

**ПРИЛОЖЕНИЕ 2. Резюмета на публикации на доц. д-р Геннадий Агре, представени на конкурс за академичната длъжност „професор” по проф. