета на Торонто, Канада
- Анелия Миткова: магистър (2004); доктор по военна психология; изследовател в Институт по военна психология, ВМА, София
- Светослав Близнашки: магистър (2007)
- Луиза Шахбазян: магистър (2008)
- Георги Николаев: магистър (2009)

*Защитили докторанти по когнитивна наука в НБУ под ръководството на Бойчо Кокинов:*

- Александър Петров: доктор (1998), постдокторски специализации в Университета Карнеги-Мелън, Университета на Калифорния в Ървин и Университета на Колорадо, САЩ; доцент в Държавния университет на Охайо, САЩ
- Неда Зарева: доктор (2003), мениджър Човешки ресурси в Кауфланд, България и Инвестор.бг АД; изпълнителен директор на „100 foxes Consulting“
- Пенка Христова: доктор (2005), асистент в департамент „Когнитивна наука и психология“ в НБУ
- Георги Петков: доктор (2005), асистент в департамент „Когнитивна наука и психология“ в НБУ
- Милена Мутафчиева: доктор (2007), асистент в департамент „Когнитивна наука и психология“ в НБУ
- Иван Ванков: доктор (2010), асистент в департамент „Когнитивна наука и психология“ в НБУ
- Светослав Близнашки: доктор (2013)

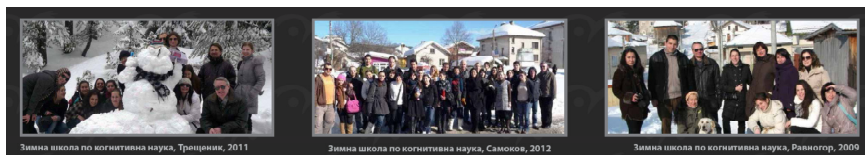
Бойчо Кокинов е главен инициатор за създаването на Международна лятна школа по когнитивна наука в НБУ през 1994. Идеята на школата е да даде възможност за среща на младите изследователи със световните достижения в областта.

Както казва самият Бойчо в едно интервю за „България сега“<sup>1</sup> (23.08.2007 г.): „Създаването на лятната школа, а преди това и създаването на програмата по когнитивна наука в НБУ, залага на основното разбиране, че трябва да се отворим за света и че не можем да правим когнитивна наука само на местно ниво, защото това би означавало да се правим, че правим такава наука. Студентите, а и самите ние, трябва да общуваме с експертите по света в тази област. Лятната школа е инструмент, чрез който всяка година тук идват едни от най-известните учени, с които нашите студенти и ние се срещаме, разговаряме с тях, чуваме новостите в

---

<sup>1</sup> 23.08.2007, България сега, Интервю с Бойчо Кокинов "За лятната школа по когнитивна наука 2007", [http://www.bulgariasega.com/lichnosti\\_usa\\_canada/59.html](http://www.bulgariasega.com/lichnosti_usa_canada/59.html)

когнитивната наука. И нещо, което е изключително важно, дипломирането на нашите студенти става винаги пред международно жури, така че тези хора участват в изпитването на нашите студенти при тяхното дипломиране. Последното не ни позволява да правим някакъв местен стандарт, а ни дава сигурност, че това, с което се занимаваме, е на световно ниво<sup>1</sup>. Школата играе особено голямо значение за създаването на кадри по когнитивна наука не само за България, но и в цяла Източна и Централна Европа. Много от младите изследователи, срещайки се с известни професори, поканени като лектори на школата, намират своето научно призвание и общ език, и по-късно завършват докторантури в престижни университети.



*Част от платното, направено от неговите студенти за аудитория 407 на НБУ,  
която носи неговото име*

## **Обществено влияние**

Още преди да започне самостоятелното развитие на различни мероприятия, популяризиращи постиженията на младата наука, като младши изследовател в секция „Математическо осигуряване“ той беше активен член на организационните, а по-късно програмните комитети на провежданата Школа по програмиране, която не след дълго прерасна в Международна конференция по информатика. Още тогава неговият задълбочен ум и ерудиция в изнасяните от него доклади или участия в дискусии правеха силно впечатление. Участвайки в организацията и провеждането на Школата, Бойчо вижда възможностите, които такива мероприятия създават за „младите български информатици да се срещат и учат от водещи международни учени, включително и отвъд Желязната завеса“<sup>1</sup>.

С развитието на идеите в световен мащаб направлението постепенно започна да добива собствен облик и той беше един от пионерите в неговото изграждане и утвърждаване.

Бойчо Кокинов винаги се бореше за качество на световно ниво. Затова, освен с чисто научните си изследвания, той активно работеше да създаде международна среда, в която българските когнитивисти да могат да общуват и да показват своите резултати.

Участва в програмните комитети, а понякога е и председател и поема домакинството на редица международни конференции в областта на когнитивната наука като: Annual Conference of the Cognitive Science Society

---

<sup>1</sup> Паметно слово от Бойчо Кокинов за проф. Бърнев

(2004, 2005, 2009, 2010, 2011), European Workshop on Context (1997), International and Interdisciplinary Conference on Modeling and Using Context (1999, 2001, 2003, 2005, 2007), Workshop on Modeling and Reasoning about Context (2008, 2010), International Conference on Cognitive Modeling (2007, 2010), International Symposium on Smart Home (2007, 2008) и др.

Активно участва в работата на Европейската конференция по когнитивна наука, която се провежда веднъж на четири години, като през 2011 г. София става домакин на тази конференция<sup>1</sup>, събирайки над 150 учени от Европа и света в момент, когато когнитивната наука е един от трите приоритетни области за научно развитие. Други събития под негова инициатива, проведени в София, са International Conference on Advances in Analogy Research (1998), International Conference on Constructive Memory (2003), International Conference on Cognitive Economics (2005), 2nd International Conference on Analogy (2009) и др.

Дългогодишен член на редакционната колегия на списанието „Cognitive Science Quarterly“, изпълнителен редактор на „Artificial General Intelligence“, съредактор на книжната серия „Atlantis Thinking Machines: Studies in Computational Cognition“.

Членува в Асоциацията по когнитивна наука (САЩ), в Американската асоциация за изкуствен интелект, а през 1997 г. самият той е основател на Централно- и Източно-европейски център по когнитивна наука<sup>2</sup> към НБУ.

## **Заклучение**

В експозето, представящо Бойчо Кокинов като лектор на TEDxNBU 2013<sup>3</sup>, студентите пишат за него:

*„Бойчо Кокинов е човек, който не признава много границите и постоянно ги пресича, защото вярва, че там се случват най-интересните неща. Занимава се с когнитивна наука, която строи мостове между природни, хуманитарни, социални и технически науки. Тя се опитва да отговори на въпроси като как хората възприемат, учат, помнят, мислят, говорят, съзнават като обединява експерименталните методи на психологията и невронауките със събирането на данни в лингвистиката и антропологията, с мисловните експерименти на философията, и с компютърните модели на информатиката, изкуствения интелект и роботиката.*

---

<sup>1</sup> Интервю в "Отпечатъци" по bTV, 29.05.2011 за Европейската конференция по когнитивна наука, проведена в София, 21-24.05.2011 - <http://www.btv.bg/video/shows/otpechataci/videos/dots-boicho-kokinov-za-kognitivnata-nauka.html>

<sup>2</sup> <http://old.nbu.bg/cogs/center/index.html>

<sup>3</sup> Представяне на Б. Кокинов като лектор на TEDxNBU 2013 - <http://tedxnbu.com/2013/03/15/%D0%BB%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%B8-2013-%D0%B1%D0%BE%D0%B9%D1%87%D0%BE-%D0%BA%D0%BE%D0%BA%D0%B8%D0%BD%D0%BE%D0%B2/>

*Той се опитва да разбере как хората решават проблеми, как вземат решения, как си спомнят отминали събития, като навсякъде му се привиждаат аналогии. Убеден е, че правенето на аналогии е фундаментална човешка способност, която участва във всички гореизброени явления. Особено са му интересни несъзнаваните процеси – как хората се влияят от контекста докато разсъждават, вземат решение или си припомнят нещо. Прави опити да изследва експериментално тези несъзнавани процеси и да открие техните механизми, както и да построи компютърни модели, които да имат поведение, сходно с човешкото и дори да предсказват нови явления и особености на човешкото поведение.“*

С особена благодарност към колегите му от Нов български университет, които ни съдействаха при формирането на този текст. И на Александър Петров, който откликна отвъд океана и ми помогна с представянето на научните резултати.

Вярно е, че в лицето на Нов български университет Бойчо намери място, където да разгърне напълно своите новаторски идеи.

Институтът по математика и информатика пък от своя страна беше мястото, където всичко това започна и той получи първите си признания като учен. По думите на самия Бойчо – той е имал уникалното щастие и късмет да бъде ученик на проф. Бърнев.

Лично аз пък имах уникалното щастие да бъдем колеги с Бойчо.

Информатик по душа и човек по сърце – истината е, че много ни липсва.

## **Приложение:**

### **Извадки от научната биография на Бойчо Кокинов<sup>1</sup>**

#### **Участие в международни проекти**

- 6th FP of the EU, ANALOGY: Humans – the Analogy-Making Species. (partners: Cambridge, UCL, CNRS, CNR, U Athens, U Heidelberg, UBC, UCD), 2006-2008, (Principle Investigator, Coordinator).
- 6th FP of the EU, Explaining Religion. (partners: Oxford, Queen Univ. of Belfast, U. Liverpool, U. Helsinki, U. Aarhus, U Groningen, CNRS, U. Salzburg, U. Zurich, U. Brunel), 2006-2008, (Principle Investigator, Coordinator), 2007-2009, (Principle Investigator).
- 6th FP of the EU, RASCALLI: Responsive Artificial Situated Cognitive Agents Living and Learning on the Internet (partners: OFAI, DFKI, ARC, Radon, Sirma), 2006-2008, (Principle Investigator).
- 6th FP of the EU, MIND-RACES: From Reactive to Anticipatory Cognitive Embodied Systems. (Partners: CNR, U Wuerzburg, IDSEA, OFAI, U Lisbon, Lund U., Noze), 2004-2007, (Principle Investigator).
- 5th FP of the EU, EUROCOG project: The Centre of Cognitive Science at NBU: Centre of Excellence, 2003-2005, (Principle Investigator).
- McDonnell Foundation, USA, „Cross-Linguistic Studies in Aphasia“, 1993-2001 (Principle Investigator), in cooperation with Elizabeth Bates (UCSD).

---

<sup>1</sup> Лична страница на Бойчо Кокинов в НБУ - <http://old.nbu.bg/cogs/personal/kokinov/>

- CNR, Italy, „Context and Emergent Cognition“, 1998-2000 (Principle Investigator), in cooperation with Cristiano Castelfranchi (Institute of Psychology-CNR, Rome)
- CNR, Italy, „Cognitive Modelling“, 1992-1997 (Principle Investigator), in cooperation with Cristiano Castelfranchi (Institute of Psychology-CNR, Rome)
- 3rd FP of the European Union, COST, „Real World Commonsense Reasoning“, 1993 (Principle Investigator), in cooperation with Christopher Habel and Christian Freksa (Hamburg University)

### **Участие в национални проекти:**

- НФНИ – „Priming and Context Effects on High-Level Cognitive Processes“, 1994-1997, (основен изследовател)
- НФНИ – „Cognitive Processing of Bulgarian Language in Norm and Pathology“, 1994-1997, (изследовател)
- НФНИ – „Cognitive Modeling by the Symbolic and Connectionist Approaches“, 1991-1994, (основен изследовател)
- МОН – „Интелигентни системи“, 1987-1991, (изследовател)

### **Редактор на книги**

- Kokinov, B., ed. (1995). Perspectives on Cognitive Science. Vol. 1. (1995), Vol. 2. (1996), Vol. 3 (1997), Vol. 4 (1999), NBU Press, Sofia.
- Holyoak, K., Gentner, D., Kokinov, B., eds., (1998). Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences. Sofia: NBU Press.
- Gentner, D., Holyoak, K., Kokinov, B., eds. (2001). The Analogical Mind: Perspectives from Cognitive Science. Cambridge, MA: MIT Press.
- Kokinov, B., Hirst, W., eds. (2003). Constructive Memory. Sofia: NBU Press.
- Dey, A., Kokinov, B., Leake, D., Turner, R. eds. (2005). Modeling and Using Context. Lecture Notes in AI, vol. 3554, Berlin: Springer Verlag.
- Kokinov, B. ed. (2005). Advances in Cognitive Economics. Sofia: NBU Press.
- Kokinov, B., Richardson, D., Roth-Berghofer, Th., Vieu, L. eds. (2007). Modeling and Using Context. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), vol. 4635, Berlin: Springer Verlag. Sofia:
- Kokinov, B., Holyoak, K., Gentner, D., eds. (2009). New Frontiers in Analogy Research. Sofia: NBU Press.
- Kokinov, B., Karmiloff-Smith, A., Nersessian, N., eds. (2011). European Perspectives on Cognitive Science. NBU Press.

### **Реферирани статии:**

1. Kokinov, B. (1988). Associative memory-based reasoning: How to represent and retrieve cases. In: T. O'Shea, V. Sgurev (Eds.), Artificial intelligence III: Methodology, systems, applications (pp. 51-58). Amsterdam: Elsevier Science Publ.
2. Kokinov, B. (1989). About modeling some aspects of human memory. In: F. Klix, N. Streitz, Y. Waern, H. Wandke (Eds.), Man-computer interaction research MACINTER-II. Amsterdam: Elsevier Science Publ. pp. 349-359.
3. Kokinov, B., V. Nikolov (1989). Associative memory-based reasoning: A computer simulation. In: I. Plander, J. Miklosko (Eds.), Artificial intelligence and information-control systems of robots. Amsterdam: Elsevier Science Publ.
4. Kokinov, B. (1990). Associative memory-based reasoning: Some experimental results. In: Proceedings of the Twelfth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum, 741-749.

5. Kokinov, B. (1992). Similarity in analogical reasoning. In: V. Sgurev, B. du Boulay (Eds.), *Artificial intelligence III: Methodology, systems, applications*. Amsterdam: Elsevier Science Publ.
6. Kokinov, B. (1992). Inference evaluation in deductive, inductive and analogical reasoning. In: *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 903-908). Hillsdale, NJ: Erlbaum.
7. Kokinov, B. (1994). Flexibility versus efficiency: The DUAL answer. In: P. Jorrand, V. Sgurev (Eds.), *Artificial intelligence: Methodology, systems, applications*. Singapore: World Scientific Publ.
8. Kokinov, B. (1994). The context-sensitive cognitive architecture DUAL. In: *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
9. Kokinov, B. (1994). The DUAL cognitive architecture: A hybrid multi-agent approach. In: A. Cohn (Ed.), *Proceedings of the Eleventh European Conference on Artificial Intelligence*. London: John Wiley & Sons, Ltd.
10. Kokinov, B. (1994). A hybrid model of reasoning by analogy. In: K. Holyoak, J. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol.2. Analogical connections* (Chapter 5, pp. 247- 318). Norwood, NJ: Ablex.
11. Kokinov, B. (1995). A dynamic approach to context modeling. In: P. Brezillon, S. Abu-Hakima (Eds.), *Proceedings of the IJCAI-95 Workshop on Modeling Context in Knowledge Representation and Reasoning*. LAFORIA 95/11.
12. Kokinov, B., V. Nikolov, A. Petrov (1996). Dynamics of emergent computation in DUAL. In: A. Ramsay (Ed.), *Artificial intelligence: Methodology, systems, applications* (pp. 303-311). Amsterdam: IOS Press.
13. Kokinov, B. (1996). Analogisches Schließen [Analogical Reasoning]. In: *Wörterbuch der Kognitionswissenschaft*. Stuttgart: Klett-Cotta.
14. Kokinov, B., M. Yoveva (1996). Context effects on problem solving. In: *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
15. Kokinov, B., K. Hadjiilieva, M. Yoveva (1997). Is a hint always useful? In: *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
16. Kokinov, B. (1997). A dynamic theory of implicit context. In: *Proceedings of the Second European Conference on Cognitive Science*. Univ. of Manchester Press.
17. Kokinov, B., K. Hadjiilieva, M. Yoveva (1997). The influence of pragmatic distance on context effects. In: B. Kokinov (Ed.), *Perspectives on cognitive science*, Vol. 3. Sofia: NBU Press.
18. Kokinov, B. (1997). Micro-level hybridization in the cognitive architecture DUAL. In: R. Sun, F. Alexander (Eds.), *Connectionist-symbolic integration: From unified to hybrid architectures*. Hillsdale, NJ: Lawrence Erlbaum Associates. (Preliminary version in *Proceedings of the IJCAI-95 Workshop on Connectionist-Symbolic Integration*)
19. Petrov, A., B. Kokinov (1998). Mapping and access in analogy-making: Independent or interactive? A Simulation Experiment with AMBR. In: K. Holyoak, D. Gentner, B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 124-134). Sofia: NBU Press.
20. Kokinov, B. (1998). Analogy is like cognition: dynamic, emergent, and context-sensitive. In: K. Holyoak, D. Gentner, B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 96-105). Sofia: NBU Press.
21. Petrov, A., B. Kokinov (1999). Processing Symbols at Variable Speed in DUAL: Connectionist Activation as Power Supply. In: T. Dean (Ed.), *Proceedings of the 16th International Joint Conference on Artificial Intelligence*. San Francisco, CA: Morgan Kaufman, p. 846-851.
22. Kokinov, B. (1999). Dynamics and Automaticity of Context: A Cognitive Modelling Approach. In: Bouquet, P., L. Serafini, P. Brezillon, M. Benerecetti, F. Castellani (Eds.), *Modeling and Using Context*. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), vol. 1688, Springer Verlag.
23. Kokinov, B., A. Petrov (2000). Dynamic Extension of Episode Representation in Analogy-Making in AMBR. In: *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Erlbaum, Hillsdale, NJ.

24. Holyoak, K., D. Gentner, B. Kokinov (2001). The Place of Analogy in Cognition. In: K. Holyoak, D. Gentner, B. Kokinov (Eds.), *The Analogical Mind: Perspectives from Cognitive Science*. Cambridge, MA: MIT Press.
25. Kokinov, B., A. Petrov (2001). Integration of Memory and Reasoning in Analogy-Making: The AMBR Model. In: D. Gentner, K. Holyoak, B. Kokinov (Eds.), *The Analogical Mind: Perspectives from Cognitive Science*, Cambridge, MA: MIT Press.
26. Kokinov, B., M. Grinberg (2001). Simulating Context Effects in Problem Solving with AMBR. In: V. Akman, R. Thomason, P. Bouquet (Eds.), *Modeling and Using Context. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence)*, vol. 1775, Springer Verlag.
27. Kokinov, B., N. Zareva-Toncheva (2001). Episode Blending as Result of Analogical Problem Solving. In: *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
28. Kokinov, B., R. French (2002). Computational Models of Analogy-Making. In: L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. Macmillan.
29. Grinberg, M., B. Kokinov (2003). Simulation of Episode Blending in the AMBR Model. In: *Proceedings of the 4th European Cognitive Science Conference*. Erlbaum, Hillsdale, NJ.
30. Kokinov, B. (2003). Analogy in Decision-Making and Social Interaction: Emergent Rationality. *Brain and Behavioral Sciences*.
31. Zareva, N., B. Kokinov (2003). Blending of Non-Similar Episodes as a Result of Analogical Mapping with a Third One. In: *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
32. Kokinov, B. (2003). The Mechanisms of Episode Construction and Blending in DUAL and AMBR: Interaction Between Memory and Analogy. In: B. Kokinov, W. Hirst (Eds.), *Constructive Memory*. Sofia: NBU Press.
33. Grinberg, M., B. Kokinov (2003). Analogy-Based Episode Blending in AMBR. In: B. Kokinov, W. Hirst (Eds.), *Constructive Memory*. Sofia: NBU Press.
34. Zareva, N., B. Kokinov (2003). Blending of Spatial Configurations – A Result of Co-Activation or Analogical Mapping? In: B. Kokinov, W. Hirst (Eds.), *Constructive Memory*. Sofia: NBU Press.
35. Nestor, A., B. Kokinov (2004). Towards Active Vision in the DUAL Cognitive Architecture. In: *International Journal on Information Theories & Applications*, 11(1), pp. 9-15.
36. Kokinov, B., P. Hristova, G. Petkov (2004). Does Irrelevant Information Play a Role in Judgment? In: *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
37. Kokinov, B., D. Raeva (2004). Can an Incidental Picture Make Us More or Less Willing to Risk? In: *Proceedings of the 1st European Conference on Cognitive Economics*
38. Kokinov, B. (2005). Can a Single Episode or a Single Story Change our Willingness to Risk? The Role of Analogies in Decision-Making. In: *Advances in Cognitive Economics*. Sofia: NBU Press
39. Petkov, G., P. Hristova, B. Kokinov (2005). How Irrelevant Information Influences Judgment. In: *TR LIP 2005/007 of the Laboratoire d'Informatique de Paris 6*.
40. Hristova, P., G. Petkov, B. Kokinov (2005). The Influence of Irrelevant Information on Price Judgments. In: *Advances in Cognitive Economics*. Sofia: NBU Press
41. Petkov, G., B. Kokinov (2006). JUDGEMAP – Integration of Analogy-Making, Judgment, and Choice. In: *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
42. Hristova, P., B. Kokinov (2006). A Common Mechanism Is Possibly Underlying the Shift in Perceptual and Conceptual Judgment Produced by Irrelevant Information. In: *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
43. Kokinov, B. (2006). Does Human Memory Keep our Past or Does it Construct it and How? In: *Proceedings of the 2nd Russian Conference on Cognitive Science*. Petersburg, Russia.
44. Kokinov, B. (2006). Analogy and Constructive Memory. In: *Proceedings of the International Conference on Cognitive Modelling (ICCM-06)*.

45. Kokinov, B., D. Raeva (2007). A Cognitive Approach to Context Effects on Individual Decision Making Under Risk. In: R. Topol, B. Walliser (Eds.), *Cognitive Economics: New Trends (Contributions to Economic Analysis)*, Vol. 280, Amsterdam: Elsevier. pp. 99-116
46. Kokinov, B. (2007). Can Mass Media and the Educational Institutions Change the Entrepreneurial Culture? A Study of How a Single Story Can Change the Risk-Taking Behavior of Individuals. In: *Global Business and Economics Review*, Vol. 9 (2,3), pp. 286-296.
47. Petkov, G., T. Naydenov, M. Grinberg, B. Kokinov (2007). Building Robots with Analogy-Based Anticipation. In: C. Freksa, M. Kohlhasse, K. Schill (Eds.), *KI 2006, LNAI*, Vol. 4314, Springer Verlag, Berlin, pp. 76-90.
48. Kiryazov, K., G. Petkov, M. Grinberg, B. Kokinov, C. Balkenius (2007). The Interplay of Analogy-Making with Active Vision and Motor Control in Anticipatory Robots. In: M. Butz et al. (Eds.), *ABiALS 2006, LNAI*, Vol. 4520, Springer Verlag, Berlin, pp. 233-253.
49. Mutafchieva, M., B. Kokinov (2007). Does the Family Analogy Help Young Children To Do Relational Mapping? In: *Proceedings of the European Conference on Cognitive Science*. Erlbaum, Hillsdale, NJ. 407-412
50. Petkov, G., K. Kiryazov, M. Grinberg, B. Kokinov (2007). Modeling Top-Down Perception and Analogical Transfer with Single Anticipatory Mechanism. In: *Proceedings of the European Conference on Cognitive Science*. Erlbaum, Hillsdale, NJ. 395-400.
51. Hristova, P., B. Kokinov (2007). Perceptual Learning vs. Context-Sensitive Retrieval: Why do people judge green lines to be shorter/longer than red lines of the same length? Do they perceive them differently or do they retrieve a biased set of alternatives in their comparison set? In: *Proceedings of the European Conference on Cognitive Science*. Erlbaum, Hillsdale, NJ. 377-382.
52. Mutafchieva, M., B. Kokinov (2007). Can Language be Replaced? Physical Representations of Relations Instead of Language Labels in Relational Mapping: Do They Help Young Children? In: *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
53. Kokinov, B., S. Bliznashki, S. Kosev, P. Hristova (2007). Analogical Mapping and Perception: Can Mapping Cause a Re-Representation of the Target Stimulus? In: *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
54. Hristova, P., G. Petkov, B. Kokinov (2007). Objective vs. Subjective Scales: The Challenge that the Scale Type Poses to the JUDGEMAP Model of Context Sensitive Judgment. In: B. Kokinov, D. Richardson, Th. Roth-Berghofer, L. Vieu (Eds.), *Modeling and Using Context. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence)*, Vol. 4635, Berlin: Springer Verlag, 263-276.
55. Kokinov, B., G. Petkov, N. Petrova (2007). Context-Sensitivity of Human Memory: Episode Connectivity and its Influence on Memory Reconstruction. In: B. Kokinov, D. Richardson, Th. Roth-Berghofer, L. Vieu (Eds.), *Modeling and Using Context. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence)*, Vol. 4635, Berlin: Springer Verlag, 317-329.
56. Mutafchieva, M, B. Kokinov (2008). Can Analogy Help Children Make Transitive Inferences? In: *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
57. Kokinov, B., M. Grinberg, G. Petkov, K. Kiryazov (2008). Anticipation by Analogy. In: G. Pezzulo, M. Butz, C. Castelfranchi, R. Falcone (Eds.), *The Challenge of Anticipation. LNAI*, Vol. 5225, Berlin: Springer Verlag, 189-218
58. Mutafchieva, M., B. Kokinov (2009). Can Actions Represent Relations? In: *Proceedings of the 31st Annual Conference on Cognitive Science*. Lawrence Erlbaum, Hillsdale, NJ.
59. Shahbazyan, L., B. Kokinov (2009). The Effect of Objecthood on Processing Efficiency. In: *Proceedings of the 31st Annual Conference on Cognitive Science*. Lawrence Erlbaum, Hillsdale, NJ.
60. Petkov, G., B. Kokinov (2009). Modeling Cued Recall and Memory Illusions as a Result of Structure Mapping. In: *Proceedings of the 31st Annual Conference on Cognitive Science*. Lawrence Erlbaum, Hillsdale, NJ.
61. Feldman, V., B. Kokinov (2009). Analogical Episodes are More Likely to be Blended than Superficially Similar Ones. In: *Proceedings of the 31st Annual Conference on Cognitive Science*. Lawrence Erlbaum, Hillsdale, NJ.

62. Kokinov, B., V. Feldman, G. Petkov (2009). Analogy-Making Automatically Produces False Memories in the Both Situations. In: B. Kokinov, B., K. Holyoak, K., D. Gentner (Eds.), *New Frontiers in Analogy Research*. Sofia: NBU Press.
63. Kokinov, B., I. Vankov, S. Bliznashki (2009). How Analogy Could Force Re-representation of the Target and Inhibition of the Alternative Interpretation. In: B. Kokinov, K. Holyoak, D. Gentner (Eds.), *New Frontiers in Analogy Research*. Sofia: NBU Press.
64. Bliznashki, S., B. Kokinov (2009). Analogical Transfer of Emotions. In: B. Kokinov, K. Holyoak, D. Gentner (Eds.), *New Frontiers in Analogy Research*. Sofia: NBU Press.
65. Feldman, V., B. Kokinov (2009). Anxiety Restricts the Analogical Search in an Analogy Generation Task. In: B. Kokinov, K. Holyoak, D. Gentner (Eds.), *New Frontiers in Analogy Research*. Sofia: NBU Press.
66. Vankov, I., B. Kokinov (2009). Grounding Relations in Action. In: B. Kokinov, K. Holyoak, D. Gentner (Eds.), *New Frontiers in Analogy Research*. Sofia: NBU Press.
67. Kokinov, B., V. Feldman, I. Vankov (2009). Is Analogical Mapping Embodied? In: B. Kokinov, K. Holyoak, D. Gentner (Eds.), *New Frontiers in Analogy Research*. Sofia: NBU Press.
68. Vankov, I., B. Kokinov (2010). The role of action in perceiving and comparing functional relations. In: *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
69. Bliznashki, S., B. Kokinov (2010). Relational Versus Attributional Mode of Problem Solving? In: *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
70. Feldman, V., P. Hristova, B. Kokinov (2010). How Does Anxiety Influence Analogical Mapping? In: *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
71. Hristova, P., B. Kokinov (2011). Anxiety Fosters Relational Encoding. In: B. Kokinov, A. Karmiloff-Smith, N. Nersessian (Eds.), *European Perspectives on Cognitive Science*. NBU Press.
72. Vankov, I., B. Kokinov (2011). Embodied Comparison of Functional Relations. In: B. Kokinov, A. Karmiloff-Smith, N. Nersessian (Eds.), *European Perspectives on Cognitive Science*. NBU Press.
73. Petkov, G., I. Vankov, B. Kokinov (2011). Deduction, Induction, and Analogy Modeled by the Same Mechanisms. In: *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Hillsdale, N.J.: Erlbaum.

# **История на обучението по информатика в Шуменския университет**

**Теодоси Теодосиев**

Шуменски университет „Епископ К. Преславски“,  
Факултет по математика и информатика, Шумен  
[t.teodosiev@fmi.shu-bg.net](mailto:t.teodosiev@fmi.shu-bg.net)

## **1. Въведение**

Имах късмета и щастието да започна преподавателската си работа, когато започна развитието на обучението по информатика в Шуменския университет (тогава Висш педагогически институт). След решението на Правителството за компютризация на образованието и създаване на новата специалност „Математика и информатика“, през 1985 година се създаде сектор „Информатика и Числени методи“ към катедра „Алгебра и геометрия“ с щатен състав от четирима преподаватели. През 1991 година, в резултат на увеличения брой студенти, секторът прерасна в катедра „Информатика и числени методи“. От 1998 година от името ѝ отпадна „...и числени методи“. През 2004 година катедрата се раздели на две самостоятелни катедри: „Компютърна информатика“, и „Компютърни системи и технологии“ с общ щатен състав над 20 преподаватели. По-нататък в текста ще използвам „катедрата“ като събирателно понятие за двете катедри.

## **2. Материална база**

Развитието на материалната база започва с осем компютъра „Правец 82“ (Фигура 1) през 1984 година. Като външна памет се използваша аудиокасети. За връзка с компютрите бяха закупени касетофони „Весна“. През следващата година се появиха и бяха закупени външни флопи-дисквени устройства и дискети 5.25“. Това позволи да започне обучението на студентите по програмиране на Паскал (средата UCSD PASCAL). Организирането на обучението при тези условия беше трудно и предизвикваше недоволството на студентите, на които не можехме да осигурим персонално работно място.

Първата практическа задача, която бе осъществена с новата техника, беше през зимата на 1985 година. Тогава, в условията на тежък „режим на електроподаването“ („2 часа има, 2 часа няма“), трябваше да се направи инвентаризация на Дълготрайните материални активи (ДМА) на Шуменския университет. В тези условия, за едно денонощие, използвайки стая в Областната болница (там токът не спираше), беше извършена автоматизирана обработка на ДМА.

През 1987 бяха закупени първите 16-битови компютри. През 1989 по проект на МОН получихме 30 компютъра „Macintosh“ (Фигура 2). С тези компютри за

първи път започна обучение на операционна система с графичен потребителски интерфейс (OS 7).

Така, започвайки от една компютърна зала с 8 компютъра, в момента Факултетът по математика и информатика разполага с 8 компютърни зали за обучение с над 100 компютъра. Трябва да се отбележи, че ръководството на ФМИ и на ШУ полагаше усилия да поддържа материалната база на добро съвременно ниво.



*Фигура 1. Персонален компютър  
„Правец 82“*



*Фигура 2. Персонален компютър  
„Macintosh“*

### **3. Обучение по информатика**

От 1985 година провеждаме обучение на студенти от специалност „Математика и информатика“, които намериха добра реализация в различни сфери. От 1993 г. във ФМИ започна обучение на студентите и по специалност „Информатика“ (в последните години профилирана като „Компютърна информатика“, „Компютърни информационни технологии“ и „Бизнес информатика“). В началото проблемите бяха основно в липсата на достатъчно компютърна техника и учебна литература, но имахме доста по-мотивирани студенти. Постепенно в годините компютрите станаха повече от студентите пред тях, литературните източници достатъчно (включително в Интернет), но мотивираните студенти, за съжаление, намаляха. Това не означава, че няма какво да се похвалим. На националните студентски олимпиади по програмиране нашите отбори са традиционно сред най-добрите. Като своеобразна оценка на това обучение може да се посочи, че в момента почти

всички преподаватели по информатика в ШУ са наши възпитаници. Голяма част от нашите бивши студенти са реализирани професионално успешно в страната и чужбина.

Не може да бъде отречено, нещо повече – с благодарност трябва да отбележим помощта, която сме получили от по-опитните ни колеги от Софийския университет и ИМИ при БАН. През изминалите години като щатни преподаватели, по съвместителство или на хонорар в ШУ са преподавали редица добре познати на колегията преподаватели като проф. Милко Петков (първи ръководител на катедрата), доц. Маргарита Бърнева, доц. Петър Стойков, проф. Атанас Раденски, проф. Павел Азълов, проф. Петър Бърнев, доц. Димитър Добрев, проф. Венцеслав Вълев, проф. Мария Нишева, Никола Манев, доц. Ангел Дичев и др. Всички те с примера си, със съвети и препоръки са помагали и на нас (преподавателите в катедрата), и на нашите студенти.

Искам специално да кажа няколко изречения в тази връзка за патрона на конференцията – проф. Петър Бърнев. В ШУ той преподава няколко години като хоноруван преподавател. За всеки е ясно, че хоноруваните преподаватели рядко и за кратко присъстват в Университета. Проф. Бърнев, въпреки тези обективни обстоятелства, намери време и сили да помогне за създаването на дружество на Асоциация „Развитие на информационното общество“ в ШУ, да не отказва помощ и консултация на колегите. Той ръководи няколко дипломанти, на които даде добър старт в професионалната кариера.

Наше възпитаници и колеги, преминали през ШУ, в момента работят в чужбина или други университети: проф. Иван Г. Иванов (СУ), д-р Анатоли Начев (Национален университет Galway, Ирландия), д-р Ваня Димитрова (Университета в Лийдс, Великобритания), Йордан Димитров (Великобритания), Никола Николов (Ирландия), Стоян Стоянов (Норвегия) и др.

През последните 30 години преподавателите от катедрата активно участват в обучението на учители по информатика и ИТ. Още през 1985 г. активно се включихме в провеждане на курсове по компютърна грамотност в близките до Шумен области. Участниците в тези курсове, провеждани по места, идваха едновременно с притеснение от новото и непознатото, но и с желание да са в крак с новите технологии. През 2000 година, като участници в проект на МОН, проведохме обучение на над 300 учители по информационни технологии за 5-8 клас. През целия период постоянно участваме в групови и индивидуални курсове по информатика и информационни технологии за повишаване на квалификацията и преквалификация.

По идея на д-р Бисерка Йовчева, ФМИ стана съорганизатор на Националния есенен турнир по информатика „Джон Атанасов“, който се превърна в традиционен и от няколко години вече е международен.

По идея и с активната работа на проф. Станимир Станев във ФМИ е създаден Учебен компютърен музей (Фигура 3). В музея са събрани много ценни експонати, показващи етапите на развитие на компютърната техника. Там студенти и гости могат да се запознаят нагледно с развитието на информатиката в България.



*Фигура 3. Изглед от Учебния компютърен музей*

#### **4. Заключение**

Това беше кратка ретроспекция на развитието на информатиката в Шуменския университет през последните тридесет години. Началото беше трудно, заради непознатото и неопитността. През годините трудностите не намаляха, заради новите предизвикателства на времето. Бързото развитие на информатиката изисква от нас големи усилия, за да се задържим на гребена на вълната.

#### **Благодарности**

Този кратък исторически преглед е частично финансиран по проект РД-08-273/11.03.2015 г.

# Републиканска студентска олимпиада по програмиране

Николай Киров

Нов български университет, София  
[nkirov@nbu.bg](mailto:nkirov@nbu.bg)

**Резюме:** Републиканската студентска олимпиада по програмиране (РСОП) е отборно състезание, в което участват студенти и преподаватели от българските университети, в които се изучава информатика и програмиране. То е стимул за добрите студенти да се усъвършенстват в учебните дисциплини Алгоритми и Програмиране, а също така е обективен критерий за нивото на преподаване на тези дисциплини във висшите ни училища.

**Ключови думи:** Програмиране, университет, състезание, история.

**ACM 1998 Classification Keywords:** K.3.2 Computer and Information Science Education – Computer science education

## 1. Въведение

От 15 години досега, всяка година у нас се провежда **Републиканска студентска олимпиада по програмиране** (РСОП). Началото на национално студентско състезание по програмиране е поставено някъде в края на 20. век, провеждано е няколко години, и след прекъсване е възстановано през 2000 година с името „Междуниверситетски турнир по програмиране“ (МТП). През 2005 година организаторите сменят името на състезанието на „Републиканска студентска олимпиада по програмиране“.

Данните в тази публикация са взети от интернет (Таблица 1), главно от [1] и [2], а за състезанията преди 2000 година не е намерена достъпна информация.

Таблица 1. Състезанията по години

	Година	Домакин	Град	Брой участващи университети	Източник
-	2000	АУБ	Благоевград	6	<a href="http://infoman.musala.com/contests/139/1145/">http://infoman.musala.com/contests/139/1145/</a>
-	2001	НБУ	София	-	-
5 МТП	2002	БСУ	Бургас	10	<a href="http://nikolay.kirov.be/2002/5cp/cp.html">http://nikolay.kirov.be/2002/5cp/cp.html</a>
6 МТП	2003	ШУ	Шумен	10	<a href="http://infoman.musala.com/contests/138/476/">http://infoman.musala.com/contests/138/476/</a>

*Национална конференция по информатика,  
посветена на 80 г. от рождението на професор Петър Бърнев*

7 МТП	2004	ПУ	Пловдив	11	<a href="http://infoman.musala.com/contests/137/475/">http://infoman.musala.com/contests/137/475/</a>
8 МТП 17 PCOП	2005	СУ	София	7	<a href="http://judge.openfmi.net/bcpc/competitors.html">http://judge.openfmi.net/bcpc/competitors.html</a>
18 PCOП	2006	НБУ	София	9	<a href="http://nikolay.kirov.be/2006/Olympiad/index.html">http://nikolay.kirov.be/2006/Olympiad/index.html</a>
19 PCOП	2007	BCY	Варна	9	<a href="http://infoman.musala.com/contests/134/472/">http://infoman.musala.com/contests/134/472/</a>
20 PCOП	2008	СУ	София	12	<a href="http://judge.openfmi.net/bcpc20/index.html">http://judge.openfmi.net/bcpc20/index.html</a>
21 PCOП	2009	ИУВ	Варна	12	<a href="http://infoman.musala.com/contests/132/470/">http://infoman.musala.com/contests/132/470/</a>
22 PCOП	2010	РУ	Русе	13	<a href="http://rsop2010.ami.uni-ruse.bg/">http://rsop2010.ami.uni-ruse.bg/</a>
23 PCOП	2011	ШУ	Шумен	11	<a href="http://www.bcpc.eu/XXIII/">http://www.bcpc.eu/XXIII/</a>
24 PCOП	2012	BCY	Бургас	12	<a href="http://bcpc.eu/XXIV/">http://bcpc.eu/XXIV/</a>
25 PCOП	2013	НБУ,СУ,ТУС	София	14	<a href="http://bcpc.eu/XXV/">http://bcpc.eu/XXV/</a>
26 PCOП	2014	ТУВ	Варна	13	<a href="http://bcpc.eu/XXVI/">http://bcpc.eu/XXVI/</a>
27 PCOП	2015	ЮЗУ	Благоевград	11	<a href="http://bcpc.eu/XXVII/">http://bcpc.eu/XXVII/</a>
28 PCOП	2016	BCY	Варна	-	-

## 2. Участници в състезанието

PCOП е единственото национално състезание по програмиране за студенти. То е независимо състезание, организирано от участниците, като правилата за провеждането му се променят и утвърждават единствено от участващите в него. Те са максимално близки до тези на Международната студентска олимпиада по програмиране на ACM ([ACM International Collegiate Programming Contest](#)). PCOП се провежда всяка година през месец май и се организира от някой български университет – домакин на олимпиадата. Състезанието е отборно. Участват студентски отбори от българските университети, където се изучава информатика и програмиране (Таблица 2). Само СУ и ШУ са участвали във всичките 15 състезания.

*Таблица 2. Участващи в състезанието университети*

	Университет	*	Брой участия	Брой домакинства
АУБ	Американски университет в България	Ч	12	1
БСУ	Бургаски свободен университет	Ч	13	2
ВСУ	Варненски свободен университет „Черноризец Храбър“	Ч	13	1
ВТУ	Великотърновски университет „Св.Св. Кирил и Методий“		10	0
ИУВ	Икономически университет – Варна	Ик	12	1
НБУ	Нов български университет	Ч	14	2
ПУ	Пловдивски университет „Паисий Хилендарски“		9	1
РУ	Русенски университет „Ангел Кънчев“		12	1
СУ	Софийски университет „Св. Климент Охридски“		15	3
ТУВ	Технически университет – Варна	Тех	4	1
ТУГ	Технически университет – Габрово	Тех	10	0
ТУС	Технически университет – София	Тех	9	1
УниБИТ	Университет по библиотекознание и информационни технологии		1	0
УНСС	Университет за национално и световно стопанство	Ик	1	0
ШУ	Шуменски университет „Еп. Константин Преславски“		15	2
ЮЗУ	Югозападен университет „Неофит Рилски“		9	1
	Общо:		-	17

\* Ик-икономически; Тех-Технически; Ч-Частен

### **3. Класиране на университетите**

Всеки участващ университет се представя от един или няколко отбора в състав: ръководител (треньор) и трима състезатели. В официалното класиране на университетите участва най-добре представилият се отбор на съответния университет (Таблица 3). РСОП се провежда на еднакви компютри, свързани в мрежа. Всеки отбор разполага с един персонален компютър. Официални езици за програмиране са Java и C/C++. Провежда се в един ден и е с продължителност 5 часа. Дават се 5-12 задачи, решенията се изпращат по

мрежата на журито за проверка, а резултатите от проверката са видими за всички отбори в реално време. Проверката на изпратените от отборите текстове на програми се извършва от тестова система. До 2015 година тестовата система беше pc<sup>2</sup> (<http://www.ecs.csus.edu/pc2/>), а тази година успешно беше използвана spoj0 (<http://judge.openfmi.net:9080/spoj0/>).

Пълният регламент на PCOP се публикува на сайта на олимпиадата – [vsrce.eu](http://vsrce.eu).

*Таблица 3. Официално класиране по университети*

година	2015	2014	2013	2012	2011	2010	2009	2008
	27 PCOP	26 PCOP	25 PCOP	24 PCOP	23 PCOP	22 PCOP	21 PCOP	20 PCOP
домакин	ЮЗУ	ТУВ	София	БСУ	ШУ	РУ	ИУВ	СУ
класиране								
1	СУ	СУ	СУ	СУ	СУ	СУ	СУ	СУ
2	АУБ	НБУ	НБУ	НБУ	НБУ	АУБ	ШУ	ШУ
3	НБУ	РУ	ТУС	ВТУ	АУБ	РУ	ВСУ	РУ
4	ШУ	ТУВ	ШУ	ИУВ	ТУС	ТУС	ТУС	ТУС
5	РУ	ШУ	РУ	АУБ	ШУ	ШУ	РУ	ПУ
6	ТУВ	ЮЗУ	ИУВ	РУ	ВТУ	ИУВ	НБУ	НБУ
7	БСУ	БСУ	ВСУ	ТУВ	БСУ	НБУ	ИУВ	ВСУ
8	ТУГ	ТУГ	БСУ	ШУ	РУ	БСУ	АУБ	БСУ
9	ВСУ	АУБ	ВТУ	ТУС	ПУ	ВТУ	ВТУ	АУБ
10	ЮЗУ	ВСУ	ЮЗУ	ВСУ	ТУГ	ВСУ	БСУ	ВТУ
11	ИУВ	ВТУ	ТУВ	БСУ	ВСУ	ПУ	ПУ	ИУВ
12		ИУВ	ПУ	ТУГ		ТУГ	ЮЗУ	ТУГ
13		ТУС	ТУГ			ЮЗУ		
14			УниБИТ					

година	2007	2006	2005	2004	2003	2002	2000
	19 PCOP	18 PCOP	8 МТП	7 МТП	6 МТП	5 МТП	?
домакин	ВСУ	НБУ	СУ	ПУ	ШУ	БСУ	АУБ
класиране							
1	СУ	СУ	СУ	СУ	СУ	СУ	СУ
2	ТУС	ШУ	ШУ	ПУ	АУБ	АУБ	PMF <sup>1</sup>

<sup>1</sup> Skopje University

3	ШУ	ПУ	НБУ	НБУ	ШУ	БСУ	АУБ
4	БСУ	ТУС	ПУ	АУБ	ПУ	ПУ	ВТУ
5	ТУГ	БСУ	РУ	ШУ	ИУВ	ТУГ	НБУ
6	РУ	БСУ	БСУ	ИУВ	ЮЗУ	ИУВ	ШУ
7	БСУ	НБУ	УНСС	БСУ	НБУ	ЮЗУ	
8	АУБ	ИУВ		ЮЗУ	РУ	НБУ	
9	ИУВ	ЮЗУ		РУ	ТУГ	ШУ	
10				ВТУ	БСУ	ВТУ	
11				БСУ			

Както се вижда, СУ е безспорен фаворит, спечелил всички първи места до 2015 г., след него е НБУ с 4 втори и 3 трети места, а ШУ и АУБ имат по 4 втори и 2 трети класирания.

#### **4. Организатори на състезанието**

Заслуга за организирането и провеждането на РСОП имат преподаватели по информатичните дисциплини в българските университети и бивши състезатели. Данни за ръководителите на отбори и участници в журито има за следните години: 2002, 2005, 2006 и 2008-2015 – общо за 11 състезания.

Общо 54 колеги са участвали в провеждането на състезанията. В Таблица 5 са имената на тези с най-много участия като треньори и/или членове на журито.

*Таблица 4. Ръководители с най-много участия*

	Име на ръководител <sup>1</sup>	Университет	Брой участия
1	Красимир Манев	СУ-8, НБУ-3, АУБ-1	12
1	Николай Киров	НБУ-9, ЮЗУ-2, БСУ-1	12
3	Бисерка Йовчева	ШУ	11
4	Валентин Бакоев	ВТУ	10
5	Стойчо Стоев	ИУВ	7
5	Димитър Минчев	БСУ	7
5	Емил Келеведжиев	АУБ-5, НБУ-1, УниБИТ-1	7
5	Иван Тренчев	ЮЗУ	7
5	Пламенка Христова	РУ	7

<sup>1</sup> Научните звания и степени са изпуснати, тъй като през годините те обикновено неколккратно са променяни.

5	Стоян Капралов	ТУГ	7
11	Каталина Григорова	РУ	6
11	Милослав Средков	СУ	6

Известни са ми и няколко ръководители на отбори, които като студенти са били състезатели: Антония Ташева – ТУС, Велислав Николов – НБУ, Кристиян Хараламбиев – СУ, Ласко Ласков – НБУ, Милослав Средков – СУ.

### **Заклучение**

Повече информация за ежегодните състезания има на сайта на олимпиадата [1], който се поддържа от 2011 година. Обобщени данни за провежданите досега олимпиади могат да се видят на сайта [3], където авторът има амбицията да събере и публикува и данни за състезанието преди 2000 година.

### **Литература**

1. Сайт на РСОР – <http://bcpc.eu/>
2. Информатически портал ИнфоМан – <http://infoman.musala.com/>
3. Републиканска студентска олимпиада по програмиране (история) – <http://nikolay.kirov.be/2016/WCP/rsop.html>

## **Republican Student Programming Olympiad**

### **Nikolay Kirov**

**Abstract:** Republican Student Programming Olympiad (RSOP) is a competition for teams from Bulgarian universities, in which informatics and programming are studied. It is a good incentive for students to improve their knowledge in the areas of algorithms and programming, but also an objective criterion for the level of teaching of these subjects in the higher Bulgarian schools.

# **Научни доклади**



# Приложение на итеративни принципи и обектно-ориентиран подход при разработката на образователен софтуер

Калинка Калоянова<sup>1</sup>, Андрей Антонов<sup>2</sup>

1: СУ „Св. Кл. Охридски“/ ИМИ-БАН, София,  
[kcaloyanova@fmi.uni-sofia.bg](mailto:kcaloyanova@fmi.uni-sofia.bg)

2: Химикотехнологичен и металургичен университет, София,  
[andrio@uctm.edu](mailto:andrio@uctm.edu)

**Резюме:** Статията представя използването на итеративен процес в комбинация с обектно-ориентиран подход и приложение на UML моделиране при разработката на софтуер. Анализирани са основни елементи на използваните методологии и предимствата на тяхното съчетаване. Посочен е пример за използването на подхода при създаването на софтуер за образователни цели – система за визуализация на 3D обекти.

**Ключови думи:** Software engineering.

## 1. Въведение

Постоянният напредък в развитието на технологиите в последните години поставя нови задачи пред разработката на софтуер. От съвременните софтуерни продукти се изисква да бъдат интерактивни, динамични, с богати възможности и ефективен интерфейс. Това се отнася в особена степен към софтуера, създаден за образователни цели. Въпреки че в световен мащаб съществуват многобройни продукти, все още тези от тях, които се предлагат и на български език, са недостатъчни.

Статията разглежда приложението на съвременни подходи за разработка на софтуерни приложения при създаването на система за визуализация и трансформация на 3D обекти. Засегнати са въпроси, свързани с основните характеристики и предимствата на използваните методологии за разработка.

## 2. Избор на методология

Изборът на методологията за разработка на софтуерния продукт е изключително важен за успешната му реализация. В последните години промените в това направление са изключително динамични.

Първоначално методологиите фокусират върху дейностите, свързани с моделирането на данните (data driven approach) – събиране и анализиране на изискванията, концептуален модел на базата от данни (пълно разбиране на структурата, връзките и ограниченията на базата от данни (БД)), преобразуване на концептуалния модел в логически, логическо и физическо проектиране,

имплементиране и настройка на БД, върху която вече се изгражда приложението.

В последните години все повече се налагат методологии, които фокусират върху цялостния процес на разработка на софтуерното приложение (process driven approaches), което включва:

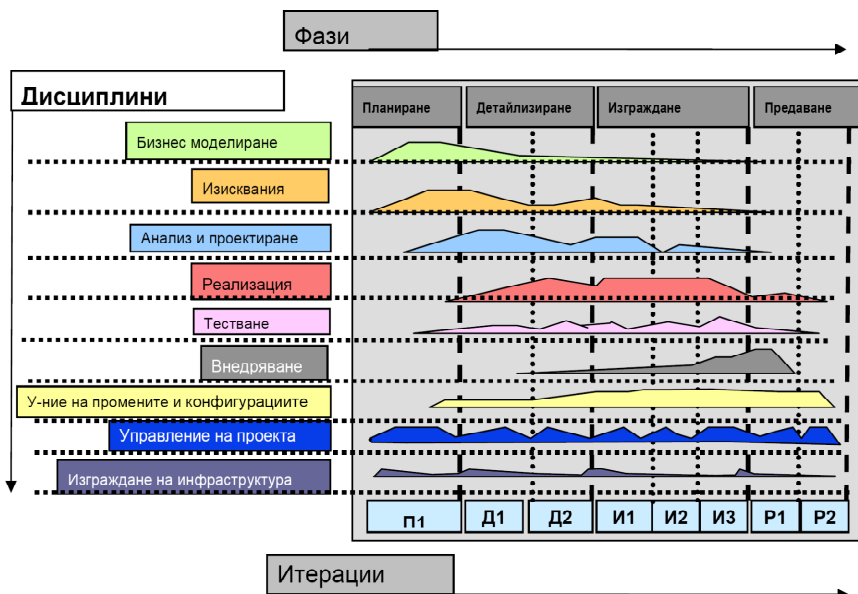
- анализ на осъществимостта;
- събиране и анализиране на изискванията;
- проектиране;
- разработка;
- валидиране и тестване;
- внедряване;
- поддържане.

Една от методологиите, която покрива детайлно всички фази при разработката на софтуерни продукти, е Унифицираният процес (УП).

## 2.1. Унифициран процес

Унифицираният Процес (Unified Process) е итеративен процес за разработка на софтуер [7]. При изграждането на софтуерния продукт чрез тази методология трябва да се премине през четири основни фази – *Планиране (Inception)*, *Детайлизиране (Elaboration)*, *Изграждане (Construction)* и *Предаване (Tranzition)*.

Всяка фаза може да съдържа в себе си една или няколко итерации.



Фигура 1. Унифициран процес – фази, итерации и дисциплини

Фигура 1 представя нагледно процеса: хоризонтално – чрез разпределението на отделните *фази* във времето и вертикално – чрез разпределението на *дисциплините* (дейностите), които поддържа [8].

*Планирането* изисква подготовката на първоначалната визия на проекта, определянето на обхвата и приблизителна начална оценка на усилията за разработката на проекта.

Във фазата на *Детайлизиране* се работи основно по изясняване на изискванията към софтуерния продукт и се създават основните елементи на системния проект, а *Изграждането* е свързано със същинската разработка на софтуера и окончателното фиксиране на функционалността. *Предаването* изисква наличието на вече готова версия на продукта, която се предоставя на клиента/потребителя.

Другата перспектива в разглежданата методология са *дисциплините* (дейностите), които се извършват в процеса на разработка. УП поддържа редица дейности, като основните сред тях са:

- Бизнес моделиране;
- Определяне на изискванията;
- Проектиране;
- Имплементиране;
- Управление на проекта;
- Тестване и вграждане на решението.

Определянето на изискванията безспорно е една от най-важните дейности при разработката на софтуерни продукти. Традиционно те са обособени в две основни групи – функционални и нефункционални [3]. Функционалните изисквания дават спецификацията на функциите, която трябва да поддържа системата, а нефункционалните изисквания – ограниченията. При Унифицирания процес функционалните изисквания се описват чрез т.нар. потребителски случаи (use cases) – формализирани текстови описания, които представят различни сценарии на работа със софтуерния продукт от гледната точка на различни потребители (актьори). Анализът на тези изисквания води до създаване на модел на потребителските случаи (use case model), който служи като модел за уточняване на потребителските изисквания от всички заинтересовани страни.

Този модел на представяне на изискванията е особено удачен, защото той е разбираем за потребителите и клиентите и дава възможност на разработчиците гъвкаво да актуализират спецификацията на изискванията, да отразяват изводите, направени по време на анализа, да преглеждат промените с клиентите и потребителите.

## **2.2. Използване на обектно-ориентиран подход и UML**

Друго важно предимство на Унифицирания процес е неговата интеграция с обектно-ориентирания подход за разработка на софтуер и използването на UML.

Обектно-ориентираните методи успешно съчетават прилагането на решенията за моделиране на приложението и на домейна [10]. Областта на приложение първо се моделира като набор от обекти и връзки. Този модел се използва от системата, за да представи концепции от реалния свят, които по-късно се преобразуват в софтуерни компоненти.

От своя страна целта на UML е да предостави стандартна нотация, която може да се използва от всички обектно-ориентирани методи и е приложима за разработката на широк кръг софтуерни приложения [4].

УП прилага горните две концепции най-вече на основата на артефакта **домейн** модел, който представя концептуалните класове в предметната област, които постепенно се детайлизират в софтуерни класове, описани чрез диаграми на класовете, а функционалността на системата се уточнява с други UML диаграми [6].

Така при разработката на софтуерния продукт се оформят три различни модела [3]:

- Функционален модел – модел на потребителските случаи (представен чрез диаграма на потребителските случаи – use case diagram), описващ функционалността на продукта от гледна точка на потребителя.
- Обектен модел, представени чрез UML диаграми на класовете (class diagrams), които описват структурата на софтуерната система в термините на обекти, атрибути, асоциации и операции. По време на определянето и анализа на изискванията при УП обектният модел започва да се описва като домейн модел, а при проектирането на системата той се прецизира и уточнява.
- Динамичен модел, представени чрез UML диаграми на взаимодействието, диаграми на състоянието и диаграми на дейността, които описват вътрешното поведение на системата. Докато диаграмите за взаимодействие (interaction diagrams) описват поведението като поредица от съобщения, разменяни сред набор от обекти, диаграмите на състоянието (state machine diagrams) описват промени в състоянията на значими обекти и възможните преходи между тях. Диаграмите на дейността (activity diagrams) описват поведението по отношение на последователността, контрола, разклоненията и паралелните дейности.

Дори и само този набор диаграми успяват в повечето случаи да моделират в достатъчна степен софтуерния продукт и са основа на успешната му реализация.

### **3. Интерактивна графична система за визуализации и трансформации на 3D обекти**

На основата на представените предимства, посочените по-горе методологии са използвани при разработката на специфична графична система за образователни цели – **Интерактивна графична система за визуализации и трансформации на 3D обекти** [5].

При прилагането на Унифицирания процес работата преминава през четирите фази, като в две от тях е направено допълнително разделение на итерации (Таблица 1):

*Таблица 1: Фази на разработката на системата*

Дисциплини	Артефакти	Плани- ране	Детайлизиране		Изграждане		Преда- ване
			И1	И2	И1	И2	
Бизнес модел	Домейн модел		с	п			
Изисквания	Модел на потре- бителски случаи	с	п	п	п	п	
	Визия	с	п	п			
	Допълнителна спецификация	с	п	п	п		
	Речник	с	п	п	п	п	
Проектиране	Системен проект		с	п	п	п	
Управление на проекта	План за разработване	с	п	п	п	п	

Разработени са основните артефакти, изисквани от Унифицирания процес, в обем и детайлност, съответстващи на приложението. Таблицата указва кога започва (стартира) работата по съответния артефакт („с“) и в кои фази продължава („п“).

Визията съдържа основната информация за целите и особеностите на системата, като акцентира на възможностите ѝ за подобряване на познанията и разбиранията на крайните потребители относно трансформациите и операциите върху основни обекти като точка, отсечка, равнина и изследването на свойствата им в 3D пространството.

Моделът на потребителските случаи обособява няколко подгрупи потребителски случаи:

- Смяна на свойствата на сцената;
- Смяна на позицията на 3D обекти;
- Създаване на динамични 3D обекти;
- Създаване на 3D конфигурации;
- Показване на допълнителна информация;
- Извършване на допълнителни операции върху 3D обектите;
- Извършване на действия с конфигурации;
- Допълнителна група за помощна функционалност.

Всяка група съдържа няколко на брой потребителски случаи, които са описани в пълен формат. При описанието им е използван итеративният подход – основен положителен сценарий, допълнителни и алтернативни сценарии. Чрез пълното описание на всички потребителски случаи, постигнато на фаза

Изграждане, се обособява и пълната функционалност на системата, която се базира на най-използваните функции, предлагани от подобни системи [2].

Създаденият домейн модел определя базовите концептуални класове, които представят информационното съдържание на системата. Този модел е основата за следващата стъпка – изграждането на класовете и връзките между тях. Основните класове и интерфейси описват различните обекти – точка, права, равнина и др. Наред с основните класове и отговорностите, които всеки 3D обект съдържа, системата борави и с редица помощни класове: за работа със сцената, за запазване и зареждане на конфигурации във файловата система и т.н.

Разработени са около няколко десетки UML диаграми – основно диаграми на последователността, диаграми на дейността и диаграми на класовете, които постепенно детайлизират проектното решение и подпомагат софтуерната реализация.

Системата е разработена на Python. Използването на итеративния подход създаде възможност за постепенно изграждане и детайлизиране на функционалността, като позволи елементи от нея да бъдат готови и използвани на по-ранни етапи.

## **Заклучение**

Съвременните методи за разработка на софтуер предлагат богат инструментариум на специалистите в областта [9], [1]. Същевременно тяхното разнообразие понякога затруднява софтуерните екипи в избора на точната методология за конкретния проект.

Използването на итеративни процеси за разработка, както е Унифицираният процес, е полезно решение в повечето случаи, защото дава възможност на разработчиците гъвкаво да реагират на промените в хода на проекта. Познаването на процесите, както и подходящата им адаптация, са ключови за успеха на проекта.

Настоящата статия дава пример за едно такова приложение, което би могло да се използва с успех и при други проекти.

## **Благодарности**

Работата по статията е частично финансирана от ФНИ на СУ „Св. Кл. Охридски“ – договор 36/2015 г.

## **Литература**

1. Ескенази А., Н. Манева: Софтуерни технологии, КЛИМН, 2006
2. Bantchev B.: A brief tour to dynamic geometry software. сп. Дидактическо моделиране, т.3, 2009/2010г., ISSN:1314-1651/(онлайн издание, <http://www.math.bas.bg/omi/DidMod>)
3. Bruegge B., A. H. Dutoit: Object-Oriented Software Engineering Using UML, Patterns, and Java, Pearson Education, 2008

4. Dunn F., I. Parberry: 3D Math Primer for Games and Graphics development, Wordware publishing, 2002
5. Kaloyanova K., A. Antonov: Object-oriented approach for developing interactive 3D graphics system in education. Proceedings of 7th Annual International Conference on Computer Science and the Education in Computer Science, Sofia, Bulgaria, July 06-10, 2011, pp. 89-100
6. Kaloyanova K.: Design from data: how to use requirements for better information system analysis and design, Proc. of the Int. Conference Informatics in Scientific Knowledge, Varna, June, 26-29, 2012, pp. 189-197
7. Kruchten P.: The Rational Unified Process: An Introduction, Pearson Education, 2004
8. Larman G.: Applying UML and Patterns: An Introduction to Object-Oriented analysis and Design and The Unified Process, 3-rd edition, Prentice Hall PTR, 2004
9. Sommerville I.: Software Engineering, 9th Edition, Addison Wesley, 2011
10. Todorova M., K. Kanev: Educational Framework for Verification of Object-oriented Programs, The Joint International Conference on Human-Centered Computer Environments HCCE'2012, Hamamatsu, Japan, March 8-13, 2012

## **Implementing Iterative Principles and Object Oriented Approach for Educational Software Development**

**Kalinka Kaloyanova, Andrey Antonov**

**Abstract:** The paper presents the use of an iterative process in combination with object-oriented approach and application of UML modeling for software development. The key elements of the methodologies are analyzed as well as the benefits of their combination. An example of using this approach for creating software for educational purposes is shown – a system for visualization of 3D objects.

# **Methodological Approach for Modelling, Simulation & Assessment of Complex Discrete Systems**

**Zlatogor Minchev**

Institute of ICT/Institute of Mathematics & Informatics  
Bulgarian Academy of Sciences, Sofia

[zlatogor@bas.bg](mailto:zlatogor@bas.bg)

**Abstract:** Today's world is constantly evolving towards improving human-environment interaction and thus achieving a better quality of life in the digital age. Apart of this, the proper understanding and perspectives for progressive development of this idea is discoverable in multiple current and future problem areas. The paper outlines a methodological approach for modelling, simulation and assessment from the complex discrete systems perspective. The presented idea is illustrated from both theoretical and practical prospects, noting numerous successful results. Some possible improvements of the proposed approach are finally discussed.

**Key words:** complex discrete systems, modelling, simulation, assessment

## **1. Introduction**

Modern digital 21st century world headlines are constantly evolving in various fields, to note: climate, society, governance, economy, security, industry, technologies, research, education & innovations [1].

The progressive and dynamic understanding of these fields from the human-environment interaction perspective is a rather challenging task, successfully outlined by the General Systems Theory [2]. Apart of this, the possible moments of unexpected changes and perturbations are definitely hard for practical handling and prediction.

A suitable approach, in support to these problems solving, is the complex systems modelling [3]. However, the real world models in the digital reality are discrete and not completely deterministic by design.

Plenty of logical and mathematical formalisms for complex systems machine representation are available in the research literature [3], [4] to mention: graph-based modelling, finite state machines interpretations, non-linear system modelling, discrete event-driven models, object-, agent- based modeling. They are providing different levels of abstraction and details observation for studied systems structure and properties.

What however stands indefinite are the system complexities, related to multiple quantified uncertainties and exploration errors that are difficult for coping in general [5].

Additional support, in this sense, is the expert knowledge usage, implementing different fusion techniques for probabilistic and non-probabilistic reasoning [6], [7].

Another added value in the modelling process is given to fuzzy logic with its modifications [8], [9].

As the examination of the created models is usually conducted in the digital space, both open and closed discrete simulations could be used [10]. These however require different human-in-the-loop roles and computational resources that vary from normal personal multi-core computers to high-performance facilities with millions computational cores.

The assessment of the obtained simulation results is the final moment that needs a suitable framework for correct and profitable evaluation.

Real-time and post-simulation multicriteria evaluations (like: performance variables observation, benchmarks, score cards, etc.) are applicable in accordance with the assessment objectives [11].

Further on in the paper an ad hoc methodological framework for modelling, simulation and assessment of complex discrete systems will be given. Some noticeable applied results and future work discussion are also provided.

## **2. Methodological Framework**

In this section a threefold approach will be considered, including: 2.1. Machine Representation; 2.2. Knowledge Extraction & Fusion; 2.3. Simulation & Results Assessment. All stages are successfully implemented in the *Intelligent Scenario Computer Interface Program for System Analysis* (I-SCIP-SA environment, see Section 3) for complex discrete systems (CDS) practical exploration.

### **2.1. Machine Representation**

The idea behind the machine representation is encompassing the application of Chen's E-R model [12] into an evolving intelligent agent concept [13], used for CDS modelling. This provides a possibility for intuitive understanding and exploration of many real life problems, based on agent modeling approach [10].

In general, all entities are presented by indexed agents –  $A_i$  ( $i = 1 \div k$ ,  $k \in N$ ), equipped with sensors –  $S$  and effectors –  $E$  for interaction with the external environment (see Figure 1). Sensors and effectors at this stage are abstractions, able to operate in physical and information world, in accordance with the exploration necessities.

The communication between agents in this complex system concept is achieved via their multiple indexed causality relations set –  $R$ .  $R$  contains all objects  $R_{ij}$  ( $i \neq j$ ,  $i = 1 \div k$ ,  $j = 1 \div k$ ,  $k \in N$ ) responsible for bilateral agents connectivity, following the model necessities. This requires also a proper understanding of the relations causality modelling role and involves both:  $W$  – weights and  $D$  – delayed transition functions sets, following the ideas of Vester [14].  $W$  and  $D$  elements are of considerable importance for the models discrete simulation implementation (see Section 2.3).

As the practical modelling realization of CDS machine representation is requiring adequate problem space expert knowledge extraction and sensors data suitable fusion, a relevant approach will be further outlined.

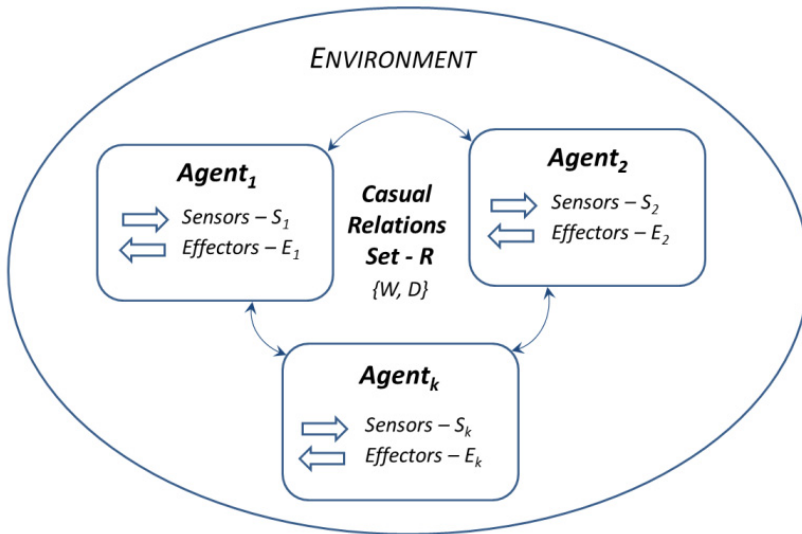


Figure 1. Graphical illustration of the idea for CDS machine representation.

## 2.2. Knowledge Extraction & Fusion

Extracting the knowledge of subject matter experts is an important step for complex environment models creation in various problem areas. The present CDS modelling approach is benefiting from different techniques, including: brainstorming, backcasting, discussions, interviews, questionnaire based surveys or other similar ones [15]. What however is important to note here is that generally knowledge extraction is always uncertain to some extent and at the same time context dependable. This practically means that knowledge extraction should be performed with uncertainty coping [16].

In this sense (see Figure 2) the challenge for experts knowledge and opinions fusion, combined with the simultaneous sensors data dynamics is of key importance.

The assumed agent-based E-R representation of CDS is providing a suitable background, using experts' opinions for initial entities and relations creation, followed by model relations weighing (see Section 2.3). Concerning the practical solution of the problem, the scenario-based contextualization was successfully implemented for both opinions and sensors data fusion [17].

However, it is worth nothing that at this stage, CDS models are providing just a priori, general understanding for the studied complex environment phenomena.

Further on, a dynamic exploration of the created CDS models is organized through simulation, combined with multicriteria results assessment.

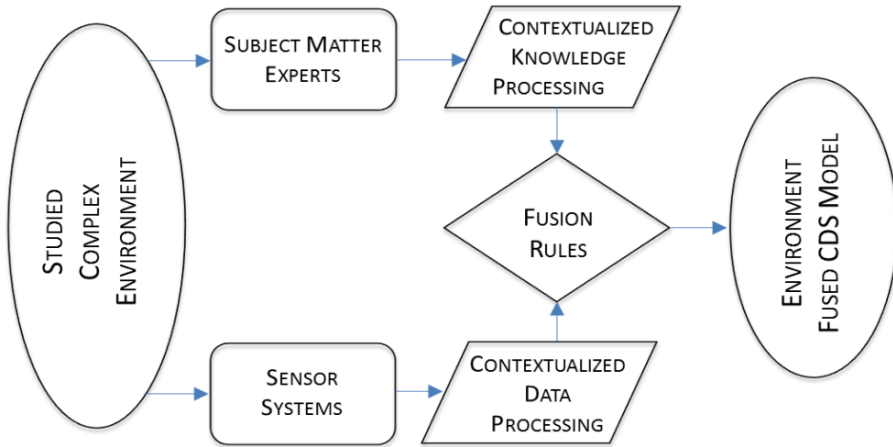


Figure 2. Knowledge extraction & fusion idea for CDS modelling.

### 2.3. Simulation & Results Assessment

Proper understanding of CDS model system properties in multiple scenario context was achieved via discrete event simulation [10]. This provides an opportunity for studying the system dynamics, using an approximation, implementing different delayed transition functions –  $D$  [18] of the relations weights –  $W$ . An important moment to note here is the practical implementation of this approach, directly connected with the model assessment.

Generally, a dual assessment of CDS models is performed for better results understanding (see Figure 3).

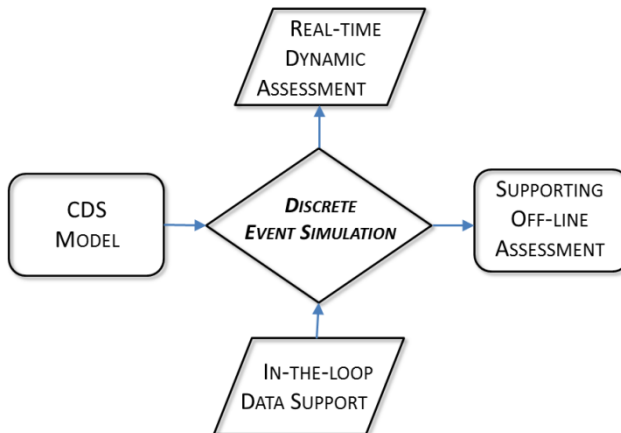


Figure 3. General idea of CDS model discrete simulation and dual assessment with in-the-loop data support.

At first experts knowledge a priori assessment is organized, using a clustered evaluation of the system dynamics trends. Next, during the simulation process, experts' beliefs for the agents bilateral relations weights –  $W$  evolutions is visualized for the entities, using the data from  $D$  functions, encompassing expert and sensor data with in-the-loop simulation support [19], [20].

The second off-line assessment component is organized by using recorded transition functions –  $D$  data trends from the simulation and probabilistic computational judgments [21], [22]. This provides a possibility for results non-linear validation and computational verification. A supplementary participation of the human factor, within the discrete simulation, could also be added with in-the-loop data support, implementing the computer assisted exercise approach in some specific tasks proper exploration [23].

## **2. Software Implementation**

The practical implementation of the presented methodological framework is initially accomplished for I-SCIP-SA, v.1 in Borland Delphi 7 environment [19]. Updated program realizations of I-SCIP-SA, v.2 [24] and v.3 [20] are also produced in Borland CodeGear Delphi 2007 & Embarcadero RAD Studio XE5 editions.

Originally, the approach is using an agent-based organization (see Section 2.1), interconnected with weighed uni- or bi- directional relations (see Figure 4a). Their weights –  $W$  (measured in integer percentages from the interval  $[0, 1]$ ) could also be presented as time delayed transition functions values (times equal to 0, concern static models with single value element array, whilst – arrays of discrete time values are referring the dynamic ones).

Graphically, the entities are noted with labeled rectangle or circle and relations, with arrows, labeled for weight (yellow) and time delayed function values array size (blue).

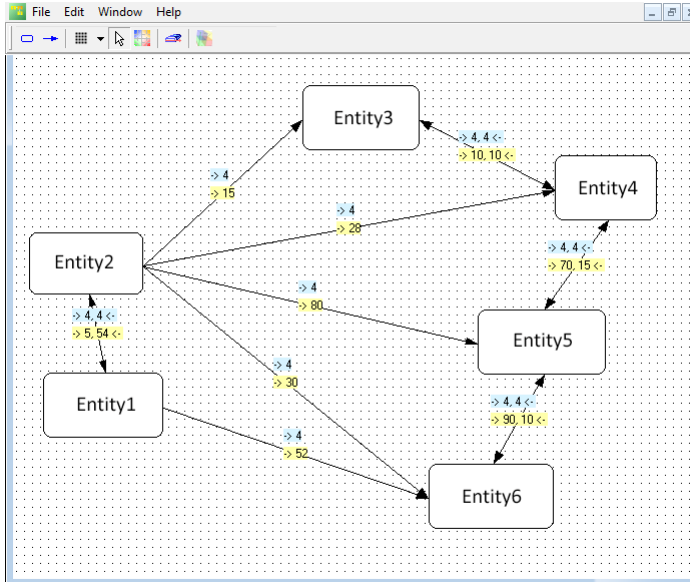
Following the idea proposed in Section 2.2, expert knowledge and sensor data are used for entities and relations definition, together with their discrete simulation realization. In-the-loop data support is providing, at the same time, a possibility for delayed transition functions modification, implementing both human factors activities and sensors information.

The dynamic assessment (after Section 2.3) is visualized into a three dimensional Sensitivity Diagram (SD), using: *Influence* (forward connectivity,  $x$  coordinate), *Dependence*, (backward connectivity,  $y$  coordinate) and *Sensitivity* (absolute difference  $z = |x - y|$ ) values.

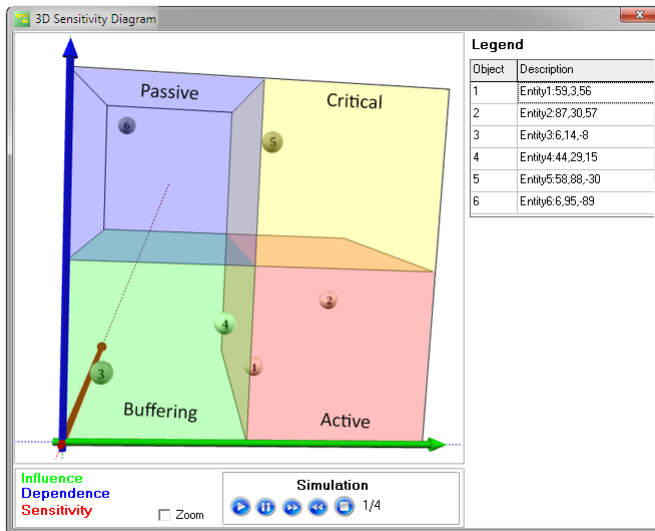
SD is providing four-sector entities classification (in accordance with  $x, y, z$  values): green – ‘buffering’, red – ‘active’, blue – ‘passive’ and yellow – ‘critical’ (see Figure 4b).

Additional, ‘active’ (white, positive  $z$  values) and ‘passive’ (grey, negative  $z$  values) reassessment for each of the entities in a certain sector is also fulfilled. This is directly related to sensitive evaluation of elements towards the  $z$  axis. All entities from the model are visualized in SD with indexed balls.

Supporting off-line results assessments are realized in separate software applications [21], [22].



(a)



(b)

Figure 4. A screenshot of I-SCIP-SA environment with sample model (a) and the resulting four zone Sensitivity Diagram classification (b).

### **3. Noticeable Applied Results**

The presented methodological approach for modelling, simulation and assessment of complex discrete systems have been successfully implemented in multiple research areas during the last ten years.

In the fields of crisis and emergency management the results were mainly used for experimental establishing of a Center of Excellence on Applied Operational Analysis & Joint Training Simulation & Analysis Center – JTSAC [25]. The accomplished research experience was also used in 2010 for the ‘Alliance Strategic Concept 2020’ preparation and is available at the first Bulgarian knowledge portal in that field [26].

Further on, the research efforts were mainly focused in JTSAC successful role for the creation of ‘European Network of Excellence on System Security’ – SysSec [27] with more than eighty partnering institutions all over the world that successfully developed ‘The Red Book: A Roadmap for System Security Research in Europe’ up to 2020 [28].

Some distinguished national results in this sense are related to social networks and smart homes, noting the human factor key digital resilience role [29] within three SysSec supporting research projects: DMU 03/22 [30], DFNI T01/4 [31] and TK 02/60 [32].

Recent achievements in the cybersecurity field are affiliating the Alliance Wales Summit Cyber Policy 2014 updates and the new hybrid challenges exploration [33], marked in the methodological and future threats analysis support to the ‘National Cyber Security Strategy 2020’ ongoing preparation.

### **4. Discussion**

Evidently exploring modern world research headlines with the presented complex discrete systems approach is producing added values to both fundamental studies and societal understandings.

Apart of this, there are some weak points that need to be addressed in the future from human factor and technological perspectives, to note: models simplification due to causal agent representation, difficulties on data fusion sources implementation due to numerous errors, uncertainties and different trend discretization dynamics, limitations of the clustering assessments and the resulting ambiguous visual representations, external data sources interoperability and synchronization problems during discrete simulation.

As the specified methodological weaknesses list is quite complex but not exhaustive at all, many of these gaps are difficult to be surmounted in general. However new results are expected to be developed, in accordance with the explored social and technological context, providing methodological and applied improvements towards the presented approach.

## References

1. HORIZON 2020 in Brief, EC, DG Research & Innovations, 2014, Available at: [https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020\\_inBrief\\_EN\\_Final\\_BAT.pdf](https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_inBrief_EN_Final_BAT.pdf)
2. Bertalanffy, L.: General System Theory: Foundation, Development, Applications, George Braziller, New York, 1968.
3. Sayama, H.: Introduction to the Modeling and Analysis of Complex Systems, Open SUNY Textbooks, Milne Library, 2015.
4. Newman, M. E. J.: Complex systems: A survey, American Journal of Physics, Vol. 79, pp. 800-810, 2011.
5. Oden, T., Moser, R., Ghattas, O.: Computer Predictions with Quantified Uncertainty, SIAM News, Vol. 43, No.9, 2010, Available at: <https://www.siam.org/pdf/news/1842.pdf>
6. Kochenderfer, M.: Decision Making Under Uncertainty: Theory and Application, MIT Lincoln Laboratory Series, 2015.
7. Smarandache, F., Dezert, J. (Editors): Applications and Advances of DSMT for Information Fusion, Vol. 3, American Research Press, Rehoboth, 2009, Available at: <http://fs.gallup.unm.edu/DSMT-book3.pdf>
8. Aliev, A. R.: Fundamentals of the Fuzzy Logic-Based Generalized Theory of Decisions, Studies in Fuzziness and Soft Computing Series, Springer Berlin Heidelberg, Vol. 293, 2013.
9. Atanassov, K. et al.: Novel Developments in Uncertainty Representation and Processing, Advances in Intelligent Systems and Computing Series, Springer International Publishing, Vol. 401, 2016.
10. Borschev, A.: The Big Book of Simulation Modeling, AnyLogic, North America, 2013.
11. Jain, R.: Art of Computer Systems Performance Analysis: Techniques for Experimental Design Measurements Simulation and Modeling, John Wiley & Sons Inc., USA, New York, 2015.
12. Chen, P.: The Entity-Relationship Model-Toward a Unified View of Data, ACM Transactions on Database Systems, 1, pp. 9-36, 1976.
13. Russel, S., Norvig, P.: Artificial Intelligence: A Modern Approach, 3rd Ed., PEARSON, 2009.
14. Vester, F.: The Art of Interconnected Thinking – Ideas and Tools for Dealing with Complexity, MCB-Verlag, München, 2007.
15. Popper, R.: Foresight Methodology, In Georghiou, L., Harper, J., Keenan, M., Miles, I., Popper, R. (Eds.), The Handbook of Technology Foresight: Concepts and Practice, pp. 44-91, Edward Elgar Publishing, Massachusetts, 2008.
16. Kochenderfer, M.: Decision Making Under Uncertainty: Theory and Application, MIT Lincoln Laboratory Series, 2015.
17. Minchev, Z.: Multiple Human Biometrics Fusion in Support of Cyberthreats Identification, International Workshop on Information Fusion, Sofia, September 25, 2015, Published by Int. Journal 'Cyberetics & Information Technologies', Vol. 15, No. 7, pp. 67-76, 2015.
18. Naim, M., Towill, D.: System Dynamics and Learning Curves. In Monaghan, C. and Wolstenholme, E. (Eds.), Proceedings of the 12th International Conference of the System Dynamics Society, Scotland, Stirling, pp. 164-173, 1994.
19. Minchev, Z.: Intelligent Scenario Development for CAX, NATO Science for Peace Security Series, D: Information and Communication Security, Amsterdam, The Netherlands: IOS Press, Vol.12, pp. 6-24, 2007.
20. Minchev, Z., Boyanov, L.: Smart Homes Cyberthreats Identification Based on Interactive Training. In D. Velev (Ed.), Proceedings of ICAICTSEE – 2013, Sofia, University of National and World Economy, pp. 72-82, 2013.

21. Minchev, Z., Shalamanov, V.: Scenario Generation and Assessment Framework Solution in Support of the Comprehensive Approach. In Proceedings of SAS-081 Symposium on Analytical Support to Defence Transformation, RTO-MP-SAS-081, Sofia, pp. 22-1-22-16, 2010.
22. Minchev, Z., Dukov, G., Boyadzhiev, D. et al.: Cyber Intelligence Decision Support in the Era of Big Data, In ESGI 113 Problems & Final Reports Book, Chapter 6, FASTUMPRINT, Sofia, pp. 85-92, 2015.
23. Minchev, Z.: Challenges to Human Factor for Advance Persistent Threats Proactive Identification in Modern Social Networks, In Proceedings of NATO Advanced Research Workshop 'Encouraging Cyber Defence Awareness in the Balkans', Skopje, Macedonia, March 17-19, 2015, Published by 'Information & Security. An International Journal', Vol.34, 2015.
24. Minchev, Z., Petkova, M.: Information Processes and Threats in Social Networks: A Case Study. In T. Atanasova (Ed.), Proc. of Conjoint Scientific Seminar Modelling and Control of Information Processes, Sofia, College of Telecommunications & Post, pp. 85-93, 2010.
25. JTSAC Web Page, [http://gcmarshall.bg/?page\\_id=341&lang=en](http://gcmarshall.bg/?page_id=341&lang=en)
26. Bulgarian Knowledge Portal on OA & CAX Web Page, [http://gcmarshall.bg/?page\\_id=265&lang=en](http://gcmarshall.bg/?page_id=265&lang=en)
27. SysSec Project Web Page, <http://www.syssec-project.eu/>
28. Markatos, E., Blazarotti, D., Minchev, Z. et al.: The Red Book: A Roadmap for System Security Research in Europe, SysSec Consortium, September, 2013, <http://www.red-book.eu/>
29. Minchev, Z.: Human Factor Role for Cyber Threats Resilience, In Handbook of Research on Civil Society and National Security in the Era of Cyber Warfare, Chapter 17, IGI Global, pp. 377-402, 2015.
30. DMU 03/22 FB Page, <https://www.facebook.com/snfactor>
31. DFNI T01/4 Web Page, <http://smarthomesbg.com/>
32. TK02/60 Web Page, <http://cleverstance.com/>
33. Minchev, Z. et al.: Hybrid Challenges in Cyberspace and the Role of the Human Factor, In Proceedings of International Conference 'New Threats for the Regional Security in SE Europe', New Bulgarian University, Sofia, June, 2015, Available at: <https://goo.gl/IXFeRz>

## **Методологичен подход за моделиране, симулация и оценка на комплексни дискретни системи**

**Златогор Минчев**

**Резюме:** В днешния свят интерактивното взаимодействие между човека и средата постоянно еволюира, повишавайки качеството на живот в дигиталната ера. От друга страна, правилното разбиране и наличието на перспективи за прогресивно развитие на тази идея може да бъде установено в множество съвременни и бъдещи изследователски области. Статията представя методологичен подход за моделиране, симулация и оценка от гледна точка на комплексните дискретни системи. Представената идея е илюстрирана както теоретично, така и практически с множество успешни резултати. В края са обсъдени и някои възможни подобрения на предложеното решение.

# Методологии за проектиране и разработване на сигурни уеб услуги: съвременно състояние

Малинка Иванова

Технически университет – София  
[m\\_ivanova@tu-sofia.bg](mailto:m_ivanova@tu-sofia.bg)

**Резюме:** Уеб услугите са софтуерни компоненти, подпомагащи комуникацията между софтуерни приложения и устройства. Широкото им приложение в Интернет поставя изисквания за защитата им от заплахи и атаки и гарантиране на сигурност на предаваните данни. Докладът представя обобщение и анализ на съвременни методологии за проектиране и разработване на сигурни уеб услуги, въз основа на което е направен избор на методология, която да бъде използвана при създаване на софтуерен инструмент.

**Ключови думи:** уеб услуги, сигурност, методологии за проектиране и разработване на сигурни уеб услуги

**ACM 1998 Classification Keywords:** A.1: Introductory and Survey, D.2: Software Engineering – D.2.1: Requirements/Specifications – Methodologies, C.2. Computer-communication Networks – Security and Protection

## 1. Въведение

Уеб услугата се разглежда като софтуерен продукт с определена функционалност, който лесно може да се добавя, премахва или заменя в софтуерно приложение [1]. Често уеб услугата работи самостоятелно или в комбинация с други уеб услуги [2]. Тя е част от архитектурата, основаваща се на услуги, част от облачни структури или други решения, използващи услуги. Установено е, че една голяма част от заплахите и атаките срещу уеб услуги се дължат на не добре проектирани и/или разработени уеб услуги, а друга част се основават на администраторски грешки и неадекватна политика на съответната организация относно сигурността [3]. Затова проектирането и разработването като един от факторите за постигане на сигурност са особено важни процеси, свързани с ефективното използване на уеб услуги и гарантирането на сигурна уеб комуникация в рамките на една организация, няколко свързани организации или в отвореното уеб пространство. Проектирането и разработването на сигурни уеб услуги трябва да се извършва въз основа на стандарти и утвърдени принципи и методи в практиката, както и адаптивно да се отчитат условията и изискванията на конкретния сценарий и контекст. Целта на статията е да се обобщи и анализира състоянието до момента в областта на методологии за проектиране и разработване на сигурни уеб услуги и да се направи избор на подходяща методология, която да се използва при изграждане

на софтуерен инструмент, представящ комплексна методология за осигуряване на сигурност на уеб услуги, отчитаща различни фактори.

## **2. Стандарти**

Уеб услугите използват набор от стандартизирани протоколи, които, правилно приложени, осигуряват необходимата сигурност в комуникационния процес. W3C (World Wide Web Consortium) и OASIS (Organization for the Advancement of Structured Information Standards) препоръчват прилагането на стандарти, свързани със сигурността на уеб услуги, по-важните от които са:

- *W3C Web Services (WS) Policy* – Спецификацията дефинира набор от конструкции/политики за описание на характеристиките на уеб услуги и изискванията към тях [4]. Една политика позволява на уеб услугите да дефинират изисквания към крайните/междинните точки. Тези изисквания включват и правила за запазване на секретните данни, правила за криптиране и използване на сигурни токъни.

- *OASIS WS Security* – Този стандарт позволява реализирането на сигурни SOAP (Simple Object Access Protocol) съобщения чрез криптиране и проверки за цялост и конфиденциалност на съобщение или на част от него [5].

- *OASIS WS Trust* – Стандартът позволява размяна на сигурни токъни за установяване на доверие в различни домейни – предоставя методи за възстановяване и валидиране на сигурни токъни, както и начини за оценяване на доверието [6].

- *OASIS WS SecureConversation* – Използва се съвместно със спецификациите WS Policy, WS Security, WS Trust с цел реализиране на сигурната комуникация между клиента и услугата [7].

- *OASIS WS ReliableMessaging* – Спецификацията дава описание на протокол за надеждно предаване на съобщения. Стандартът позволява на уеб услугите и клиентите да си имат доверие, когато съобщението е изпратено и трябва да бъде доставено до съответната страна [8].

- *OASIS WS Atomic Transaction* – Спецификацията описва три специфични протокола за договаряне при осъществяване на атомична транзакция. Стандартът установява възможност за обратно връщане на транзакциите в случай на неуспешно предаване на съобщението [9].

## **2. Методологии за изграждане на сигурни уеб услуги**

Тъй като уеб услугите са софтуерни системи, то се предполага, че тяхното изграждане може да се извърши по някоя от известните методологии за разработване на софтуер (например обектно-ориентирано проектиране и анализ). На практика се оказва, че услугите притежават редица особености и изследователите търсят специални методологии за тяхното реализиране чрез три подхода – (1) използване на съществуваща методология и усъвършенстването ѝ, (2) комбиниране на няколко методологии, (3) създаване на нова методология. Много добър обзор на известните методологии за

проектиране на уеб услуги е направен от Ramollari, Dranidis и Simons [10], които посочват наличието на: (1) методология на IBM – анализ и проектиране, ориентирано към услуги, (2) процес на IBM – моделиране, ориентирано към услуги и архитектура, (3) процес на IBM – унифициран процес, ориентиран към услуги, (4) методология на Sun Microsystems – непрекъснато качество, (5) методология на Everware-CBDi – архитектура, ориентирана към услуги и инженерство, (6) методология, предложена от Erradi, Anand и Kulkarni – архитектурна рамка, ориентирана към услуги, (7) методология на Papazoglou и van den Heuvel – проектиране и разработване, ориентирано към услуги, (8) методология на Erl – проектиране и анализ, ориентирани към услуги, (9) методология за моделиране на бизнес процеси, (10) методология на Jones – архитектура, ориентирана към услуги. В статията на Ramollari, Dranidis и Simons нищо не се съобщава за сигурността на уеб услуги и дейностите, които трябва да бъдат предвидени в тези методологии.

Stojanovic, Dahanayake и Sol представят подобрен метод на компонентно-базирано разработване, съобразен с особените характеристики на уеб услугите [11], макар че методът не отчита важноста на сигурността при създаване на уеб услуги. Компонентно-базираното разработване предлага разделяне на вътрешната и външната перспектива, но е ограничено от гледна точка на гъвкавост, композиране и преизползване. Изборът на услуга се извършва динамично, въз основа на набор от политики. Използването на инсталирани компоненти не позволява същия начин на преизползване и динамично поведение. При компонентното разработване трудно се използва обектното моделиране и тази трудност се увеличава с нарастване на сложността на модела.

Друг метод за проектиране, ориентирано към услуги въз основа на бизнес процеси е представен от Lo и Yu, но също не се препоръчват дейности по сигурността [12]. Моделирането на бизнес процеси разглежда бизнес процеси, идентифицира начини за подобряването им и определя бариерите, които спъват постигането на бизнес целите. Резултатът от моделирането е практически план за действие с подробно описание за постигане целите на организацията и реализиране на промяна, съобразно стратегическата ѝ мисия.

Понастоящем изследователски групи от академичната общност и от софтуерната индустрия продължават да търсят подходи за проектиране и реализиране на сигурни уеб услуги. Налице са няколко публикации, в които по-пълно или частично са описани дейности, свързани със сигурността на уеб услуги или са представени методологии.

Microsoft предлага набор от инженерни дейности, класифицирани в четири групи, които да подпомагат софтуерните разработчици, които трябва да се съобразят с изискванията за сигурност [13]: (1) *Идентифициране на целите на сигурността* – определяне на изискванията за сигурност; (2) *Разбиране на заплахите* – установяване на възможните заплахи за конкретния сценарий и контекст чрез прилагане на техниката моделиране на заплахите. Поставените цели допринасят за приоритизиране на заплахите и уязвимостите. Чрез изградения модел на съответната заплаха, софтуерните разработчици се стараят

да избегнат уязвимостите.; (3) *Прилагане на установени принципи, шаблони и практики* – така могат да се елиминират редица проблеми на сигурността. Адаптивното им прилагане, в зависимост от конкретния сценарий и случай, е фактор за реализиране на ефективна сигурност.; (4) *Прилагане на инженерни техники, касаещи сигурността през целия жизнен цикъл на уеб приложението* – това са специфични дейности, чрез които се осигурява реализирането на целите по сигурността. За сигурността трябва да се мисли през целия жизнен цикъл на продукта, като се започне от дефиниране на целите за постигане на сигурност. Преминава се през моделиране на заплахите, което подпомага проектирането на услугата. Достига се до проверка и контрол при проектирането, кодирането, внедряването и тестването на сигурността на услугата, които са задължителни мерки, подобряващи цялостната реализация на сигурността. Microsoft използва инженерен процес за разработване на уеб услуги, включващ следните фази: Планиране, Изисквания и анализ, Архитектура и проектиране, Разработване, Тестване, Внедряване, Поддръжка [13], през които фази са предвидени дейности по реализиране на сигурността.

Joо и Woo прилагат добре известната методология на обектно-ориентирано проектиране и анализ при разработване на сигурни уеб приложения и по този начин за сигурност се мисли не само в крайната фаза на разработване на софтуера, но и през целия жизнен цикъл на създаването му [14]. При обектно-ориентирания анализ и проектиране разработчиците са фокусирани върху създаването на класове и обекти, които съответстват на реални или концептуални обекти в дадена организация. Нивото на гранулираност се определя от класа, който се характеризира с ниска степен на абстракция в сравнение с моделирането на услуги. Обектно-ориентираното проектиране и анализ е итеративен процес, състоящ се от фазите: Бизнес моделиране, Изисквания, Анализ и проектиране, Реализиране, Тестване и Внедряване. Авторите предлагат усъвършенствана методология на обектно-ориентирано проектиране и анализ с фокус върху сигурността, включваща три фази: Анализ на изискванията, Анализ и проектиране на системата и Реализация. Изискванията по сигурността са добавени в трите фази под формата на нефункционални изисквания.

От друга страна, автори като Parazoglou и van den Neuvel считат, че при разработване на уеб услуги и на уеб приложения, ориентирани към услуги, не е възможно директно да се приложат отделни методологии като: обектно-ориентиран анализ и проектиране, компонентно-базирано разработване и моделиране на бизнес процеси [15]. Тези три методологии не съдържат три ключови елемента, характерни за приложения, ориентирани към услуги: услуги, възможност за създаване на композиции от услуги и изграждане на компоненти, реализиращи услуги. Те не дават възможност и за разработване на разпределени и хибридни услуги, доставяне на услуги и управление на услуги. Всяка една от тези три методологии, приложени по отделно при разработване на уеб услуги се оказва неуспешна. Parazoglou и van den Neuvel считат, че е необходим нов подход при разработване на уеб услуги и те предлагат методология за проектиране и разработване, базирана на услуги [15], чрез комбиниране на концепции от обектно-ориентиран анализ и проектиране,

компонентно-базирано разработване и моделиране на бизнес процеси. Методологията за проектиране и разработване, базирана на услуги се състои от фазите: Планиране, Анализ и проектиране, Конструирание и тестване, Обезпеченост, Внедряване, Изпълнение и наблюдение. Методологиите от [13], [14] и [15] са показани чрез Фигури 1, 2 и 3, като са отчетени и дейностите по реализиране сигурност на уеб услугите.

Дейности	Поддейности	Дейности по сигурността
Планиране		
Изисквания и анализ	<ul style="list-style-type: none"> <li>- Функционални изисквания</li> <li>- Нефункционални изисквания</li> <li>- Технологични изисквания</li> </ul>	- Цели на сигурността
Архитектура и проектиране	<ul style="list-style-type: none"> <li>- Указания за проектиране</li> <li>- Архитектура и преглед на проектните решения</li> </ul>	<ul style="list-style-type: none"> <li>- Проектиране на указания</li> <li>- Моделиране на заплахи</li> <li>- Проверка и контрол върху проектирането на сигурността</li> </ul>
Разработване	<ul style="list-style-type: none"> <li>- Тестване на отделни части</li> <li>- Преглед на кода</li> <li>- Ежедневно изграждане</li> </ul>	<ul style="list-style-type: none"> <li>- Проверка и контрол върху кода относно сигурността</li> </ul>
Тестване	<ul style="list-style-type: none"> <li>- Тестване при интегриране</li> <li>- Системно тестване</li> </ul>	- Тестване на сигурността
Внедряване	<ul style="list-style-type: none"> <li>- Преглед на внедряването</li> </ul>	<ul style="list-style-type: none"> <li>- Проверка и контрол върху внедряването</li> </ul>
Поддръжка		

*Фигура 1. Методология на Microsoft [13]*

Дейности	Поддейности	Дейности по сигурността
Анализ на изискванията	<ul style="list-style-type: none"> <li>- Функционални изисквания</li> <li>- Нефункционални изисквания</li> </ul>	<p>-----</p> <ul style="list-style-type: none"> <li>- Привилегии на администратора</li> <li>- Сертифициране</li> <li>- Оторизация на потребител</li> <li>- Конфиденциалност и цялост на данните</li> </ul>
Анализ и проектиране на системата	<ul style="list-style-type: none"> <li>- Функционални изисквания</li> <li>- Нефункционални изисквания</li> </ul>	<p>-----</p> <ul style="list-style-type: none"> <li>- Права на достъп</li> </ul>
Реализиране	<ul style="list-style-type: none"> <li>- Функционални изисквания</li> <li>- Нефункционални изисквания</li> </ul>	<p>-----</p> <ul style="list-style-type: none"> <li>- Контрол на достъпа</li> </ul>

*Фигура 2. Методология на Joo и Woo [14]*

Дейности	Поддейности	Дейности по сигурността
Планиране по проекта	<ul style="list-style-type: none"> <li>- Цели, правила, процедури</li> <li>- Формулиране на изискванията</li> </ul>	
Анализ и проектиране	<ul style="list-style-type: none"> <li>- Сценарии на бизнес процеси</li> <li>- Идентифициране и определяне на бизнес процеси и услуги</li> </ul>	<ul style="list-style-type: none"> <li>- Политики за сигурност</li> </ul>

Конструирание и тестване	<ul style="list-style-type: none"> <li>- Кодирание на уеб услуги и бизнес процеси</li> <li>- Тестване</li> </ul>	<ul style="list-style-type: none"> <li>- Мерки за сигурност при разработването</li> <li>- Функционално тестване, тест за сигурност</li> </ul>
Обезпеченост	<ul style="list-style-type: none"> <li>- Оценяване, класифициране, остойностяване на услуги</li> </ul>	
Внедряване	<ul style="list-style-type: none"> <li>- Регистриране в цифрово хранилище</li> <li>- Рекламирање</li> </ul>	
Изпълнение и наблюдение	<ul style="list-style-type: none"> <li>- Свързвания и обръщения към услугите</li> <li>- Управление и наблюдение</li> </ul>	

*Фигура 3. Методология на Papazoglou и van den Heuvel [15]*

### **Дискусия и заключение**

От направения литературен преглед се вижда, че до момента са налице редица методологии за проектиране и разработване на уеб услуги. Методологии за създаване на сигурни уеб услуги са разработени въз основа на съществуващи подходи, неориентирани към услуги, чрез адаптирането им към спецификата на уеб услугите (методологията на Joo и Woo, методологията на Papazoglou и van den Heuvel) или чрез добавяне на дейности по сигурността към методологии, ориентирани към услуги (методологията на Microsoft). Методологията на Microsoft е проверен в практиката инженерен процес за реализиране на сигурни уеб услуги, с доказан успех при изграждане на софтуер, ориентиран към услуги. Методологиите на Joo/Woo и Papazoglou/van den Heuvel са по-скоро концептуални, предлагащи постигане на определена сигурност на уеб услугите. Методологията на Joo/Woo е разработена въз основа на методологията на обектно-ориентирано проектиране и анализ, и макар, че тя е верифицирана чрез един пример, показващ разработване на онлайн система за банкиране, дейностите по сигурността са свързани само с нефункционалните изисквания на софтуерната система. Методологията на Papazoglou/van den Heuvel комбинира няколко методологии, неориентирани към услуги. Тя не е верифицирана в практиката и показва в общ план в кои фази при разработване на софтуера трябва да се отразят факторите по сигурността. Сравнение на дискутираните методологии за реализиране на сигурни уеб услуги е показано чрез Таблица 1.

Таблица 1

Методология за реализиране на сигурни уеб услуги	Използвана методология	Верифициране на практика	Добавени дейности по сигурността
<i>Microsoft</i>	Инженерен процес за разработване на уеб услуги	Да	По пет основни дейности
<i>Joo/Woo</i>	Обектно-ориентирано проектиране и анализ	Пример	По три основни дейности
<i>Papazoglou/ van den Heuvel</i>	Комбиниране на концепции от обектно-ориентиран анализ и проектиране, компонентно-базирано разработване и моделиране на бизнес процеси	Не	По две основни дейности

В контекста на софтуерния инструмент, който трябва да се разработи и който трябва да представя комплексно решение за доставяне на сигурни уеб услуги (тук се включва методология за разработване на сигурни уеб услуги, правилно администриране, съответни политики на организацията) се избира методологията на Microsoft за разработване на уеб услуги, с предвидени дейности по сигурността поради няколко причини:

- Това е методология, отчитаща спецификата на уеб услугите като софтуерни системи и е създадена като инженерен процес за разработване на уеб услуги (а не за разработване на обекти/класове, компоненти или процеси);
- Включва дейности по сигурността през целия жизнен цикъл на разработване на софтуерния продукт (а не само на етап проектиране, или през отделни фази на разработване);
- Прилага се подход на разбиране на заплахите и атаките чрез тяхното моделиране, приоритизиране и комплексно действие;
- Базирана е на натрупана практика, експерименти и ефективно действащи решения (използва предишен опит, знания и научни разработки);
- Методологията е разработена от софтуерни инженери в Microsoft с натрупан опит и непрекъснато се прилага при решаване на практически проблеми (а не се верифицира чрез няколко примера или чрез концепции).

В заключение трябва да се отбележи, че сред изследователите няма единно мнение относно подходяща методология за проектиране и разработване на сигурни уеб. Въпреки многобройните разнородни предложения от страна на академичната общност и софтуерната индустрия, тази тема продължава да бъде дискутирана и изследвана, тъй като се търси ефективно решение.

## Литература

1. W3C Working Group: Web Services Architecture, 11 February 2004, <http://www.w3.org/TR/ws-arch/>.
2. Claro, D., P. Albers, J. Hao: Web Services Composition, Semantic Web Services, Processes and Applications, Volume 3, Editors: J. Cardoso, A. P. Sheth, ISBN: 978-0-387-30239-3 (Print), 978-0-387-34685-4 (Online), 2006, pp. 195-225.
3. Gordeychik, S.: Web Application Security Statistics, The Web Application Security Consortium, 2010, <http://projects.webappsec.org/w/page/13246989/Web%20Application%20Security%20Statistic>.
4. Web Services Policy 1.5 – Framework W3C Recommendation, 04 September 2007, <http://www.w3.org/TR/ws-policy/>.
5. Web Services Security: SOAP Message Security 1.0., March 2004, <http://docs.oasis-open.org/wss/2004/01/oasis-200401-wss-soap-message-security-1.0.pdf>.
6. WS-Trust 1.4, OASIS Standard, 2 February 2009, <http://docs.oasis-open.org/ws-sx/ws-trust/v1.4/os/ws-trust-1.4-spec-os.pdf>.
7. WS-SecureConversation OASIS Standard, 1 March 2007, <http://docs.oasis-open.org/ws-sx/ws-secureconversation/200512/ws-secureconversation-1.3-os.html>.
8. Web Services Reliable Messaging Version 1.2 OASIS Standard, 2 February 2009, <http://docs.oasis-open.org/ws-rx/wsrml/200702/wsrml-1.2-spec-os.html>.
9. Web Services Atomic Transaction Version 1.2 OASIS Standard, 2 February 2009, <http://docs.oasis-open.org/ws-tx/wstx-wsat-1.2-spec.html>.
10. Ramollari, E., D. Dranidis, A. J. H. Simons: A Survey of Service Oriented Development Methodologies, In Proceedings of the 2nd European Young Researchers Workshop on Service Oriented Computing, 11-12 June 2007, Leicester, UK, pp. 75-80.
11. Stojanovic, Z., A. Dahanayake, H. Sol: An Evaluation Framework for Component-Based and Service-Oriented System Development Methodologies, Chapter III, Advanced Topics in Database Research, Volume 3, 2004, IGI Global, pp. 45-69.
12. Lo, Amy, E. Yu: From Business Models to Service-Oriented Design: A Reference Catalog Approach, In Proceedings of 26th International Conference on Conceptual Modeling, Auckland, New Zealand, November 5-9 2007q. ISBN: 978-3-540-75562-3 (Print), 978-3-540-75563-0 (Online) pp. 87-101.
13. Microsoft: Chapter 1: Security Fundamentals for Web Services, <https://msdn.microsoft.com/en-us/library/ff648318.aspx>.
14. Joo, K., J.-W. Woo: Development of Object-Oriented Analysis and Design Methodology for Secure Web Applications, International Journal of Security and its Applications, Volume 8, No.1, 2014, pp. 71-80.
15. Papazoglou, M. P., W.-J. van den Heuvel: Service-Oriented Design and Development Methodology, International Journal of Web Engineering and Technology, Volume 2, Issue 4, July 2006, pp. 412-442.

## **Methodologies for Design and Development of Secure Web Services: The Current State**

**Malinka Ivanova**

**Abstract:** Web services are software products that support communication among different software applications and devices. They are wide used via Internet and this fact leads to emerged requirements about their defence against vulnerabilities, threats and attacks and guarantees for secure data transactions. The paper summarizes and analyses the contemporary methodologies for design and development of secure web services and it shows a choice of a methodology for realization of secure software systems during their whole lifecycle. The chosen methodology will be used in development of a software tool.

# **Password Manager with 3-Step Authentication System**

**Zhelyazko Petrov, Razvan Ragazan**

University of Westminster, London

[z.petrov@my.westminster.ac.uk](mailto:z.petrov@my.westminster.ac.uk), [razvan.ragazan@my.westminster.ac.uk](mailto:razvan.ragazan@my.westminster.ac.uk)

**Abstract:** A big factor in online protection is the password complexity. The strength of this variable is determined by how fast an attacker can guess it. To increase the effectiveness of this vector the proposed password manager design has been successfully implemented and tested and have proven to be effective against main stream brute force techniques. The 3 step authentication system guarantees 100% access to physical assets in certain scenarios which makes it a great candidate for establishing a safer environment for home and business users. We have achieved successful results in terms of encryption strength and log-in procedure complexity, but more work needs to be done on system footprint and specifically memory dump based attacks.

**Keywords:** security, passwords, encryption, Hash, SMS Gateway, Java

## **1. Introduction**

In the modern cyber world data protection is essential part of a system. The equivalent of our physical key locks in the digital world is the password, and at home or at work the user is interacting with many doors. At home we have google, facebook, twitter, ebay, paypal, online banking etc and just like in the physical world we need different password for each one or we doom our privacy and security. This fact escalates another issue which is the user memory limitation. It is known that modern password guessing techniques are very efficient when it comes to guessing dictionary based passwords such as “thebigbrownfox”. To overcome this issue we need to use some randomness in our passwords and this makes them very hard to remember like the password: “GLx1lt)\*&5“. In a productive environment the administrators goal is to limit the user access only to the resources they have been allocated and passwords are one of them. In the office an employee would have to remember password to access his terminal, work email, ftp server, web server, VPN connection etc. This limits the possible complexity of the password and poses a great risk from offline based attacks such as brute force attack with rainbow table or hardware based attacks. We have designed and develop a java based application which takes care of everything for the user. Our password manager have 3-step authentication mechanism which includes password hashing, asymmetric data encryption and token based authentication. Because we have used only java in-built features the program is platform independent and can be run on any platform that supports java. The attack vectors are kept to minimum by enforcing strong anti-brute force policies during log-in procedures and offline based attacks.

## 2. Concept

The main focus is to establish a safe and easy to use environment. A sequence of events unique to the user have to complete successfully before the user is granted access to his password database. The first step is to authenticate with a *master password*. The master password is salted and hashed and then the result is kept in the configuration file along with the username. In order to keep the user passwords protected we have to encrypt the file were they are stored (there is a separate file for each user). To do that we are using asymmetric encryption and we keep the private key on a removable memory device such as a USB key. The third and final step is to successfully complete a randomly generated number request send to the user's mobile phone. The numbers are generated each time the user tries to log in. Once the 3 steps are complete the user is presented with an easy to use database editor where he can add and modify description, links, account names and passwords.

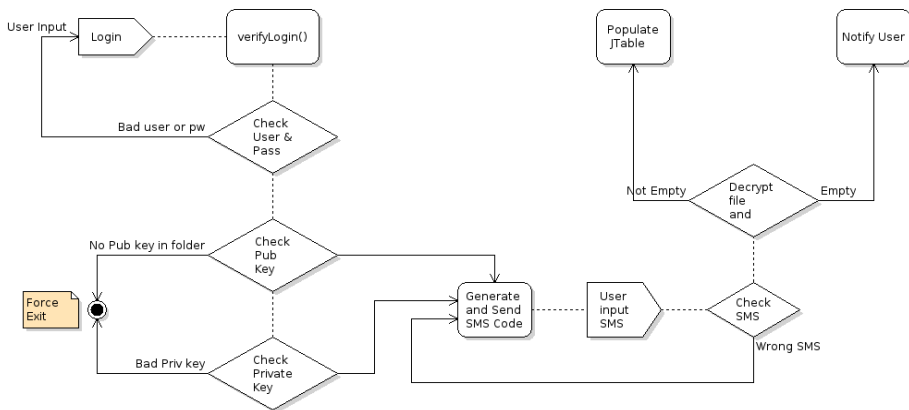


Figure 1. Authentication logic

### 2.1. Implementation

The GUI significantly increases the ease of use compared to some other solutions like CLI based software which aims at speed. The *Register* form is very self-explanatory, and in order to increase the security level a password policy is in place. The admin can easily set the required strength which is hard-coded and have to be done before compilation. Once registered and logged in, an encrypted database file is created. The main working window is a simple database editor with some extra capabilities such as the *Open URL*, *Copy Password* and *View Log* buttons.

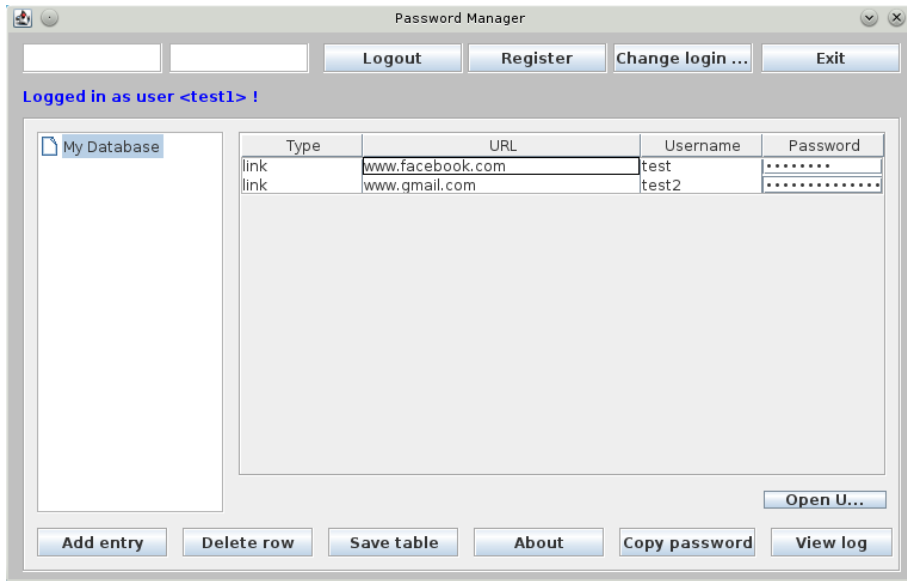


Figure 2. User space interface screenshot

## 2.2. Password Strength

The password strength policy is implemented in the *checkStrength* function which uses 5 different variables to determine the security level:

- length
- number of digits
- number of upper case characters
- number of lower case characters
- number of special characters

There are 40 special hard-coded characters (more can be added), 26 letters twice and the 10 digits which makes it a total of 92 characters. When the minimum password length is set to 10 this makes it a total of  $10^{92}$  possible combinations. If an attacker is trying to brute force a plain MD5 hash at a rate of 1 million keys per second he will need 2 million years to complete all combinations. The function increments the number of characters used and presents the user with a *total score*.

## 2.3. Random password generator

Inside the *new entry form* we have implemented a *Random passes* button which uses an array of 92 characters and a random number generator to create a password. This method of password generation is superior to using combinations of words, digits and symbols and is the recommended way of using the program. The new generation of rainbow tables include music lyrics, quotes, literature, forum posts and pretty much every kind of written content from the web because of modern web scraping software such as *Burp Suite*[1] or *Crawler4j*[2] and can make password guessing much more

efficient. The use of rainbow tables significantly increases the chances of the password being cracked in a feasible period of time using FPGA chips[3], GPU clusters [4] or botnets.

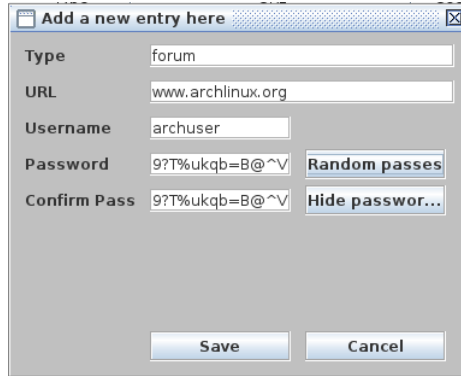


Figure 3. User data-base editor screenshot

## 2.4. Files

The program is using a single configuration file *passMan.conf* which stores the username and the hash of the master password. This file have to be manually created. For each user a new database file is created with the name of the user followed by *.enc* extension. The file is encrypted using the public and private key of the user. The only way to open this file is to use the set of keys. We are using RSA[5] algorithm and the default value is set to 4096 bit long key in order to minimise the chances of any brute force attempts.

## 2.5. Logging

We have implemented a logging system which is intended for administration only. It logs every significant user action such as date, username, account creation, log in attempts etc. The information is currently saved in a plain text file, but is only at testing stages. In a productive environment it will be available only to admins and the logging information will be saved on a *Command and Control Server (C&C)*. In case of malicious activities on the account, the administrators will have the ability to analyse the events and evaluate the damage.

## 2.6. Password cloaking

Inside the database the user cannot see the passwords. They have to use the *Copy Password* button in order to copy the password to the clipboard and *Ctrl + V* into the login form. This concept allows the user to use randomly generated passwords and even if someone have a direct sight at the user monitor they won't be able to read the password. In a productive environment administrators might choose not to reveal the password to the user and the *Open URL* button is a concept of a launcher which will eliminate the need for the user to actually know the password. It is using a system

call to load the link from the database into a browser. Username and passwords auto-fill can be implemented.

### 3. Components

#### 3.1. Bcrypt

To hash the master password we are using *bcrypt*[6]. This is a cross platform file encryption utility which implements the *Blowfish Algorithm*. It supports passphrases between 8 and 56 characters and hash value is up to 448 bit long. On top of that it can add *salt* and *slowness* to the hash which is exactly what we have been looking for. There is other solutions which can be implemented such as *PBKDF2*, *scrypt* and *Simple S2K*, but the ease of use of *bcrypt* and the *slowness* factor which can be manually selected are the main factors behind our choice. The function have java implementation and will work out of the box without the need to download and install anything. We just have to add the java file to our project and call the hashing function from anywhere. There are 2 parameters which we can adjust, the salt length and the slowness rounds which can be anywhere between 4 and 31 (the default is 10). If we increase the slowness factor above 15 there will be a noticeable delay in user space, so we have decided to leave the default parameters. We are using *bcrypt* to hash the master password and add salt and to also match the password of the user. Every time there is a log in request *Bcrypt.checkpw* will compare the inputted password with the one in the configuration file.

#### 3.2. SMS Gateway

The sms feature is implemented by using an sms gateway server and a dongle. The gateway is called *gammu*[7] and is a free to use program. At this stage this functionality is implemented only on the local machine and the daemon is listening for any incoming requests to send a message. The *generateAndSendCode* function is using an array of 10 digits and the size of the message can be adjusted. When the functions is called an RNG produces and stores the digits into a string. A second string with the phone number of the user is fetched and the two are send along with a system call which generates the call that injects new sms request.

```
String toShell=String.format("gammu-smsd-inject TEXT %s -text\n%s\"",phoneNumber,codeMessage);  
Process runningP = Runtime.getRuntime().exec(new String[]{"bash", "-c",toShell});1
```

The user have only 1 attempt to enter the correct code, if the authentication fails a new code will be generated and send. Although this feature significantly slows down the user experience, it infinitely increases online brute force attempts. We have seen

---

<sup>1</sup> Source code snippet – server have to use the inject command to send an sms.

similar token based authentication mechanisms in online banking and various other platforms and it has proven to be efficient. In working environment this feature will be implemented in a C&C with encrypted connection and a second pair of keys for the number generation. In the website you can find detailed information on how to setup the sms server and in case you would like to try the program without this feature we have provided a second version without it.

### **3.3. Encryption**

*Java.Security.\** allows us to use native RSA encryption. The obvious choice over here is the asymmetric encryption method. By encrypting the database file with 4096 bit long key we leave no practical chance for a successful brute force attempt. The recommended key size by *NIST[8]* is 2048 bits, but since 1024 bit keys are now *crackable* we decided to go for the maximum possible value. The only drawback of this is that it is a CPU intensive task and the user have to wait for the key to be computed, however on most modern systems using Intel i5+ it takes less than 5 seconds to complete on creation and less than 3 seconds on login. Once created the public key is saved in the user folder, in a working environment it is recommended that is not saved in the same location as the encrypted file (location can be adjusted before compilation). When generating the keys for first time, the user have to choose where to keep the private key. This key have to be secure at all times and it is a bad idea to keep on the workstation. On log in the user is presented with a file browser in order to locate the private key. If this fails a force shutdown will occur and login have to start again. This is a another online brute force attempt elimination. Once the user is successfully logged we read the data from the encrypted file in bytes format and use *doDec* function to decrypt the information and feed it into a string which displays the information in the database. When we save data into the file we call the *doEn* function to the encryption and save the bytes into the file.

### **3.4. USB Storage**

Since we have to keep the private key of the user in a secure location, the best way is to not have it inside the system. An external storage is the ideal candidate for this authentication step. When we keep the key on a USB memory for example we leave no trace of it inside the machine. This significantly increases the protection level because even a core dump will fail to allocate the key. In a productive environment extra security mechanisms can be implemented such as *USB keyboard integration* and encryption on the USB memory itself. This method of authentication requires physical access to the USB storage and significantly decreases the attack vectors.

## **4. Testing**

We have conducted several tests on the program resilience. There are 3 factors we have tested: master password hash decryption time, encrypted file decryption file and offline password extraction methods. To test the hash we have compared against similar password storage implementation method in Firefox Browser. The browser stores hashes of the passwords, key and data in the same folder. If an attacker finds

this files using free tools he can extract the passwords *instantly*. With our implementation if an attacker runs the hash of the password in *John*[9] without dictionary or rainbow tables or character selection the task cannot be completed in a life time because of the added salt and slowness to the hash. We have also attempted to decrypt the user database file but with the 4096 bit encryption the task seems to be impossible. The only method that compromises the password is to make a memory dump. After the user have logged in and the decryption have taken place the data can be traced if we make a process dump with *gcore*[10] or *pmap*. *Unfortunately* we have been able to extract the database information, but not the master password, however this requires privilege escalation on Windows machines and root access on Unix based systems and if this happens the security have already been compromised and almost nothing will keep the user data secure.

## **5. Applications**

### **5.1. Home users**

In future editions we are planning to support the sms gateway for home users with a remote server implementation. We can provide this function using Amazon EC2[11] services or some similar cloud based solution. Of course there are other implementation such as *KDE Wallet* for Linux based systems and *KeePass* for Windows based system, but they both lack the sms feature. The source code is well commented and a basic knowledge about programming is needed in order to customise the application. A detailed explanation and source code can be found at: <https://github.com/dreadz/password-manager>

### **5.2. Enterprise**

In a productive environment the password manager have many applications and can be customised and tailored to have full control and access over the user activity. Thanks to the password strength and cloacking and application start capabilities, administrators can select different levels of security. With this methods users can have access to content without knowing the password only after successful 3-step authentication which is forcing physical access to the facility.

## **6. Conclusion**

We all know that we can't have 100% secure environment, but we can introduce and establish a sequence of protection mechanisms which will certainly decrease the possible attack vectors in our system. A password manager can be a great tool to keep our security credentials safe, but let's not forget that even a 4096 bit key can be cracked[12] under special circumstances. In comparison to similar software our solutions is in a early alpha stage, but implementation of remote sms server and memory flushing ca be created. Keeping our passwords secure and most important – complex enough will give us peace of mind and knowing the possible attack vectors furthermore reduces the possibility of successful malicious activity.

## Bibliography

1. *Burp suite download page*, <https://portswigger.net/burp/>, 2014
2. *Java web crawler download page*, <https://code.google.com/p/crawler4j/downloads/list>, 2014
3. *FPGA Fundamentals*, <http://www.ni.com/white-paper/6983/en/>, 2014
4. *GPU Cluster Architecture and Management Software*, [http://www.ncsa.illinois.edu/People/kindr/papers/ppac09\\_paper.pdf](http://www.ncsa.illinois.edu/People/kindr/papers/ppac09_paper.pdf), 2014
5. Ireland D. *RSA Mechanics*, [http://www.di-mgt.com.au/rsa\\_alg.html](http://www.di-mgt.com.au/rsa_alg.html), 2014
6. Shelley J. *Bcrypt source code home*, <http://bcrypt.sourceforge.net/>, 2014
7. Gihar M. *Gammu Manual*, <http://wammu.eu/docs/manual/>, 2014
8. *Recommendation for Key Management – Part 1*, [http://csrc.nist.gov/publications/nistpubs/800-57/sp800-57\\_part1\\_rev3\\_general.pdf](http://csrc.nist.gov/publications/nistpubs/800-57/sp800-57_part1_rev3_general.pdf)
9. *How to use John the Ripper*, <http://www.openwall.com/john/>, 2015
10. *How to use gcore*, <http://man7.org/linux/man-pages/man1/gcore.1.html>, 2015
11. *Amazon EC2*, <https://aws.amazon.com/ec2/>
12. Genkin D., A. Shamir, E. Tromer. *RSA Key extraction via Low-Bandwidth Acoustic Cryptanalysis*, <http://www.tau.ac.il/~tromer/papers/acoustic-20131218.pdf>, 2015
13. Baker M., H. Mel. *Cryptography Decrypted*, Addison-Wesley, 2001
14. Sierra K., B. Bert. *Head First Java: 2nd Edition*, O'Reilly Media, 2005

## Мениджър на пароли с тристъпкова система за удостоверяване

Желязко Петров, Разван Рагазан

**Резюме:** Важен фактор в онлайн защитата е сложността на паролата. Силата на тази променлива се определя от това колко бързо един хакер може да я познае. За да се повиши ефективността на този вектор, предложеният дизайн на мениджър за пароли е успешно имплементиран и тестван и се е доказал като ефективно устройство срещу модерни *brute force*<sup>1</sup> техники за налучкване. В определи сценарии системата с 3 стъпки за удостоверяване гарантира 100% нуждата за достъп до материалните активи, което я прави сериозен кандидат за създаване на по-безопасна среда за домашни и бизнес потребители. Ние постигнахме добри резултати по отношение на здравината на криптиране и сложността по процедурата, макар че предстои доработване на отпечатъка в системата и по-конкретно – по отношение на атаки, базирани на *memory dump*<sup>2</sup>.

---

<sup>1</sup> Brute force – метод за налучкване на пароли чрез итерация.

<sup>2</sup> Memory dump – метод за извличане на информация от работната памет

# Компютърно творчество – вчера, днес и утре

Евгения Сендова<sup>1</sup>, Христо Стайков<sup>2</sup>

1. Институт по математика и информатика при БАН, София

[jenny@math.bas.bg](mailto:jenny@math.bas.bg)

2. МГ „Д-р Петър Берон“, Варна

[Hristo.Staykov@gmail.com](mailto:Hristo.Staykov@gmail.com)

*No computer has ever been designed that is ever aware  
of what it's doing; but most of the time, we aren't either.*

Marvin Minsky

**Резюме:** В статията са представени идеи, свързани с компютърното творчество, и по специално – с първите опити за генериране на компютърна музика в световен контекст и в България. В първата част авторите се спират на основните принципи за създаване на алгоритмичната музика. Разгледан е т. нар. *музикален микросвят*, създаден с образователна цел от първия автор в програмната среда Лого с типичните за музикалното композиране обекти и операции. Във втората част е представен съвременен подход за генериране на джазова импровизация с помощта на невронни мрежи – проект, разработен от втория автор в рамките на научно-изследователска международна школа за ученици. Накрая авторите се спират на смисъла от подобни експерименти като възможност повече хора да разберат кои аспекти от творческия процес подлежат на формализиране и да навлязат в сфери, до които иначе не биха се докоснали освен като пасивни консуматори...

**Ключови думи:** алгоритмична музика, творчество, микросвят, изкуствено изкуство, невронни мрежи.

## 1. Въведение

*Компютърното творчество* е сравнително нов термин, свързан с теоретични и практически въпроси при изучаване на творческите процеси. Изследването на естеството на тези процеси и стремежът да се даде подходяща дефиниция на „творчеството“ става успоредно с внедряването на системи, които имат „творчески“ характер. На английски въпросният термин може да се срещне в няколко варианта – *computational creativity*, *artificial creativity*, *artificial art*, *mechanical creativity* и *creative computation*. Можем да смятаме, че става дума за област, която е сечение на изкуствения интелект, когнитивната психология, философията и изкуството.

В най-общи линии компютърното творчество се състои в моделиране и симулиране на творчески процеси с помощта на компютър с цел:

1. да се създаде програма, достигаща равнището на хората-творци;
2. да се разбере по-добре спецификата на човешкото творчество като процес;
3. да се създаде програма, която подпомага човека-творец.

Нека илюстрираме някои от тези идеи в контекста на компютърното композиране на музика.

## **2. Алгоритмично композиране на мелодии**

Да се открият и изучат закономерностите, спазването на които позволява на композиторите да сътворят издържана в естетическо отношение музикална форма, е едновременно сложна и привлекателна задача. Сложността идва от това, че моделирането на този процес предполага познаване на музикалната организация, а малко такива закони са известни. Относително изключение представляват строгият контрапункт и додекафонията, поради което композирането в тези два стила се поддава сравнително по-лесно на алгоритмизиране.

При тоналната музика се говори по-скоро за господстващи тенденции. Как може да се моделира процесът на композиране на музика? Първа стъпка в това отношение е въвеждането на насочена случайност – изборът на едно случайно число съответства на избора на даден музикален елемент, който се приема или отхвърля съгласно зададени правила. Като се зададе разпределение на вероятностите за избор на музикални елементи от едно и също йерархично равнище, може да се предопредели общият характер на композицията, без да се предвиди точният ѝ вид.

Всяка композиция се характеризира с някакъв брой параметри, например тактов размер, лад, музикална форма и т.н. За да се алгоритмизира процесът на композирането, трябва:

1. да се определят правилата, които ще се спазват безусловно;
2. да се изберат параметри, чиито стойности ще се получават по случаен начин;
3. да се зададе разпределение на вероятностите за тези параметри;
4. да се установи редът, по който ще се избират стойностите на параметрите.

Натрупаният опит и получените резултати от традиционното музиковедение се съчетават със сериозен математически анализ. Изследванията от този род стават особено интензивни с привличането на компютрите в експерименталната дейност.

### 3. Кратък исторически преглед

За машинни модели на музика става дума от най-ранните дни на компютъра. Да направим кратък преглед на първите опити в това направление.

#### 3.1. По-известни идеи за компютърно композиране в световен контекст

Още първата програмистка, лейди Ада Лъвлейс, вижда в компютъра средство, с което да се разработват и проверяват музикални теории [1].

Основен принцип на първите експерименти е изследването на локалните връзки между тоновете. Този принцип се наблюдава в опитите на американските математици Brooks, Hopkins, Neumann и Wright от 1957 г. [2]. В паметта на компютъра се въвеждат и анализират известен брой мелодии. Пресмята се честотното разпределение на всички съчетания от  $n$  ноти ( $n=1, 2, \dots, 8$ ). На следващия етап се построяват редици на Марков от  $(n-1)$ -ви ред. Както се очаква, генерираните мелодии *се приближават локално* до входните данни, но тяхна обща слабост е липсата на глобална структура.

Други експерименти се основават на формализирането на правила и закономерности на музиката. При тях идеята за композиране се състои в избиране на ноти по случаен начин и записване на онези от тях, които удовлетворяват множество от правила, определени чрез музикален анализ на пиесите, които се използват като модел. Първите експерименти в тази посока са на Hiller и Isaacson през 1957 г. [3], които формализират някои правила на строгия контрапункт.

Характерно за всички ранни опити в областта на компютърното композиране е, че стъпват на теорията на информацията и на теорията на вероятностите.

Сериозен успех в моделирането на тонална музика представляват работите на Зарипов. През 1960 г. той публикува първите си резултати по компютърно композиране на еднотонни музикални пиеси във формата *ABA*, в тоналност *до мажор*, с дължина четири осемтактови периода [4]. В следващите му разработки тричастната форма е заменена с еднотонна, като са отразени повече закономерности на мелодията [5].

Интересен опит за формализиране на езика на тоналната музика прави френският математик Varbaud върху основата на свойствата на една крайна адитивна група [6].

#### 3.2. Изкуствен интелект и музика – една нова научна област

През 1980 г. се заговаря за *изкуствен интелект и музика*. Това словосъчетание е *изковано* от Curtis Roads и е свързано с изследвания, при които музиката се разглежда като когнитивен процес или като множество от дейности, моделирани с помощта на компютър [7]. Основните разглеждани въпроси включват:

- *Що е музикално знание/интелект?*
- *Представява ли музикалното знание автономна човешка компетентност или е съвкупност от исторически събрани, но несвързани компоненти?*
- *Различна ли е музикалната компетентност от други видове компетентности, например лингвистичната?*
- *Каква е зависимостта между знанието и действието в областта на музиката?*

Теорията на музиката и музикознанието се занимават основно с музиката като културен артефакт и не се интересуват от композирането като дейност или като мисловен процес. Психологията на музиката съсредоточава вниманието си върху локални явления и възприятия, без да отчита управляващите структури в музикалните процеси като експертни дейности. Именно изучаването на такива управляващи структури е една от основните цели на областта *изкуствен интелект и музика*. Работата в тази област предполага нов и по-задълбочен поглед върху музиката. Традиционното музикознание е пренебрегвало или се е спирало съвсем бегло на епистемологичните загадки, които музиката поставя като човешки способности и опит. Крайната цел на областта *изкуствен интелект и музика* според Марвин Мински е да се създаде композитор, а може би и слушател на музика [8].

### **3.3. Първи стъпки за компютърно композиране у нас**

Първият експеримент за моделиране на музикални творчески процеси у нас е върху алгоритмичното описание и програмиране на процеса на хармонизирането и е направен от Стефка Димова през 1963 г. Алгебричната формализация, предложена от Warbaud, се използва в по-късен вариант на програмата [9].

Следваща стъпка в изследванията у нас прави първият автор на настоящата статия с експериментиране на предложения от Зарипов алгоритъм за композиране на мелодии и модифицирането му с добавяне на каденци, определени след музикален анализ на моделираните мелодии [10]. В този алгоритъм са отразени закономерностите на четирите основни страни на мелодията: ритъм, музикално-височинни отношения, ладово-хармонична структура и мащабна структура. Процесът на композиране на мелодията се състои от следните етапи:

- *Формиране на общата структура на мелодията* – избира се по един тип от мащабната, ритмичната и мелодичната структури, с което на следващите етапи се осъществява разчленяване на периода на по-малки отделни построения;
- *Построяване на ритъма* – съчетанията от нотни трайности се избират по случаен начин в съответствие с честотата на разпределението им в избраната за мелодията ритмична група. Ритъмът на конкретния такт се формира, като се получават последователно съчетания от нотни трайности, докато сборът им съвпадне с тактовия размер;

- *Определяне на мелодичната структура* – организира се повторението на мелодичните фигури като избор от няколко възможни типа, извлечени въз основа на музикален анализ. Например структурата 1 1 1 1 2 2 3 3 означава, че мелодичните линии на втория, третия и четвъртия такт са построени въз основа на мелодичната фигура на първия, мелодичната линия на шестия е като на петия, но е различна от първите четири и т.н. Има се предвид евентуално видоизменено повторение на мелодичната фигура на друга височина при запазване на интервалните отношения (*секвенция*);
- *Определяне на хармоничната структура* – организира се последователността от акорди като избор от няколко възможни типа структури. Например, ако означим с T, S и D тониката, субдоминантата и доминантата, възможна хармонична структура е: T T S D S S D T. Ако два такта имат еднаква мелодична структура и ако опорните (първите) им ноти принадлежат съответно на T и S, стъпката на секвенцията се избира измежду интервалите, на които отстоят основните тонове на T и S (*до и фа в гама до мажор*) в допустимия диапазон от две октави;
- *Интервално-височинна обработка на определения ритъм в съответствие с мелодичната и хармоничната структура*:
  - a. Височината на началната нота  $W_0$  се избира по случаен начин като елемент от първия акорд на хармоничната структура (T);
  - b. Височината на текущата нота  $W_k$  се образува като алгебричен сбор от височината на предишната нота  $W_{k-1}$  и текущия интервал  $\phi$ , чийто знак определя дали движението е възходящо или низходящо;
  - c. Големината на  $\phi$  се избира по случаен начин в съответствие с разпределението на честотата на допустимите в мелодичната структура интервали;
  - d. Височината на последните ноти на всяка фраза се строи в съответствие на основните видове каденци.

При хармонизацията на 8-тактовите мелодии са формализирани основни принципи, като *икономия за движението в музиката*, изключване на явните и скритите паралелизми. На фигура 1 са показани няколко произведения, получени с програма за машината Минск-2 [10].

Получените резултати потвърждават редица достойнства на използваните алгоритми – може да се говори за *общ стил* на всички получени композиции. Сравнително големият брой правила силно увеличаваше времето за генериране на мелодия, която да ги удовлетворява. – повечето примери биваха отхвърлени, преди да се получат осем такта поради невъзможност да се избере нота, която да удовлетворява всички алгоритмични изисквания.

От информатична гледна точка е интересно да се отбележи, че грешка в програмата, чието откриване отне месеци, се дължеше на грешка в документацията. Тъй като се работеше на език за програмиране от ниско ниво (МИД-2), огромни усилия се полагаха за преодоляване на проблеми, които не бяха съществени от гледна точка на моделирания процес. Много по-естествено

се оказа изграждането на т. нар. *микросвят* с типичните за областта обекти и операции, което първият автор реализира като елемент от дисертацията си [11].



Фигура 1: Мелодии, получени с програмата на Минск-2

#### 4. Образователен потенциал на микросветовете

Компютърните микросветове и модели не претендират за представителност на това, което вече се постига в моделирането на творчески процеси.

Понятието *модел* в контекста на творчеството, на което се опираме в разработването на музикален микросвят, е в съзвучие с дефиницията на Марвин Мински [12]: *За даден наблюдател В обектът А\* е модел на даден обект А до толкова, доколкото В може да използва А\*, за да си отговори на въпроси, които го интересуват за А.*

Това, което компютрите са постигнали като *композитори, художници, писатели*, все още е предмет на оживени дискусии и нееднозначни критики [13]. Разработените в [11] микросветове и модели по-скоро могат да се използват в учебния процес за мотивиране на младите хора да погледнат произведенията на музиката, литературата, живописата, математиката с очи на изследователи.

При разработката на един микросвят се концентрираме върху сравнително малко и затворено множество от процедури и структури от данни. В този микросвят можем да експериментираме идеи и теории за дадено явление или процес, които ще моделираме с цел да ги разберем по-добре. Проверката на

идеи във формата на програми е една от най-важните характеристики на изкуствения интелект, а вече и на *изкуственото творчество*. Благодарение на компютърните микросветове могат да се тестват теории от света на изкуството от гледна точка на пълнота и непротиворечивост.

При моделирането на дадено явление често ни се случва в процеса на експериментиране с програмата да получим по-задълбочено разбиране за същността на явлението, а това на свой ред води до нов етап на формализиране и т.н. Когато придадем конкретен характер на това, как възприемаме дадено произведение на изкуството, кое за нас е съществено, кое сме пренебрегнали при създаването на компютърния модел, превръщаме отношението си към моделираното явление в обект, над който можем да разсъждаваме. Или в синхрон с изказването на Марвин Мински [14], *не само учим нови неща, но научаваме как да мислим за тези неща, след това се научаваме да мислим за самото мислене...*

При подходящо проектиран микросвят една програма може да служи като точно описание на общ метод, който е достъпен именно защото е изпълним с различни входни данни, с различни дефиниции на понятия от изкуството, които трудно се поддават на формализиране (например *близост, равновесие между очакваното и неочакваното*). Често си служим с т. нар. *работни дефиниции* – правим начално приближение на дадено понятие в конкретен контекст (например дефинираме като *близки тонове* в музикалната скала онези, които отстоят на интервал секунда). Експериментираме програмата с избраните работни дефиниции и изследваме генерираните с тяхна помощ произведения. Ако мелодиите звучат монотонно, можем да коригираме работните дефиниции и отново да изследваме ефекта от редакцията.

С други думи, освен да генерират определен тип произведения, микросветовите са ново мощно средство за изразяване на методологични идеи, а програмите, реализирани в тях, могат да служат като средство за формализиране на идеи. Фактът, че на тези идеи може да се вдъхне живот чрез експерименти, прави от програмите ненадминат източник на хипотези, свързани с анализ на дадено явление (закономерност в музиката, в езика, в абстрактната живопис), както и за откриване на нови явления, които може да се родят като страничен ефект от направените експерименти.

Ценното на програми, написани в специализирани за дадена област микросветове (в нашия случай музиката), е, че могат да се използват не като *черни кутии*, а като *прозрачни системи*. Освен че са сравнително достъпно формализирано описание на процеса на генериране на дадено произведение, човек може *да си играе* с тях достатъчно дълго, за да разбере доколко са гъвкави, т.е. кои са инвариантните елементи в тях. В този смисъл можем да говорим за тях като за среди, които са *редактор на идеи*. Идеи от изследваната област могат да се облекат в подходяща форма, а именно като програма, която представлява *материализирана хипотеза* за това какво представлява дадено творческо явление. При изгълнение на програмата се получава обратна връзка, от която получаваме информация относно хипотезата си и можем да я редактираме в смисъл на коригиране, модифициране, прецизиране,

усъвършенстване и обогатяване. Ето защо моделирането на творчески процеси с помощта на информатични средства играе огромна роля в учебния процес. Човек не само може да провери валидността на заложените в хипотезата правила (*debugging* на идеи), но и да преформулира някои от поставените цели (*degoaling*) в резултат от положително възприемане на странични ефекти (които биха могли да са резултат от недостатъци в модела). Един явен и формализиран подход ни дава възможност да изразим фундаментални принципи с по-голяма яснота. Нещо повече, свързването на дадена теория с програмиране е конструктивен подход. Изучаването на даден музикален стил например може да се сведе до проектирането на подходяща система, която да реализира схема, основана на този стил. Авторът на програмата за композиране може да използва техниката на откриване на грешки в нея, за да анализира собствената си работа, като постепенно стига до все по-дълбоко прозрение за структурните характеристики на изучавания стил [15].

Модели за композиране на музика са строени не само върху традиционни музикални принципи. Интересна е идеята на някои учени да използват език, предназначен за моделиране поведението на колония от мравки, в музикален контекст [16]. Правени са и опити да се намери музикално съответствие на различни рекурсивни структури – в частност криви, свързани със системите на Lindermaier, проектирани за описание на растежа на биологични организми [17]. При използването на фрактали идеята е да се постигне едновременно компактност и повтаряемост на структурата. Интересен факт е, че при анализ на известни творби, се натъкваме на рекурсивни механизми, затова изглежда смислено да се търси връзката между математиката и музиката и в двете посоки [18].

Изложението дотук се отнася до частта *Вчера* и се основава на [11] и на публикацииите [19-21].

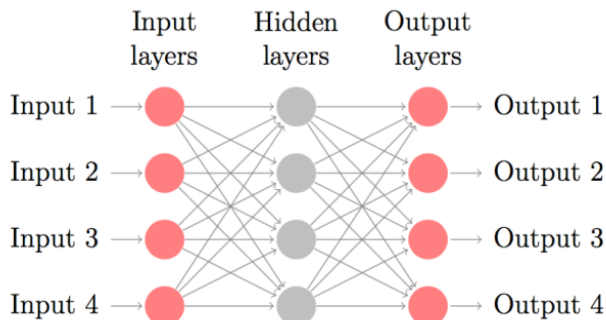
А сега да преминем към *Днес*, като хвърлим поглед към една разработка на втория автор [22]. В рамките на изследователската програма RSI [23] за ученици той получи от ментора си (Джерemi Сейгън) задачата да автоматизира вход на програма, представляваща последователност от акорди, с цел на изхода да се получи джазова импровизация. Програмата обаче бе предоставена като „черна кутия“ и затова младежът реши да направи свой вариант на джазова импровизация – с помощта на невронни мрежи. Да му дадем думата за подробности.

## **5. Невронните мрежи – един нов механизъм за композиране**

Невронната мрежа е структура, вдъхновена от биологичните невронни мрежи. Тя се състои от идентични единици, наречени *неврони*, които са опростен модел на истински неврон. Невроните са свързани в мрежа, като всеки събира подадената му информация, обработва я и я предава на съседите си (Фигура 2).

Част от невроните са обособени в слоеве за вход и изход, а останалите между тях са в т.нар. *скрит слой*. Чрез променяне на теглата на връзките,

мрежата може да се *обучи* на дадено поведение. Обучението на невронна мрежа става с трениращ алгоритъм, на който се представя тренировъчен набор от данни, които посочват какъв изход се очаква при даден вход.



Фигура 2: Примерна топология на невронна мрежа

В музикален контекст, невронна мрежа може да бъде обучена да изпълнява музикални задачи, без да е необходимо да се формализират правилата, по които да става това. Първите стъпки в използването на невронни мрежи за създаване на музика са свързани с обучаването им да възпроизвеждат мелодии [24]. Мозер [25] създава невронна мрежа с рекурентни връзки, довършва научена едногласна мелодия при предоставен откъс от началото ѝ. Нотите от откъса се въвеждат една по една в мрежата, а за всяка се извежда предположение за следващата нота. Така мрежата на Мозер е обучена да възпроизвежда мотиви от менуети на Бах при подадени първите няколко ноти. Когато се въведе откъс от мелодия, която не е тренирана, мрежата продължава да извежда ноти. Благодарение на обучението си тя създава добри мотиви на локално равнище, но е неспособна да работи върху по-дълги мотиви със сложна организация.

Следващи разработки въвеждат по-сложни структури в топологията на мрежите, наречени *Long Short-Term Memory* (LSTM) клетки. Чрез рекурентни връзки в тези структури, мрежите са способни да запомнят свои предишни състояния и така да обработват по-сложни, изменящи се във времето структури. Франклин [26] използва *LSTM* клетки в невронна мрежа, за да запомня и възпроизвежда мелодии във форма *AABA*.

Основните проблеми при използването на невронни мрежи са в определянето на подходящи топология, представяне на данните и множество от тренировъчни примери. Колкото по-голям е броят на невроните в една мрежа, толкова по-сложна информация е способна да обработва тя, но в същото време е по-трудно да се открият оптималните стойности на параметрите за обучението ѝ. Винаги съществува риск и от т.нар. *overfitting* и *underfitting* — това е когато мрежата е съответно с твърде сложна или твърде проста структура за обработваните данни, което може да доведе от „плагиатстване“ от тренировъчните примери до извеждане на случайни данни. Диапазонът на



Напоследък се популяризират и т.нар. *Deep Belief Networks* — невронни мрежи със специфична топология, които имат повече от един междинен (скрит) слой с неврони, което позволява обработването на още по-сложни структури. В последните месеци са публикувани няколко опита с такива мрежи в създаването на музика, които включват хармонизиране на едногласни мелодии [28] и композиране на ирландска фолклорна музика [29].



Фигура 4: Акомпанимент и импровизация, създадена от невронната мрежа

## Заклучение

И накрая да погледнем към бъдещето, като се опитаме да си представим как биха се отнесли към днешното компютърно творчество композиторите от времето, в което компютрите още *не са били навлезли* в тяхната територия.

Още през 1959 г. Стравински пише, че всяко изследване, наивно или не, има значение, защото води до по-големи въпроси, всъщност – до евентуална математическа формулировка на теорията на музиката и до емпирично изучаване на музикални факти, и по-точно – на изкуството на комбинирането, каквото е музикалната композиция. Развитието на новите технологии и на изкуствения интелект вече превръщат мечтата на Стравински в реалност, защото не само е възможно да се формулират модели на композиторския процес, но да се подложат тези модели на строго оценяване.

Целта на тези експерименти естествено не е да се изместят творците в изкуството, а да се даде на много повече хора шанс да разберат това, което може да се формализира и алгоритмизира, да навлязат в сфери, до които иначе не биха се докоснали освен като пасивни консуматори... Дали резултатите от тези опити ще съответстват на това, което досега се е наричало изкуство, или ще дефинират нещо, което отсега нататък ще се нарича изкуство, е спорен въпрос. Обикновено музикалният анализ се занимава със структурата на произведението, а изкуственият интелект и когнитивната наука – със структурата на процеса. А най-същественният въпрос е дали може да се опише връзката между тях и по какъв начин. Този въпрос е важен впрочем във всяка човешка дейност. Окуражаващото е, че дори малките стъпки в моделирането на специфични музикални проблеми с техники, типични за математиката, програмирането, изкуствения интелект и музикознанието, могат да допринесат не само за разбирането на музикалното творчество, но и за творчеството въобще.

Може би най-важният резултат е, че процеси, които днес изглеждат достъпни само за човек творец, в бъдеще ще могат да бъдат предоставени на компютъра, като по този начин хората ще бъдат свободни да преследват още по-предизвикателни творчески цели.

## **Благодарности**

*На първия ми учител по програмиране, проф. Бърнев, на първия български изследовател на компютърното творчество в музикален контекст, проф. Стефка Димова, и на първия български ученик, който поема щафетата с помощта на невронни мрежи, Христо Стайков, изказвам огромната си благодарност: Е. Сендова*

## **Литература**

1. Ada A. (1980) *Scientific Memoirs*, iii. p.694 (цитирана от Howe, H., Kassler M. в „Computers and Music“ in new Grove Dictionary of Music and Musicians (vol. 4), London: Macmillan Publishers Limited, p. 613).
2. Brooks, F. P., Hopkins, A. L., Neumann, P. G., Wright, W. V. (1993) An experiment in Musical Composition. In Schwanauer & Levitt (Eds.), *Machine Models of Music*, MIT Press, Cambridge, MA
3. Hiller, L. A., Isaacson, L. M. (1959) *Experimental music*. New York, McGraw-Hill Company
4. Зарипов, Р. Х. (1962) О программировании процесса сочинения музыки. *Проблемы кибернетики*, 7, 151-160
5. Зарипов, Р. Х. (1965) О моделировании мелодий заданного стиля на цифровых вычислительных машинах. *Проблемы кибернетики*, 15, 157-200
6. Barbaud, P. (1968) *La musique, discipline scientifique*, Paris, Dunod
7. Machover, T. (1985) Thoughts on computer music composition. In Roads, C. (Ed.) *Composers and the Computer*. William Kaufmann, Inc. Los Altos, CA
8. Minsky, M. (1969) *Matter, Mind and Models*. Memo 77, MIT, AI Laboratory

9. Димова, С., Коларова, Е. (1972) Алгоритмично хармонизиране, *Вселена '71*, София, Наука и изкуство
10. Димова-Боршукова, С., Коларова, Е. (1971) Моделиране на музикални творчески процеси, *Вселена '70*, София, Наука и изкуство, с.65-74
11. Сендова, Е. (2001) *Компютърни микросветове и модели за интегриране на учебния и творческия процес*, Дисертация, ИМИ-БАН, София
12. Minsky, M. (1986) *Society of Mind*, Simson&Schuster, NY
13. Hofstadter, D. (1995) *Fluid concepts and creative analogies: Computer models of fundamental mechanisms of thought*, Basic Books
14. Minsky, M (1981) Music, mind and meaning, in Clynes, M. (Ed.) *Music, Mind and Brain: The Newropsychology of Music*, Plenum, NY
15. Arveiller, J. (1982) Comments on University Instruction in Computer Music Composition, in *Computer Music Journal*, vol. 6, No. 2
16. Brooks, S., Ross, B. (1996). Automated composition from computer models of biological behavior. *Leonardo Music Journal*, vol. 6 pp. 27-31
17. Mason, S., Saffé, M. (1994). L-Systems, melodies and musical structure. *Leonardo Music Journal*, vol. 1 pp. 31-38
18. Johnson, T. (1996) *Self-similar melodies*. Two\_Eighteen Press, NY, USA
19. Сендова, Е. (2001) Може ли компютърът да композира музика? *Математика и информатика*, кн. 1, стр.10-30
20. Сендова, Е. (2001) Компютри и творчество – оптимизъм или скептицизъм, *Математика и информатика*, кн. 2, стр. 25-31
21. Sendova, E. (2001) Modelling Creative processes in Abstract Art and Music, *Eurologo 2001*, Proceedings of the 8 th European Logo Conference 21-25 August, Linz, Austria
22. Staykov, H. (2015) *Computer-generated Improvisations in Jazz and Blues Music*, a project developed and presented at RSI'15
23. Research Science Institute (RSI) – <https://www.cee.org/> (последно посещение: 01.12.2015)
24. Todd, P. M., Loy, D. G. (1991) *Music and connectionism*, MIT Press
25. Mozer, M. C. (1994) Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multiscale processing, *Connection Science*
26. Franklin, J. A. (2004) Computational models for learning pitch and duration using LSTM recurrent neural networks. In *The Proceedings of the Eighth International Conference on Music Perception and Cognition (ICMPC8)*, Adelaide, Australia. Causal Productions.
27. Franklin, J. A. (2006) Recurrent neural networks for music computation. *FORMS Journal on Computing*
28. Sun, F. (2015) *DeepHear – Composing and harmonizing music with neural networks*, <http://web.mit.edu/felixsun/www/neural-music.html> (последен достъп: 1-ви декември 2015)
29. Strum, B (2015) *The Infinite Irish Trad Session*, <http://www.eecs.qmul.ac.uk/~sturm/research/RNNIrishTrad/> (последно посещение: 01.11.2015)

## **Computer Creativity – Yesterday, Today, and Tomorrow**

**Evgenia Sendova, Hristo Staykov**

**Abstract:** The paper deals with ideas, concerning the computer creativity, and more precisely – the first experiments in generating computer music in a national and international context. The first part considers the basic principles for creating algorithmic music. A Logo musical microworld, developed by the first author with educational goals is described with emphasis on its objects and operations typical for composing music. The second part deals with alternative approaches to generating music involving the use of artificial neural networks (ANN). An ANN system for creating improvisations over given accompaniments has been proposed and experimented with by the second author as a research project within the RSI international summer school for high school students.

# Challenges to the Support of Bulgarian Open Institutional Repositories

**Juliana Peneva<sup>1</sup>, Stanislav Ivanov<sup>1</sup>,  
Krassimira Ivanova<sup>2</sup>, Peter Stanchev<sup>2,3</sup>**

1. New Bulgarian University, Department of Informatics, Sofia  
[jpeneva@nbu.bg](mailto:jpeneva@nbu.bg), [sivanov@nbu.bg](mailto:sivanov@nbu.bg)
2. Institute of Mathematics and Informatics – BAS, Sofia  
[kivanova@math.bas.bg](mailto:kivanova@math.bas.bg), [stanchev@math.bas.bg](mailto:stanchev@math.bas.bg)
3. Kettering University, Flint, USA  
[pstanche@kettering.edu](mailto:pstanche@kettering.edu)

**Abstract:** Open access academic archives are the proper instrument to make visible the majority of the works done during the years and to propagate the results of already paid research activities, thus raising their value. This is the way to achieve an improved management of intellectual assets of the whole university and research community. This paper examines the indicators of the success as well as the challenges in the support of Bulgarian institutional repositories. The findings are based on a comparative case study of the six repositories in research institutes and universities and the local repository of the Department of Informatics at New Bulgarian University.

**Key words:** institutional repositories, sustainability, Open Access, OpenAIRE2020 project, Responsible Research & Innovation, Access to Knowledge

**ACM 1998 Classification Keywords:** H3.5 On-line information services–Data sharing; H3.7 Digital libraries-Collection

## Introduction

In today energetic business environment the proper management of digital resources is vital for any organization in response to the changing business conditions. Nowadays digital resources are increasingly being recognized as a very important organizational asset au par with finance and human resources. For higher education and research institutions digital resources can be strategically used to expose intellectual assets such as learning content and research papers. Because of the very rapid development in information and communications technologies it becomes possible to manage the academic assets more effectively than in the past two decades. Usually digital resources are organized as institutional repositories, including long-term preservation and distribution.

During the last five years different types of repositories ranging from digital libraries through various institutional collections and e-journals up to collaborative learning environments have been built. In this perspective, universities and scientific institutions demonstrate a remarkable activity. The main reason for this perpetual interest is the huge diversity of purposes, deposited resources, services and potential

users. Universities and scientific institutions need to exhibit and deploy different kinds of their intellectual assets. It is a matter not only of user's convenience, but of representativeness and prestige as well. In this plan, it is quite natural that the main share of active repositories belongs to countries with advanced higher education and science – about 20% of them in the North America, 47% in Europe shared among United Kingdom, Germany, Spain, France and Italy.

Six Bulgarian digital archives of open access have been reported in the OpenDOAR Database. New Bulgarian University (NBU) and the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences (IMI-BAS) were among the first academic institutions in the country that developed and maintain three of the registered open digital archives. These archives represent a cardinal digital environment for long-term preservation of the electronic scientific output of the academic staff. Apart from this, at NBU a departmental repository to assist all non-auditoria activities was developed. This repository was designed to deploy content not covered by the university infrastructure: investigations, learning resources, theses, students' projects and papers.

The success of institutional repositories (IR) depends on different factors and there is no agreement whether these factors are general for all IRs or a local occurrence. This justifies the purpose of our research effort, namely to examine the use and the life cycle of Bulgarian IRs. It appears that the initial verve throws the years. During the last three years no more attempts concerning the creation of new archives have been reported.

The goal of this paper is to study the indicators of the success as well as the challenges in the support of Bulgarian IRs. Our findings are based on a comparative case study [1] of the six IRs in the research institutes and universities and the local repository of the Department of Informatics at NBU. We argue that the success of a repository depends not only on the functionalities for members of the community, content recruitment and interoperability but also it depends on the user acceptance and their voluntary participation as well as of the attitude of the institutional government body.

## **Overview**

Usually digital resources are organized as institutional repositories, ensuring long-term preservation and distribution thus permitting to derive their maximum business value. The collection of digital content into a repository enables research and higher education institutions to support investigations, teaching, learning, and administrative processes.

Institutional repositories are described as “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members”. Lynch [2] further notes that essential services include “the management of technology changes and the migration of digital content from one set of technologies to the next as part of the organizational commitment to providing repository services”.

There is no straightforward answer to the question “what it is a successful repository”. In [3, 4, and 5] content recruitment has been outlined as a critical success factor. Blythe and Chachra [6] argue that the IRs “will be successful only when they achieve broad and voluntary participation by individuals in the communities they serve”. Several frameworks for success have also been developed. For example Thibodeau [7] proposes that “a framework for organizing information needed to evaluate the success of digital repositories can be articulated along five dimensions: service, orientation, coverage, collaboration, and state”. Service applies to the functionalities for members of the community, orientation refers to the place between preservation and access where the repository operates, coverage relates to content, collaboration indicates whether the IR works alone or might best collaborate in some IR functions, and state is the maturity in the development of the IR. All of these success factors are internally-driven. Generally speaking the success can be defined and measured in terms of the libraries and institutions larger goals.

## Examination of Bulgarian Open IRs

The number of the digital repositories worldwide increases successively very quickly. Open access academic repositories exhibited a boost of 300 during the mid of 2006. Since the beginning of year 2007 the growth of such repositories listed in the OpenDOAR Database [8] shows a constant increase of 100 repositories per year up to its present number of over 2900 (Fig.1).

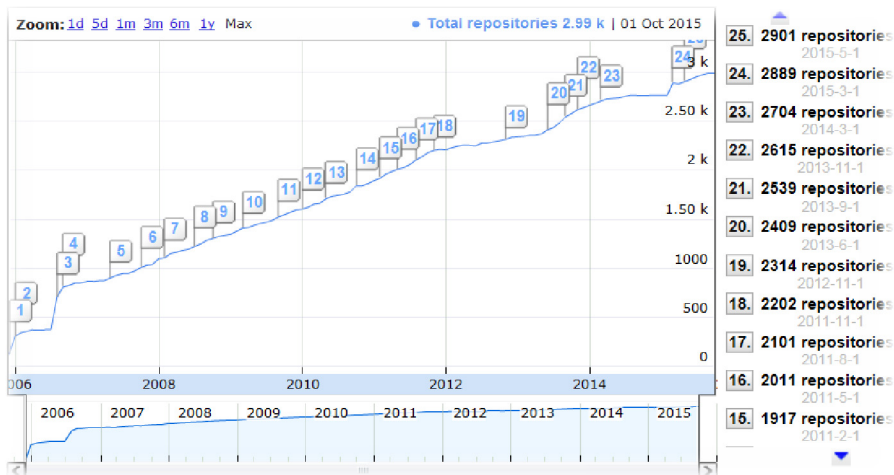


Fig.1 The changing size of the OpenDOAR Database over time

Fig.2 represents the shares of different types of repositories worldwide. Repositories are:

- Institutional or departmental – 84%;
- Disciplinary (a cross-institutional subject repository) – 10%;

- Aggregating (collecting data from several subsidiary repositories) – 3%;
- Governmental – 3%.

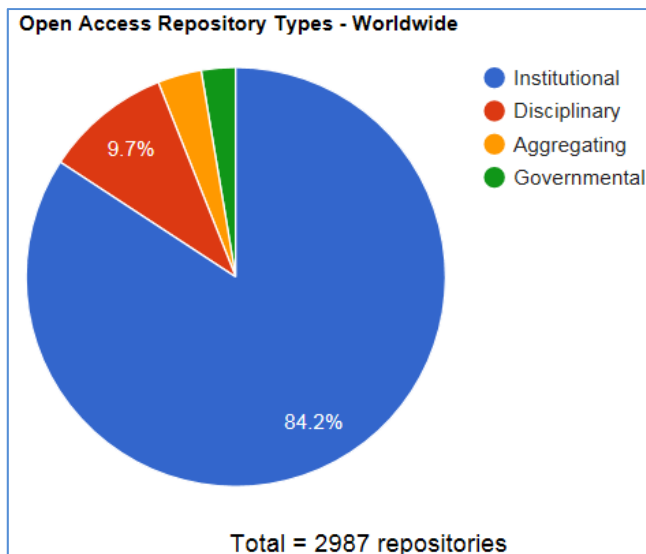


Fig.2 Open access repository types worldwide

In Bulgaria there are six open repositories listed in the OpenDOAR Database [8]:

- New Bulgarian University Scholar Electronic Repository with 1775 items consisting of the research output of the institution. The subjects are multidisciplinary; the content comprises articles, conferences, books and learning objects [9].
- BulDML at IMI with 2115 items consisting of the journal output of the institution. The subjects are multidisciplinary; the content comprises articles, conferences and books [10].
- Academic research repository at Burgas Free University with 483 items consisting of the research output of the institution. The subjects are multidisciplinary; the content comprises articles, conferences and books [11].
- Electronic Repository – Central Medical Library – MU, Sofia with 576 items consisting of the research output of the institution. The subjects are multidisciplinary; the content comprises articles, theses, books and learning objects. Broken and not accessible.
- Research at Sofia University with 877 items consisting of the research output of the institution. The subjects are multidisciplinary; the content comprises articles [12].
- Bulgarian OpenAIRE Repository with 118 items consisting of all peer-reviewed publications resulting from ERC funded FP7 projects in Bulgaria. The subjects are multidisciplinary; the content comprises articles [13].

Compared to an earlier review [14] the findings are similar – no new attempts to build repository have been reported.

### **The repositories of NBU**

The Scholar Electronic Repository of the New Bulgarian University is governed by a supervising board called Evaluation Commission. The Head of Library and Information Services Department is acting as a repository editor thus enforcing compliance with certain rules when an item is going to be deposited. Academic staff and university PhD students are allowed to deposit their documents that might be not published or published via electronic or traditional means. In the case of documents published elsewhere, the depositing author has to hold the copyrights. If the copyrights belong to a publisher or other organization, permission has to be granted. Academic staff can submit unpublished documents as well. In this case authors should respect ethical standards and ensure quality content. The document types are listed in the user interface of the software. The repository is build using E-Print [15], which allows depositing articles, books, conference items, theses, artifacts, images, compositions, audios, videos etc.

In addition we consider the local repository of the Department of Informatics at NBU used to deploy content not covered by the university infrastructure i.e. the repository and the LMS of the university. That content comprises unpublished gray and supporting material: investigations, learning resources, theses, students' projects and papers. Key users of the departmental repository are students and faculty In this way the department's intellectual production is stored and represented. The open access is encouraged, so publishing students' works could restrain mutual plagiarism.

### **The repositories of IMI-BAS**

The BulDML at IMI becomes a part of the global EuDML aggregator. Currently, BulDML contains more than 1500 publications from six journals and several proceedings. BulDML is based on the DSpace [16] version 1.52, with slightly modified source code for practical reasons to keep consistency with future software releases and make more easily migration from old to new software versions. Mainly for submitting content and metadata in BulDML, DSpace native web interface without any difficulties is used. The repository fulfills minimal metadata requirements for its content to be presented in the unique framework of the European Digital Mathematics Library.

IMI-BAS is the contact point for the European Commission Projects OpenAIRE, OpenAIRE+ and OpenAIRE2020. One of the main goals of those projects is to create infrastructure focused on open access publications and scientific data deposited in institutional repositories, thus creating a common scientific platform for the everyday research life of European scientists. OpenAIRE2020 continues and extends OpenAIRE's scholarly communication infrastructure to manage and monitor the outcomes of EC-funded research. It combines its substantial networking capacities and technical capabilities to deliver a robust infrastructure offering support for the Open Access policies in Horizon 2020, via a range of pan-European outreach

activities and a suite of services for key stakeholders. It provides researcher support and services for the Open Data Pilot and investigates its legal ramifications. The project offers to national funders the ability to implement OpenAIRE services to monitor research output, whilst new impact measures for research are investigated. OpenAIRE2020 engages with innovative publishing and data initiatives via studies and pilots. By liaising with global infrastructures, it ensures international interoperability of repositories and their valuable open access contents.

### **Benefits of the Institutional Repositories**

The benefits of publishing in an institutional archive of open access can be summarized from different points of view:

*For the university as an educational and research institution*

- access to the intellectual output is provided;
- enhanced scholar communication becomes possible;
- shared learning and teaching materials otherwise locked in Virtual Learning Environments.

The institutional repositories represent the intellectual product created by the community members thus increasing the institution's visibility and its public value. The open archive, by capturing and preserving collective intellectual capital, increases the overall institution's academic quality. There exists organizational support towards dissemination of innovative research results. Open archives become a feasible way to share ideas and know-how's and to communicate the research activities. This way collaborative research is promoted. In addition an easy access to faculty papers is achieved. The demonstration of value can attract tangible benefits including project funding from both public and private sources. Documents are searchable via the Internet as they are indexed by search engines and made accessible to a wider audience. Finally, the production of the individual authors and the institution as a whole is promoted.

In the traditional system of scholarly communication much of the research findings are dispersed through different journals. However nowadays scholars use Internet intensively to disseminate their achievements. There is growth in the open access availability of research publications, both gold (author pays for publication) and green (self-archiving by the researcher). Approximately 30% of all articles are thought to be available as open access, two-thirds in green and one-third gold [17]. The institutional repository preserves and make accessible the staff intellectual output in a straightforward manner. This is a lightened publishing model. Journal published articles can also be placed in the repository to attain a larger audience, thus enforcing completed publications. Global access to research literature is achieved.

Publication of educational resources in addition to the management of the curriculum is particularly important for dynamic areas of human knowledge. Provision of copyright curriculum in such areas is also a kind of advertising on campus. Making a learning content visible (open courses) increases the potential reuse of the materials.

*For the academic staff*

- improved research knowledge management;
- broad dissemination of published research findings;
- increased citation of some papers;
- indexing.

Sharing of research outputs, unpublished ideas and know-how facilitates their public visibility. In fact researchers manage and store digital content connected with their investigations including research data. Thus collaborative work on institutional projects is promoted. Community members have a place where their scholarly works are permanently exposed. The available scientific results give rise to the preparation of new joint projects. Knowledge sharing and reuse is facilitated.

Because of the short period after uploading the document, authors could achieve a fast dissemination of their results and a greater impact. In rapidly developing areas of scientific knowledge e.g. computer science or communications researchers can offer preprints via the open access archive in order to claim priority and to get a fast feedback.

Open access favors the dissemination of published research in the archives and contributes to the growth of citing of the scientific production. Citation analysis demonstrates that research papers that are freely available are easier to cite.

Open access archives are indexed by search engines, which promote both the production of the individual authors and the institution as a whole. This allows for the creation of personalized publication lists and increases the citation. Via the usage of different metrics the researchers obtain hit rates on specific papers. The impact factor is also derivable.

## **Challenges in Sustaining Institutional Repositories**

There are different factors that contribute for the success or the decline of an IR.

### **1. The population of IRs**

The answer of the question “how to achieve staffs’ participation in a repository?” concerns the contents’ acquisition of any IR. There exists a psychological resistance of users to deploy the content in the digital archive. The staff doesn’t see the value of depositing their work or it is not eager to learn yet another tool that seems to be used rarely. To overcome this opposition different incentives are to be proposed. The new role of IRs as an alternative to publishing i.e. the content recruitment is to be clearly stated.

### **2. The cost**

Building an institutional repository is not a cost free activity. Costs are influenced by different factors e.g. the chosen technology, the provided services, the price of data preservation, the expenses for digitizing content etc. One of the first decisions for an institution is to choose the type of hardware and software for the implementation of the archive. The choice of open source software reduces cost and allows for customization. The use of commercial software cuts the maintenance costs but also can limit the system tuning to the proper needs. Once the software platform

is chosen, the institution has to determine the personnel needed. Generally 2-3 employees are sufficient for metadata management, advertisement activities and community members' induction. In addition, some operational costs e.g. for hardware replacements, for scanning materials, are to be considered. Nevertheless that open access institutional repositories represent an alternative to the purchase of expensive scholarly journals, repositories come with their own price.

### **3. The copyrights**

Usually it is assumed that the depositor possess the rights on digital copy of a completed manuscript or other own work to be exposed in the repository. This makes sense when preprints, not yet published articles, and working papers are considered. However, when institutions decide to post published content, copyright permissions have to be respected. Reasonably, the depositors are often unaware of whether or not they can post an article in an institutional repository. Authors (and IR managers) are confused by the diversity of publisher permissions policies. Some publishers allow deposit of the final published version of the article, others allow the deposit of the author's pre-copyediting version, and some allow no deposit at all. Investigations show that about 80% of publishers allow some version to be used.

### **4. The quality of the deposited materials**

The quality of documents submitted to the open access archives are estimated by the supervisors following formal criteria. The author's contributions might appear obsolete and/or disputable. It is the author's personal responsibility for the flaws in the content e.g. plagiarized texts, faults or junk. Author's produced digitization can also be a quality problem, e.g. badly scanned texts or images. Another concern is that depositors may not know how to describe their work in a way that will increase the chances of the article to be discovered by search engines. It is the institution liability to take care such materials to be improved and to ensure a quality repository that others will use.

### **5. The strategic vision of the governing board**

Most of the repositories are created via targeted projects which execution takes the mainstream of efforts and resources. After acquiring some critical mass of deposited items and opening for public use, the repository is said to be functional. Projects concerning the development of IRs after their ending cannot further provide for a stable business plan. Our experience shows that the initial enthusiasm associated with the creation and accumulation of repository content decreases. Meanwhile expenditures persist as well as organizational support should do. Therefore, a continuous attention from the institution governing body is needed to ensure a long-term life of any IR, here included a stable framework for use and maintenance.

## **Conclusion**

Open digital archives are of great importance for the public visibility and recognition of research institutes and universities. By allowing an improved management of intellectual outputs and freeing up the process of dissemination, their main purpose is to rise up the representativeness of the institution. The rationale behind the

implementation of an institutional repository is to preserve the intellectual assets of members' community.

In this paper we discussed the academic use of open digital archives. An attempt to summarize the benefits and to analyze the usability and the factors leading to the success of Bulgarian IRs is presented. The IR is an open e-space which provides long-term preservation of electronic documents and assists in the dissemination of research findings both at public and local level. However staff is not necessarily interested in depositing their work and it is up to the managerial board to stimulate proper supporting activities. The presence of well configured framework and the allocation of sufficient cash resources for maintenance are often neglected. The success of an IR also depends on the integration with the overall librarian collection of books, journals, digital files, etc.

Based on analysis of current state of open access, made here, as well as in [18], we suggest some recommendation for further work:

- To be defined clear policies for the dissemination of and open access to scientific publications resulting from publicly funded research. It should be on the basis of the green model, within which quality is ensured by scientific publications. This should embrace all research institutions which perform and/or disseminate fully or partially state-funded research. Access to the results of state-funded research should be provided to the greatest possible extent;
- To be ensured that research funding institutions responsible for managing public research funding and academic institutions receiving public funding implement the open access policies.

## References

1. Goodrick, D. (2014). Comparative Case Studies, *Methodological Briefs: Impact Evaluation 9*, UNICEF Office of Research, Florence.
2. Lynch, C. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *ARL Bimonthly Report*, 226: 1–7.
3. Shearer, K. (2003). Institutional repositories: Towards the identification of critical success factors. *Canadian Journal of Information and Library Science*, 27(3): Available: <http://hdl.handle.net/1880/43357>, accessed on October 4, 2015.
4. Bell, S., Foster, N. F., Gibbons, S. (2005). Reference librarians and the success of institutional repositories. *Reference Services Review*, 33(3), 283-290.
5. Ferreira, M., Rodrigues, E., Baptista, A. A., Saraiva, R. (2008). Carrots and sticks: Some ideas on how to create a successful institutional repository. *D-Lib Magazine*, 14(1/2). Available: <http://www.dlib.org/dlib/january08/ferreira/01ferreira.html>, accessed on October 4, 2015.
6. Blythe, E., Chachra, V. (2005). The value proposition in institutional repositories. *EDUCAUSE Review*, 40(5), 76-77.
7. Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository, *Journal of Digital Information*, 8(2). Available: <http://journals.tdl.org/jodi/article/view/197/174>, accessed on October 4, 2015.
8. [www.openoer.org](http://www.openoer.org), accessed on October 4, 2015.
9. <http://eprints.nbu.bg/>, accessed on October 4, 2015.

10. <http://sci-gems.math.bas.bg/jspui/> accessed on October 4, 2015.
11. <http://research.bfu.bg/>, accessed on October 4, 2015.
12. <http://research.uni-sofia.bg/>, accessed on October 4, 2015.
13. <http://www.bg-openaire.eu/>, accessed on October 4, 2015.
14. Simeonov, G., Stanchev, P. (2011). Open Access and Institutional Repositories in Bulgaria. *Proc. of the 1st Int. Conf. "Digital Preservation and Presentation of Cultural Heritage"*, V.Tarnovo, Bulgaria, 165-170.
15. <http://eprints.nbu.bg/>, accessed on October 7, 2015.
16. <http://www.dspace.org/>, accessed on October 7, 2015.
17. Poynder, R. (2011). Open Access by Numbers, Open and Shut, <http://poynder.blogspot.com/2011/06/open-access-by-numbers.html>, accessed on October 7, 2015.
18. Stanchev, P. (2014). Business Models for Open Access. Copyright and Licensing under Open Access. *Proc. of UNESCO Conf. „Digital Presentation and Preservation of Cultural and Scientific Heritage“*, V.Tarnovo, Bulgaria, p. 324.

## **Предизвикателства към поддръжката на институционалните хранилища с отворен достъп в България**

**Юлиана Пенева, Станислав Иванов,  
Красимира Иванова, Петър Станчев**

**Резюме:** Академичните архиви с отворен достъп са подходящо средство за представяне и популяризиране на научните постижения на университетите и изследователските институти. По този начин се постига по-добро управление на интелектуалните активи на научната общност. Настоящата статия изследва изградените български институционални хранилища и анализира факторите, допринасящи за техния успех както и предизвикателствата, свързани с поддръжането им.

# Онтологичен модел за музейна документация: пример от българската практика

Калина Сотирова

Институт по математика и информатика при БАН,  
[kalina@math.bas.bg](mailto:kalina@math.bas.bg)

**Резюме:** Информационният мениджмънт на дигиталното културно наследство включва управление на документирането, обмена, представянето и съхранението на артефактите на материалната култура, както и научно-изследвателския дискурс, който интерпретира миналото. Изискванията за документиране и описание на движимите културни ценности е предизвикателство за архитектурата и дизайна на ефективни информационни системи за представяне на знания и семантично търсене. Бързото развитие на стандартите и приложенията на WEB 3.0 (RDF, SPARQL, OWL, SKOS) предлага иновативни приложения за каталогизация, търсене и виртуално експониране на музейни обекти. Настоящата статия предлага анализ на музейната документация у нас и типологизация на 40<sup>-те</sup> съставни полета на паспорта с оглед създаване на концептуален модел на онтология, като се основава на (1) националната нормативна база, (2) музейната практика и (3) изискванията на web 3.0

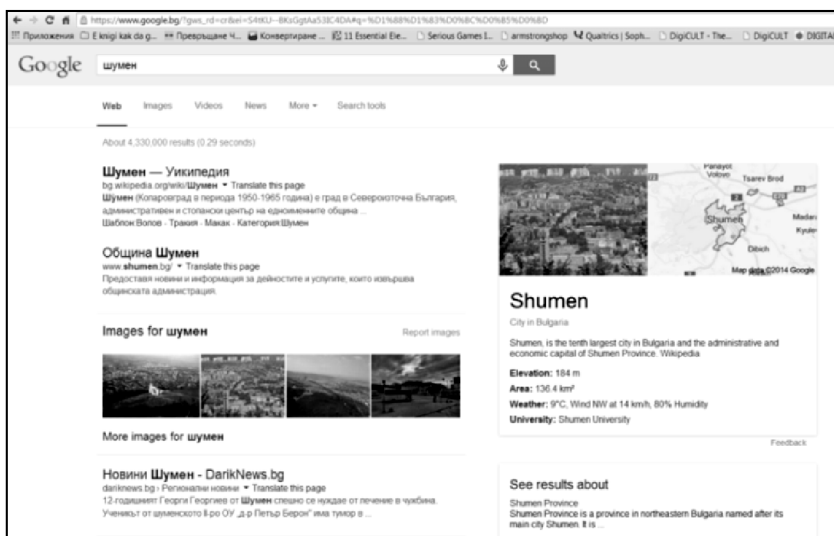
**Ключови думи:** семантична мрежа, онтологии, музеен паспорт, OWL, CIDOC-CRM, Dublin Core

**ACM 1998 Classification Keywords:** H.1. Information Systems: Models and Principles; H.2.5 Heterogeneous Databases; H.3.1 Information Storage and Retrieval: Content Analysis and Indexing

## Увод в тематиката

Информационният мениджмънт на дигиталното културно наследство включва управление на документирането, обмена, представянето и съхранението на артефактите на човешката материална култура, както и научно-изследвателския дискурс, който реконструира и интерпретира миналото. Особено важни при това са анализът и задоволяването на изискванията на потребителя и на съвременното управление на музейни колекции. [Dobрева & Duff, 2015]. Дигиталното културно наследство (ДКН) изисква интердисциплинарен подход и многостранна експертиза. То е обект на изследване от дисциплини като „дигитална хуманитаристика“ (Digital Humanities), „Компютърни науки в областта на наследството“ (Heritage Computing), „Информатика в културната област“ (Cultural Informatics) и „Наука за наследството“ (Heritage Science). Настоящата статия е в областта на ДКН, като се фокусира върху създаването на модел за онтология, представяща разнородната информация от различните типове музейна документация у нас. Представени са начини за типологизация на информатичните полета с цел да бъдат представени със стандартите на Семантичната мрежа (RDF, OWL) и

стандартите за културна информация (Dublin Core, CIDOC-CRM). Терминът „семантична мрежа“ се отнася към визията на консорциума W3C за създаване на хранилища за хетерогенни данни, тезауруси и индексирани правила за обработка. Семантичните технологии и приложения се развиха силно през 2013 г., във връзка с напредъка на изследванията в областта на изкуствения интелект. [Issac & Baker, 2015]. От април 2014 г. търсещата машина на Google започна да ползва „граф-на-знанието“ Hummingbird, като за пръв път резултатите представят семантично свързана информация в отделен блок на страницата. (фиг. 1) [Szekely et al, 2014], [Nisheva et al, 2014], [Minelli et al, 2014]. Дотогава търсенето се извършва на базата на съответствие на низове. Това развитие може качествено да промени достъпа до семантично богатата информация на дигиталното наследство.

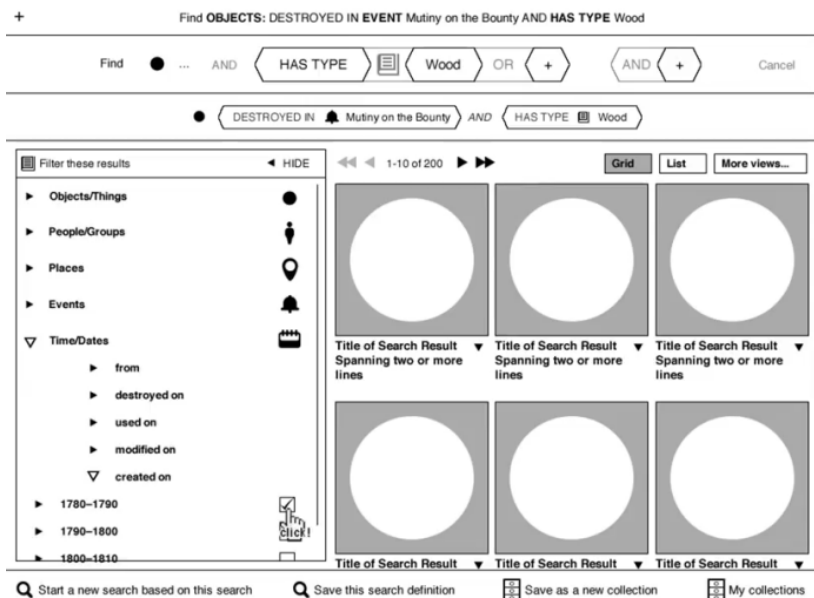


*Фигура 1. Визуализация на индексирането от графа на знанието на Гугъл – (Hummingbird)*

Семантичната интерпретация на обектите в онлайн достъпни музейни колекции има потенциала да „създаде ново знание“, което подлежи на машинна и експертна проверка и обогатява възможностите за бързо контекстно търсене. Това обаче е предизвикателство за архитектурата и дизайна на ефективни информационни системи за институциите за културна памет. Това предизвикателство вече се посреща с различен темп и успех в институциите за културна памет.

Стандартите на WEB 3.0 ползват съвременни принципи и алгоритми за структуриране и индексирание на данни (Resource Description Framework<sup>1</sup>, Ontology Web Language<sup>2</sup>, Linked Open Data<sup>3</sup>, SKOS<sup>4</sup>, SPARQL<sup>5</sup>).

Добър пример за използване на тези стандарти в контекста на ДКН са успехите на проекта *Research Space*<sup>6</sup> на Британския музей. Към февруари 2016 г. на портала на проекта са представени 2.5 милиона обекта като отворени свързани данни.



Фигура 2. Снимка на екрана от представяне на информационната система на *Research Space*, в която се търси (1) обект, направен от (2) дърво и свързан/разрушен от (3) събитие/война от (4) определен период

<sup>1</sup> Рамка за описание на ресурси (RDF). <http://www.w3.org/RDF/>. Последно посетен 4.11.2015

<sup>2</sup> Език за уеб-онтологии (OWL). <http://www.w3.org/TR/owl-features/>. Последно посетен 4.11.2015

<sup>3</sup> Отворени свързани данни (LOD). <http://linkeddata.org/>. Последно посетен 4.11.2015.

<sup>4</sup> Simple Knowledge Organization System (SKOS). [www.w3.org/2004/02/skos](http://www.w3.org/2004/02/skos). Последно посетен 4.11.2015.

<sup>5</sup> SPARQL Query Language for RDF. <http://www.w3.org/TR/rdf-sparql-query>. Последно посетен 4.11.2015.

<sup>6</sup> Проект *Research Space* и ползването на CIDOC-CRM в информационната система. <http://www.researchspace.org/>. Последно посетен: 6.02.2016.

Целта на Research Space е създаване на *Виртуална изследователска среда за изкуство*. Към момента това е единствената работеща в музейни условия среда, която предоставя контекстно търсене, базирано на стандарта за културна информация **CIDOC-CRM**, признат от Международната организация на музеите ИКОМ.

Създаденото през 2016 г. приложение в рамките на проекта ползва семантичната база данни OWLIM, като обектите са представени като отворени свързани данни (ОСД). Стандартите и технологиите за ОСД имат амбицията да революционизират *извличането на информация* чрез недвусмислено търсене и свързване на релевантна информация от различни източници при постоянно динамично обновяване в уеб среда. Контекстното търсене в посоченото приложение става по 6 критерия: хора/групи, обекти, понятия, събития, места и периоди/време (фиг.2).

Стандартите на WEB 3.0 правят бързо откриваеми музейните обекти в една информационна система чрез структуриране на стандартизирано представена и многократно използвана информация.

*Стандартите за културна информация*, които описват колекциите и са одобрени от ИКОМ и музейната практика, са също толкова важни, колкото и технологичните стандарти на W3C за семантична интерпретация. Тук ще споменем няколко, поради значението им за прехода от типология на артефактите в най-общ план към онтологии на наследството: CIDOC-CRM<sup>1</sup>, SPECTRUM<sup>2</sup>, LIDO<sup>3</sup>, DUBLIN CORE<sup>4</sup>. [Woolgar & Lezaun, 2015].

Семантичният уеб помага в разпространението на знанието чрез инструменти, които частично автоматизират процеса на семантична интерпретация. Семантичната интерпретация е асоцииране между структурирани подмножества от данни (синтаксис) спрямо множество обекти в тясна област от знания. Онтологиите, базирани на дескриптивната логика са такива инструменти. Те автоматизират семантичната интерпретация чрез модели от тройки: обект, субект, предикат.

---

<sup>1</sup> CIDOC-Conceptual Reference Model [[www.cidoc-crm.org](http://www.cidoc-crm.org)] е стандарт за културна информация, подкрепен от ИКОМ и одобрен като ISO 21127:2006 през 2006; Неговата шеста версия към май 2015 може да се ползва и като формална онтология, създадена за да улесни интегрирането и обмена на информация от хетерогенни източници от областта на КИН. Вж. проекта на Британския музей с българско участие Research Space.

<sup>2</sup> SPECTRUM стандарт за управление на музейни колекции [<http://www.collectionstrust.org.uk/collections-link/collections-management/spectrum>]. Последно посетен 4.11.2015.

<sup>3</sup> LIDO е XML базиран стандарт за описание на музейни артефакти [<http://network.icom.museum/cidoc/working-groups/lido/what-is-lido>]. Последно посетен 4.11.2015.

<sup>4</sup> DUBLIN CORE е стандарт за описание на музейни артефакти, ползван от проекта Европейска [<http://dublincore.org>]. Последно посетен 4.11.2015.

## Анализ на изследванията в областта и българска музейна практика

Българската музейна практика на описание на обектите, структурираните подмножества от данни, която е обект на нашия интерес, може да се разгледа в музейните инвентарни книги, паспорти и досиета. Това са три типа документи със сходни, но различни полета за попълване. Тяхното попълване е свързано с познания на специализирана терминология от профила на музея – антропология, археология, етнография. Тези познания обаче остават „скрити“ за масовата публика поради липсата на унифициран музеен софтуер за документация, базиран на стандарти, специализирани тезауруси и съвременни функционалности.

Паспортите и инвентарните книги масово се попълват на хартиен носител, без възможност за автоматизирано търсене, допълване с нови полета и редакция на съществуващи. Предложените в статията нови полета са взети от практиката и съществуват дефакто като неформален „стандарт“. Промяната се затруднява и поради нормативни пречки: заличаването и вмъкването на нови полета в инвентарните книги и паспортите е забранено от Закона за културното наследство.

При създаване на рамка за онтологичен модел на музеен паспорт и инвентарна книга е необходимо съобразяването им както с нормативната уредба, музейната практика и спецификата на конкретната колекция, така и с *формален стандарт за културна информация*. Широко ползван е опростеният стандарт DUBLIN CORE, но международната организация на музеите ИКОМ препоръчва ползването на комплексния стандарт CIDOC-CRM. Неговото модифициране за нуждите на конкретен музей обаче изисква сериозно научно и приложно усилие, както и обучен IT екип, с каквото българските музеи не разполагат.

## Музейна документация у нас

Научният паспорт е част от научната обработка на движимата културна ценност. Научната обработка включва както паспортизация, така и електронна обработка. Електронната обработка според Закона за културното наследство трябва да се извършва по *единен софтуер за управление на музейните фондове, одобрен от министъра на културата* (към ноември 2015 г. такъв не съществува). Паспортът е документ с формат А5, със 17 съставни полета от двете му страни, вкл. поле за диапозитиви и рисунки. Полетата на паспортите също са *категорийни*, т.е. едно такова поле включва около четири допълнителни полета. Всички паспортни полета трябва да се включат на съответното им място в онтологичния модел. Анализът на практиките на документиране в нашите музеи води до няколко извода:

- Музейната документация в електронен вид (там, където я има) е слабо структурирана, рядко попълвана в цялост и не следва утвърдени технологични стандарти. Създаването на *модел* на музеен паспорт е част от решението на проблема, заедно със структуриране на полетата и

допълването им с нови. Направени са множество интервюта, анкети и анализи на добри европейски практики.

- Решение, което би работило у нас, и то в съответствие с тенденциите за развитие и управление на съвременния музей, е интеграцията на полетата чрез предметна онтология и речник на предметната област (тезаурус от типа на ААТ на института GETTY) в информационна система с добър потребителски интерфейс. Не е достатъчно премахването на структурната нееднородност на информацията. За да има коректна семантична интерпретация на различни по вид музейни колекции е необходимо ползването на съществуващ тезаурус, или създаването на нов такъв.
- Наличието на (1) краен разширяем речник е едно от трите задължителни свойства на една онтология, която да е в основата на музейна информационна система, заедно с (2) възможността за еднозначна интерпретация на класовете и релациите и (3) йерархична структура на системата от класове (наличие на релация подклас–клас или клас–суперклас, която определя йерархията между класовете). Съществува разлика в използвания речник и терминология дори в рамките на един музей, което предполага съчетаване на няколко речника от различни предметни области (ако това е технически възможно).

Необходимата промяна е свързана с новите възможности на семантичните технологии за създаване на нови или модифициране на съществуващи информационни каталогизиращи системи, които да предложат опростен потребителски интерфейс, удобен за музейния служител. Такава система трябва не само да следва изискванията на нормативната база и музейната практика, но и да включва функционалности за семантично търсене и извеждане на нови знания. Това изисква създаване на обща и/или предметна онтология според класификацията на *Гуарино* и *Оберле* [Guarino et al, 2009], действащи заедно с тематични речници/тезауруси.

## **Типологизация на информацията в музейния паспорт**

Настоящият текст е съсредоточен върху типологизация на съставните полета на музейния паспорт според характеристиката, която описват. Оформени са три типа полета: **административни**, служещи за отчетност и контрол, **дескриптивни**: описват вида, съдържанието и „значенията“ на движимата културна ценност с цел групиране в колекции и изложби, и **технически**. Пример за административни съставни полета (метаданни) са: наименование на музея, каталожен № или инвентарен №, наименование, фонд. Пример за описателни съставни полета са: форма, материал, размер, тегло, местопроизход, период, функция и др. Пример за технически полета са: файлов формат, резолюция на изображението, местонахождение на файла и тип електронен достъп до него, и др. Тези три типа полета показват различни „същности“ на предмета и са косвено свързани помежду си чрез свойствата им (неявно подразбиране).

На научна паспортизация подлежат основният и обменният фонд на всеки български музей (вж.фиг.3).

**Д-р П. Бърнев**      Селище **Младина**      (653) 1455

Адрес гр. **Сабот Тръст - 1010 - 70 1880 Санданово - 15.9.1966**

Източник на творбата, година - **1966 г.**      Измерения - **10x10**

Техника, материал - **л. в. на картон**      Измерения - **10x10**

Датум, година - **15.9.1966**

Идентификация на творбата в регистъра

Класификация	Крайна цена	Класификация	Пазар	С всички данни	Датум
2	500		1060		

Описание на творбата:  
Копие с оригинала на К. М. П. на 1966 г. в гр. Сабот Тръст. Изготвянето е свършено в гр. Сабот Тръст. Височина на оригинала е 10 см. Широчина на оригинала е 10 см. Цвят на оригинала е черен.

Състояние: **добра**

Реставрация: **да** и **не** за протозоните

Дата на творбата	Дата на събирането	Идентификация	Крайна цена	Пазар	Дата на събирането
1966	10.1962	19	500	1060	10.1962

Фигура 3. Научен паспорт на движима културна ценност

Според нормативната база<sup>1</sup> научният паспорт разкрива историческото, природното и художественото съдържание на движимата културна ценност за научни, образователни и възпитателни цели. Научните паспорти и техните електронни носители се съхраняват в научния архив на музея. Предвидено е

<sup>1</sup> НАРЕДБА № Н-6 от 11.12.2009 г. за формиране и управление на музейните фондове, издадена от министъра на културата, обн., ДВ, бр. 2 от 8.01.2010 г.

създаване на „научно-справочни картотеки“, но там, където има такава, достъпът е ограничен. Нормативните предписания за музейна документация в Закона за културното наследство и свързаните наредби не са полезни за създаването на детайлна рамка за унифициран електронен вариант. Към това се добавя и липсата на национална стратегия за дигитализация и културна политика в областта. Естественият резултат е, че българският музей „решава“ неунифицираната задачата за електронната каталогизация на своите фондове.

Особено важен недостатък на нормативната база е, че тя не дефинира процедура как и по какви критерии да бъдат организирани артефактите в наличните описания – във фондове (музеи, архиви), в колекции (музеи), в сбирки (библиотеки). Не се дефинира принципът на описание на обектите – според местоположението или според реда на постъпване, или нещо друго. Анализът на съществуващите практики за организация на музейните фондове ги свежда до три: (1) *хронологично*, по исторически периоди, (2) *по тип документ* или (3) *по ред на постъпване*. Организацията по който и да е от тези критерии в информационна система, базирана на онтологии (и терминологични речници), ще позволи лесно сортиране на записите, търсене и извличане на данни по критерии, определени от потребителя.

### **Онтологичен модел за научен паспорт на движима културна ценност**

Описаният по-долу модел няма претенция да реши наболелите проблеми и задачи с е-документацията в нашите музеи. Това е въпрос на национална политика. Тук се предлага начална стъпка за анализ на различни типове полета, препоръчани от музейни специалисти, полета от трите типа документи, които да са в основата на разширяем концептуален модел на онтология. Моделът и реалната онтология могат да се ползват в бъдеще от музейна информационна система, като улеснят както въвеждането на данни, така и контекстното търсене.

Дейността на разработчиците на онтологии е затруднена от обстоятелството, че „на практика почти липсват утвърдени общи и верифицирани методологии, които да определят какви процедури трябва да се изпълнят в процеса на създаването на една онтология“ [Нишева, 2013]. Понастоящем съществуват малък брой предметно-независими методологии (подходи) за построяване на онтологии, най-популярна е METHONTOLOGY [López et al, 1999].

При предварителната работа по анализ на всички полета от трите типа документи в София, Пазарджик, Шумен, Дупница и Пловдив се стигна до извода, че инвентарната книга предлага твърде малко полета. Нейната цел не е да опише в пълнота обекта, а да улесни неговата инвентаризация. В модела са включени и полетата на музейния паспорт, с което описанието е по-пълно, а така също и нови полета, предложени от специалистите по места. Определихме 25 полета (класове, категории), които структурираме в таблица според техния

вид и съответствие с категории на CIDOC-CRM. Таблицата е повлияна от онтология, създавана с OWL редактора на Станфордския университет *Protege*<sup>1</sup>.

Предложеният комплект (set) от 25 съставни „полета“ не включва всички 90 категории на CIDOC-CRM. 25-те полета са разделени на три типа – дескриптивни, технически (структурни) и административни.

Дескриптивните метаданни описват ресурса с цел откриване и идентифициране (Пример: заглавие, анотация, автор и ключови думи). Структурните метаданни посочват как съставни обекти могат да бъдат описани заедно. Административните метаданни предоставят информация, подпомагаща управлението на ресурса (Пример: кога е бил създаден, тип на файла/документа и др.).

Тези три типа са съотнесени спрямо съдържанието на съответните категории на CIDOC-CRM и Dublin Core (DC).

## **Източници**

Работата по онтологичния модел се улеснява от онтологичното представяне на CIDOC-CRM като свободно достъпен OWL документ, както и техническата документация на Британския музей по проекта Research Space, който единствен от европейските музеи към днешна дата въвежда този стандарт в своите каталози. Целта на големия екип на проекта Research Space е създаване на *виртуална изследователска среда за изкуство*. 2.5 млн. обекта на Британския музей са представени като отворени свързани данни.

M. Doerr [Doerr, 2015] е посочил *mapping* метод (proof of concept) за картиране на DC към CIDOC. В онтологичния модел ще ползваме DC, като имаме предвид такава възможност в бъдеще, макар и трудна. Изборът ни е мотивиран и от факта, че към ноември 2015 г. само Британският музей, в сътрудничество с екип от IT експерти, работи със CIDOC-CRM. Това е първата онтология, представена на Международния комитет по стандартизация (ISO), за да стане международен стандарт за трансфер на данни от областта на културното наследство (*не* стандарт за метаданни!), а DC е популярен и опростен библиографски стандарт за метаданни, който съдържа минимален комплект от семантични полета (access points) и е масово ползван в европейските музеи, а дори и в някои български.

Елементите на DC<sup>2</sup> са: (1) DC: contributor, (2) DC: coverage, (3) DC: creator, (4) DC: date, (5) DC: description, (6) DC: format, (7) DC: identifier, (8) DC: language, (9) DC: publisher, (10) DC: relation, (11) DC: rights, (12) DC: source, (13) DC: subject, (14) DC: title, (15) DC: type. В музейната практика тези елементи се разбират неунифицирано, поради което са добавени множество речници.

---

<sup>1</sup> Protégé е редактор за онтологии с отворен код на Станфордския университет.  
<http://protege.stanford.edu>. Последно посетен 4.11.2015.

<sup>2</sup> Dublin Core [<http://dublincore.org/documents/dces/>]. Последно посетен 4.11.2015.

*Таблица 1. Данни в музейния паспорт на ДКЦ и съответствията им с полетата в стандартите CIDOC-CRM и Dublin Core*

№	НАИМЕНОВАНИЕ НА ПОЛЕТО	Категории метаданни Д - Дескриптивни Т - Технически А - Административни	CIDOC-CRM ENTITY	DUBLIN CORE
1	НАИМЕНОВАНИЕ НА МУЗЕЯ	(А)	E45, E47, E48	DC:contributor
2	НОМЕР ПО РЕД	(А)	служебно попълване	служебно попълване
3	КАТАЛОЖЕН № / ИНВЕНТАРЕН № <sup>1</sup>	(А)	E15, E49	DC: identifier
4	НАИМЕНОВАНИЕ	(А)	E35	DC: title; DC: subject
5	ФОНД / МЕСТОЛОЖЕНИЕ	(А)	E45, E47, E48, E53	-
6	ДИАПОЗИТИВ №, НЕГАТИВ №, ФОТОГРАФИЯ	(Д)	E38, E36, E62	-
7	МЕСТОПРОИЗХОД <sup>2</sup>	(Д)	E12, E2, E22/24, E39, E65	-
8	ПРОИЗХОД <sup>3</sup>	(А)	E15, E49, E63/64, E8	DC: creator
9	СЪСТОЯНИЕ ПРИ ПРИЕМАНЕТО	(А)	E14, E3	-
10	ОЦЕНКА (ЦЕНА В ЛВ.)	(А)	E13	-
11	ФОРМА	(Т, Д)	E13, E46	DC: format (?)
12	МАТЕРИАЛ	(Т, Д)	E13, E57	DC: format (?)
13	РАЗМЕР	(Т)	E13, E16, E54, E46, E58	DC: format (?)
14	ТЕГЛО	(Т)	E13, E58	DC: format (?)
15	НАДПИСИ, ДАТИ, ПОДПИС, ГОДИНА	(Д)	E13, E34, E33, E37	DC: date DC: coverage

<sup>1</sup> Инвентарните номера се поставят хронологично, по реда на постъпването на обекта във фонда, а не по местоположението му към момента. Това създава проблеми при при търсенето по местоположение.

<sup>2</sup> Местопроизход: къде, как, кога и от кого е намерен и изработен.

<sup>3</sup> Произход: стар инвентарен №, дата и номер на приемо-предавателния протокол, КП №. Как е придобит предметът: собствен (лева); купен (лева); подарен (лева); от кого е откупен или подарен (трите имена и адрес).

16	ВИД	(Д)	E13, E55	DC: type
17	БРОЙ	(Т)	E13, E60	-
18	РЕСТАВРАЦИЯ	(Т)	E81, E79, E11	-
19	ИСТОРИЧЕСКА СПРАВКА	(Д)	E62, E63/64	DC: description DC: source DC: relation
20	ПРАВА	(Д)	E62	DC: rights
21	ДВИЖЕНИЯ НА ДКН	(А)	E78, E87, E9, E10	-
22	АРХИВНИ ДАННИ	(Д)	E31	-
23	ЛИТЕРАТУРНИ ДАННИ	(Д)	E62	DC: publisher (?) DC: language
24	ИМЕ И ДЛЪЖНОСТ НА ПОПЪЛВАЩИЯ	(А)	E21	DC: contributor
25	ДАТА НА ПОПЪЛВАНЕ	(А)	E21, E50	DC: contributor

За съжаление, DC елементите не покриват съдържателно всички необходими за музейния паспорт полета. Табличното представяне показва невключените полета с тирета. Включването им е въпрос на бъдещо проучване на изискванията на представителни типове музеи у нас, за да се унифицира подходът. Предвижда се и включване на *музейни процедури*. Моделът няма претенциите да е детайлен, той „рамкира“ първи стъпки за създаване на онтология за документиране и описание на движими културни ценности. Изборът на подходящ стандарт за нея – Dublin Core или CIDOC-CRM – трябва да отразява дългосрочна национална политика, но препоръката на ИКОМ за CIDOC-CRM трябва да се има предвид. Нашата крайна цел е на базата на предлагания модел да се създаде работеща в български музейни условия информационна система с функционалности за контекстно търсене поне по три критерия: номер на паспорта, име на обекта и исторически период. Това ще бъде предмет на бъдеща изследователска и приложна работа.

## Литература

- [Dobrova & Duff, 2015] M. Dobrova, W. Duff. The Ever Changing Face of Digital Curation: Introduction to the Special Issue on Digital Curation. In: Springer Science+Business Media Dordrecht, 2015, 15(2), pp. 97-100.
- [Doerr, 2015] M. Doerr. Mapping Dublin Core Metadata Set Elements to CIDOC-CRM. [http://www.cidoc-crm.org/docs/dc\\_to\\_crm\\_mapping.pdf](http://www.cidoc-crm.org/docs/dc_to_crm_mapping.pdf). Last accessed: 5.11.2015.
- [Guarino et al, 2009] N. Guarino, D. Oberle, St. Staab. What Is an Ontology? In: Handbook on Ontologies, Springer, 2009, pp. 1-17.
- [Issac & Baker, 2015] A. Isaac, T. Baker. Linked Data Practice at Different Levels of Semantic Precision: The Perspective of Libraries, Archives and Museums. In: Bulletin of the American Society for Information Science and Technology, 41(4), 2015, pp. 34-39.

- [López et al, 1999] M. F. López, A. Gómez-Pérez, J. P. Sierra. Building a Chemical Ontology Using Methontology and the Ontology Design Environment. In: IEEE Intelligent Systems, 14(1), Jan. 1999, pp. 37-46.
- [Minelli et al, 2014] S. Minelli et al, MOVIO: A Semantic Content Management and Valorisation Approach for Archives and Cultural Institutions. doi:10.1016/j.procs.2014.10.006.
- [Nisheva et al, 2014] M. Nisheva-Pavlova, N. Spyrtos, P. Stanchev. Museum Collections and the Semantic Web. In: Digital Presentation and Preservation of Cultural and Scientific Heritage, issue: IV, 2014, Sofia, pp. 33-39.
- [Szekely et al, 2014] P. Szekely, C. Knoblock, F. Yang, E. Fink, S. Gupta, R. Allen, G. Goodlander. Publishing the Data of the Smithsonian American Art Museum to the Linked Data Cloud. In: International Journal of Humanities and Arts Computing. Vol. 8, Issue supplement, pp. 152-166.
- [Woolgar & Lezaun, 2015] S. Woolgar, J. Lezaun. Missing the (question) mark? – What is a turn to ontology? In: Social Studies of Science 45.3 (2015), doi:10.1177/0306312715584010.
- [ЗКН] Закон за културното наследство и НАРЕДБА № Н-6 от 11.12.2009 г. за формиране и управление на музейните фондове, издадена от Министъра на културата, обн., ДВ, бр. 2 от 8.01.2010.
- [Нишева, 2013] М. Нишева-Павлова. Създаване на хетерогенни цифрови библиотеки, основано на онтологии. Хабилизационен труд във връзка с конкурс за професор по информатика и компютърни науки, ФМИ при СУ, 2013.

## **Ontological Model of Museum Documentation: Bulgarian Practice Example**

**Kalina Sotirova**

**Abstract:** Digital cultural heritage (DCH) requires interdisciplinary approach and rich expertise. Information management in DCH includes management of documenting, transfer, presenting and long term preservation of the original artifacts of human culture (tangible and intangible). High requirements of museum sector for documenting, description and search of movable cultural assets is a big challenge for creating effective information systems with rich semantic functionalities. Rapid development of standards, technologies and applications of WEB 3.0 (RDF, SPARQL, OWL, SKOS) offers new possibilities in contextual presentation of memory institutions artifacts. The current paper offers analysis of museum documentation in Bulgaria, resuming legislation requirements for description and documenting in field typology. In Bulgarian museum practice „museum passport“ of each object has numerous fields to describe it. We create conceptual model of ontology, creating new typology of these fields, based on W3C standards of Semantic Web.

## Цифровизирани колекции на пловдивски културни институции

Владимир Балчев<sup>1</sup>, Димитър Минев<sup>2</sup>, Георги Врагов<sup>3</sup>

1. Общински институт „Старинен Пловдив“  
[baltchevvlad@abv.bg](mailto:baltchevvlad@abv.bg)

2. Регионална народна библиотека „Ив.Вазов“, Пловдив  
[dimin@abv.bg](mailto:dimin@abv.bg)

3. Институт по математика и информатика при БАН, София  
[vragov@math.bas.bg](mailto:vragov@math.bas.bg)

**Резюме:** В доклада се прави обзор на дейностите, извършени през последните няколко години от пловдивски културни институции съвместно със специалисти от Института по математика и информатика при БАН за цифровизиране на обекти от техните колекции и поместването им в цифрови библиотеки. По-конкретно са представени постигнатите резултати през 2015 г. Отделено е внимание на виждането на специалистите от тези институции за насоките на развитие и на започналата работа по проекта „Дигитално културно-историческо наследство на община Пловдив“.

**Ключови думи:** цифровизирани обекти, цифрови библиотеки, метаданни, агрегатор.

### 1. Въведение

През 2015 г. започна работа по проект „Дигитално културно-историческо наследство“ на Община Пловдив. Основните дейности по проекта са свързани с дигитализацията на обекти от културното и историческо наследство на града – съдържание, съхранявано в пловдивски културни институции. Подготовката на обектите и описанието на метаданни за тях става от тези институции (Регионална народна библиотека „Иван Вазов“, Регионален исторически музей, Регионален археологически музей, Общински институт „Старинен Пловдив“ и Градска художествена галерия, Пловдив) – партньори на общината. Резултатите ще се представят в обща цифрова библиотека, а някои от тях ще допълнят представянето на града и в тази на Европейския съюз – <http://www.europeana.eu>.

Работата по проекта се основава на натрупания опит през последните няколко години от съвместната работа в тази съвременна технологична област.

В доклада се обобщава извършеното от културните институции в досегашната им дейност и тяхното виждане за развитието на проекта.

### 2. Регионална народна библиотека „Иван Вазов“

Дигитализирането на информация, улеснените способности за пренос и търсене на данни са предпоставки за решения за по-пълноценното им използване.

Специалистите от Пловдивската библиотека през последните години постигнаха сериозни резултати при създаване на цифровизирани колекции за опазване на библиотечните фондове, които са на разположение на потребителите при ползването на автентични документи. Тези резултати са следствие на самостоятелната работа на сътрудници от библиотеката и компетентни в областта специалисти от други институти, а също и резултат на участие в национални и международни проекти за дигитализация. Поставена бе целта за създаване на цифрова библиотека като част от Европейската електронна библиотека Europeana[1][2]. Това предизвика запознаването със съвременни методи и технологии за съхраняване и опазване на националните книжовни и други паметници и осигуряване на възможности за бърз достъп на потребителите до цифровите ресурси в помощ на образованието, научните изследвания, културата и други области.

Цифровизирането на значимите и ценни фондове на Регионална народна библиотека „Иван Вазов“ започна от 2008 г. Чрез участие в съвместни проекти с други библиотеки и с помощта на специалисти от Института по математика и информатика при БАН бе натрупан значителен масив от електронни пълнотекстови бази данни. Интерес представляват електронните каталози „Старопечатни, редки и ценни издания“ от фондовете на отдел „Специални сбирки“, „Периодичен печат в България от възникването му до 1944 г.“, обхващащ целия източнорумелийски печат, голяма част от Възрожденския печат и редица ежедневни издания от началото на XX в., уникалните колекции фотографии и картички, носещи специфичния облик на времето, за което се отнасят. В периода до 2012 г. новите проекти на библиотеката за опазване на културното и историческо наследство позволиха създаването на професионален работен екип, обучен за работа със съвременните информационни и комуникационни технологии.

Проектът за съхраняване на старопечатните, редки и ценни издания [3] позволи въвеждането на специализиран модул за създаване на електронен каталог за съхраняване и описание на старопечатни, редки и ценни книги чрез информационно-търсещата система E-Lib, тяхното представяне в Интернет чрез програмния продукт I-Lib с възможност за съвременен и удобен интерфейс на основата CDS/ISIS.

Основни критерии за подбор при формиране на дигиталните колекции са: значението на материалите за културното наследство на общността, региона, страната; уникалността на оригинала; физическото състояние на материалите и нуждата от опазване на оригиналите чрез създаване на дигитални версии като алтернативи за ползване.

Дигиталният център на библиотеката е създаден през 2008 г. като продължение на проект за опазване на културно-историческото наследство. Някои от проектите и резултатите от тях, предоставени на широкия кръг потребители чрез цифрови образи на особено ценните писмени документи, са:

- Своден електронен каталог „Старопечатни редки и ценни книги“. Към него са добавени в пълен текст – петнадесет редки и ценни книги с 4219

сканирани страници и двадесет и девет старопечатни книги с обем от 6004 сканирани страници.

- Цифровизиран каталог „Портрети и снимки“, финализиран през 2015 г., като към него са добавени снимките, влизащи в състава на Българския исторически архив на библиотеката в обхват 3446 записа и 4373 сканирани снимки.
- Участие в EuropeanaPhotography – международен проект, финансиран от Европейския съюз, в който участват деветнадесет институции от тринадесет европейски държави. Началото му бе поставено през 2012 г., а продължителността му бе до 2015 г. Наименованието на проекта е абревиатура на израза „Европейски хранилища на колекции от дигитализирани стари фотографии с историческа стойност“(EUROPEAN Ancient PHOTOgraphic vintaGe repositoRies of digitized Pictures of Historical qualitY). Целта му е попълване в Европейската дигитална библиотека Europeana на уникални и качествени цифрови изображения от първите сто години на фотографията (1839–1939).
- Участие в Европейския проект EMBARK с четири ръкописа. Manuscriptorium е европейска дигитална библиотека за документално културно историческо наследство, която предоставя достъп до споделени ресурси на културни институти на страни от и извън Европейския съюз. Услугата осигурява непрекъснат достъп до повече от пет милиона цифрови изображения. Manuscriptorium е подагрегатор на Europeana в сферата на историческите ресурси.
- Изготвен е цифровизиран каталог на ръкописите от Агушевата библиотека.
- Изготвен е цифровизиран каталог на Кукленските ръкописи.
- Допълва се цифровата колекция „Обединен каталог на български периодични издания до 1944 г.“.
- От 2012 г. Регионална народна библиотека „Иван Вазов“ е асоцииран член на Фондация „Национална академична библиотечно-информационна система“ (НАБИС).
- Програмата за цифровизация на славянската ръкописна колекция в библиотеката започна през 2013 г. За тригодишния период бяха дигитализирани пълнотекстово деветдесет и пет славянски ръкописа (XII-XVIII век), които са достъпни към момента на съвършите на цифровата библиотека на Фондация НАБИС. Предстои работа по заснемането на последните четиредесет кодекса от тази уникална ръкописна сбирка.
- Участие в цифровата библиотека Plovdivartefacts от колекции на пловдивски културни институции.

През март 2015 г. беше спечелен проект по Програма БГ08 „Културно наследство и съвременни изкуства“, мярка 2 „Документиране на културната история“ за дигитализация на културно-историческо наследство, който се изпълнява в партньорство между Община Пловдив и пловдивски културни институции. Планираният финансов ресурс за дейностите по дигитализация на

културно-историческо наследство в културните институти – партньори по проекта е в общ размер от 586779 €. В настоящия момент протича процесът за оборудването на цифровите лаборатории със съвременни технически и програмни ресурси.

Дигитализацията позволява на библиотеката да обслужва традиционните си ползватели по нов начин. Модерните технологии предлагат нови подходи към осъществяването на социално значими цели – получаване на образование, извършване на научни изследвания, културно израстване и пр. Усилията ни се насочват и към представянето на съвременните писмени постижения с отчитане на ограниченията и трудностите, предизвикани от Закона за авторското право и сродните му права.

Отчитаме повишените изисквания в епохата на информационното общество към квалификацията на библиотечните специалисти и най-вече при ползването в текущата им работа на най-модерни технологии и предвиждаме мерки за тяхната квалификация.

### **3. Процесът на цифровизация в Регионален исторически музей – Пловдив през 2015 г.**

За опита на Регионалния исторически музей в цифровото представяне на свои колекции свидетелстват доста публикации – [9][10] и др. Добри резултати през последните няколко години са постигнати в сътрудничество със секция СофтИС на Института по математика и информатика при БАН.

През 2015 г. музеят стана част от Google Cultural Institute. Представянето му е вече достъпно за всички потребители [11]. Тази онлайн платформа функционира като глобална мрежа, обединяваща световните музеи, галерии и архиви. Артефактите могат да бъдат разглеждани в изключително висока резолюция и представляват интерес както за любителите на историята, така и за изследователите, които имат професионално отношение към съхраняването във фондовете.

Схемата на метаданни включва най-важните характеристики на обекта – `itemid`, `title`, `description`, `creator`, `location:placename`, `date:start`, `date:end`, `filetype`, `filespec`.

На този етап Регионален исторически музей – Пловдив участва със 150 артефакта и две онлайн изложби – „България в Първата световна война“ и „Душата на японската кукла“ [11]. Предстои допълването им, както и създаването на нови изложби.

Музеят продължава дигитализацията на своите фондове, като ще подобри участието си в Регионалната цифрова библиотека на културните институции. В тази насока се разчита и на резултатите от работата по проекта за дигитализиране на културно-историческото наследство на Пловдив по програмата на Европейското икономическо пространство.

#### **4. Процесът на цифровизация в Регионален етнографски музей – Пловдив през 2015 г.**

Работата на пловдивския Етнографски музей в областта на цифровото представяне на свои колекции започна с подготовката и представянето на над 400 обекта от колекцията на фонд „Занаяти“ в инициативата на ИМИ при БАН за създаване на агрегатор на цифрови колекции в рамките на проекта Д002-308 към НФНИ „Автоматизирано генериране на метаданни за спецификации и стандарти на е-документи“ [5][6][7].

През 2015 г. РЕМ-Пловдив изгради устойчиви партньорски отношения с Google Ireland Limited. Във връзка с това бе използвана платформата Google Cultural Institute за дигитализиране на движими културни ценности и изложби на музея. До този момент има дигитализирани две етнографски изложби – „Знаци на времето“/“Signs of Time“, представяща историята на времеизмервателните уреди от древността до наши дни, и „Български традиционни накити“/“Bulgarian Folk Jewels“, представяща традиционни накити от епохата на Възраждането [12]. В тези изложби са включени 129 движими културни ценности със снимки на висока резолюция и с метаданни за всеки предмет – на български и английски език. В момента работим върху дигитализацията на най-новата изложба на РЕМ-Пловдив – „Романтика от кутията за спомени“/“Romance from the Memory Box“. Освен галерия с изображения с добра резолюция от фондовете на музея, виртуалните изложби са снабдени и с текстови пана, които дават допълнителна информация по темата, представяна в изложбата.

Засега платформата GCI не поддържа българска версия, но дигитализираните материали и данните за предметите са подготвени и на български език. Изображенията са защитени от копиране или изтегляне, като авторските права принадлежат изцяло на РЕМ-Пловдив. Дигиталните изложби са достъпни за вграждане в други сайтове след съответното съгласуване. Двете изложби са включени и в сайта на РЕМ-Пловдив ([http://www.ethnograph.info/front/virtual\\_exhibitions.php](http://www.ethnograph.info/front/virtual_exhibitions.php)), а също така могат да бъдат разгледани направо от страницата на Google Cultural Institute.

Във връзка с партньорството си с Google Ireland Limited, РЕМ-Пловдив включи в проекта на Google Inside Street View. Интериорът на РЕМ-Пловдив беше заснет с модерни 3D устройства. Визуалните материали са обработени и свързани с някои от дигитализираните обекти, така че да може да се разглеждат отблизо в експозицията с подробна информация. На 21 януари 2016 г. дигитализацията на интериора на РЕМ-Пловдив ще бъде онлайн представена заедно с дигитализацията на други партньори на Google по проекта Inside Street View. Тогава е и началото на новата дигитална изложба „Romance from the Memory Box“/“Романтика от кутията за спомени“.

Регионалният етнографски музей – Пловдив е партньор по проект „Дигитално културно-историческо наследство на Община Пловдив“ (2015-2016) по мярка 2 „Документиране на културната история“ по програма БГ08 „Културно наследство и съвременни изкуства“, финансирана от ФМ на ЕИП. По този проект РЕМ-Пловдив предвижда дигитализиране на около 9000 броя

движими културни ценности от фондовете на музея, представлящи етнографското културно-историческо наследство на Пловдив и региона. Също така се предвижда дигитализация на съвременни движими културни ценности и нематериално културно наследство на различни етноси, във връзка с проектираните теренни етнографски експедиции, които ще бъдат проведени на територията на град Пловдив.

## **5. Цифрова библиотека на пловдивските културни институции**

През април 2012 г. стартира съвместната работа по цифровизация на колекции и обекти, свързани с културно-историческото наследство на Пловдив и региона. В рамките на тази идея се подготвят и публикуват цифровизирани обекти, обединени по определени тематички с достъп през общ портал. Инициатори на проекта са Пловдивската народна библиотека в сътрудничество с Регионалния исторически музей, Регионалния археологически музей, Регионалната художествена галерия и Регионалния държавен архив. Системата е представена като агрегатор на цифрови културно-исторически артефакти, базиран на стандарта, използван от европейското цифрово хранилище Europeana. За целта е ползвана технологичната разработка на секция СофтИС на Института по математика и информатика при БАН [5][6][7].

Необходимо е развитие на съществуващия проект за представяне на пловдивските културни институции в обща цифрова библиотека и това е още по-актуално с обявяването на Пловдив за Европейска културна столица през 2019 г.

Първите стъпки и съответните изводи от натрупания опит във взаимодействието на пловдивските културни институции са налице [4][8].

Нашето разбиране за развитието на този проект е както в количествено отношение – увеличаване на колекциите и обектите, включени в тях, така и добавянето на нови функционални възможности според очакванията на съвременните потребители на цифровизирано съдържание.

Едно изследване на д-р Милена Добрева ([http://www.um.edu.mk/maks/las/staff/dr\\_milena\\_dobрева](http://www.um.edu.mk/maks/las/staff/dr_milena_dobрева)), направено съвместно с нейни студенти от ПУ „Паисий Хилендарски“, представя някои от тези изисквания към цифровата библиотека:

- да позволява достъп до книгите и другите архивни единици на фондовете;
- да разполага с копие от всяка картина, фотография, филм или музикално произведение;
- да поддържа аудио- и видеосъдържание от радио- и телевизионни програми;
- Интернет страниците вече имат своята немалка история, съдържат огромна по обем информация и също трябва да се намерят гъвкави технологични решения за архивирането и ползването им;
- мултимедийното, интерактивно съдържание е предпочитано пред хартиения си вариант и пр.

Напредването на Web технологиите, възможностите за извличане на знания и семантично търсене изискват подобряване на софтуерните решения. Споделянето на цифрово съдържание със съвременни мултимедийни средства е все още недостигната задача в този проект.

## **Заклучение**

Първите стъпки за дигитализиране на богатото културно и научно съдържание във фондовете на пловдивските културни институции са направени. Предстои още много самостоятелна и съвместна работа за агрегиране на тези огромни количества информация и създаването на удобен и ефективен достъп до ресурсите. Осъзната е отговорността за цифровизиране на нашето културно и историческо наследство и интегрирането му в глобалните библиотеки на съвременното общество.

## **Литература**

1. European Digital Library – [www.europeana.eu](http://www.europeana.eu).
2. Минев, Д.: Доклад на международната конференция „Стратегическият план за Европеана 2011- 2015“, Брюксел, 29 март 2011 г.
3. [http://www.libplovdiv.com/index.php?option=com\\_content&view=article&id=132&Itemid=174&lang=bg](http://www.libplovdiv.com/index.php?option=com_content&view=article&id=132&Itemid=174&lang=bg)
4. Минев, Д.: Инициатива за създаване на Регионален културен цифров портал, Конференция „Европеана и българските институции“, Пловдив, 3-4.04.2012.
5. Ivanova, K., Dobрева, M., Stanchev, P., Totkov, G. (editors): Access to Digital Cultural Heritage: Innovative Applications of Automated Metadata Generation, University Publishing House „Paisii Hilendarski“, (2012), Plovdiv, Bulgaria, book site: [www.math.bas.bg/infres/book-ADCH](http://www.math.bas.bg/infres/book-ADCH).
6. Сомова, Ел., Врагов, Г., Тотков, Г.: Към регионален агрегатор на цифровизирани културно-исторически обекти, Сборник с доклади на конф. „Образованието в информационното общество“, Пловдив, май 2010.
7. Hadjikolev, E., Vragov, G., Totkov, G., Ivanova, K.: REGATTA – Regional Aggregator of Heterogeneous Cultural Artifacts, Review of the National Center for Digitization, Belgrade, Serbia, Issue: 21, (2012), pp.8-18.
8. Balchev, V., Vragov, G., Ovcharova, R., Staneva M.: Collection of Digital Objects for the Balkan War (1912-1913), Proc. of the Fifth Nat. Conf. „Education in the Information Society“, 2012, Plovdiv, pp.171-177.
9. Шивачев, Ст., Карадечева, М.: Опитът на Регионалния исторически музей в дигитализация на фондовете, Конференция „Европеана и българските институции“, Пловдив, 3-4.04.2012.
10. Karadecheva, M.: Online Catalogue – Relics from the Bulgarian Museums, DiPP2012 Conference: V.Turnovo, September, 2012.
11. Регионален исторически музей в Google Cultural Institute: <https://www.google.com/culturalinstitute/collection/regional-museum-of-history-plovdiv>; <https://www.google.com/culturalinstitute/exhibit/bulgaria-in-world-war-i/QQMvZ5Fw>
12. Регионален етнографски музей в Google Cultural Institute: <https://www.google.com/culturalinstitute/exhibit/signs-of-time/gQznE1AX?hl=en>; <https://www.google.com/culturalinstitute/exhibit/bulgarian-folk-jewels/gQxkKRvN?hl=en>

## **Digitized collections of Plovdiv cultural institutions**

**Vladimir Balchev, Dimiter Minev, Georgi Vragov**

**Abstract:** The article provides an overview of activities undertaken in recent years by Plovdiv cultural institutions together with specialists from the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences to digitize objects in their collections and their incorporation into digital libraries. The results achieved in 2015 is presented in details. The experts from these institutions shared their vision about the directions of development of the recently started project "Digital cultural heritage of Plovdiv Municipality".

## **Моделиране на бизнес-обекти чрез СУБД с външна логика на данните**

**Анани Ризов, Илия Митов, Красимира Иванова**

Институт по математика и информатика при БАН, София  
[a.rizov@math.bas.bg](mailto:a.rizov@math.bas.bg), [imitov@math.bas.bg](mailto:imitov@math.bas.bg), [kivanova@math.bas.bg](mailto:kivanova@math.bas.bg)

**Резюме:** В статията се разглежда процесът на изграждане на качествени бизнес модели. Дадени са определения на основните елементи при моделирането на бизнес обекти. Посочено е значението на използването на метаданни за наличието на проблеми при моделирането на бизнес обекти и изграждането на качествено информационно осигуряване на процесите в тези обекти. Дефинирано е наличието на нов тип нерелационни бази от данни – СУБД с външна логика на данните. Посочени са предимствата пред използваните досега СУБД с вътрешна логика на данните при изграждането на качествени бизнес модели.

**Ключови думи:** Бизнес модел, бизнес обект, бизнес процес, СУБД, метаданни, СУБД с вътрешна логика на данните, СУБД с външна логика на данните.

**ACM 1998 Classification Keywords:** H.2.0 Database Management

### **Въведение**

Едно от важните приложения на информатиката е в автоматизирането на управлението на фирмените дейности. Формално тези дейности можем да обобщим като управление на бизнес процеси. Изучавайки от информатична гледна точка последователността от дейности, закономерности и проблеми при изграждането на автоматизирани информационни системи, обслужващи фирменото управление, можем не само да повишим качеството на обслужването на тези процеси, но и да открием нови възможности и направление за развитие в конкретната приложна област и в информатиката като цяло. Управлението на фирмените дейности е подходяща за изследване област, защото:

- има натрупан богат опит на информационното осигуряване в тази сфера;
- правени са опити за формалното описание и изграждането на информационни модели на различни по характер и обем бизнес процеси;
- поддават се на изучаване сравнително лесно (от ресурсна гледна точка);
- новополучените резултати могат бързо да се реализират и внедрят в практиката.

Изследването, описано в тази статия, има за цел да покаже някои от недостатъците на досегашните подходи за изграждане на модели на фирмено управление и да представи възможни нови подходи за преодоляването на забелязаните неудобства при изграждането на информационни модели на фирмените дейности и съответно да предложи нови, по-ефективни механизми

за реализирането на програмни продукти за информационното осигуряване на създадените модели.

## **Бизнес обекти**

Въпреки че в началото обектът на нашето изследване бяха фирмените бизнес структури, с напредване на работата се установи, че тези структури са част от един по-голям клас обекти, към който можем да отнесем не само фирмените бизнес структури, но и обществени, държавни и т.н. структури. По нататък в работата тези обекти ще наричаме „Бизнес обекти“. Може да дадем следното формално определение: **Бизнес обектът е организирана единна структура, която осъществява външни и/или вътрешни съзидателни дейности.** Тези обекти имат еднакво поведение и близки по характер външни (спрямо обкръжаващата среда) и вътрешни (спрямо единната структура) реакции при еднакви външни и/или вътрешни въздействия.

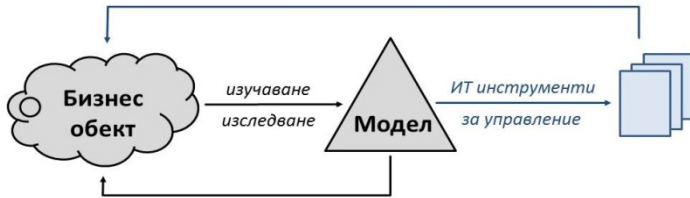
При изследване на дейностите на бизнес обектите следва да отбележим, че за тях е характерно почти пълното отсъствие на надеждни математически модели, които да ги описват. При това положение водещи позиции при моделирането на дейностите и процесите на бизнес обектите заемат ергономични съображения: естественост на модела от гледна точка на „човешкото възприятие“; надеждност при съпровождането и постоянно налагащо се модифициране. Друга характерна особеност на тези дейности и процеси е тяхната пирамидална структура. Тази структура и строгата субординация в организацията на работата в тях също трябва да се вземат под внимание при изграждането на моделите, подпомагащи информационното обслужване на бизнес обектите.

## **Информационни модели на бизнес обекти**

Концепцията за бизнес модела става популярна в края на 90-те години на миналия век [Lo & Yu, 2007] и се счита за базова при обсъждането и реализирането на информационното осигуряване на бизнес обекта. Често успешното автоматизиране на дейностите на бизнес обекта се приписва на жизнеспособността на изградения модел. Има различни аспекти при изграждане на бизнес модела. В зависимост от целите, които са поставени, можем да разграничим модели на:

- методи за правене на бизнес [Rapra, 2010];
- бизнес архитектура на компанията;
- информационно осигуряване на бизнес дейностите и т.н.

Различни са и средствата, с които се реализират съответните модели. В нашата работа сме се ограничили само до обсъждане на проблемите, техниките и средствата за моделиране на работата на бизнес обектите от гледна точка на информационното им осигуряване.

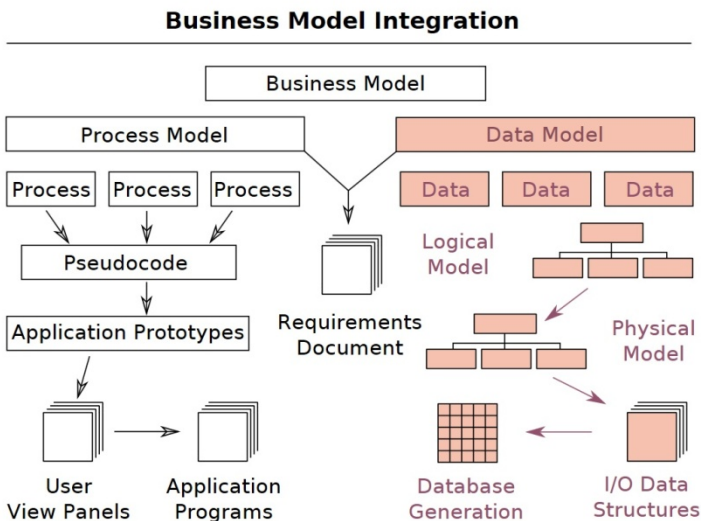


Фиг.1. Процес на информационното осигуряване на бизнес обекти.

Процесът на информационното осигуряване на бизнес обекта е цикличен (фиг. 1) и в това отношение за него важат правилата на традиционните подходи и технологии за разработване на софтуер (Extreme Programming – XP, Dynamic System Development Method – DSDM, Adaptive Software Development – ASD, Feature Driven Development – FDD, Kanban, Scrum). От показаната схема на фиг. 1 се вижда, че колкото по-добър е моделът на обслужвания бизнес обект, толкова по-качествени информационни инструменти за управление ще се получат, затова и изследването на възможностите за създаването на добри бизнес модели е предмет на настоящата работа.

Всеки модел на бизнес обекта се състои от две относително самостоятелни части (фиг. 2):

- модел на бизнес процеси;
- модел на данни.



Фиг. 2 Схема на бизнес модел и връзка между модела на процесите и модела на данните [Smith & Sarfaty, 1993]

## Моделиране на бизнес процеси

*Бизнес процесът* представлява свързан набор от повтарящи се в определен ред действия, които преобразуват входящи и/или вътрешни ресурси в съответствие с предварително установени правила с цел създаване на ползи за бизнес обекта.

Често бизнес обектът се отъждествява с фирма. Тогава бизнес процесите обхващат потока от стопански и търговски дейности, извършвани от и във фирмата с цел постигането на някаква търговска цел.

Съществуват различни методологии и стандарти, с които се описват бизнес процесите [Тужаров, 2009]: Unified Modeling Language (UML)<sup>1</sup>; Business Process Modeling Notation (BPMN)<sup>2</sup>; Business Process Execution Language (BPEL)<sup>3</sup> и т.н. Изграждането на точен модел на бизнес процесите в бизнес обекта е важно, защото при изграждане на информационното му осигуряване моделът на бизнес процесите пряко влияе върху вида и функционалността на потребителския интерфейс на програмните продукти, изградени на негова база (фиг. 2).

Колкото и добре да бъде изграден моделът на бизнес обекта, той е такъв към някакъв момент. С течение на времето условията и параметрите се променят и първоначалният модел вече не съответства напълно на действителността. Освен че имат пирамидална структура, дейностите и процесите на бизнес обекта са и динамични, т.е. те се променят във времето поради различни причини: финансови, технологични, законови, цели и т.н. Тази динамика също трябва да се има предвид, когато се моделират процесите и трябва да се търси съответното ИТ решение, което най-добре може да реализира изградения модел при автоматизация на информационното обслужване на бизнес обекта.

Успешното решаване на проблеми, породени от динамичните промени на външни и вътрешни за обекта условия, е една от основните слабости на съществуващите досега информационни технологии за изграждане на информационни системи за обслужване на бизнес обекти. Интегрирани системи от типа на ERP, CRM, HRMS и т.н. са относително статични. След като вече са изградени, извършването на промени или развитие в тези системи е много трудно и скъпо от гледна точка на използваните за това финансови и човешки ресурси.

В голяма степен проблемите с динамичното развитие на интегрираните информационни системи са концептуално заложили с използването на неподходящи модели на данните при информационното моделиране на бизнес обектите.

---

<sup>1</sup> <http://www.uml.org/>

<sup>2</sup> <http://www.bpmn.org/>

<sup>3</sup> [http://www.service-architecture.com/articles/web-services/business\\_process\\_execution\\_language\\_for\\_web\\_services\\_bpel4ws.html](http://www.service-architecture.com/articles/web-services/business_process_execution_language_for_web_services_bpel4ws.html)

## Модел на данните

Неразделна част от модела на бизнес обекта е моделът на данните. Моделът на данните определя структурата и семантиката на данните в общия бизнес модел [West, 2011]. Той предоставя детайлна информация за това какви са данните, какви са отношенията между тях, как ще бъдат съхранени и е от първостепенно значение, когато крайната цел на изграждането на модела е реализирането на програмен продукт, подпомагащ информационното обслужване на бизнес обекта.

Има разработени много средства за описание на модела на данните, например: UML Profile for DDL, Entity Relationship Diagrams (ERD), Integration DEFinition for Information Modeling (IDEFIX), EXPRESS-G и др. Правени са опити и за стандартизиране на моделите на данни [West, 2011], според които качествените модели на данни трябва:

- да отговарят на изискванията за необходимото количество и многообразие на данни;
- да бъдат ясни и недвусмислени за всички (а не само на авторите на модела);
- да бъдат стабилни от гледна точка на променящите се изисквания към данните;
- да бъдат гъвкави спрямо променящите се бизнес практики;
- да могат да се използват от други модели;
- да могат да интегрират данни от различни други модели на данни.

Освен това, моделите на данни трябва да могат да се разработват бързо и евтино.



Фиг. 3. Процес на създаване на модел на данни [West, 2011]

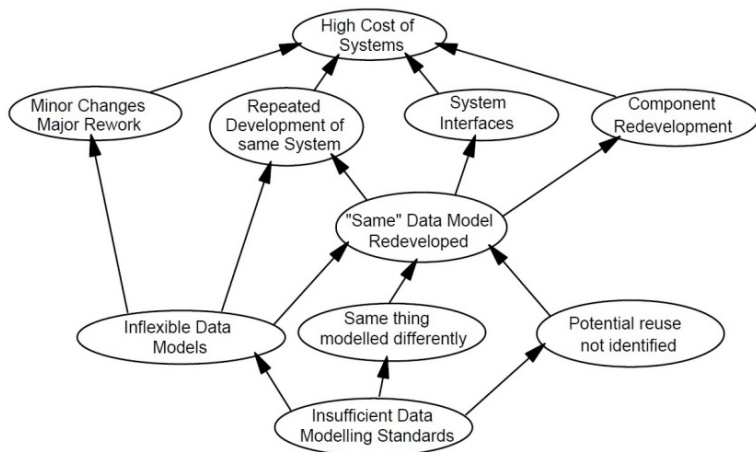
На фиг. 3 е илюстриран начинът на разработването и използването на моделите на данни.

Традиционно моделът на данните има три нива на абстракция [Sparg Systems, 2011]:

- концептуално – тук се определят основните същности на описвания бизнес обект и връзките между тях;
- логическо – на това ниво същностите се дефинират по-подробно, определят се атрибутите им и техните взаимоотношения, без да се описват подробности по изпълнението на тези взаимоотношения. Работата с данните не е обвързана с конкретна СУБД;
- физическо – тук се определя подробно структурата на данните според типа на конкретната СУБД, която ще се използва при управлението на данните в модела.

Натрупаният през годините опит показва, че прилаганите подходи при изграждането на модели на данните са успешни само при първоначалното изграждане на модела на бизнес обекта. Прилагането им за описване на бизнес дейности във времето и/или при динамична промяна на условията води до значителни проблеми. Един от тях е сериозното поскъпване на процеса на изграждане и поддръжка на програмните продукти (фиг. 4), а в някои случаи и невъзможността за надстройка на изградените програмни продукти. Отчита се, че качеството на моделите на данни, на базата на които се разработват приложни информационни системи и съответните потребителски интерфейси, не е добро, защото:

- правилата на бизнес процесите, описващи как се правят нещата на определено място в бизнес обекта, често са фиксирани в структурата на модела на данните. От това следва, че малки промени в начина, по който бизнес процесът се извършва, водят до големи промени в интерфейсите на информационните системи;
- видовете същности, често не са идентифицирани, или са неправилно идентифицирани. Това може да доведе до дублиране на данни, структури на данни и функционалности, както и на разходите, съпътстващи такова повторение при разработването, коригирането и поддръжката;
- моделите на данни за различните информационни системи са съществено различни. Резултатът от това е, че е необходимо изграждането на допълнителни сложни интерфейси между системите, които трябва да споделят данни. Тези интерфейси може да струват от 25 до 70% от стойността на конкретните системи;
- данните не могат да бъдат споделяни по електронен път с други бизнес обекти, защото структурата и смисъла на данни не са стандартизирани.



Фиг. 4. Някои проблеми, произтичащи от сегашния подход при описание на модела на данните [West, 2011]

### СУБД с вътрешна логика на данните

Много от проблемите, които възникват между модела на данните и информационното реализиране на общия модел на бизнес обекта са породени от това, че същностите на моделирания обект се представят с реляционни модели (например чрез ER диаграми). Това е следствие на факта, че до началото на 21. век реляционните системи за управление на бази от данни бяха единствената възможност за практическа работа с данни в световен мащаб. С развитието на информационните технологии и появата на нови класове задачи стана ясно, че механизмите, заложи в реляционния модел, не могат изцяло да отговорят на постоянно повишаващите се изисквания към системите за управление на бази от данни. В някои случаи с тази технология дори е невъзможно да се изгради модел на данните. В търсене на решения на възникващите проблеми, се появиха нови и различни по тип СУБД – системите за управление на бази от данни от нереляционен тип.

Въпреки успешното решение на някои от проблемите за реляционните бази от данни (в областта на големите бази от данни и WEB приложенията), от гледна точка на изграждането на добър модел на данните при моделиране на работата на бизнес обектите, нереляционните бази от данни също не са очакваното добро средство за моделиране, защото:

- има много видове нереляционни бази от данни, които поддържат различни модели на данните. Те са предметно ориентирани и с тях не могат да се описват универсални процеси, съответно не може да се осъществи интеграцията на два модела, изградени с различни типове нереляционни бази от данни;
- нямат задължително предварително точно разписване на информационната структура на базата, тя може да се развива динамично

без особени проблеми, но поради голямата си специализация трудно се описват сложни процеси;

- съдържат голямо количество метаданни, в различни случаи от 2 до 10 пъти повече от обема на реалните данни;
- те са предназначени да работят в разпределена среда (на множество сървъри) и за тях важи теоремата на Брюър [Brewer, 2000] (т. нар. CAP theorem: Consistency, Availability, Partition tolerance). Технологичното решение на повечето видове нерелационни бази от данни е да изберат Наличност и възможност за Разделяне на части. Оттук следва, че при тях не може да се гарантира Консистентност на данните, което за процесите в Интернет може и да не от голямо значение, но когато става въпрос за финансова информация е от жизнена важност.

В db-engines.com има описани повече от 270 бази от данни с различни модели на управление на данните. Подробното им анализиране в хода на нашата работа показва, че независимо от типа на базата от данни (релационен или нерелационен), всички те имат няколко еднакви характеристики:

- задаването на типовете на данните и връзките между тях на логическо ниво става с помощта на метаданни;
- метаданните се съдържат във файловата структура на базата от данни, т.е. те са неразделна част с реалните данни;
- метаданните са в основата на осъществяване на процесите по управление на данните.

Ролята на метаданните в управлението на базите от данни е основополагаща. Те участват и в логическото, и във физическото ниво на работа с базите от данни [Garcia-Molina et al, 2009]. В зависимост от това как метаданните описват данните и връзките между тях, се формират и различните типове бази от данни – релационни и нерелационни.

Горните три характеристики се отнасят за метаданни, използвани в логическото ниво на базите от данни. На база на това наблюдение определяме:

**Дефиниция:** Система за управление на база от данни, която на логическо ниво на работа с данните използва метаданни, които се съдържат във файловата структура на базата от данни, ще наричаме **система за управление на данни с вътрешна логика на данните**.

Съпоставяйки:

- проблемите, съществуващи при изготвянето на модела на данните;
- концепция за СУБД с вътрешна логика на данните;
- и наблюдението, че при програмната реализация на модела на бизнес обекта работата на СУБД с вътрешна логика на данните обхваща логическото и физическото ниво на модела на данните (фиг. 5),

се вижда, че метаданните са това, което ги свързва. Оттук следва изводът, че ако може да се избегне ползването на метаданни при работата с логическото

ниво на базите от данни, голяма част от проблемите при моделирането на данните в модела на бизнес обекта ще отпаднат.



Фиг. 5. Действие на СУБД с вътрешна логика на данните в модела на данните.

### СУБД с външна логика на данните

След като бяха посочени основните недостатъци на съществуващите релационни и нерелационни бази от данни при моделирането на данните в общия модел на бизнес обекта, стигнахме до заключението, че ползването на метаданни на логическото ниво на модела е генераторът на тези недостатъци. Възниква въпросът:

*Може ли да се избегне използването на метаданни при логическото ниво на модела на данните и има ли такива системи за управление на база от данни, които поддържат такива модели?*

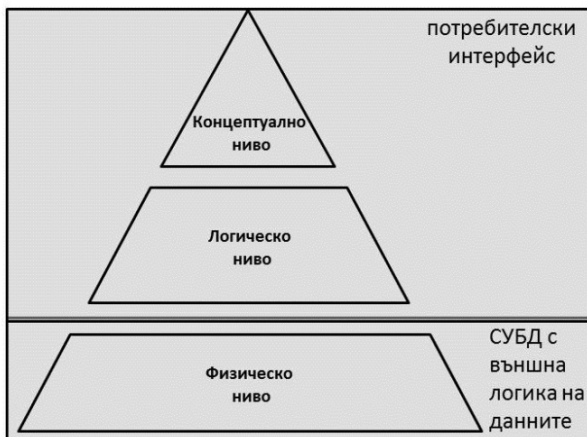
На практика такава СУБД е реализираната на основата на многомерния информационен метод на достъп [Марков, 1984, 2004] нерелационна СУБД, наречена АгМ32. АгМ32 е от съвършено нов тип и съответно предлага нови възможности за изграждане на по-ефективни модели на данните. От гледна точка на изграждането на качествен модел на данни на бизнес обекта АгМ32 притежава следните характеристики:

- във файловете на базата липсват метаданни, които да описват типовете на реалните данни и връзките между тях;
- няма връзки, които да са определени предварително (схеми на таблици и релации, или JSON документи и ключове);
- семантиката и връзките между данните зависят от потребителския интерфейс, който ги използва в съответните програмни приложения, реализиращи модела.

- връзките между данните могат да се променят динамично в зависимост от конкретното интерфейсно решение, без да се променя физическото състояние на данните.

**Дефиниция:** СУБД, притежаваща горните четири характеристики, ще наречаме **система за управление на база от данни с външна логика**.

СУБД с външна логика на данните работят изцяло на физическото ниво на модела на данните при описание на бизнес обектите (фиг. 6).



Фиг. 6. Обхват на СУБД с външна логика на данните върху модела на данните.

За да се получи достатъчно добра представа за технологичните и организационни (от гледна точка на работата с данните) характеристики на **Arm32 DBMS**, ще изброим още някои от свойствата на базата:

- според вида на достъпа принадлежи към категорията на вградените СУБД. Това са СУБД, които влизат в съставната част на приложението. Обикновено са реализирани като библиотеки, които потребителят включва в приложението си. Различава се от тях по това, че Arm32 може да работи с обеми от разпределени данни и поддържа многопотребителско ползване;
- позволява лесното изграждане на OLAP върху данните;
- работи с атомарни структури – елементи. Елементите могат да имат произволна структура. Този механизъм е подобен на идеята на документните бази от данни, които могат да съхранят целия документ в един JSON обект, но е по-гъвкав и позволява да се прилагат различни математически методи за обработка на данните.

Формално за Arm32 данните са просто поредица от байтове, съдържащи се в елементите на съответните информационни пространства. Основните операции за работа с данните в базата се извършват от функции, за които типът

на данните няма значение. Всички възможни елементи от всички възможни информационни пространства съществуват виртуално (тези, в които има данни, съществуват реално) и върху тях може да се изпълняват всички функции, обслужващи работата с данните. В AgM32 има специални функции, които подпомагат обхождането на елементите с данни в конкретните информационни пространства.

С използването на база от данни с горните характеристики успешно се преодоляват голяма част от проблемите при изграждането на модели на бизнес обектите. Динамичното развитие на модела във времето и обединяването на различни модели може да става лесно – отгук следва и лесната практическа реализация на модела на бизнес обектите в програмни продукти, независимо от вида, структурата и характера на данните, които се обработват. С такива бази от данни може да се осигуряват информационно по-голям клас приложни задачи, включително и такива, свързани с динамично променящи се параметри и характеристики. Голямата степен на абстракция позволява при обработка на данните да се използват математически функции.

## **Заклучение**

Изследваните в тази работа бази от данни с външна логика на данните притежават голям потенциал за изграждането на качествени модели на бизнес обекти. Практическото им приложение е голямо и от гледна точка на факта, че базата от данни не е строго специализирана и може да се използва в различни области.

Обосновано може да се твърди, че използването на СУБД с външна логика на данните за реализирането на интегрирани програмни продукти за бизнес управление има съществени предимства пред използваните досега СУБД с вътрешна логика на данните, а именно:

- водещ при изграждането на модела на бизнес обекта е потребителският интерфейс;
- позволява изграждането на динамични модели на данните на бизнес обекта;
- възможно е прякото обединение на два и повече бизнес модела;
- при обработката на данните позволява лесно да се прилагат мощни математически алгоритми;
- намалява времето за практическа реализация и поддръжка на програмните системи, независимо от вида, структурата и характера на данните и сложността на модела на бизнес обекта.

Поле за бъдещо изследване е постигане на по-голяма сигурност на данните в бизнес модела, използвайки нестандартните методи на работа на тези СУБД.

## **Благодарности**

Част от изследванията, представени в статията, са реализирани при изпълнението на проект № 7ИФ-02-5/25.07.2014 „Нов подход за изграждане на софтуерни системи за бизнес управление и контрол, предназначени за масово

ползване: концепция, методология, програмни средства, експерименти и анализ“, финансиран от Националния иновационен фонд чрез Програмата за насърчване на малките и средните предприятия (Седма конкурсна сесия – 2014).

## **Литература**

- Brewer, E. (2000). Towards robust distributed system (invited talk). Principles of Distributed Computing, Portland, Oregon, July 2000.
- Garcia-Molina H., J. D. Ullman, J. Widom (2009). Database Systems – The Complete Book. Pearson Education Inc., 1248 pp.
- Lo A., E. Yu (2007). From business models to service-oriented design: a reference catalog approach. In Proceedings of the 26th international conference on Conceptual modeling (ER'07), Christine Parent, Klaus-Dieter Schewe, Veda C. Storey, and Bernhard Thalheim (Eds.). Springer-Verlag, Berlin, Heidelberg, 87-101.
- Markov Kr. (2004). Multi-Domain Information Model. Int. Journal „Information Theories and Applications“, 11/4, pp.303-308.
- Markov, Kr. (1984). A Multi-domain Access Method. Proceedings of the International Conference on Computer Based Scientific Research, Plovdiv, pp. 558-563.
- Rappa, M. (2010). Business Models on the Web.  
<http://digitalenterprise.org/models/models.html>
- Smith P., R. Sarfaty (1993). Creating a strategic plan for configuration management using Computer Aided Software Engineering (CASE) tools. Paper For 1993 National DOE/Contractors and Facilities CAD/CAE User's Group.
- Sparx Systems (2011). From conceptual model to DBMS,  
[http://www.odbms.org/wp-content/uploads/2013/11/Data\\_Modeling\\_ConcepttoDBMS.pdf](http://www.odbms.org/wp-content/uploads/2013/11/Data_Modeling_ConcepttoDBMS.pdf)
- West, M. (2011). Developing High Quality Data Models. Morgan Kaufmann, 408 pp.
- Тужаров, Хр. (2009). Методологии и стандарти за управление на ИТ услуги.  
<http://tuj.asenevtsi.com/BIS09/BIS47.htm>

## **Business Objects Modelling Using DBMS with External Data Logic**

**Anani Rizov, Iliya Mitov, Krassimira Ivanova**

**Abstract:** The article examines the process of building quality business models. The definitions of key elements in the process of modeling business objects are given. The existence of a new type of non-relational databases – DBMS with external data logic, is defined. The advantages of these DBMS against currently used DBMS with internal data logic in the processes of building quality business models is mentioned.

# Variable Ordering Based on Distinguishing Sets

Vesela Angelova

Institute of Mathematics and Informatics – BAS, Sofia  
[vaa@math.bas.bg](mailto:vaa@math.bas.bg)

**Abstract:** Finding of an appropriate variable ordering is considered as important for many applications, in particular for the graph representation of Boolean function. This sequencing is associated with a binary decision tree which optimal construction is known to be NP-complete and may be regarded as a compressed representation of the data structure. There are many heuristics to find a graph with minimal number of nodes and they are typically associated with adding a single parameter. Another application of ordering is sequential fault analysis. Our approach is based on classification concepts related to the sets of parameters and does not depend on the initial order.

**Terms:** classification test, representative sample, binary decision diagram (BDD), fault diagnosis.

**ACM 1998 Classification Keywords:** I.5 Pattern Recognition, E.4 Coding Theory

## 1. Introduction

An attribute-value system is a basic experience representation in tabular form, where the objects (rows of the table) are described by the values of particular attributes (columns of the table). When these objects are collected in groups by given type, they form a class and we can think of it as an additional attribute for a function output (the range has two or more than two values which show belonging to a class) or for identification of a specific situation in a system (state). The process in which objects are recognized, differentiated and understood is called classification and also categorization. One of the most important issues for such descriptive tables is to find the shortest route to reach each class. In this context we build partial order of attributes, which induces a graph representation with minimal number of nodes. The main idea in this paper is to select only attributes which distinguish one or a few objects from other classes and these attributes as a whole group, according to their occurrences and cardinality, participate in a sequence for the shortest route in the graph representation. The second section describes a partial function, determined only for a few objects belonging to two classes (but in common case more than two classes) and this is the base for our considerations about ordering of attributes. The introduced concepts of distinctness enable rearrangement of attributes unlike most examples for minimizing of graph at a fixed order of attributes. The origin of these special “distinguishers” of different classes is classification or logical pattern recognition. We shall apply this approach for different types of functions, which output values show distinctness of the classes. Section 3 is for totally determined Boolean function where the range is again Boolean. Section 4 discusses an identification function, where the range has more than two values.

## 2. Distinguishing sets for a partially determined function

The distinguishing sets are described in an earlier work [1] and they are good for classification when there is not enough material for statistics. The main concept is **irreducible classification test**, a subset of attributes which distinguishes all classes (in particular two classes) and which is irreducible. The very way of getting all the tests give ordinance of parameters, according to the test length and the participation in them.

*Table 1: An example with two classes*

Class		x1	x2	x3	x4	x5	x6	x7	x8	x9
K1	E1	1	0	1	0	0	1	0	0	1
	E2	1	1	0	1	0	1	1	0	1
	E3	0	0	0	0	0	0	1	1	0
K2	E4	0	0	1	0	1	1	0	1	1
	E5	1	1	0	1	1	0	0	1	1
	E6	0	1	1	0	1	1	0	0	0

Our table consists of six input elements and we may perform classification (value of the function) only by checking the fifth parameter. Thus the only node of the graph will be  $x_5$  and depending on value of parameter  $x_5$ , the result is K1 or K2. But, the function is determined only by six combinations of parameters. For greater reliability due to unknown elements, we may continue classification by parallel checking OR-node for next classification test. However, checking parameters  $x_8$  and  $x_9$  is more complex as to be performed consequently with AND. In this case, the distinguishing sets may have different values in one class; they defer the class as a whole from the other class. For example, the set  $\{x_8, x_9\}$  has values '01' two times and '10' once for class K1. Further, the links between the parameters are complicated and the contribution for determining class of object decreases. That's the reason the terminal tests are convenient to use for classification – each of them alone shows belonging to class (regardless of accepted values). If these six elements are the only known elements of classes, the function for a new element may be guessed through parallel checking of the parameters, forming irreducible distinguishing sets. This is heuristic classification by voting (eventually weighted depending on the length) from such sets. We see that any classification test cuts a short description (columns of the table) sufficient to determine the value of the function. There are a total of 11 classification tests for this example –  $\{x_5\}$ ,  $\{x_8, x_9\}$ ,  $\{x_1, x_6\}$ ,  $\{x_1, x_2, x_7\}$ ,  $\{x_1, x_3, x_7\}$ ,  $\{x_1, x_3, x_8\}$ ,  $\{x_1, x_4, x_7\}$ ,  $\{x_1, x_7, x_8\}$ ,  $\{x_2, x_3, x_8\}$ ,  $\{x_2, x_7, x_8\}$ ,  $\{x_2, x_4, x_6, x_8\}$ .

But for graph representation the values of distinguishing sets for different classes are important because these values don't fill all combinations of attribute values for particular distinguishing set. For example, the value '1' for  $x_7$  shows belonging to class K1 regardless  $\{x_7\}$  is not alone a classification test. That's why, the same procedure we may perform using more specific subsets – **irreducible representative**

**samples.** They are distinguishing subsets for one element and all elements of the other classes (that is, the combination of corresponding values of attributes is unique and it is not found in other classes). For example, the value '1' for  $x_7$  is representative sample for class K1 with 2 occurrences; the values '11' and '00' for parameters  $x_8$  and  $x_9$  are representative samples for class K2.

The classification tests are more common concept; they include a few representative samples for one class even shorter ones. The tests give distinctness, the samples give descriptiveness. The problem of tests finding is equivalent to the set cover problem (set of all pairs from different classes). It also relates to finding of representative samples for each element. But representative samples are with different length, values and occurrences. They are different for different classes.

Our goal is to take only few of them, which cover all elements of the class and all attributes. This choice gives heuristic rule for graph representation first by the shorter and more frequent combined values of parameters possibly overlap with the next sample parameters selected [2].

*Remark 1.* In the following example (Section 3) describing one class (value) by a graph, obviously the other class is obtained in the opposite cases because all input cases are included in the table. This way we get a total order of parameters, which induces a minimal graph representation of the function.

*Remark 2.* When we arrange the parameters involved in a particular classification test, we arrange a presentation describing one route to reach each class (state) only in this segment (cut set), what is a partial order of parameters inducing a partial graph representation (Section 4).

### **3. BDD for a full determined function with two values**

A Boolean function can be represented as a rooted, directed acyclic graph which consists of several decision nodes and terminal nodes. There are two types of terminal nodes called 0-terminal and 1-terminal. Non-terminal nodes are labeled with Boolean variables. A path from the root node to the 1-terminal represents a (possibly partial) variable assignment for which the represented Boolean function is true (output of  $f$  is 1), these paths leading to 1 represent models of the formula, while paths leading to 0 represent counter-models. As the path descends to a low (or high) child from a node, then that node's variable is assigned to 0 (resp. 1). There are two kinds of lines – dashed for 0 and solid for 1.

The binary decision tree of the Figure 1 can be transformed into diagram BDD (Figure 2) by maximally reducing it. There are two kinds of reduction – removing redundancy and sharing identical nodes [3].

The size of the BDD is determined both by the function and the chosen ordering of variables.

For Ordered Decision Tree (OBDT) from root to leaves variables are encountered always in the same order without repetitions along paths. Changing the ordering of variables  $x_1 < x_2 < \dots < x_n$  may increase or decrease the size of the tree and this is important for checking of validity or satisfactoriness.

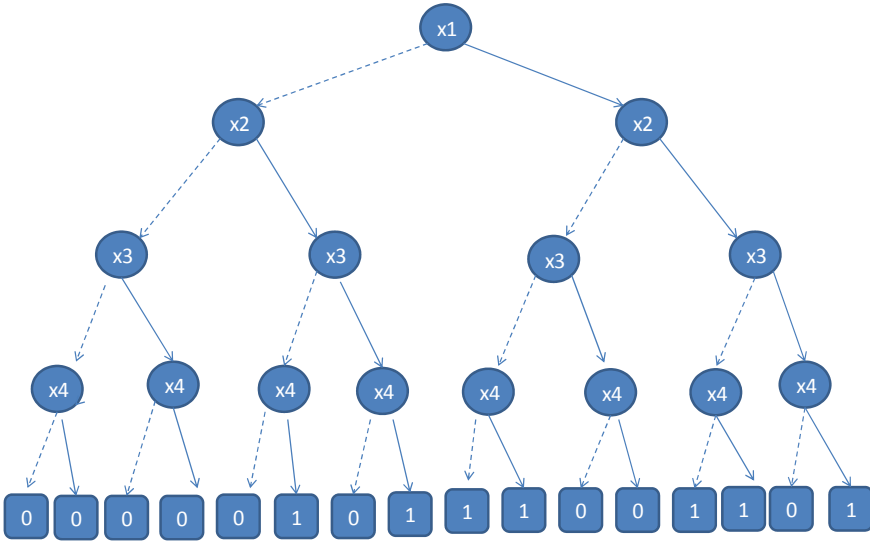


Figure 1. Binary decision tree for a function

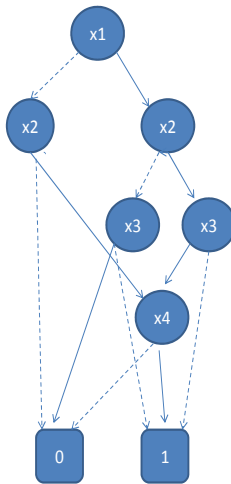


Figure 2. BDD for the function

Let's look at the tabular function as two classes according to its output. Generally, this function does not have its own subset of parameters differing two function outputs, because it would be a function of a smaller number of variables. In this case we search distinguishers of each input vector yielding to 1 with all input vectors yielding to 0. They are representative samples (special associating rules) for the separate vector of one class. The same procedure is used for other class. Next, we describe more convenient class.

The idea is to arrange the parameters in a sequence such that the shorter distinguishing subsets of attributes appear earlier in the sequence and when the length is equal, more frequent appear earlier. The next subset is chosen to cover other objects of the same class. This arrangement of parameters would provide a compact graph representation of the function.

For example, our function, producing two classes is represented in tabular form as follows:

*Table 2: Tabular form sorted by function values*

class= value	inputs	x1	x2	x3	x4
f=0	E0	0	0	0	0
	E1	0	0	0	1
	E2	0	0	1	0
	E3	0	0	1	1
	E4	0	1	0	0
	E6	0	1	1	0
	E10	1	0	1	0
	E11	1	0	1	1
f=1	E14	1	1	1	0
	E5	0	1	0	1
	E7	0	1	1	1
	E8	1	0	0	0
	E9	1	0	0	1
	E12	1	1	0	0
	E13	1	1	0	1
	E15	1	1	1	1

For the second class we find that the part of parameters  $\{x_2, x_4\}$  with values '11' is a representative sample of E5, E7 and E15; another group of parameters  $\{x_1, x_3\}$  with values '10' is representative for E8, E9 and E12. The row E13 has both of them. In this case the order of parameters must be a sequence of these two subsets with equal frequency and length –  $\{x_1, x_3\}$  followed by  $\{x_2, x_4\}$  are one proposal (Figure

3),  $\{x_2, x_4\}$  followed by  $\{x_1, x_3\}$  are another proposal. So, the shorter and more frequent subsets are ahead of ordering.

Obviously, if an element has particular representative sample, it is not from opposite class. If an element is from particular class, it has at least one representative sample, which is chosen to cover parameters (otherwise these parameters would not participate in the function as essential parameters).

Generally, the procedure for graph representation is to add in the order parameters participating in the most frequent representative sample for the remaining objects in the more convenient class from two classes. In [4] this procedure is connected with mining frequent item sets for customer database, but our representative samples are special type association rules where right-hand side (consequent) is the class, support (frequency) is with respect for a class and an item set involve values '1' and '0'.

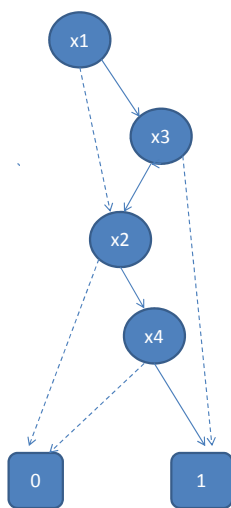


Figure 3. Reordered BDD

#### 4. Sequential analysis for a function with more values and probabilities

By using a similar technique, we propose a decision for the weighted graph compression problem, where the classes are a few specific faults of electronic system. In this case we will call the class "state" which must be identified. Less nodes and edges would give a better idea of the structure. For illustration of this task we refer the readers to sequential fault diagnosis [5].

In its simplest form, the test sequencing problem (also termed the test planning problem) is defined by the tetrad  $(S, p, T, c)$ , where  $S = \{s_0, s_1, \dots, s_m\}$  is a finite set of system states with  $s_0$  denoting the "fault-free" state of the system and  $s_i$ ,  $(1 \leq i \leq m)$  specifying the different faulty states of the system;  $p = [p(s_0), p(s_1), \dots, p(s_m)]^T$  is the *a priori* probability vector of the system states;  $T = \{t_1, t_2, \dots, t_n\}$  is a finite set of  $n$

available tests and  $c = [c_1, c_2, \dots, c_n]^T$  is the vector of test costs measured in terms of time, manpower requirements, or other economic factors. Each test  $t_j$ ,  $1 \leq j \leq n$ , is represented by a binary column vector  $d_j$ , of dimension  $(m+1)$  such that test  $t_j$  detects a failure state  $s_i$  if the  $i$ -th component of the test vector  $d_j$ , is 1. Clearly,  $d_{i0} = 0$  for  $1 \leq j \leq n$ . We impose the standard assumption that only one of the system states  $s_i$  ( $0 \leq i < m$ ), occurs. The goal is to develop an efficient test sequencing procedure to minimize the expected cost of testing, while taking into account the test outcomes, test costs, and failure probabilities. In order to differentiate these single tests  $t_i$  (one-parametric) from aforementioned classification tests (sets), we will call them **checks** (Table 3).

*Table 3: An example from [5]*

class=fault	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	Prob $s_i$
K1 = $s_0$	0	0	0	0	0	0,70
K2 = $s_1$	0	1	0	0	1	0,01
K3 = $s_2$	0	0	1	1	0	0,02
K4 = $s_3$	1	0	0	1	1	0,10
K5 = $s_4$	1	1	0	0	0	0,05
K6 = $s_5$	1	1	1	1	0	0,12
check cost	1	1	1	1	1	

Our proposal is to use only test sets which distinguish all states of the system with evaluated probabilities of these states. Sometimes these sets are also called “minimal cut sets” especially for fault detection [6]. This is in contrast to the existing step-by-step optimization bottom-up (dynamic programming) and top-down (by greedy heuristics) and to approaches that use a few evaluation functions (also heuristic) with respect of Huffman code.

For this example there are six cut sets with minimal length:  $\{t_1, t_2, t_3\}$ ,  $\{t_1, t_2, t_4\}$ ,  $\{t_1, t_3, t_5\}$ ,  $\{t_1, t_4, t_5\}$ ,  $\{t_2, t_3, t_5\}$   $\{t_2, t_4, t_5\}$ . The decision from the cited paper is the second set  $\{t_1, t_2, t_4\}$  among the six sets. The normal approach is to cycle through  $2^3$  options for identifying the fault. The goal is not to find minimal cut sets, but the ordering of the parameters for minimization of the search tree. Our proposal is first to assess participation in the set under the most frequently occurring object  $s_0$  in relation to all the others and all  $s_{1-5}$  together. The representative sample is for  $\{t_2, t_4\}$  and we get  $s_0$  for 2 steps. Thus, the frequency is determined by a priori probabilities, unlike the number of occurrences in the objects from section 3. Next parameters are added to distinguish all other objects. In our example one solution is after  $\{t_2, t_4\}$  to add  $\{t_1\}$  (Figure 4). As this triad is a classification test, after the ordered sequence of checks, all states will be detected.

The same procedure can be carried out for the other classification tests. If the checks have different costs, we choose the set with minimal total cost. Here, all sets have equal total weights, but the second one has the least weighted similarities compared to the most probable state  $s_0$ , namely the shortest representative sample

$\{t_2, t_4\}$ . Therefore, a similar decision is received for the sixth classification test  $\{t_2, t_4\}$  followed by  $t_5$ .

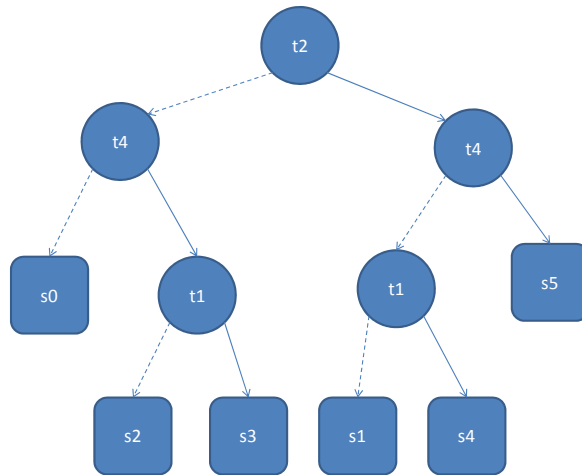


Figure 4. Optimal order of checks

## 5. Conclusion

Ordering of the input parameters is important for cost and security assessment of the described systems. There are several applications of compression by ordering – logic synthesis, formal verification, fault diagnosis, product configuration, information retrieval. The basic idea for our ordering is to add frequent parameters from a set as an ensemble to shorten the path to the terminal nodes in the describing graph. This new order is achieved by taking into account the length and frequency of the distinguishing sets. The goal is this new order of attributes to specify quickly the same values of function of belonging for all defined objects. To use this approach easily, we have presented a methodology for finding aforementioned sets of attributes by SAT solvers [7] rather than an own program.

The new model for compressing of data structure needs refinement and more experiments. The problem is the large amount and different number of representative samples which induces heuristics to select the next parameter group and future work in detail.

## Acknowledgments

The author would like to thank the anonymous reviewers for their helpful comments which improve understandability of the paper.

## References

1. Журавлев Ю.: Об алгебраическом подходе к решению задач распознавания или классификации. *Проблемы кибернетики*, т. 33, Москва, Наука, 1978, стр.5-68
2. Angelova V.: Weight of Tests. *International Journal "Information Models and Analyses"*, (ISSN 1314-6416), Vol. 1, No. 2, 2012, 193-199
3. Akers S. B.: Binary Decisions Diagrams. *IEEE Transactions on Computers*, Vol. C-27, No. 6, June 1978, 509-516
4. Ayadi W, K. Arour: A Binary Decision Diagram to Discover Low Threshold Support Frequent Itemsets. *Proceedings of the 18th International Conference on Database and Expert Systems Applications*, DEXA Workshops, IEEE Computer Society, 2007, 509-513
5. Pattipati K., M. G. Alexandridis: Application of Heuristic Search and Information Theory to Sequential Fault Diagnosis, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 4, July/August 1990, 872-887
6. Rauzy A.: Mathematical Foundations of Minimal Cutsets. *IEEE Transactions on Reliability*, Vol. 50, No. 4, December 2001, 389-396
7. Angelova V.: Early Concept of Classification Test and Contemporary Approach to Diagnostic Test and Combinatorial Test, *Proceeding of the 15th International Conference on Computer Systems and Technologies (CompSysTech'14)*, ACM New York, USA, 2014, ISBN: 978-1-4503-2753-4, p.240-247

## **Подредба на променливи чрез различаващи множества**

### **Весела Ангелова**

**Резюме:** Намирането на подходяща подредба на променливите се счита за важно при много приложения, по-специално при представянето на булева функция чрез граф. Оптималната подредба е свързана с минимизацията на BDD (Binary Decision Diagram), чието намиране е NP-пълна задача и може да се разглежда като компресирано представяне на структурата на данните. Има много евристични методи за намиране на оптимално решение, което обикновено се свързва с последователно добавяне на следваща променлива или с първоначално известна наредба. Друго приложение за търсене на подредба е последователен анализ на грешки, където състоянията на дадена система имат определени предварително известни вероятности. Нашият подход се основава на понятия, свързани с различаващи множества от параметри – класификационни тестове и представителни набори. Основната идея е в рамките на класификационен тест, който „отсича“ достатъчно количество параметри, да се използват най-често срещаните представителни набори, започвайки от по-късите, които се добавят последователно като цялостна група и така бързо да се стигне до крайните възли в графа (функционалните стойности).

# SDBC-Based Technology-Independent Functionality Models (Applications in Emergency Medical Response)

**Boris Shishkov**

1. Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia
2. IICREST, Sofia  
[b.b.shishkov@math.bas.bg](mailto:b.b.shishkov@math.bas.bg)

**Abstract:** Technology-independent functionality models of the software system-to-be are crucial for adequately aligning the software to its real-life environment. This is achieved, by deriving such models based on underlying real-life enterprise models, such that consistency and traceability is guaranteed between enterprise level and software level. In the current paper, we explore the strengths of the SDBC approach in this regard and partially illustrate this by means of a real-life case example from the healthcare domain.

**Keywords:** Functionality models; Enterprise-IT alignment; Traceability; Emergency medical response.

## 1 Introduction

Technology-independent functionality models of the software system-to-be are crucial for adequately aligning the software to its real-life environment. This is achieved, by deriving such models based on underlying real-life enterprise models, such that consistency and traceability is guaranteed between enterprise level and software level. In the current paper, we explore the strengths of the SDBC approach in this regard and partially illustrate this by means of a real-life case example from the healthcare domain.

The remaining of the paper is organized as follows: in Section 2, we briefly introduce the SDBC approach and discuss its relevant strengths. In Section 3, we present an emergency medical response scenario. In Section 4, we partially apply SDBC, reflecting the scenario. Finally, in Section 5 we present conclusions and outlook for further work.

## 2 SDBC

*SDBC* stands for: *Software Derived from Business Components* and the SDBC approach (“SDBC” for short) has been introduced in [5], being further discussed in [3,15].

SDBC facilitates the early stages of the software development lifecycle [2] and particularly – software specification, by providing mechanisms for deriving (re-

usable) software specification constructs as methodologically mapped from corresponding enterprise modeling building blocks.

## 2.1 Modeling Perspectives

The following 4 modeling perspectives are relevant to SDBC:

- Structural Perspective that reflects entities and their relationships;
- Dynamic Perspective that reflects the overall business process and corresponding to this – the states of each entity, evolving accordingly;
- Data Perspective that reflects the information flows across entities and within the business process;
- Language-action perspective that reflects real-life human communication and expression of promises, commitments, etc. as also relevant to soundly building an exhaustive enterprise model.

The current paper would not go discussing those perspectives in more detail, for the sake of brevity. Interested readers are referred to [5].

## 2.2 Theories Behind SDBC

Among the theories behind SDBC are LAP -Enterprise Ontology [6], Organizational Semiotics [4], and Workflow Management [9]. Further, SDBC is in line with the Unified Modeling Language – UML [12], Model-Driven Engineering [7], Service-Oriented Computing [8,10,13], and Context-Aware Computing [12]. Those are briefly discussed below:

### 2.2.1 LAP – Enterprise Ontology

This theory emphasizes the importance of interaction and communication (“LAP” stands for: *Language-Action Perspective*), recognizing that language is not only used for exchanging information, as in reports (for example), but that language is used also to perform actions, as in promises or orders (for example). Such actions are claimed to represent the foundation of communities and organizations / enterprises. This points to the *white-box model* of an organization that is of key importance for building valid enterprise ontologies – this model acknowledges actors (the entities fulfilling corresponding actor-roles) who may be executing not only *production acts* (for example: deliver a pizza) but also *coordination acts* (for example: promise a delivery), and those acts may be of relevance to three perspectives of an organization, namely: *documental* (documents being created and used, for example), *informational* (customer enters his or her PIN in order to realize a bank transaction, for example), and *essential* (the bank transaction itself, for example). Finally, Enterprise Ontology considers a generic interaction atomic pattern, claiming that any complex interaction can be decomposed in such pattern primitives and there are always two roles, namely customer (the one who initiates anything, for example – order something) and producer. There is a request-promise-execute-state-accept actions sequence between them and it can be reflected in a success layer and also a failure layer, as well as

discussion layer, in between. For more information on Enterprise Ontology, interested readers are referred to [6].

### 2.2.2 Organizational Semiotics

Organizational Semiotics (OS) considered a number of concepts, such as *sign* and *affordance*, as essentially useful in modeling a (real-life) system and adequately reflecting relationships and meanings. Often what we observe goes beyond the primary appearance – for example, one could hold a Rolex pen not only as a means of writing but also as a way to demonstrate wealth (this is a sign). As for the affordance concept, it relates to potential abilities (for example: a book affords to be borrowed). Those concepts and also other OS concepts, allow for building complex models that reflect both semantics and norms (rules), and that is reflected in the widely popular OS norm pattern:

```
whenever <condition>  
if <state>  
then <agent>  
is <deontic operator>  
to <action>
```

The OS norm pattern is considered useful in modeling relationships among entities, in the context of a business process. For more information on OS, interested readers are referred to [4].

### 2.2.3 Workflow Management

It is claimed that any business process can be viewed as a collection of processes, where a *process* can be described as “a set of identifiable, repeatable actions which are some way ordered and contribute to the fulfilment of an objective”; typical process patterns are sequence, parallelism, split, and so on.

Workflows play useful role in modeling business processes and those models can be enriched in terms of OS norms and / or entities information. For more information on Workflow Management, interested readers are referred to [9].

### 2.2.4 MDE and UML

MDE reflects a way of modeling where *computation-independent* models are identified and reflected in *platform-independent* models which in turn are reflected in *platform-specific* models, such that technical details are introduced step-by-step allowing for re-use and traceability. Such models can be expressed by means of UML – the *de facto* standard modeling language – through diagrams, such as use case diagram, class diagram, activity diagram, state diagram, and so on. For more information on that, interested readers are referred to [7].

## **2.2.5 Service-Oriented Computing and Context-Aware Computing**

Web services appear at high-level to be (dynamically) composed by users, hiding thus their underlying technical complexity – this complexity is with the software components who are implementing the corresponding service(s). Composability, traceability, and interoperability are hence of crucial importance in web service provisioning. Further, context-awareness is often desired and this means adapting the delivered service (software) behaviour depending on the particular context state which in turn is to be sensed (by sensors, for example). More information on those issues can be found in [8,9,10,11,12,13,14,15].

## **2.3 Concepts**

Among the main SDBC concepts are the following:

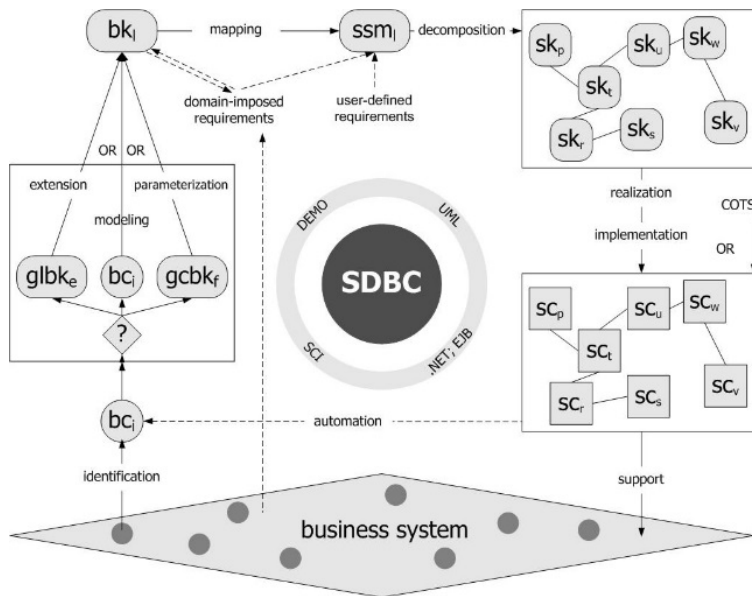
- Component vs CoMponent: while components represent part of the whole, coMponents reflect a model of a component adequately elaborated in all 4 perspectives (see Sub-Section 2.1), and we could thus have business components (business sub-systems) and software components (pieces of implemented software) as well as business coMponents and software coMponents, respectively;
- General vs Generic: these concepts are both about re-use, still – general is about re-using an abstract core (a general reservation engine, for example) while generic is about parameterizing something that is multi-specific (a car system to be adjusted to automatic or gear regime, for example);
- Software Specification Model – this is a technology-independent functionality model of the software system-to-be.

## **2.4 Outline**

The current Sub-Section briefly summarizes the outline of SDBC, with the help of Figure 1.

As seen from the figure, we consider a Business System from which a Business Component(s) is to be identified and then reflected in a relevant model – a Business CoMponent. Another way for arriving at a Business CoMponent is by applying re-use: either extending a general Business CoMponent or parameterizing a generic one. Then, the Business CoMponent should be elaborated with the domain-imposed requirements, in order to add elicitation on the particular context in which its corresponding Business Component exists within the Business System. Then, a mapping towards a software specification model should take place and the user-defined requirements are to be considered, since the derived software model should reflect not only the original business features but also the particular requirements towards the software system-to-be. The software specification model in turn needs a precise elaboration so that it provides sufficient elicitation in terms of structure, dynamics, data and language-action –related aspects. It needs also to be decomposed into a number of Software CoMponents reflecting functionality pieces. Those CoMponents then are to undergo realization and implementation, being reflected in this way in a set of Software Components. Some Software Components could also be

purchased. The Software Components are implemented using Software Component technologies, such as .NET or EJB, for instance. Finally, the (resulting) component-based ICT application would support informationally the target Business System, by automating anything that concerns the considered Business Component (identified from the mentioned system).



Abbreviations:

- |                                   |                                    |
|-----------------------------------|------------------------------------|
| bc – Business Component           | ssm – Software specification model |
| bk – Business CoMponent           | sc – Software Component            |
| glbk – General Business CoMponent | sk – Software CoMponent            |
| gcbk – Generic Business CoMponent |                                    |

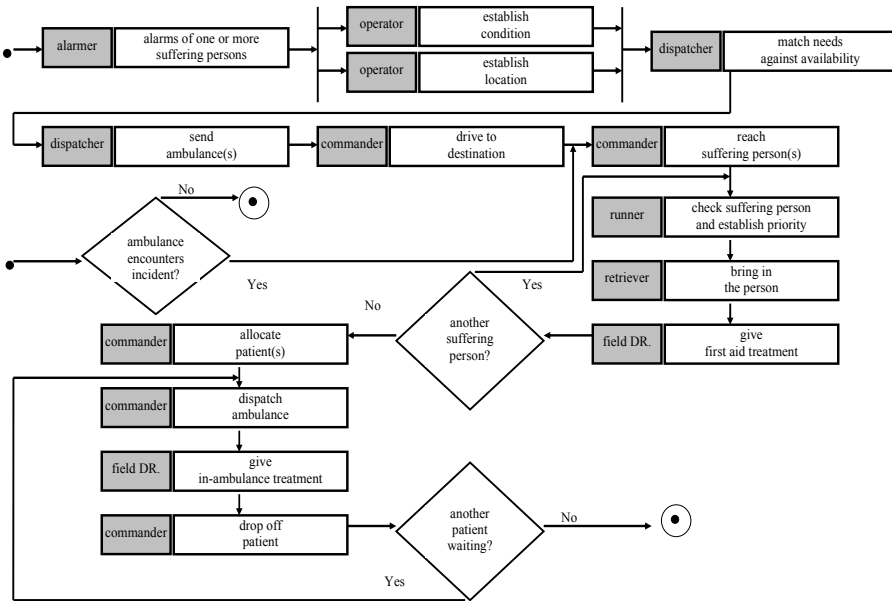
*Figure 1: SDBC – outline*

SDBC was just briefly introduced in the current section, for the sake of brevity. Still, interested readers can find more information on SDBC in [5].

As mentioned in the Introduction, the remaining of this paper contains a particular real-life scenario which is reflected in a partial application of SDBC.

### 3 The Emergency Medical Response Case

Serious emergency events often cause high mortality. To improve emergency management and reduce the damages to public health it is important to optimise the coordination of medical emergency response as a building block of the overall emergency response. A typical emergency response scenario is outlined on Figure 2.



*Figure 2: The emergency medical response scenario*

As seen from the figure, somebody fulfilling the role “alrmer” calls 112 and alarms of suffering person(s). The one receiving the call who hence would be fulfilling the role “operator”, is to do in parallel two things: (i) (s)he should establish the person’s condition – the best that can be derived from the call; (ii) (s)he should establish the person’s location. This information is automatically forwarded to somebody fulfilling the role “dispatcher” who should match the needs (immediately sending an ambulance to a particular location, for example) to availability (how many ambulances are available in relevant proximity, for example). Then, if necessary, the Dispatcher has to send ambulance(s) accordingly. Then one of the persons inside the ambulance – the one fulfilling the role “commander” must drive to the corresponding destination. Then the ambulance is at the destination and another possible way of its appearing in the same situation would be if the ambulance has encountered an incident by chance. Then the Commander has to reach the suffering person(s). Another (or the same) person in the ambulance, the one fulfilling the role “runner” should then check each suffering person and establish priority. Then a person fulfilling the role “retriever” brings in the suffering persons to the ambulance where each suffering person must receive first aid treatment by somebody fulfilling the role “field doctor”. This is the moment when the suffering person starts fulfilling the role “patient” and the Commander should allocate each patient (for example: one should reach Surgery, another one should reach Dermatology, and so on) and dispatch ambulance(s) accordingly. While it is on the run, the Field Doctor must give in-ambulance treatment to the corresponding Patient. Finally, the Commander drops

off the Patient at the corresponding hospital whose emergency department take over the treatment.

#### 4 Applying SDBC

Because of the limited scope of this paper, we only consider in this section a partial illustration of the applicability of SDBC, in reflection of the emergency medical response scenario, and in particular, we are only focused on the Dispatcher's matching needs against availability, needed by the corresponding Commander, for delivering service in turn. What the Dispatcher does (for the benefit of the Commander) is illustrated in Figure 3.

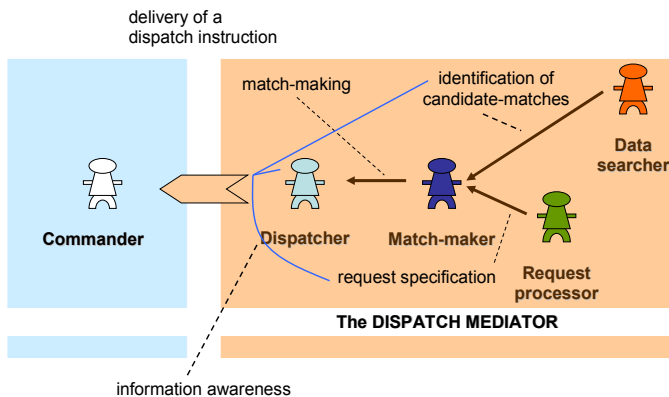


Figure 3: General business process model for the dispatch mediation case

As seen from the Figure, a further “zoom in” needs to be applied with regard to the scenario (that is outlined in Fig. 2), such that this actually complex behaviour is adequately grasped and this assumes also the identification of additional role types in addition to Commander and Dispatcher, namely: (i) Match-Maker reflecting the one who is actually matching what is needed against what is available; (ii) Data Searcher who searches through the available information for the Match-Maker; (iii) Request Processor who processes all input information accordingly such that it is guaranteed that all is formulated in the right way.

To summarize: the Commander requests and the Mediator delivers a dispatch instruction – if we put things as simple as that, we may have these two actor roles, namely Commander and Mediator. We need however to ‘peek’ inside the Mediator, in order to be able to specify its functionality, and as seen from the figure, we go to another level of granularity where we can model actor roles which are internal with regard to the Mediator. Considering requirements and other issues (not mentioned for brevity), we arrive at a model where the following 4 actor roles are presented (these are 4 major roles and the model does not claim exhaustiveness): (i) Dispatcher (D) – delivering the dispatch instruction to the Commander (C) (on behalf of the Mediator) on where and how urgently an ambulance is to go; (ii) Match-maker (MM) – realizing match between what has been captured (through data searching) and

analyzed (through reasoning), on one hand, and what are the dispatch specifications (as it is according to the law and on the basis of the received request(s)), on the other hand; (iii) Request processor (RP) – taking raw data and realizing interpretation through reasoning techniques; (iv) Data searcher (DS) – considering all available relevant information.

Hence, the Mediator can deliver the dispatch instruction (through D) to C, only after MM has delivered to D, which in turn needs to be preceded by the deliveries of RP and DS.

We then go to the SDBC Entity Model, as exhibited in Figure 4, where we have the dashed line delimiting what is inside the system from the rest, and we have all actor roles discussed so far presented. Hence, the actor roles represent the entities in this model and there are some lines connecting to entities – this reveals the structure (how the entities are related to each other). Those relations (representing the interactions) follow the ‘deliveries’ already discussed: D delivers to C, MM delivers to D, RP delivers to MM, and DS delivers to MM (the grey boxes indicate the one who is delivering).

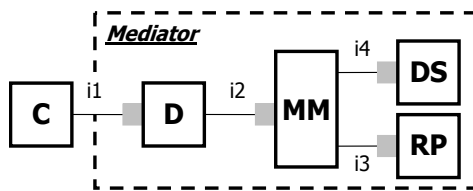


Figure 4: Entity model for the dispatch mediation case

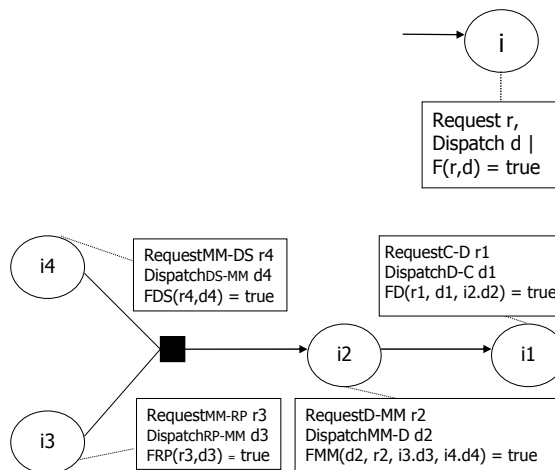
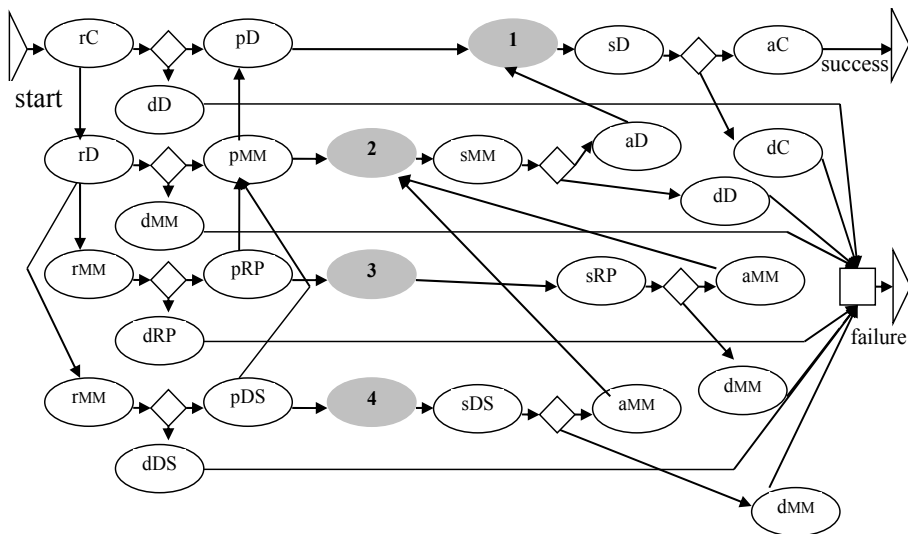


Fig. 5: Interaction model for the dispatch mediation case

The Entity model is then straightforwardly reflected in an Interaction model (Fig. 5) where the emphasis is on the interactions which are modeled using the ISDL technique [9] where workflow details, such as sequence, synchronization, and so on, are reflected: as it can be seen: i3 and i4 should both be completed (synchronization) before i2 is completed, and i1 would be completed after i2 has been completed (sequence). On the top right corner of the figure is the model at the higher granularity level (as already discussed) where we look at all as one interaction (a delivery between the Mediator and the Commander).

Finally, we elaborate each interaction, by applying the request-promise-state-accept LAP pattern, as it is shown in Figure 6. The 4 grey ovals correspond to the 4 interactions' production facts and those are elaborated with corresponding coordination acts (such as request, promise, state, and accept); thus, the first 'line' from top to bottom is to be read as follows: Start, request by C, promise by D, production fact, state by D, accept by C, and so on. There are decision points – allowing not only for promise but also for decline, and these in turn allow for failure scenarios as well. We are not going to explain the model in more detail – interested readers can find more in-formation on LAP in the references already mentioned, and this would help understanding the model better.

This is considered as a 'service-level' model since it is useful for deriving (web) services.



*Figure 6: Service-level model for the dispatch mediation case*

Deriving service specifications based on such a model is nevertheless not trivial and requires resolving issues, such as tight coupling (possibly through introducing an orchestrating entity) which means introducing new entities and new interactions in turn. Resolving other issues, such as context-awareness, for instance, may also require introducing new entities and interactions. Still, this way of progressing from very general real-life information to a well-structured and theory-rooted model is claimed to be beneficial. Such a model can be elaborated further, by adding for example semiotic norms.

However, for the sake of brevity, we are not going in further details with regard to this case. The purpose of this section was to just give an impression on how the SDBC can be useful in modeling particularly emergency care processes, aiming at reflecting them in the specification of a supportive information system.

## **Conclusions**

In this paper, we have discussed the specification of information systems that is based on underlying enterprise models and we have demonstrated how an emergency care scenario can be reflected in semi-formal models that are useful for specifying software/web services. In this, we have followed the SDBC approach, demonstrating its strengths and potentials accordingly. As further research, we plan projecting this work to a lower level, establishing adequate relations between technology-independent models (as focused in this work) and technology-specific solutions while also capturing context-awareness –related features

## **References**

1. EHST (Home), <http://www.is-ehst.org>
2. Software Engineering Institute (Home), <http://www.sei.cmu.edu>
3. Shishkov, B., Dietz, J. L. G., Liu, K.: Bridging the Language-Action Perspective and Organizational Semiotics in SDBC. In ICEIS'06, 8<sup>th</sup> Int. Conf. on Enterpr. Inf. Systems (2006)
4. Liu, K.: Semiotics in Information Systems Engineering. Cambridge University Press, Cambridge (2000)
5. Shishkov, B.: Software Specification Based on Re-usable Business Components. Delft, The Netherlands: Sieca Repro (2005)
6. Dietz, J.: Enterprise Ontology, Theory and Methodology. Springer-Verlag, Berlin Heidelberg (2006)
7. Object Management Group (OMG): MDA – Guide, V1.0.1, omg/03-06-01 (2003)
8. Papazoglou, M. P.: Web Services: Principles and Technology. Pearson Pr. Hall (2008)
9. Shishkov, B., Van Sinderen, M. J.: Service-Oriented Coordination Platform for Technology-Enhanced Learning. In: I-WEST'09, 3<sup>rd</sup> International Workshop on Enterprise Systems and Technology. INSTICC Press (2009)
10. Leymann, F.: Combining Web Services and the Grid: Towards Adaptive Enterprise Applications. CAISE Workshops (2005)
11. Van Sinderen, M. J.: From Service-Oriented Architecture to Service-Oriented Enterprise. In: I-WEST'09, 3<sup>rd</sup> International Workshop on Enterprise Systems and Technology (2009)

12. Shishkov, B., Van Sinderen, M. J.: On the Design of Context-Aware Applications. In: IWEST'08, 2<sup>nd</sup> Int. Workshop on Enterprise Systems and Technology. INSTICC Press (2008)
13. Alonso, G., Casati, F., Kuno, H., Machiraju, V.: Web Services, Concepts, Architectures and Applications. Springer-Verlag, Berlin-Heidelberg (2004)
14. Bosworth, A.: Developing Web Services. In: 17th Int. Conference on Data Engineering (2001)
15. Shishkov, B., Van Sinderen, M. J., Quartel, D. A. C.: SOA-Driven Business-Software Alignment. In: ICEBE'06, 2<sup>nd</sup> Int. Conference on e-Business Engineering. IEEE Computer Society (2006)

## **Концептуални функционални модели, базирани на SDBC (приложение в Спешна медицинска помощ)**

**Борис Шишков**

При изграждането на софтуерни системи, концептуалните функционални модели са важни за гарантирането на адекватна връзка между софтуера и обкръжаващата го среда. Това се постига, като софтуерът се специфицира на база на съответни бизнес модели, установявайки консистентност и проследимост между софтуерното ниво и бизнес нивото. В тази статия ние изследваме силните страни на методологията SDBC по отношение на този проблем, като използваме за илюстрация казус от областта на здравеопазването.

# A Decision Making Support Method in Agile Software Development

**Neli Maneva**

Institute of Mathematics and Informatics, BAS, Sofia  
[neman@math.bas.bg](mailto:neman@math.bas.bg)

**Abstract:** The paper describes how a method, supporting decision making, can be used during software development in an agile framework. The method of Comparative analysis is briefly presented. An incremental approach to object modeling is explained and some models, constructed for a number of intermediate products (requirements, user stories, designs, programs and increments) are shown. The application of the method is illustrated by a few examples.

**Ключови думи:** agile software development, decision making, quality object models

**ACM 1998 Classification Keywords:** D.2: Software Engineering, D.2.8 Metrics, D2.9 Management

## 1. Introduction

Agile software development appears as an alternative to conventional development, trying to overcome some of its stated deficiencies and responding to the need of a new approach. Nowadays there is a big family of many and different agile methodologies, sharing the key ideas of the signed in 2001 Manifesto for Agile software development [1], but proposing some specific guidelines for their implementation.

In Agile software development requirements and solutions evolve by collaboration between members of self-organizing, cross-functional teams. Adaptive planning, evolutionary development, early delivery, continuous improvement and encouragement of rapid and flexible response to change are promoted [2]. This new development paradigm influences the content and performance of any activity, including decision making [3]. From agile perspective, decisions have to be made just in time, quickly, on the base of all available information, using appropriate methods and tools.

In this paper we describe how a rigorous decision support method can be used during agile software development. The rest of the paper is organized as follows. Section 2 describes the method of Comparative analysis. In Section 3 the most difficult part of method application – the object modeling – is presented and some generic quality models for basic objects are given. In Section 4 there are a few examples of cases, describing the context of the Comparative analysis. Some ideas for future research and development work are mentioned in Conclusion.

## 2. The method of Comparative Analysis

Let us present briefly a formal method for a reasonable choice. It supports the multi-criteria decision making and has been developed, taking into account the best achievements in this area [4].

**Comparative analysis** (CA) is a study of the quality content of a set of homogeneous objects and their mutual comparison so as to select the best, to rank them or to classify each object to one of the predefined quality categories.

The compared objects can be products, processes or resources, identified as significant for the activity under consideration. For CA practical use, we distinguish two main players: the **Analyst**, responsible for all aspects of CA implementation, and a **CA Customer** – a single person or a group of individuals, who should make a decision in a given situation and would like to use the CA as a supporting method.

Depending on the identified problem to be solved by a CA Customer at a given moment, a **case** should be opened to determine the context of the desired comparative analysis. Each case is specified by the following six elements:

case = {View, Goal, Object, Competitors, Task, Level}

The **View** describes the CA Customer's role and the perspective from which the comparative analysis will be performed. The Analyst has to identify the potential decision makers. For example, in Scrum software development the following roles of CA Customer can be considered: Product owner, Scrum master, Member of the development team, Stakeholders and Agile mentor. Taking into account the responsibilities and typical tasks of these players, several cases will be defined, considering their respective point of view to the analyzed problem.

The **Goal** expresses the main Customer's intentions in CA accomplishment and can be to describe, analyze, estimate, improve, or any other, formulated by the Customer, defining the case.

The **Object** represents the item under consideration. For each object in CA application a *quality model* should be created – a set of characteristics, selected to represent the quality content in this context, and the relationships among them.

According to the stated Goal, the instances of the objects to be compared are selected and described in the set of **Competitors** –  $C = (C_1, C_2, \dots, C_n)$ .

The element **Task** of a case can be *Selection* (finding the best), *Ranking* (producing an ordered list), *Classification* (splitting the competitors to a few preliminary defined quality groups) or any combination of them.

The Depth **Level** defines the overall complexity (simple, medium or high) of the CA and depends on the importance of the problem under consideration and on the resources needed for CA implementation.

A step-wise procedure for the CA accomplishing has been proposed and successfully used in a number of application areas. Next follows a brief description of the steps.

The first step is *Initialization*. When a problem, for which the CA seems to be useful, appears, the Customer creates a CA request – an informal description of the problem situation, moment and area of consideration. This *CA Request* is studied by

the *Analyst*, who evaluates its conceptual and technical feasibility. When the *CA request* is approved, a new case is opened and its elements *View*, *Goal*, *Object*, *Competitors*, *Task* and *Level* are defined, using some relevant sources of information.

During the *Construction* a hierarchical object model is created and the weights of all included in the model quality characteristics are defined. Then the method for accomplishing the *Task* (*selection, ranking or classification*) is selected, taking into account the defined *Task* and depth *Level*.

During the *Execution* stage the *Analyst* evaluates each object from the set of *Competitors* and fills the objects-factors table, applies the selected method for implementing the *CA* task and documents the obtained results.

The *Completion* stage comprises preparation of the Final report, describing and saving some elements for re-use and collecting historical data to be available for further processing.

Each step covers a number of activities. For each activity its aim, the results to be obtained and how the activity must be performed have been specified. In this way it is clear who has to do what and when, thus making the comparative analysis manageable and properly utilized.

### **3. Quality models for some intermediate products in agile software development**

The case-driven *CA* begins with construction of the hierarchical *object quality model*. The first level comprises  $m$  quality factors  $F_1, F_2, \dots, F_m$ . They characterize the local state of the object as regards the defined case. Each factor  $F_j$  must be weighed in accordance with its relative importance. Depending on the defined depth level and the cognitive complexity of the factors, the latter can be further decomposed. The obtained hierarchical structure describes which quality characteristics will be considered at different levels and what the relationships among them will be.

One of the most challenging problems in *CA* implementation is the construction of a quality model for a given object. Our suggestion is to apply an incremental approach. When a new object appears in a case definition, the *Analyst* has to create its first model, using different sources of information. The Customer's requirements for the studied object should be stated and further analyzed so as to be reformulated in terms of quality characteristics, described in a hierarchical model. Then the created model is saved as a generic (basic) model for the object under consideration. When a request for *CA* with the same object arises, the generic model is found and is made available for modification in two directions – expanding the generic model and constructing its derivative image, relevant to the defined case. First, there should be an attempt to find some new quality characteristics to be added at some levels of the hierarchy so as to reflect a new view or goal of the current case. If the attempt is successful, the expanded hierarchical structure will be stored as a generic model. The second step is to decide whether some characteristics can be ignored as irrelevant to the current case. So a derivative of the model is obtained. A special pattern,

describing the deviation from the basic model is created and saved for future re-use with reference to the case.

In order to facilitate the CA application, a few quality models for intermediate products in agile development have been proposed. They can be used as generic ones and further modified to reflect the peculiarities of the defined situation.

### **3.1. Quality model for object “requirement”**

Agile Requirements Engineering comprises elicitation, validation, estimation and prioritization of requirements at different levels [5]. The Comparative analysis can support decision making at some situations appeared when these activities have been performed.

A simple linear model for object “requirement” has been created. According to this model the requirement should be:

**Unambiguous** – to possess the capability to have a unique interpretation,

**Testable** – the fulfillment of each requirement is measurable or testable,

**Consistent** – there no contradictory statements in requirement description,

**Traceable** – each requirement has a rationale,

**Understandable** – requirement is fully understood when used for developing software,

**Complete** – it is possible to make references to precisely identified entities,

**Detailed** – the description of requirement is given in detail and it is traceable to user needs.

### **3.2. Quality model for object “set of requirements”**

A simple linear model for object “set of requirements” has been created. The selected characteristics are:

**Completeness** – i.e. the set **R** comprises *all* requirements, necessary to describe the chosen functionality of a system or a system’s compound element;

**Consistency** – the involved in the set **R** requirements are in reasonable and logical harmony, without any contradictions in their content;

**Relevance** – all included in **R** requirements should possess a direct and clearly identified connection to the studied system’s functionality.

### **3.3. Quality model for object “user story”**

In agile software development user stories are small units of work. They are a few sentences in simple language that outline the desired functionality. The goal of a user story is to deliver a particular value back to the customer.

Usually the user stories are sketched out by the Customer/Product owner. Having such initial version, the entire development team collectively decides about more detailed descriptions. They are the granular pieces of work that help define the implementation items for the story and the upcoming iteration.

A two-level quality model for object “user story” has been constructed (see Table 1).

*Table 1. Quality model for a user story*

<b>Factors</b>	<b>Criteria</b>
Accuracy	
Trackability	
Adequacy	Robustness Completeness Correctness
Understandability	Self-descriptiveness Conciseness

### **3.4. Quality model for object “design of increment”**

One of the most important characteristics of agile development is its incremental nature – working software is delivered fast and frequently. As noted in [6] such short, focused iterations make the entire agile development cycle much more like a maintenance phase.

Bearing in mind this, for object “design of increment” we propose a linear model, comprising 4 characteristics: conceptual integrity, effectiveness, maintainability and correctness. Their explanation is given in Table 2.

*Table 2. Quality model for design of increment*

<b>Criterion</b>	<b>Explanation</b>
Conceptual Integrity	Conceptual integrity defines the consistency and coherence of the overall design. This includes the way that components or modules are designed.
Effectiveness	The degree to which a design is able to achieve the desired functionality and behavior.
Maintainability	Maintainability is the ability of the system to undergo changes with a degree of ease. These changes could impact components, services, features, and interfaces when adding or changing the functionality, fixing errors, and meeting new business requirements.
Correctness	Extent to which the design conforms to its specifications and declared objectives.

### **3.5. Quality model for object “intermediate product”**

The agile development process is iterative with several releases of intermediate products approaching the final one. Each iteration produces intermediate deployable code that has been discussed, designed, implemented, and tested. An increment is a

subset of the final product under development and builds on the functionality of the prior iteration, reflecting the revised set of requirements. That is why for the object “intermediate product” we propose a quality model, comprising the characteristics from the Product Revision perspective (ability to undergo changes) in the McCall Quality model as described in [7]. The selected quality factors and criteria are shown in Table 3.

*Table 3. Quality model for an intermediate software product*

Quality factors	Explanation	Quality criteria
Maintainability	The capability of the software product to be modified so as to incorporate corrections, improvements or adaptations to changes in environment and requirements.	Simplicity Conciseness Self-descriptiveness Modularity
Flexibility	Ease of changing the software to meet revised requirements.	Self-descriptiveness Expandability Generality
Testability	Ease of validation that software meets requirements.	Simplicity Instrumentation Self-descriptiveness Modularity

### 3.6. Quality model for object “program”

It is well known that quality is not something that can be easily added later. So it is better to ensure quality from the very beginning, writing quality code. Analysis of the code quality can be done on the base of different quality models. As an example, let us describe a number of code characteristics, proposed in [8]. Next follows a brief explanation of each code property and a comment about its measurability:

**Conciseness:** A program is said to be concise when its meaning is expressed in a small volume of text.

To estimate conciseness, one can start with a simple metric of program size, such as number of tokens in a program. But conciseness is essentially a relative measure: to tell if, or how much concise a program is, one has to compare sizes of different programs that solve the same problem by possibly using different methods and tools, including different programming languages.

**Straightforwardness:** Straightforward expression means that the programming language is used in a directly understandable way, according to generally accepted rules, including idioms.

As the notion of straightforwardness reflects the users' opinion, it can only be measured indirectly, through asking experts and using a scale for evaluation.

**Adequacy:** Faithful representation of the abstract model of a solution by the (text of the) program.

Unless a formal modelling tool exists in which the abstract solution is described, and a method that compares the program solution against its model, adequacy can only be estimated vaguely.

**Explicitness:** A property of a program object, or a relation between such objects is explicit when it is expressed within the program (as opposed to only assumed).

Cannot be measured in general, since explicitness, as well as the inverse property – non-explicitness – is something that relates a program text with the programmer's intention. However, expert estimation can be used in place of formal measurement method. And true measurement is still possible for several isolated manifestations of this property.

**Apparency:** Visual information is apparent when it can be readily grasped. Besides visual clarity of the information itself, certain preparedness (w.r.t. knowledge) is implied on behalf of the observer.

Not measurable (except by means of expert estimation).

**Memory load:** With respect to certain program fragment, this is the amount of 'memory' required, related to mentally interpreting the fragment, as well as knowing of other elements of the program, required to understand the fragment.

Applicable to small fragments, and can be measured based on the internal structure of the text, along with its relations with the rest of the program. For example, a highly hierarchical construct requires more memory load in order to be interpreted, as opposed to a 'flat', or sequentially interpreted construct.

**Notationality/verbality:** This concerns whether a text is primarily symbolic or verbal.

Can be measured by comparing the number of symbolic or verbal tokens to the total number of tokens in a program. It is more a characteristic of the programming language than of particular programs written in it.

**Mnemonicity:** The ability of program tokens to convey meaning by making use of associations with something known.

This property, although rather important in practice, is too vague and abstract to admit measuring.

**Uniformity:** Sameness or similarity of expressing similar ideas in a language.

This too characterizes not only particular programs, but the programming language as a whole. It is hard to devise a general measure but measuring with respect to simple benchmarks is feasible.

**Visual characteristics and attributes:** May include form, layout, special graphics, colour, fonts and typography, etc.

As should be apparent, these are not meant to be measured even informally.

Some of the above described characteristics can be studied as significant for understandability and maintainability of a program. Our suggestion for a code quality model is to use a linear model, comprising 5 characteristics, namely **Conciseness, Straightforwardness, Adequacy, Explicitness and Uniformity.**

#### **4. How the Comparative Analysis can support the decision making in agile software development**

Generally speaking, CA can be used in any situation, in which a problem is identified and a case can be defined, describing the context of the CA – when, who and how makes a decision.

##### **Example 1: How the CA can be used for requirements prioritization**

One difficult problem in requirement engineering is to create an ordered list of all elucidated requirements. Such prioritization will facilitate further requirements analysis so as to be accomplished in a more systematic and efficient way.

A number of cases for this problem can be defined as a combination of the following elements:

**View** – that of the Product Owner, a stakeholder, a member of the development team;

**Goal** – to compare the requirements, selected for prioritization;

**Object** – a requirement;

**Competitors** – all requirements, chosen to be prioritized;

**Task** – ranking so as to produce an ordered list of the compared requirements;

**Level** – any chosen – simple, middle or high.

##### **Example 2: How the CA can be used for analysis of a set of requirements**

**View** – that of the Product Owner, a customer or a member of the development team;

**Goal** – to analyze a set of requirements before a new iteration;

**Object** – a set of requirements

**Competitors** – only one element – the constructed set;

**Task** – evaluating the quality of the set of requirements;

**Level** – simple

##### **Example 3: How the CA can be used for evaluating the effect of refactoring**

Refactoring is an agile practice which deals with changing the design or structure of the code without changing its results. Refactoring involves rewriting the code to improve its understandability and maintainability, while explicitly preserving its behavior.

**View** – that of the developer, coding the program

**Goal** – to compare code quality before and after refactoring;

**Object** – a program

**Competitors** – two versions of the same program – before and after refactoring;

**Task** – ranking

**Level** – middle.

#### **Example 4: How the CA can be used for communication strategies selection**

The Product owner wants to compare different forms of communication with stakeholders to be used during an agile project.

**View** – that of the Product owner;

**Goal** – to compare the communication strategies, identified as appropriate

**Object** – a communication strategy

**Competitors** – 9 communication strategies, chosen to be compared [9], namely Face to face (F2F), F2F at Whiteboard, Email, Online chat, Teleconference calls, Videoconferencing, Overview diagrams, Overview documentation, Detailed Documentation.

**Task** – ranking so as to produce an ordered list of the compared strategies;

**Level** – simple.

A linear model for the object “communication strategy” has been created, comprising only two characteristics – *effectiveness* and *efficiency*.

#### **Conclusions**

The main goal of the present paper is to describe the method of Comparative analysis and how it can support the decision making during the agile software development. One of the most difficult activity – object quality modelling – has been illustrated by presenting some models for a number of objects – requirement, user story, set of requirements, design, program and product. As example, four CA cases in Scrum framework for software development have been given.

Some possible directions for future research and development work can be:

- To identify and describe entirely some other CA situations during agile software development, in which the decision making is crucial;
- To examine the usefulness of the CA in a real-life agile project, expanding the re-use library with many different items – cases, models of other objects, etc.;

#### **References**

1. <http://agilemanifesto.org/> Agile Manifesto
2. <http://www.successfulprojectmanager.com/agile/>
3. Drury, M., K. Conboy, K. Power: Obstacles to decision making in Agile software development teams. *Journal of Systems and Software*, 85(6), 2012, pp. 1239-1254.
4. Maneva, N.: Comparative Analysis: A Feasible Software Engineering Method. *Serdica Journal of Computing*, 1(1), pp.1-12 (2007).
5. Helmy, W. at el.: Requirements Engineering Methodology in Agile Environment, *Int. Journal of Computer Science Issues* (IJCSI), Vol. 9, Issue 5, Number 3, September 2012, pp. 293-300.

6. Boehm, B., R. Turner: Management challenges to implementing agile processes in traditional development organizations. *Software*, IEEE, 22(5), pp. 30-39.
7. Wang, S., D. Samadhiya, S. Chen: Software Quality: Role and Value of Quality Models. *Int. Journal of Advancements in Computing Technology*, 3(6), July 2011, pp. 65-74.
8. Bantchev, B.: *Issues of explicit expression in program texts*. PhD thesis, Institute of Mathematics and Informatics, Sofia, 2014 (in Bulgarian).
9. <http://www.agilemodeling.com/essays/communication.htm#sthash.TLa5qtKf.dpuf>

## **Метод, подпомагащ вземането на решения при гъвкаво разработване на софтуер**

**Нели Манева**

**Резюме:** Статията описва как метод, подпомагащ вземането на решения, може да се използва при гъвкаво разработване на софтуер. Методът на Сравнителния анализ е представен накратко, заедно с процедура за използването му. За една от най-трудните дейности – моделиране на изследваните обекти, е обяснен предлаганият инкрементален подход и са описани модели, конструирани за основни междинни продукти: изисквания, потребителски истории, проекти, програми и частични продукти. Няколко полезни приложения на метода са илюстрирани с примери.

# **Оптимизиране на процесите по интеграция между информационните системи и изграждане на единен потребителски профил**

**Младен Георгиев**

Институт по математика и информатика при БАН, София  
[mladen@math.bas.bg](mailto:mladen@math.bas.bg)

**Резюме:** В статията е разгледан проблем, пред който е изправен ИТ отделът на повечето средни и големи фирми. Когато е необходимо персоналът на компанията да има достъп до информационни системи и ресурси в организацията, а този персонал надвишава 200-300 души, от гледна точка на ефективността започва да става належащо да се обърне повече внимание върху начина на споделяне и достъпване на информацията и оптимизиране на процеса по управление на потребителски групи с този размер.

## **Въведение**

При наличието на голям брой служители в дадена фирма, текучеството на персонал и нарастващите изисквания за повишаване на сигурността в организациите, изпълнението на задълженията по администриране на потребителски акаунти става изключително трудоемко, което носи със себе си опасност от неволни грешки, пропускане на някои моменти от политиките за сигурност и един по-висок риск от допускане на грешка. От друга страна, квалифициран персонал е натоварен с рутинна работа, която отнема време, което от своя страна е недобра парична инвестиция. Всичко това създава едно допълнително натоварване на персонала в отдела, понижаващо ефективността на служителите.

Също така възникват и проблеми в комуникацията между отделните служители, звена и ръководството. Липсата на удобни и лесни начини за комуникация подтикват всеки служител сам да търси начини за връзка с колегите си и често това са външни за организацията пощенски сървъри, сайтове за споделяне на файлове и други подобни. Много служители съхраняват служебна информация в личните си електронни пощи или „USB“ преносими памети, които губят или се повреждат. Това се дължи на факта, че обикновено служителите не са информирани достатъчно за това какво може да се случи с ценните данни. Контролът на работа на отделните звена е много труден, различните отдели имат изградени собствени стереотипи на работа и някои от тях са изостанали, а други – напред в работата. Въвеждането на нови организационни правила на ниво организация е изключително трудоемко и затруднено.

Всичко това са проблеми, за които е необходимо да се помисли своевременно, за да се избегнат непредвидени обстоятелства, които биха могли да се изразят както във финансови загуби, така и в загуба на ценна корпоративна информация. Ще представим един модел на управление на потребителски профили и изграждане на информационен портал, който е внедрен и специфициран изцяло за нуждите на СБАЛ по онкология ЕАД, гр. София [1]. Това включва дефинирането на процеса по управление на акаунтите на служители при назначаване, смяна на статута и напускане с ясно специфициране на правата на отделните еднотипни групи от служители спрямо длъжността и принадлежността им към определено звено. Това ще допринесе за установяване на правила за менажиране на потребителските акаунти. Тези правила се дефинират поотделно за всяка една от софтуерните системи, с които разполага компанията, за да се постигне максимално прецизиране на процеса. Следващ етап е автоматизирането на процеса по създаване на потребители на базата на дефинираните правила. Това осигурява единен акаунт за достъп на всеки служител до всички електронни ресурси и системи на болницата, независимо от софтуерния продукт.

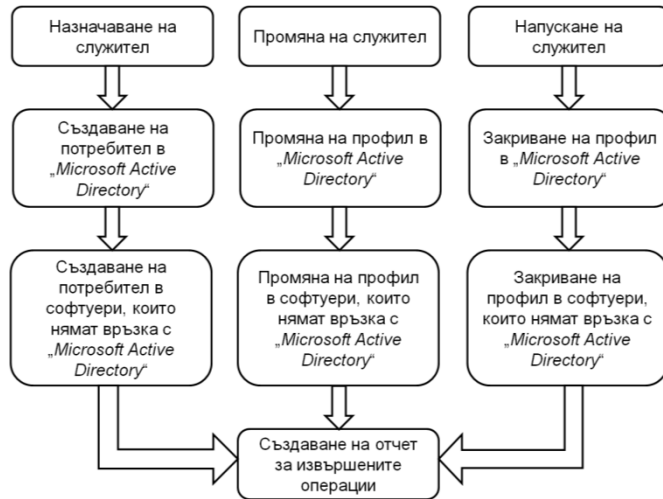
Изграденият информационен портал осигурява средата, където се осъществява комуникацията и споделянето на информация между персонала. Това е системата, която централизира всички информационни ресурси на едно място. Освен това така се повишава информираността на служителите. Интранет порталът дава способите, чрез които може много по-лесно да се организира работата по общи работни групи. Всеки служител може да получава актуални новини и инструкции за работа със специализирания приложен софтуер. Предимството на портала е, че е гъвкаво решение, което дава възможност на потребителите сами да го запълнят със съдържание, да го подредят по собствено желание и да интегрират бизнес процесите си в него.

### **Анализ на решенията**

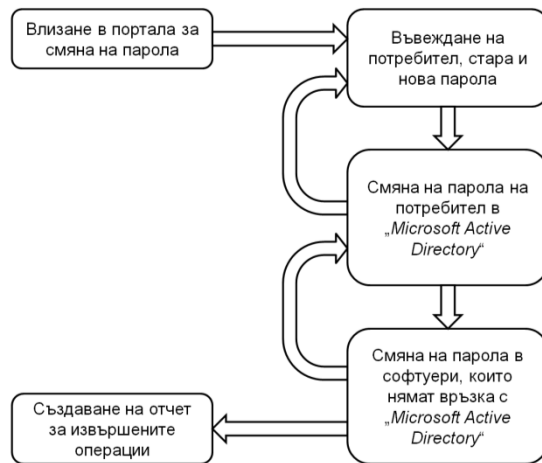
Резултатът от разработката на системата за управление на потребителски акаунти следва да бъде набор от скриптове, чрез които, когато служител е назначен, с променен статус или напуснал, да се създава потребителски профил спрямо информацията, която се съдържа за него в базата от данни на отдел Човешки ресурси. Концептуалният модел е показан на фигура 1.

Всяка смяна на статута на даден служител, трябва да се отразява в „Microsoft Active Directory“, което е централният контейнер за съхранение на акаунти и след това същите промени да се извършват във всички други софтуери, които нямат разработена връзка с „Active Directory“.

Всички действия, извършени от скриптовете върху потребителските профили, се записват във файл. Историята е нужна от гледна точка на сигурността – за проследяване и откриване на случаи на злонамерени действия върху потребителски акаунти, а и за отстраняване на проблеми със системата, които може да възникнат.



Фигура 1. Концептуален модел на процесите по управление на потребителски профили



Фигура 2. Концептуален модел за смяна на парола на потребител

Тук има и още един процес, който трябва да се разгледа. Освен промените, инициирани от софтуера за управление на човешките ресурси, единственото нещо, което потребителят има възможност сам да променя по профила си, е своята парола. За да не се получава объркване по системите и ситуации, в които един потребител има различни данни за достъп в различните системи, е необходимо да се създаде портал, където да бъде извършвана смяната на паролата и тази смяна да бъде отразявана във всички софтуери, в които е необходимо. За тази цел на потребителите трябва да се забрани смяната на

парола през всички програмни продукти, където това е възможно, освен през портала за смяна на парола. На фигура 2 е показан и този концептуален модел.

Тук, също както и при предния модел, е много важно да имаме отчетност. За тази цел потребителите, които са сменили успешно паролата си, получават писмо на електронната си поща, а администраторите имат достъп до отчет с история на действията в портала – успешни и неуспешни.

Всички процеси, и при двата модела, е необходимо да бъдат имплементирани по такъв начин, че ако имаме неуспешно действие някъде по веригата, процесът да бъде прекъснат, да се върнат всички промени в обратен ред и да се запише какъв е проблемът в историята на действията.

Що се отнася до концепцията за интернет портал, то той трябва да има три основни части – централна страница, подсайтове за отделните звена или групи и „self-service“ част.

**Централна страница:** това е мястото, където са ситуирани ресурсите, достъпни за всички служители на болницата. Дизайнът трябва да е съобразен с корпоративната идентичност на организацията, т.е. да се използват официално лого, цветове и елементи. Достъпът се осъществява автоматично според потребителя, който ползва компютър чрез „Single Sign On“ процес. Ако компютърът, от който потребителят се опитва да влезе, не е част от организацията, то потребителят сам трябва да въведе данните си за достъп. Целта на страницата е да е основна за достъп до всички възможни ресурси, които предлагат информационните системи на СБАЛО. Това са:

- новини;
- указател;
- инструкции;
- вътрешно-нормативни документи;
- бланки;
- Болнична информационна система (БИС) – инструкции и уеб достъп;
- секция „За мен“ – попълване на електронна молба за отпуск, смяна на парола, създаване на собствена работна област (подсайт) и други.

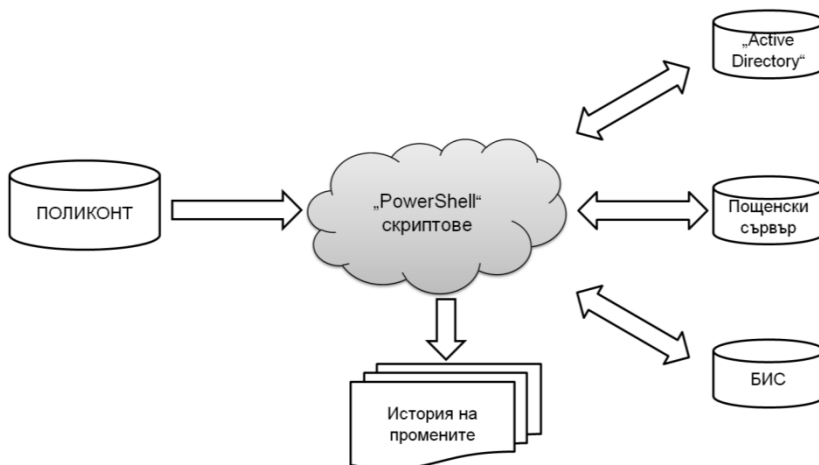
*Подсайтове за отделните звена или групи:* Това са области, създадени от администраторите за целите на различни отдели или групи потребители. Например сайт на информационен отдел, на всички старши сестри, на счетоводство и т.н. Тези сайтове са създадени от администратора, но контролът по управлението на съдържанието е поверен изцяло на съответното звено и е съобразен със структурата му. С цел олекотяване и опростяване на съдържанието в портала, тези подсайтове са видими само за потребителите, които имат достъп до тях.

*„Self-service“ част:* Страница, през която всеки служител може да създава собствени работни области, да ги споделя с когото пожелае и сам да има пълен контрол върху съдържанието, което се публикува.

## Архитектура

### Система за управление на потребителски акаунт

На фигура 3 е показан моделът на работа на скриптовете за управление на потребителски акаунти. Софтуерът на човешки ресурси е избран за началото на процесите по управление на служителските профили, защото той най-пълно и точно отразява реалното им състояние. В него е задължително да са въведени всички договори на персонала.



Фигура 3. Обща архитектура на системата за управление на потребителски профили

Всеки ден се генерира XML файл, в който се съдържат промените, направени в данните за служителите през последното денонощие. Този файл се прочита с „PowerShell“ скрипт. Във файла може да има три типа промени: назначаване, напускане и промяна на служител. Според типа промяна се задейства функция, която изпълнява необходимите действия по управлението на потребителските профили. Всички извършени действия се запаметяват във файл с история на промените. „PowerShell“ скриптовете са много добър избор за реализирането на тези операции, защото скриптовият език е част от пакета с програми на всички операционни системи и продукти на „Microsoft“ от момента на създаването му. Той е интегриран в тези софтуерни продукти и дава възможност за автоматизация на процесите по конфигуриране [2]. Това е и целта на тази система: да автоматизира процесите на работа и да премахне човешкия фактор при повечето операции по менажиране на акаунти. Тези скриптове могат да се изпълняват автоматично чрез конфигурация в „Task Scheduler“ на „Windows“ [3]. Имат възможност да бъдат подписани със сертификат, за да се гарантира, че първоначалният код не е променен и да се избегнат злоупотребите.

Друг „PowerShell“ скрипт проверява всички акаунти на потребители, премества ги в определения „organizational unit“ в „Active Directory“ и им назначава нужните права според длъжността и звеното. Той се изпълнява на всеки няколко часа. Това е нещо като свързване на потребител към определена роля според неговата принадлежност към звено и длъжност. Този скрипт е нужно да бъде постоянно променян, за да отговаря на променящата се структура в организацията.

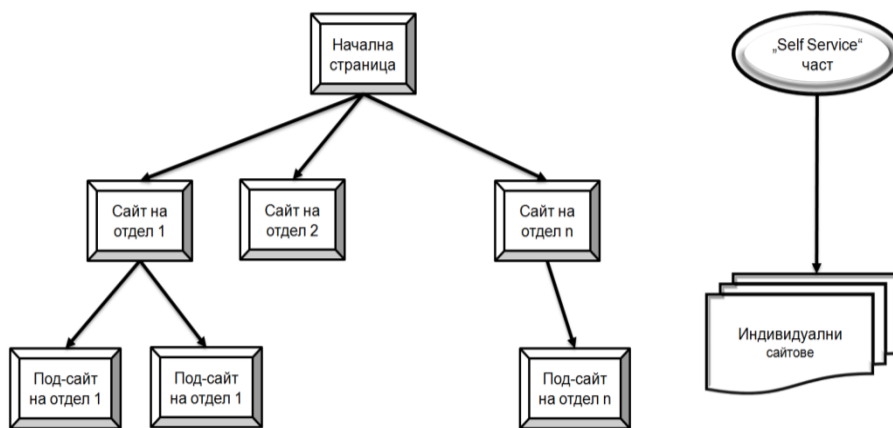
Архитектурата на уеб страницата за смяна на парола не е много сложна. Тя се състои от две обособени части. Едната е уеб интерфейсът, който се визуализира на клиента, достъпващ уеб страницата. Той е разработен на „ASP.NET“. Технологиата е много мощна, с добре разработена функционалност. Това я прави чудесен избор за разработване на решения, които да са лесно поддържани и внедрени. Тя изпълнява най-важната цел на интерфейса за смяна на парола – лесно достъпна е за потребителите. Реализацията като уеб страница дава възможност за интеграция вътре в интранет портала на организацията. Потребителите имат достъп от всеки компютър, без да е необходимо да се инсталират допълнителен софтуер. Лесно може да се контролира откъде може да се достъпва. Клиентите комуникират с уеб частта, а всички действия се изпълняват в безопасната среда на сървъра. За да бъде комуникацията между клиента и сървъра защитена, е необходимо да се използва криптирана връзка посредством сертификат. Така пренасянето на данни е напълно защитено в двете посоки – клиент-сървър и сървър-клиент.

Втората част е „PowerShell“ скрипт, който се стартира на сървъра, когато клиентът подаде команда за смяна на парола. „PowerShell“ средата за скриптове ни дава съвместимост с управление на „Microsoft“ софтуерните продукти и операционни системи. Също така има възможност за осъществяване на връзка и администриране на други видове софтуерни системи посредством модули. От гледна точка на сигурността, защото говорим за работа с потребителски данни за достъп до важни системи за организацията, е създаден потребителски акаунт по модела за най-малко право. На потребителя са дадени възможно най-малко права, за да изпълнява задачата си, а именно – да може да изпълнява „PowerShell“ скрипта на уеб сървъра и да има право да сменя потребителски пароли в „Active Directory“ и другите информационни системи. В това число не се включват сервизни, вградени, администраторски и други важни акаунти в „Active Directory“. Тук, както и при системата за управление на акаунти, типът на този потребител е „Group Managed Service Account“.

## **Интранет портал**

Порталът може да се визуализира като дървовидна структура от сайтове или групи от сайтове. На фигура 4 е визуализирана карта на наследяването. Всички подсайтове могат да наследят нивата на достъп от сайта родител или да имат свои собствени уникални. Различните цветове на визуализация означават различни нива на достъп в отделните сайтове. „Началната страница“ е основната и до нея имат достъп всички служители. Към нея са свързани подсайтове, които са създадени статично, за да обслужват различните отдели

или обособени групи от служители. Тези сайтове се създават от администратор при възникване на необходимост. Те се виждат само от потребителите, на които е разрешен достъпът до тях. „Self Service“ частта е отделена от общата структура и там всеки служител може сам да създава собствени сайтове и области.



Фигура 4. Структура на интранет портала

**Дизайн** – Той отговаря на корпоративната идентичност на болницата. Тъй като СБАЛО има интернет страница, то вътрешният портал е с максимално близък до нейния дизайн, за да може служителите да гледат на вътрешния портал като на нещо познато, а не като на още една нова система, която трябва да използват. Затова се използват графични елементи и цветове от интернет страницата на болницата. Подредбата и логиката на структуриране на отделните елементи в централната страница също следва логиката на досегашните модели на работа. За пример може да се даде съществуващият хартиен телефонен указател на болницата, с който хората са свикнали да работят до момента. Електронният вариант отговора на общоприетата специфика на подредба. Освен всичко това тук важат и стандартните изисквания за всяка уеб страница – да бъде интуитивна, бърза и максимално опростена, за да може желаната информация да се достига бързо.

**Съдържание** – Проблемите, които много фирми имат при внедряването, са, че е почти невъзможно да се каже как „трябва да изглежда интранет порталът на дадена организация“. „SharePoint“ не трябва просто да бъде имплементиран, а трябва да се използва. Много важно е да се идентифицират специфичните и критични бизнес процеси вътре в организацията. Критични са процесите, които правят организацията това, което е. Необходимо е пълно разбиране на тези процеси и имплементирането им в „SharePoint“. Освен това, внедряването предполага основно познание на самия продукт. Трябва да се

проучат възможностите на „SharePoint“ относно конкретните процеси в организацията. След това е необходимо да се намерят допирните точки между функционалност и вече проучените бизнес процеси. Трябва да се мисли и напред. Важно е да се отсеят нещата, които наистина има смисъл да се имплементират. Ако нещо няма да бъде широко използвано, то е по-добре да се наблегне на друга част. Често служителите трудно приемат промените и новостите. За да се грабне вниманието им е необходимо нещо, което да ги заинтригува. Това може да бъде нещо изключително дребно, но което е полезно в ежедневната им работа. Например, възможност да правят справки за оставащия им годишен отпуск и електронно попълване на молба за отпуск. Нещо, което ще им спести време и е лична облага. Това ще ги накара чрез ползването на една функционалност да открият и други полезни за тях неща.

**Търсене** – Възможността за търсене е ключова. Целта на решението е да бъде централно място за съхранение на информация. Това предполага прогресивно натрупване на големи количества данни с течение на годините. С времето ще се окаже проблем намирането на релевантна информация. Имплементирането на качествено и бързо търсене е от висока важност. За постигането на по-голяма ефективност е вградена възможност за търсене в съдържанието на файловете.

**Кошче** – Възможност за възстановяване на изтрити елементи на системата от потребителите със съответните права. „Кошчето“ е реализирано на две нива. Първото ниво пази съдържание, изтрито в рамките на текущия сайт. То може да се възстановява – това означава да бъде връщано на мястото, откъдето е било премахнато, в състоянието, в което е било. След изтичане на конфигуриран период, съдържанието от първо ниво на „кошчето“ отива във второ ниво. Това ниво дава възможност на администраторите да възстановяват съдържание, както това става в предното ниво. След изтриването от второ ниво съдържанието не може да бъде възстановено, освен ако не се съдържа в някой архив на системата.

**История на версии** – Може да се конфигурира съдържанието в сайтовете да има история на промените, за да се проследява „Кой?“, „Кога?“ и „Какво?“ е променил по елементите и всичко това да може да се възстанови до предишен етап.

## **Обобщение и бъдещо развитие**

Желанието за оптимизиране и подобряване на ежедневната работа доведе до имплементирането на такова решение. Служителите в болницата са слабо компютърно грамотни, но в днешно време информационните системи са неразделна част от всеки сектор. Извършването на ежедневната работа е немислимо без тях. Изключително важно е да се създаде максимална леснота за работа, а в същото време и достатъчно функционалност, за да не се налага служителите да я търсят по начини, които могат да застрашат сигурността на компанията. Системата се опитва да намери перфектния баланс между функционалност и ниво на сигурност. Наученото от този проект със сигурност

ще бъде полезно и приложимо като стратегия в много организации, като, разбира се, е необходимо да се съобрази със спецификите на всяка от тях. Това решение не възникна изведнъж, визията за него бе променяна многократно, докато се достигне до този избран подход. Истинският критерий за успеха на такова начинание е именно дали проектът е полезен и се използва от хората. Персоналът много бързо свикна с портала, като всеки намери по нещо полезно за себе си, което улеснява ежедневната му работа. В момента служителите активно дават препоръки и изказват желаниа за внасяне на още подобрения в работата им.

Доразработката и развитието на интранет портала определено е един непрекъснат процес, който със сигурност ще продължи и занапред. Няма как точно да кажем в каква посока ще се насочи, защото, въпреки че са създадени някои базови функционалности и са автоматизирани част от процесите, именно съдържанието и потребителите на системата са тези, които движат развитието напред. Целта е всеки да успее да оптимизира и подобри работата си. Има някои ключови неща, които се планират в момента и се очаква да се реализират в близко бъдеще:

- Отваряне на портала към интернет, за да бъде наистина достъпен отвсякъде. За целта е необходимо закупуването на електронен сертификат, за да се постигне защита на информацията при нейното пренасяне;
- Автоматично създаване на личен сайт за всеки служител по зададен шаблон, до който само той да има достъп. След това, чрез групова политика, настройка за синхронизиране на „Desktop“ и „My documents“ папките му върху този сайт. Така ще се осигури достъпност до важните документи, централизирано управление на съдържанието и запазване на история на версиите.

Системата за управление на потребителски акаунти е изключително зависима от другите информационни системи и е пряко свързана с променящото се щатно разписание. Това я определя като непрекъснато изменяща се. Най-важното е постигането на добра обратна връзка за възникнали проблеми и запис на подробна история на действията, за да може своевременно да се предприемат промени. Това са и нещата, на които се очаква да се обръща най-много внимание и занапред.

## **Литература**

1. Георгиев, Мл.: Решение, внедрено в „СБАЛ по онкология“ ЕАД, за повишаване ефективността на работа на ИТ отдела и подобряване на информираността и комуникацията между служителите в рамките на организацията. Дипломна работа за придобиване на магистърска степен във ФМИ при СУ, София, 07.10.2015.
2. Windows PowerShell,  
<http://www.computerperformance.co.uk/powershell/index.htm> (accessed 20.08.2015)

3. Windows TaskScheduler,  
[https://msdn.microsoft.com/en-us/library/windows/desktop/aa383614\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/aa383614(v=vs.85).aspx)  
(accessed 20.08.2015)
4. Conceptual Overview of SharePoint Foundation. Microsoft MSDN, 5.01.2011,  
[https://msdn.microsoft.com/en-us/library/office/ee537319\(v=office.14\).aspx](https://msdn.microsoft.com/en-us/library/office/ee537319(v=office.14).aspx)  
(accessed 20.08.2015)

## **Optimizing the Integration Processes between Enterprise Information Systems through Single Sign On**

**Mladen Georgiev**

**Abstract:** The article examines a problem facing the IT department of most medium-sized and large companies. When the employees of the company who need to have access to information systems and resources in the organization exceed 200-300 people, from the perspective of efficiency and productivity it is indispensable to pay more attention to the way of sharing and accessing valuable business information in the company and optimising the processes of administering user groups of this size.

# Параметрични вектори за разпознаване на емоциите в говора

Александър Илиев

Институт по математика и информатика при БАН, София  
[al.iliev@math.bas.bg](mailto:al.iliev@math.bas.bg)

**Резюме:** В това проучване са представени група от параметри, необходими за създаване на система за разпознаване на емоциите в говора на свободна реч. То обхваща набор от параметрични елементи, необходими за съставянето на вектори от данни, които могат да бъдат използвани за обучението на разнородни класификатори. Съставянето на модели, съдържащи параметри на емоции, е първата и най-важна стъпка за създаването на система за разпознаване на различни емоционални състояния. Емоциите са пряко свързани с езика и бита на различните общности и следва моделите на емоционални състояния да бъдат пряко зависими от този факт. Поради това, целта на проучването е да предостави параметрите, които да представляват градивните елементи на различни модели на емоции в говора. За целта подробно е описан моделът на изговаряне на речта, а след това е обсъдена методиката на извличане на отделните параметри от говора. Предоставени са и блок-диаграми на създадените методики.

**Ключови думи:** емоции, говор, разпознаване, параметри, класификация

**ACM 1998 Classification Keywords:** D.3.3: Language Constructs and Features – Patterns; I.5: Pattern Recognition – I5.2 Design Methodology – Feature evaluation and selection

## 1. Въведение

Информацията, съдържаща се в речта, е богата, многообразна и многостепенна. Докато съобщението може да бъде ортографски транскрибирано, то съдържа допълнителна информация, която показва идентичност, пол, възраст, физическото и емоционалното състояние на говорещия. Автоматичното извличане на емоционалното състояние може да бъде много полезно в сектора на услугите и има приложение в сферата на сигурността, поради което то получи значително внимание в наши дни. Емоцията се съдържа в прозодичните модели на словото. Тоналните и прекъсващи индекси на системата Тоби (Tonal and Break Indices – ToBI), която е стандарт за анотация, се превърна в стандарт на прозодична транскрипция на интонационни модели, като резултат от сътрудничеството на различни научни области като езикознание, психология и електротехника, търсещи общи прозодични елементи за транскрипция. Тоби се използва предимно за синтез на говор, където прозодията трябва да бъде описана по някакъв начин. Тази система се състои от няколко области, всяка от които описва различни параметри в различни моменти от времето. Четирите основни области с

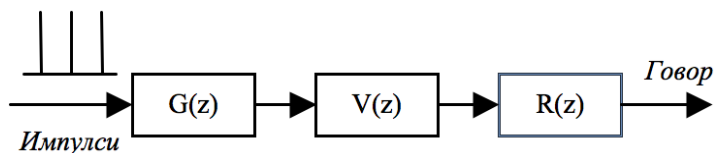
параметри са: *тонална, прекъсваща, правописна и обща*. В това изследване изразяването на специфичните езикови особености е извършено въз основа на два интонационни класа: *интонационни фрази (IP)* и *междинни фрази (intermediary phrasing – ip)*. В тази работа за изобразяване на различните емоционални състояния бяха използвани само елементи от тоналното ниво. Въз основа на правилата за генериране на Тоби, както и на елементите, базирани на интонация на американски английски език, предоставени от Jilka и др. и Stibbard, бе разработен метод за автоматично извличане на Тоби маркери. Бяха използвани 16 параметъра от *тоналната* област на Тоби, които според вида си бяха разпределени в пет различни категории, показани в таблица 2. След детекция и разпознаване на параметрите, те бяха записани в матричен вид за анализ. Извлечените параметрични вектори бяха използвани за разработване на Гаусов Микс Модел (GMM) като класификатор на разпознаване на емоционалното състояние в системата. Бяха използвани четири различни класа параметрични вектори: (1) сигнал, съдържащ класически параметри (11 параметъра), (2) Тоби (16 параметъра) и (3) комбинация от двете (27 параметъра общо). Също така бяха извлечени и кепстрални коефициенти посредством мел скалата, а именно (4) MFCCs за съставянето на още един параметричен вектор от данни за разпознаване на емоциите в говора. Този метод бе приложен върху сигнала на говора, както и на сигнала, получен от епиглотиса.

В тази работа шестте емоции за класифициране и разпознаване бяха *радост, гняв, тъга, страх, изненада* и *неутрална*. Корпусът на речта, който бе използван за тези емоционални състояния, бе ръчно маркиран. Системата бе обучена върху 80% от пробите на речта в корпуса. Дължините на фразите варираха от 1 секунда до 12 секунди. Останалите 20% бяха използвани за тестване.

## 2. Извличане на параметри от сигнала

### 2.1. Модел на гласовото възпроизвеждане

На гласната реч може да се гледа като на изходен сигнал от възпроизвеждаща система, състояща се от три линейни и променливи във времето подсистеми, както е показано на фигура 1.



Фигура 1. Модел на изговаряне на речта

По време на речевото изразяване възбуждането на системата се произвежда предимно в глотиса и има квазипериодичен характер. Импулсните серии от

въздушно напрежение, насочени към глотиса, са последователни във времето и се формират като резултат от трептенията на гласните струни по време на производството на вокален сигнал. Глоталната вълна определя основната честота на говора. Спектърът на получената реч, измерен при устните,  $S(Z)$ , може да се изрази в комплексната честотна област като:

$$S(z) = G(z)V(z)R(z), \quad (1)$$

където  $G(Z)$  е глотален модел,  $V(Z)$  е функцията за трансфер на вокалния тракт и  $R(Z)$  е ефектът на излъчване на речта от устните.

За да получим спектъра на глоталната вълна от възпроизведената реч трябва да са налични системата на вокалния тракт, както и тази на устната кухина. За  $R(Z)$  имаме прост филтър от първа степен, а именно:

$$R(z) = 1 - \alpha z^{-1}, \quad (2)$$

където  $0.95 \leq \alpha \leq 1.0$ . Ако моделът на излъчващия ефект от устните, както и този на вокалния тракт, са адекватно моделирани, тогава методът на обратното филтриране на глоталния сигнал е възможен. По време на гласово възпроизвеждане вокалният тракт може да се моделира като многополюсен филтър, изразен като:

$$V(z) = \frac{1}{1 + \sum_{i=1}^p a_i z^{-i}}, \quad (3)$$

Коефициентите на  $V(Z)$  филтъра могат да бъдат лесно получени посредством анализ по метода на линейното предсказване (LP), от типа на автокорелация или ковариантен метод (Rabiner и Schafer, 1978), където  $p$  е от степента на предсказване. Решението на глоталния модел  $G(Z)$  от формула (1) може да дойде от обратното филтриране на глоталния сигнал, а именно:

$$G(z) = \frac{S(z)}{V(z)R(z)}, \quad (4)$$

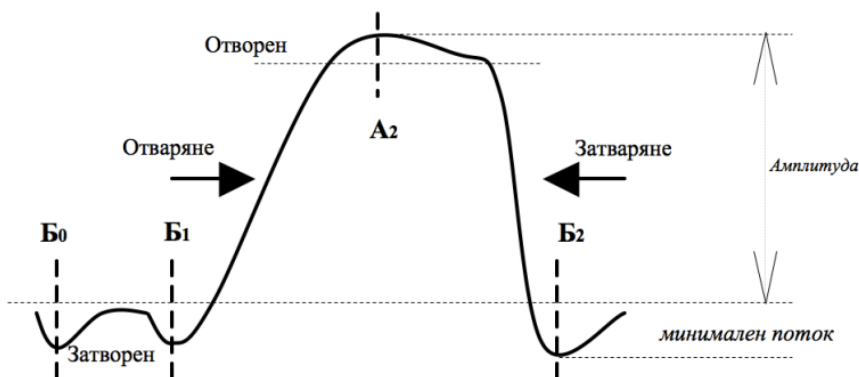
Ако моделът на вокалния тракт е определен с точност в кратковременните интервали на сигнала на речта, тогава обратното филтриране на глоталния сигнал може да се осигури, при условие, че съставните части на модела на гласовото възпроизвеждане са линейно отделими и не си взаимодействат една с друга. В действителност, вокалното възпроизвеждане се влияе от вокалния тракт, което води до различия в гласовия обем и енергия като променя и структурата на звуковата форма на гласовата вълна. Освен това, вариациите на глоталния сигнал, както се наблюдава при използване на ларингограф, не винаги отразяват подобни вариации в глоталния въздушен поток (Flanagan, 1972; O'Shaughnessy, 2000; Quatieri, 2002). Независимо от това, ако бъдат

намерени само главните характеристики на гласовия сигнал, като моментите на отваряне и затваряне на глотиса и съотношението между фазите на отваряне и затваряне (глотална симетрия), обратното филтриране може да бъде ефективен метод в осигуряването на гласова информация, идваща директно от епиглотиса. В тази работа обратното филтриране осигурява нивото на точност, изискващо се за ефективно използване на гласовата информация в задачата за класификация на емоциите.

Експериментално глоталният сигнал бе получен посредством две контактни пластинки от EG2-PC ларингограф на Glottal Enterprises, поставени на гърлото на говорители-доброволци. В пълна симетрия с това бе записана и кореспондиращата реч посредством микрофон и дигитално записващо устройство със семплираща честота  $F_S = 22,050$  Hz. Времени съответствия между двата сигнала се постигна чрез разглеждане на времевата разлика между гласовите събития в епиглотиса и в речта, като последните бяха записани от микрофон, разположен на разстояние  $d$  от устата на говорителя с вокален тракт дължина  $l$ . Тогава разликата във времето в периоди на вземане на пробата  $n_0$  между източника и записа на речта бе определена като:

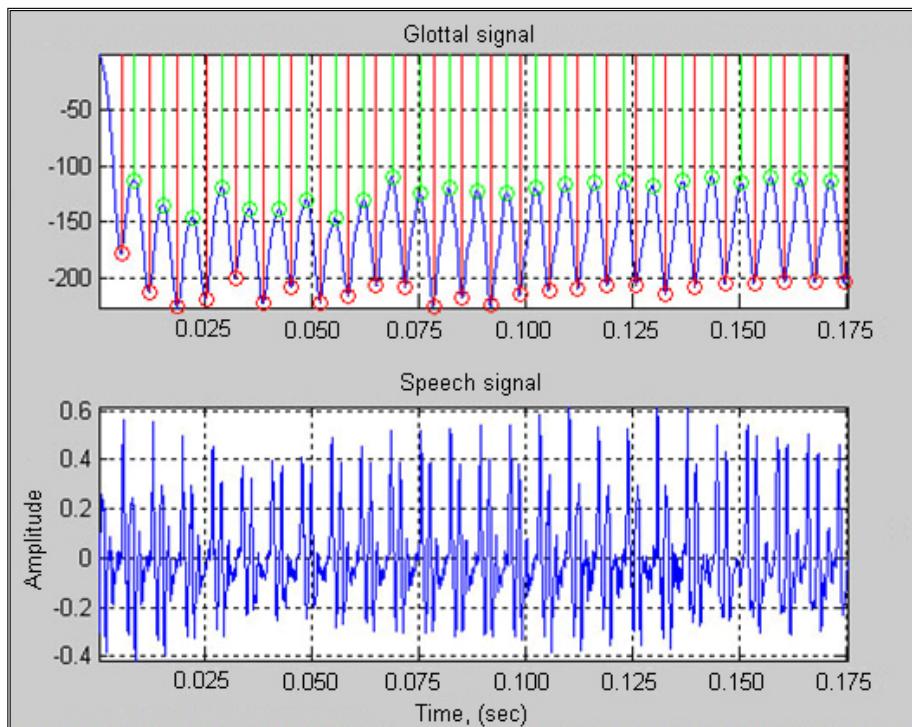
$$n_0 = \frac{(d+l)f_S}{c}, \quad (5)$$

където  $c = 343$  м/с е скоростта на звука във въздуха. По-детайлно разглеждане на двата сигнала потвърждава връзката между момента на гласовото затваряне и региона на максимално смущение в сигнала, записан от речта.



Фигура 2: Глотален импулс с четирите си фази: Отваряне, Отворен, Затваряне и Затворен

Типичен гласов импулс е показан на фигура 2. Точки А и Б показват най-високите и най-ниските стойности във формата на вълната. В горната част на фигура 3 е показан глотален сигнал, извлечен с помощта на обратно филтриране, а в долната е съответната реч.



Фигура 3: Извлечен глотален сигнал и кореспондиращата му реч.

## 2.2 Обратно филтриране на глоталната вълна

Обратното филтриране може да бъде ефективен и практичен метод за изчисляване на глоталната звукова вълна от речта. То е необходимо поради това, че в свободната реч глотален сигнал не може да бъде записан и използван по друг начин. Качеството на получения резултат е с критична зависимост от точността на оценката на глоталната част от системата за възпроизвеждане на говора. Веднъж записана, речта се филтрира, за да се добие възможно най-близко копие на глоталната форма на вълната. Важен фактор, който влияе на формата на гласовия въздушен поток, е типът на фонация.

Много методи са свързани с обратното филтриране на говора (Rothenberg, 1973; Wong и др. 1979; Moore и др. 2003; Brooks и др. 2006), но основните проучвания в областта се основават на две основни процедурни групи в зависимост от метода на записа: в самата устна кухина, предложено от Rothenberg (1973), и извън нея, като при този начин се отчита влиянието на излъчването от устните, Wong (1979). В подробно проучване, Rothenberg (1973) използва специално проектирана маска за записване на звука при устните. Анализът се ограничава до честота от 1 kHz. Формантните честоти и техните честотни ленти са изчислени с помощта на теснолентова спектрограма и са използвани за филтриране на вокалния тракт, така че може да се извлече

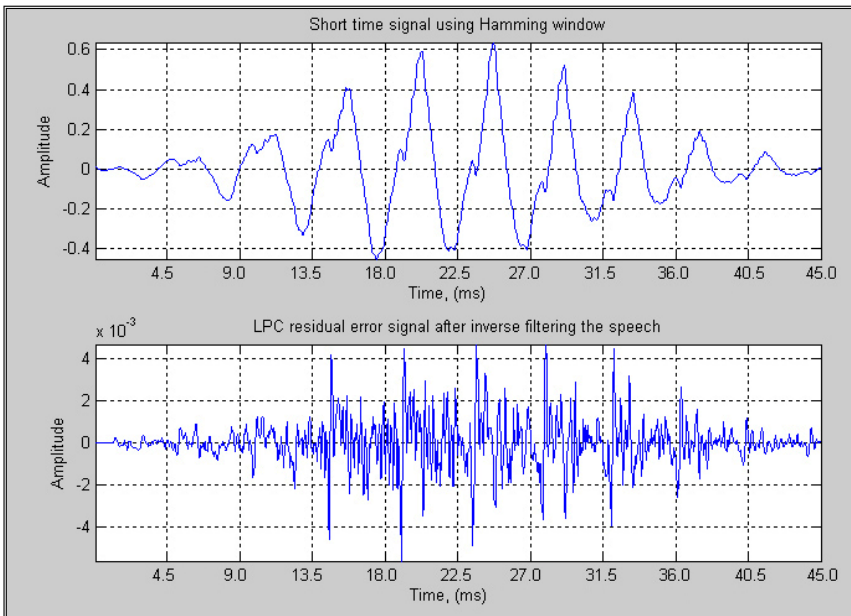
глоталният сигнал. Въпреки че настройката, използвана в това изследване, е доста рестриктивна и не може да се прилага върху предварително записани речеви бази данни, то все пак предлага повече детайли върху формата на глоталната вълна. Тъй като тази техника използва маска за измерване на потока, естествено е един от най-важните приноси да е, че получената глотална вълна носи полезна информация за амплитудата. Ако приемем, че записващото устройство е подходящо калибрирано, тогава и *минималният дебит*, и *амплитудният поток* (виж фигура 2) на сигнала, идващ от епиглотиса, могат да бъдат надеждно получени след обратно филтриране. Както бе посочено от Rothenberg, на практика обратното филтриране се ограничава до леко назални или неназални гласни. Резонансите на вокалния тракт са представени чрез комплексни двойки полюси и са посочени като *форманти*. Ефектът, създаден от формантите, трябва да бъде анулиран (обратно филтриран) чрез въвеждане на комплексна нула до всеки един комплексен полюс на вокалния тракт, както и полюс или резонанс от първи ред на нулевата честота. Предимствата на този метод са: непроницаемост на нискочестотен стаен шум; полученият сигнал постига точност до нулева честота; и по-добро калибриране на нивото на амплитудата, използвайки постоянен въздушен поток. Основният недостатък е, че тя се извършва в лабораторни условия, използвайки специализирани методи, като по този начин тя не е практична за обработка на вече записана реч при нормални условия.

### 2.3 Автокорелация и линейно-предсказващ метод (LPC)

За да направим този модел за разпознаване на емоциите в говора по-устойчив към синхронизацията между аналитичните прозорци и циклите на ларинкса, както и да се получи по-добра устойчивост към шума, за анализ на речевия корпус се прилага автокорелация LPC. При сравняване на ковариацията и автокорелацията на линейно предсказване има три основни въпроса: какъв е броят на умноженията за изчисляване на матрицата на корелацията и като следствие да се намери решение на матричното уравнение; колко е количеството използвана памет; и каква е стабилността на системата. Всички тези стойности са добре обобщени от Rabiner и Schafer (1978). За метода на ковариация броят на умноженията на корелационната матрица е  $* p$  и за да се намери решение за неговото уравнение се изискват  $\frac{(p^3+9p^2+2p)}{6}$  умножения,  $p$  разделяния и  $p$  коренувания. За сравнение, методът на автокорелация се нуждае от същото количество умножение за корелационната матрица или  $N * p$ , но решението на матричното уравнение се нуждае от много по-малко изчислителната мощ, или по-точно казано – от  $p^2$  умножения. Количеството използвана памет, необходимо за данните в метода на ковариация, съответства на броя на точките за анализ  $N$  и за корелационната матрица техният брой е  $\frac{p^2}{2}$ .

За метода на автокорелация тези цифри са:  $N$  за броя на данните и  $p$  за автокорелационната матрица, която отново е по-малка от тази, необходима в метода на ковариация. И накрая, стабилността на метода на автокорелация

почти винаги е гарантирана, когато тя се изчислява с достатъчна точност, което от своя страна означава, че трябва да се използва достатъчно висока степен на прогнозиране. Полиномите на предсказването обикновено остават стабилни, когато се използва филтър за преекспонация на речта. Въпреки това, стабилността на полиномите на предсказването не може да бъде гарантирана в метода на ковариация. Като цяло, ако броят на елементите в аналитичния прозорец е достатъчно голям, двата метода ще доведат до подобно решение. Като се имат предвид характеристиките на двата метода на ковариацията и автокорелацията на линейно предсказване, последните остават във фокуса на анализа, направен върху базата с разговорна реч, използвана тук. Фигура 4 показва остатъка от линейно предсказване върху откъс от сигнал на дадена реч. Моментите на глотално затваряне се виждат ясно, тъй като те са представени от върховете на остатъчния сигнал.



Фигура 4.: Прозорец от времевата област на Хеминг (горе) и сигналът LPC, останал след обратното филтриране на говора (долу).

### 3. Параметрични вектори

#### 3.1 Класически и Тоби прозодични параметри

Класическите прозодични параметри са: тоналност и енергия във времевата област. Автоматизираното извличане на параметри от Тоби пространството бе реализирано и изпълнението му бе сравнено с класическите прозодични

параметри. Използването на елементи от тоналната област води до разпознаване на различните емоционални състояния. Тяхното откриване се базира на разпознаване на тона посредством използването на просто проследяване от обратния филтър (Simple Inverse Filter Tracking – SIFT алгоритъм (Markel, 1972)). Класическият прозодичен метод използва тоналност и енергията от времевата област за извличане на полезни данни за разпознаване на емоции. Фина настройка на подбора на Тоби елементите и класическите прозодични елементи бе предложена чрез изчисляване на взаимната информация (Mutual Information) на всички параметри от всяка област. Подборът от параметри се намалява допълнително след преоценката с метода на последователния напредващ подбор (Sequential Forward Selection). Тук е разработена една нова и подобрена параметрична област като комбинация от двете, която демонстрира по-стабилно представяне в чиста реч. Всички тестове в корпус 1 се провеждат в текстово- и говорител-независима среда.

Параметричният вектор може да се представи като  $F_c$ :

$$\underline{F_c} = (p_1, \dots, p_6, e_1, \dots, e_3, d_1, d_2) \quad (6)$$

където  $p$  представлява тоналните елементи,  $e$  представлява енергийните характеристики и  $d$  представляват параметри с продължителност от таблица 1.

Формулировката на специфичните езикови правила се основава на два класа: *интонационни фрази* (IP) и *междинни фрази* (ip). Всяко дадено IP се състои от няколко по-малки ip-та. Основният тон на всяко IP се извлича в началото, използвайки просто проследяване от обратния филтър (SIFT алгоритъм на Markel, 1972), след което започва разпознаването на параметри за формиране на параметричните вектори. Подбраните класически прозодични параметри бяха 11 на брой. Първите шест елемента бяха извлечени от основния тон (в честотната област) на текущия аналитичен прозорец. Бяха съхранени вариациите на тоналността във всеки IP през целия проверяван аналитичен прозорец и след това бе изчислено съотношението на покачване-спадане на тона. Следващите пет параметъра бяха получени посредством използването на енергията на сигнала във времевата област. Средната продължителност на паузата бе определена с помощта на кръстосана корелация или честотата на пресичане на нулата. На тяхна база бе изчислена средната стойност на всички регионални паузи в едно IP. Темпът на говорене бе представен от съотношението на броя и продължителността на всички гласови сегменти (Wang, 2 004).

*Таблица 1: Класически прозодични параметри*

Тонални елементи	
1	средна
2	медиана
3	стандартно отклонение
4	максимум

5	възход-спад
6	максимален диапазон на спад
<b>Енергийни елементи</b>	
7	средна
8	стандартно отклонение
9	максимум
<b>Елементи с продължителност</b>	
10	средна дължина на паузите
11	скорост на изказа: брой/дължина на гласовите сегменти

Подбраните Тоби параметри бяха 16 на брой. Според вида си тези параметри бяха разпределени в пет различни категории, показани в таблица 2. Тоби системата е доста динамична, което поражда някои трудности при автоматизираната детекция на параметри. Някои Тоби елементи, след като са открити, но са в комбинация с предшестващи или следващи ги такива, може да бъдат конвертирани и да представляват други елементи. Един пример може да бъде  $\mathbf{H}^*$  и  $\mathbf{L}+\mathbf{H}^*$ . В случаите, когато в рамките на едно  $\text{ip}$  има присъствие на  $\mathbf{H}^*$  и  $\mathbf{L}$ , и  $\mathbf{H}^*$  предшества, нарастващ пик с би-тонално фразово ударение  $\mathbf{L}+\mathbf{H}^*$  е открит и горната линия  $\mathbf{H}^*$  в този момент е презаписана на  $\mathbf{L}+\mathbf{H}^*$ . Ако в рамките на дадено  $\text{ip}$  е имало наличие на  $\mathbf{L}^*$  и  $\mathbf{H}$ , и  $\mathbf{L}^*$  предшества, спадащо би-тонално фразово ударение  $\mathbf{L}^*+\mathbf{H}$  е открито и базовата линия  $\mathbf{L}^*$  в този момент е презаписана.

Важно е да се отбележи, че някои Тоби елементи са взаимно изключващи се. Групиране от вида  $!\mathbf{H}+\mathbf{H}^*$  не е възможно. Друг пример е крайната ниска граница на тона  $\mathbf{L}\%$  и крайната висока тонална граница  $\mathbf{H}\%$ . Те винаги се показват в комбинация с предходната тонална информация. Една от причините Тоби системата да бъде избрана като метод за извличане на отделните части на прозодията е, че тя осигурява допълнително и задълбочено взаимоотношение между различните елементи в тоналната област. Посредством Тоби бе установено, че глоталната форма на вълната е пряко свързана с емоционалното състояние на говорещия.

Всички Тоби параметри съставят вектора от елементи  $F_T$ , който изглежда така:

$$\underline{F}_T = (t_1, \dots, t_{16}) \quad (7)$$

където,  $t$  – представлява всички Тоби елементи, показани в Таблица 2. Комбинираният вектор от параметри би изглеждал така:

$$\underline{F} = (p_1, \dots, p_6, e_1, \dots, e_3, d_1, d_2, t_1, \dots, t_{16}) \quad \text{или} \quad \underline{F} = [F_c F_T] \quad (8)$$

Таблица 2: Целеви тонални Тоби елементи за извличане

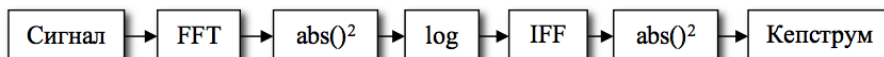
Символи:		Тоби елементи:	
1. Тонални акценти:	Монотонални акценти	L*	най-нисък тон (монотонално РА)
		H*	най-висок тон (монотонално РА)
	Би-тонални акценти	L*+H	увеличаващ се акцент (би-тонално РА)
		L*+!H	увеличаващ се акцент със занижено H (би-тонално РА)
		L+H*	нарастващ пиков акцент (би-тонално РА)
		L+!H*	нарастващ пиков акцент със занижен висока граница H* (би-тонално РА)
		H+!H*	спадащ акцент (би-тонално РА)
2. Гранични тонове	L%	финален нисък граничен тон	
	H%	финален висок граничен тон	
	%H	първоначален граничен среден-тон	
3. Изразни акценти	L-	нисък акцент на фраза	
	H-	висок акцент на фраза	
4. Изразни акценти и комбинации гранични тонове	L-L%	много ниска точка в диапазона на говорителя	
	L-H%	нисък акцент на фраза следван от висок граничен тон	
	H-L%	висок акцент на фраза леко намален в края на фезата	
	H-H%	Изключително висок акцент в диапазона на говорителя (покачване)	
	!H-L%	спадащ високо-изразен акцент с лек спад на края	
5. Спад	!H*	спад - компресия на тоналният диапазон или занижена висока граница	

### 3.2 Екстракция на MFCC

В това изследване бяха използвани четири стъпки за изчисляване на MFCCs:

1. Изчислява се FFT – трансформация в ред на Фурие за всеки прозорец от дадения сигнал;
2. Чрез използването на триъгълни прозорци се извършва транскрибиране от Hertz на Мел скала на честотния спектър;
3. Изчислява се log за всеки от новополучените MEL честоти;
4. Изчислява се дискретна косинусова трансформация DCT на MEL log честотите, като по този начин се получават MFCCs, които са представени от амплитудите на новополучения спектър.

За да се получи „спектър на спектъра“ на сигнал, или *кепструм*, бе използвана следната последователност, описана във фигура 5:



Фигура 5: Блок схема за получаване на кепструм сигнал

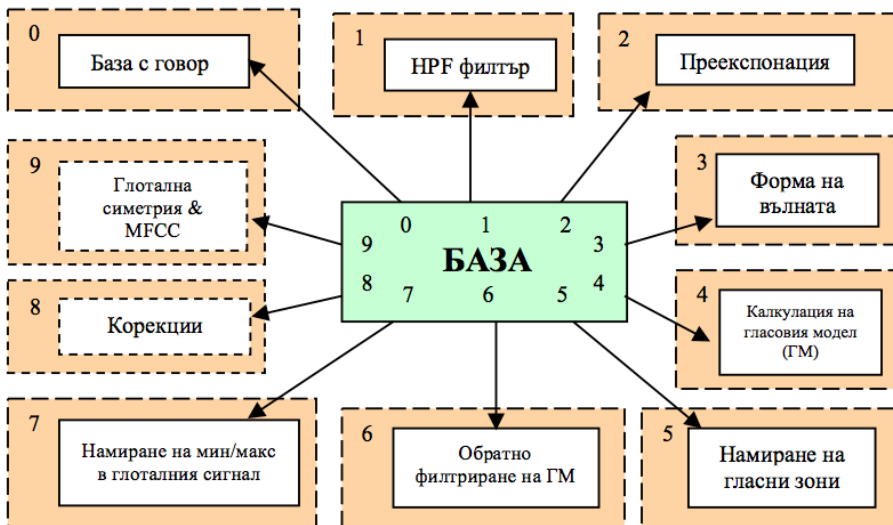
За анализа на MFCC, извършен на втория кръг от експерименти, проведени в безехова камера на звукозаписно студио, броят на параметрите варираше между 4-10 за глоталния сигнал и 6-12 за сигнала на речта, както е обобщено в Таблица 3.

*Таблица 3: MFCC степени, използвани при глотални и речеви сигнали*

Вид на сигнала	Степен на MFCC				
	4	6	8	10	12
Глотален	√	√	√	√	-
Речеви (говор)	-	√	√	√	√

#### 4. Архитектура на системата

Детайлно описание на системата за извличане на глоталната симетрия от речта визуално е изобразено на фигура 6.



*Фигура 6: Извличане на глотален сигнал и MFCCs*

Предварително записаната реч е заредена в блок 0, сигналът се филтрира посредством високо-честотно-пропускателен филтър (НРФ), за да бъдат премахнати ниските честоти под 40 Hz. За да се запази същото съотношение сигнал-шум (SNR) през целия честотен спектър, сигналът се подлага на преекспонация посредством допълнително филтриране. Формата на вълната в блок 3 се изчислява с помощта на реалната част на трансформацията на



След разделяне на емоциите, сигналът се предава на SIFT алгоритъма за извличане на основния тон, както и за да се коригират грешките. За постигане на целите бяха използвани 3 корпуса с говор. За разпознаването на параметри за корпус 1 бяха използвани класически прозодични параметри и Тоби; за корпус 2 те бяха MFCC на речта и глоталния сигнал; а за корпус 3 бяха включени всички упоменати параметри. При корпус 1 данните варираха в по-голям диапазон за различните параметрични области, следователно нормализирането им бе належащо. След като бяха извлечени всички параметри, те бяха проверени за грешки и коригирани, а в края на този процес, за да бъдат съставени параметричните вектори, те бяха нормализирани. Това се наложи, тъй като данните имаха разнороден характер. Параметричните вектори се разделиха на четири категории: Тоби (Т), класически (F), комбинирани (С), и мел честотни кепстрални коефициенти (MFCC).

## **5. Резултати и обобщение**

След анализиране на данните, събрани от експериментите, може да се заключи с голяма увереност, че глоталната симетрия се характеризира с богато емоционално съдържание, което я прави ефективно средство при изпълнение на задачата за разпознаване на емоциите в говора. Освен това бе установено, че глоталната информация е значително по-богата от класическите прозодични параметри и е по-устойчива при шумни условия, отколкото всички други анализирани параметри. Ниската честота е в естеството на глоталния сигнал и подсилва неговата способност да оцелява при тежки условия на разнороден шум и филтриране. Проблем в системата може да бъде определянето на точните моменти на глотално отваряне и затваряне, което бе решено посредством използването на подходящо групово-честотно забавяне (group-delay). Това е особено важно, когато системата се изпитва за издръжливост при шумни условия. Освен това бе доказано, че, въпреки че мел честотните кепстрални коефициенти (MFCC) не работят така добре в корпус 3, както в корпус 2 в чиста реч, като цяло MFCC показват по-добра здравина в шумни условия, отколкото класическите параметри в същите условия до SNR = 10 db. И накрая беше показано, че тоналните и прекъсващи индекси от Тоби пространството могат да допринесат за подобряване на разпознаване на емоциите. Двете области, Тоби и класическите прозодични параметри, бяха подложени на допълнителен подбор, за да бъдат възможно най-ефективни след подбора си преди съставянето на параметричните вектори за класификация. Това е показано на два пъти в корпус 1 и корпус 3. Изчислената статистическа значимост потвърждава верността на получените резултати.

## **Литература**

1. Brooks M., Naylor P., and Gudnason J., 2006. A Quantitative Assessment of Group Delay Methods for Identifying Glottal Closures in Voiced Speech. *IEEE Transactions On Audio, Speech and Language Processing*, Vol. 14, No. 2, pp. 456-466.
2. Flanagan J., 1972. *Speech Analysis, Synthesis and Perception*. Springer-Verlag.

3. Jilka M., Moler G., and Dogil G., 1999. Rules for the generation of ToBI-based American English intonation. *Speech Communication*, Vol. 28, pp. 83-108.
4. Markel J., 1972. The SIFT Algorithm for Fundamental Frequency Estimation. *IEEE Transactions on Audio and Electroacoustics*, December 1972, Vol. 20, Issue 5, pp. 367-377
5. Moore E., Clements M., Peifer J., and Weisser L., 2003. Investigating the Role of Glottal Features in Classifying Clinical Depression. *25<sup>th</sup> Annual International Conference of the IEEE EMBS*, pp. 2849-2852.
6. O'Shaughnessy D., 2000. *Speech Communications – Human and Machine*. IEEE Press.
7. Quatieri T., 2002. *Discrete-Time Speech Signal Processing Principles and Practice*, Prentice Hall.
8. Rabiner L. and Schafer R., 1978. *Digital Processing of Speech Signals*, Prentice Hall.
9. Rothenberg M., 1973. A New Inverse-Filtering Technique for Deriving the Glottal Air Flow Waveform During Voicing. *Journal of the Acoustical Society of America*, Vol. 53, pp. 1632-1645.
10. Stibbard R., 2000. Automated Extraction of ToBI Annotation Data from the Reading/Leeds Emotional Speech Corpus. *ITRW on Speech and Emotion, ISCA, in Speech Emotion*, pp. 60-65.
11. Wang Y. and Guan L., 2004. An Investigation of Speech-Based Human Emotion Recognition. *IEEE 6<sup>th</sup> Workshop on Multimedia Signal Processing*, pp. 15-18.
12. Wong D., Markel J., and Gray A., 1979. Least Squares Glottal Inverse Filtering from the Acoustical Speech Waveform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-27, No. 4, pp. 350-355.

## **Feature Vectors for Emotion Recognition in Speech**

**Alexander Iliev**

**Abstract:** In this study a group of parameters are presented to create a system for recognizing emotions in free speech. It covers a range of parametric elements necessary for compiling feature data vectors that can be used for the training of various classifiers. Compilation of models containing parameters of emotions is the first and most important step in creating a system for recognizing different emotional states. Emotions are directly related to the language and culture of different communities and therefore patterns of emotional states are directly dependent on this fact. Therefore, the aim of this study is to provide parameters to represent the building blocks of various models of emotion in speech. To this end, the model of the utterance of speech is explained in details, and then the method of feature extraction of speech parameters is discussed. Block diagrams of the newly created methods are provided as a result.

## Реклами, реклами, ... – намиране без взирание

Мирослав Иванов<sup>1</sup>, Красимира Иванова<sup>1</sup>, Илия Митов<sup>1</sup>,  
Евгения Великова<sup>2</sup>

1. Институт по математика и информатика при БАН, София  
[mivanov@math.bas.bg](mailto:mivanov@math.bas.bg), [kivanova@math.bas.bg](mailto:kivanova@math.bas.bg), [imitov@math.bas.bg](mailto:imitov@math.bas.bg)
2. ФМИ при Софийски университет „Св. Климент Охридски“, София  
[velikova@fmi.uni-sofia.bg](mailto:velikova@fmi.uni-sofia.bg)

**Резюме:** Обект на изследването са възможностите за облекчаване на търсенето на повтарящи се реклами в пресата при осъществяването на медия мониторинг. Анализира се възможността за бързо намиране на повторенията чрез използване на дескриптор на изображенията Average Hash (aHash). Предлага се оригинален подход за организация на етикетирания екземпляр и намерените повторения в една обща база от данни.

**Ключови думи:** image authentication, aHash, multimedia data base, ArM32.

**ACM 1998 Classification Keywords:** H.5.1 Multimedia Information Systems

### Въведение

В наши дни информационните и клипинг агенции добиват особена значимост при анализите на социалното присъствие на институциите и фирмите. Чрез обработване на постъпващата информация от централните и регионални електронни и печатни медии, те предоставят на своите клиенти задълбочени анализи от вида [1]:

- оценка на медийния отзвук на кампания или събитие. Този вид анализ показва коя информация е предизвикала интерес в медиите, кои комуникационни и PR дейности са били успешни. Анализът служи за оценка, оптимизиране и оформяне на бъдещите отношения между медиите и PR дейността;
- Benchmark анализ. Служи за сравняване на медийното отразяване на две или повече компании и е подходящ за проучване на медийното присъствие на конкуренти и партньори. Той показва медийното позициониране на една компания и може да служи за определяне на степента на постигане на комуникационните цели;
- изчисляване на рекламната стойност на реклами и PR публикации. Той може да се използва за остойностяване на публикувани PR материали на базата на рекламните тарифи на медиите, както и да се изчисли стойността на дадена рекламна кампания. Рекламният анализ е приложим за изчисляване на собствени материали, както и за материалите на конкуренти и партньори.

Един от основните изходни материали, предоставян на клиентите на такива компании, е бюлетинът, който съдържа както обобщени количествени данни,

оценки и анализи, така и преки референции към материалите, съдържащи името на клиента или следените от него конкуренти.

Голяма част от материалите, които са във фокуса на разглежданията, са отпечатаните реклами. Характерно е, че обикновено в рамките на период от време компанията или институцията използва една и съща реклама. В момента търсенето на срещанията на рекламите и класифицирането им по видове се осъществява ръчно от оператори, което е времеемка и уморителна работа. Цялостният процес започва със сканирането на страниците на печатните материали, след което се извършва „нарязване“ и съхраняване на отделните части в различни контейнери. Текстовите участъци се подлагат на OCR обработка, ръчно почистване и последващо категоризиране по набор от ключови думи, а изображенията се насочват към специализиран преглед за наличието на реклами на следените институции, брандове или кампании. На пазара съществуват софтуерни продукти, които се занимават с управлението на колекции от изображения, но само като пример ще посочим, че продуктът MatchEngine на канадската фирма TinEye започва от \$200 месечно за ограничен обем от изображения и стига до \$1500 (и нагоре) за корпоративни клиенти [2].

В настоящата работа се предлага един подход за организиране на процеса на преглеждане на изображенията за наличие на търсени реклами, който е полуавтоматизиран. Постъпващите изображения се сравняват с налични в базата етикетирани първообрази и при достатъчна близост те директно се категоризират в същата група. При срещане на изображение, което системата не може да причисли към някое от етикетираните, се извършва ръчно категоризиране от оператора или отхвърлянето му като нерекламен материал.

По-долу организацията е както следва: В първа точка се разглеждат фамилия дескриптори, използвани като „отпечатащи“ (fingerprint) на изображенията, като се прави оценка на приложимостта им в решаването на тази задача. Втора точка съдържа експериментални данни и анализи на избора на конкретния дескриптор и оптималните параметри за използването му. Трета точка се спира на едно оригинално предложение за организиране на базата от данни, при което етикетираните първообрази и последващо намерените близки до тях изображения се съхраняват в обща база. Заключението обобщава получените резултати и предлага насоки за бъдещо развитие.

## **1. Фамилията Nash дескриптори, използвани като отпечатък на изображения**

В практиката има редица приложения, където се получава размножаване на цифрово изображение чрез преоразмеряване, разтегляне, промяна в контрастите и форматите. Като резултат се появяват нови, подобни на първоизточника, изображения. Въпреки че изображенията не са идентични, те продължават да бъдат доста близки. И често се налага решаването на обратната задача – да се намерят близките изображения, които вероятно са трансформации от един и същ първоизточник. Повечето от изследванията са върху търсенето на сродни изображения, които са претърпели различни

цифрови трансформации [3], други изследвания са фокусирани върху проблемите на разпознаване на близост на цифров първообраз, който впоследствие е преминал през печатане и обратно сканиране [4]. Ние в случая се интересуваме от задачата за намиране на близостта на изображения, които са сканирани от различни печатни източници.

Един от предлаганите методи използва хеш функции, построявани на базата на съдържанието на изображенията. За разлика от криптографските методи (като MD5 и SHA1), при които след построяването и сравнението на стойностите на две хеш функции може да се каже само, че ако хешовете са различни, то данните са различни, а ако хешовете са еднакви, то може би данните са еднакви (поради възможността за наличие на колизии), фамилията разглеждани функции дават възможност за сравняване и търсене на близост между изображенията, чиито хеш функции се сравняват. Детайлни описания на алгоритмите са представени в [5,6]. Специално за дескриптора pHash има библиотека с отворен достъп на фирмата Aetilius Inc. (pHash е регистрирана марка от Evan Klinger) [7].

В разглежданата фамилия от хеш функции спадат **Average Hash (aHash)**, **Difference Hash (dHash)** и **Perceive Hash (pHash)**. **aHash** генерира отпечатък на изображението спрямо средната светлота на изображението. **dHash** следва подобен алгоритъм, но генерира стойността на поредния пиксел в зависимост от това дали е по-ярък от левия си съсед, т.е. той стъпва на оценка на градиентите. **pHash** е на пръв поглед най-финият метод, но за сметка на това и най-бавният. Той се основава на оценка на честотните модели чрез прилагане на дискретна косинусова трансформация на матрицата.

И в трите случая първоначално се извършва редуциране на размера и цвета на изображенията.

**Редуцирането на размера** до определени ширина  $l$  и височина  $h$  (в повечето случаи равни) е най-бързият начин за премахване на детайлизацията на изображенията, игнориране на оригиналните размери и съотношение. Всички резултатни изображения стават с едни и същи  $l \times h$  пиксела.

Следващата стъпка е **редуцирането на цвета**. При нея полученото изображение се преобразува в сивата гама (grayscale). За тази цел използваме по-финото преобразуване чрез формулата, по която се намира luminosity в цветовия модел YCbCr [8]:  $Y = 0.299 * r + 0.587 * g + 0.114 * b$

От получените еднакви по размер обезцветени изображения се изчисляват съответните дескриптори по следния начин:

- Average Hash: За всяко от изображенията се намира средната стойност на сивото в изображението. Дескрипторът aHash представлява поредица от нули и единици с дължина  $l \times h$ , определящи дали поредният пиксел от линейното разгъване на сивото изображение е по-светъл (стойност „1“) или по-тъмен (стойност „0“) от изчислената средна стойност на сивото за даденото изображение (на фиг.1 са визуализирани стъпките, през които преминава изображението до получаване на дескриптора aHash);



Сравнението става на базата на оценка на Хеминговото разстояние между двата дескриптора:

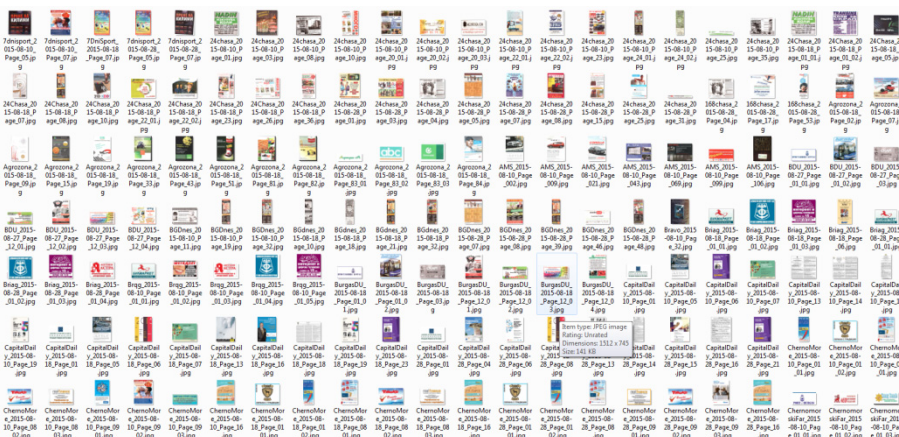
$$d(p, q) = \sum_{i=1}^{l*h} (p_i - q_i)$$

Практическата задача предполага автоматизирана обработка на изображенията поради честата поява на изображения, които са нови реклами и от една страна следва да бъдат етикетирани от оператора, а от друга възниква необходимостта той лично да прегледа верността на класифицираните екземпляри. Затова в този случай, алгоритъм, който бързо изчислява избрания дескриптор, а след това позволява достатъчно силно стесняване на потенциалните кандидати за съвпадение с постъпващите изображения, е достатъчно приемлив. По тази причина по-нататъшните експерименти ще проведем с първите два дескриптора – aHash и dHash.

## 2. Резултати от експеримента

### 2.1. Тестово множество

Тестовото множество се състои от 900 изображения, които представляват изображенията на потенциални реклами, изрязани от около 40 издания от три последователни дни (фиг.2). Имената на самите файлове носят информация за изданието, деня и страницата, където са били, както и пореден номер (в случай че има няколко реклами от една и съща страница). Част от изображенията са текстови съобщения, съдържащи информации за обявяване на търгове и други подобни. Те следва да се класифицират отделно, но също подлежат на обработка.

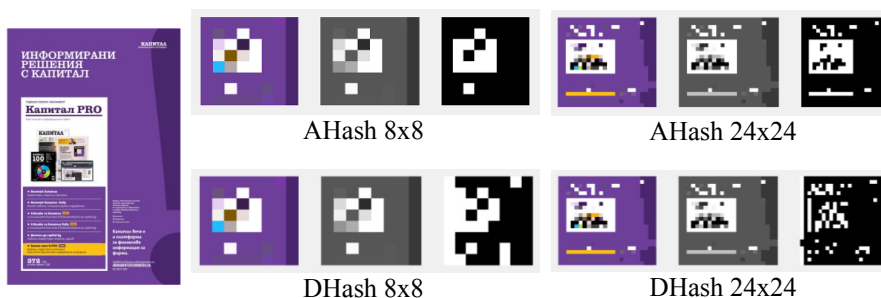


Фигура 2. Отрязък от изображенията от тестовото множество

## 2.2. Инструментариум

За целта беше изградена експериментална система, която позволява изчисляването на съответния дескриптор с различни размери на трансформираното изображение (задават се параметрично), изчисляване на разстоянията между изображенията и на тази база последващ анализ на процентите на различие между изображенията, които да се вземат като ограниченител при показването на евентуалните кандидати.

Фигура 3 показва резултати от визуализацията на двата дескриптора при различни размери за една от рекламите, съдържащи се в тестовото множество.



Фигура 3. Визуализация на двата дескриптора при размери 8x8 и 24x24  
(бяло = „1“, черно = „0“)

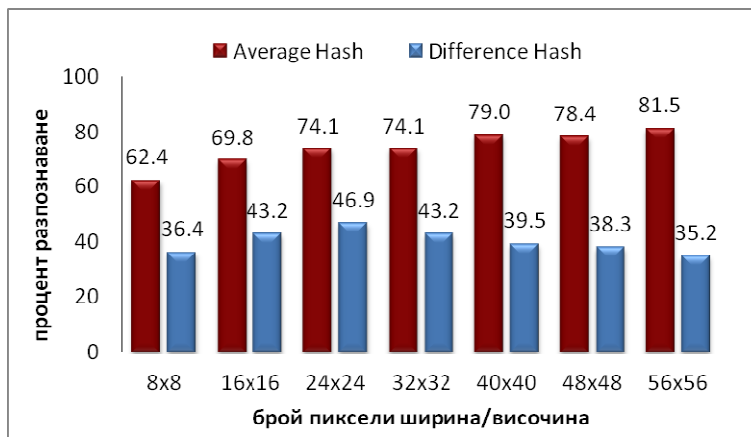
Проведените експерименти за оценка на точността при класификация бяха направени с използването на класификатора LB1 от Weka [9], който се базира на оценка на различията между отделните атрибути, което в случая съответства на Хеминговото разстояние между дескрипторите.

## 2.3. Експерименти и анализ

Първата група експерименти бяха проведени с цел изследване на точността на разпознаване при използване на aHash и dHash за различни размери.

В експеримента участваха общо 260 етикетирани изображения, като 98 от тях бяха в обучаващата извадка, а останалите 162 – в тестовата извадка (практическата задача произвежда голямо количество класове с немного на брой представители).

Резултатите показаха, че за разлика от твърдението в [6], че dHash дава по-добри резултати от aHash, в конкретния случай със сканираните изображения dHash показва лоши резултати – под 50% разпознаваемост, за разлика от aHash, който даже и при малките размерности дава добра разпознаваемост на класовете (фиг. 4).



Фигура 4. Резултати от разпознаването при прилагане на aHash и dHash

Във втората група експерименти бяха включени допълнителни 50 текстови съобщения (обяви, които също се третират като рекламни съобщения) и още 590 неетикетирани изображения (нови реклами).

При класифицирането с aHash (в случая беше използвана 10-стъпкова крос валидация при размер 40x40) коректно бяха класифицирани 73.44 %, което за целите на намаляване на извадката за операторски избор е достатъчно добър резултат.

Допълнителен анализ (със собствени софтуерни средства) показва, че при неразпознатите класове изображението с верен клас се намира в първите шест най-близки класа в 43.2 % от случаите (границата б в случая е избрана предвид психологическата възможност за възприятие на човека до седем обекта).

### 3. Програмен проект и организация на базата от данни

Анализът на резултатите от проведените по-горе експерименти доведе до избор на използването на aHash дескриптора като оптимален спрямо получената точност и времето за изчисление. Размерността, до която да се свиват изображенията, беше оставена като параметър, който операторът да избира с цел последващо натрупване на информация относно качеството на класифициране на изображенията. Друг такъв параметър е броят етикетирани изображения, близки до постъпващото ново изображение, от което операторът да потвърждава класа му или да поставя нов етикет.

Самият алгоритъм за изчисляване на aHash дескрипторите на постъпващите изображения беше оптимизиран посредством прилагане на многонишкова обработка на изображенията, чрез която значително се повишава скоростта на работа. Например, при проведен тест на компютър с процесор с 8 логически ядра (4 физически, всяко от тях с по 2 нишки) времето за обработка намаля около 3 пъти.

Съхраняването на данните за изображенията съдържа както информация за източника на изображението (издание, дата, страница, място на съхранение на отрязъка), така и информация за това как е класифицирано изображението. От гледна точка на разпознаването на класа има два основни типа изображения (данни за тях). Единият тип са т. нар. *първообрази* – изображението се появява за пръв път в базата и тогава операторът ръчно му задава клас, към който принадлежи. Като първообраз може да попадне и изображение, каквото вече е имало в базата, но разпознаването му е било твърде лошо и не е било предложено сред най-близките класове. От гледна точка на задачата това няма значение, освен че операторът трябва да го етикетира още веднъж. Вторият тип са т. нар. *повторения* – системата е предложила първообразите на най-близките класове спрямо дескриптора на постъпващото изображение, от които операторът е посочил първообраза, към чиито клас принадлежи постъпващото изображение, или потвърдил избора на системата като първи най-близък клас.

При организацията на базата се предложи оригинален подход, при който данните за първообразите и последващите повторения се съхраняват в обща база, като първообразите съхраняват векторите, с които се сравняват новите изображения, и таговете, въведени от оператора при първото срещане, а повторенията само сочат към идентификаторите на съответните първообрази. За целта се използва ArM32 като многомерен метод, позволяващ компактното съхраняване на разнородна информация в обща база.

## **Заклучение**

Проведените експерименти показаха, че прилагането на разглеждания дескриптор aHash успешно може да се приложи и в разглеждания случай на разпознаване на сканирани от различни източници изображения. Експериментална софтуерна система, която да обедини отделните стъпки и да позволи последващ анализ на резултатите, е вече поектирана и в начален етап на разработка. Задачи, които екипът си поставя са бъдещо разрешаване, са:

- намиране на подходящ индексен механизъм с цел оптимизиране на търсенето по близост;
- изследване на възможностите за допълнително ускорение чрез търсене първо в същите издания, клъстера от издания и накрая в останалите;
- оценка на времевия интервал, в който средно се задържа дадена реклама, с цел оптимизиране на търсенето.

## **Литература**

1. Медиа анализ, <http://mediazoom.bg/>
2. TinEye – MatchEngine, <http://www.tineye.com/>
3. Miller M.L., I.J. Cox, J.-P.M.G. Linnartz, T. Kalker: A Review of Watermarking Principles and Practices. Marcell Dekker Inc., New York (1999) pp. 461–485 (Chapter 18)
4. Wua D., X. Zhou, X. Niuc: A novel image hash algorithm resistant to print–scan. Signal Processing J. Special Section: Visual Information Analysis for Security. Volume 89, Issue 12, Dec. 2009, pp.2415-2424.

5. Krawetz N., <http://www.hackerfactor.com/blog/?archives/432-Looks-Like-It.html>
6. Krawetz N., <http://www.hackerfactor.com/blog/?archives/529-Kind-of-Like-That.html>
7. pHash – The open source perceptual hash library, <http://www.phash.org/>
8. JPEG File Interchange Format, Version 1.02, <http://www.w3.org/Graphics/JPEG/jfif3.pdf>
9. Weka Data Mining Software, <http://www.cs.waikato.ac.nz/ml/weka/index.html>

## **Adverts, adverts,... – finding without staring**

**Miroslav Ivanov, Krassimira Ivanova,  
Iliya Mitov, Evgenia Velikova**

**Abstract:** The goal of this research is to examine the opportunities to facilitate the discovery of equal advertisements extracted from the press, which is one of the goals of media monitoring. The possibility to find duplicates quickly using the descriptor Average Hash (aHash) is analysed. Original approach for organization of the metadata keeping labeled specimens and found repetitions in a common database is proposed also.



## **Индекс на авторите**

Весела	Ангелова	21, 185
Андрей	Антонов	95
Владимир	Балчев	165
Петър	Бойваленков	64
Евгения	Великова	239
Георги	Врагов	165
Младен	Георгиев	215
Александър	Геров	30
Людмила	Димитрова	39
Аврам	Ескенази	11, 21
Мирослав	Иванов	239
Станислав	Иванов	143
Красимира	Иванова	68, 143, 173, 239
Малинка	Иванова	111
Александър	Илиев	225
Калинка	Калоянова	95
Николай	Киров	87
Нели	Манева	21, 205
Димитър	Минев	165
Златогор	Минчев	102
Илия	Митов	173, 239
Мария	Нишева	30
Юлиана	Пенева	143
Желязко	Петров	121
Разван	Рагазан	121
Румен	Радев	21
Анани	Ризов	173
Анна	Самева	58
Евгения	Сендова	129
Калина	Сотирова	58, 153
Маргарита	Спиридонова	30
Христо	Стайков	129
Петър	Станчев	11, 143
Теодоси	Теодосиев	83
Борис	Шишков	194

**Национална конференция по информатика, посветена на 80-годишнината  
от рождението на професор Петър Бърнев – сборник доклади**

Редактори: Красимира Иванова и Нели Манева

Технически редактор: Анна Салева

Издател: Институт по математика и информатика, БАН

Със съдействието на: Асоциация за развитие на информационното общество

Печат: „Ракурс“ ООД, гр. Пловдив

формат: ISO B5

ISBN 978-954-8986-45-8

© 2016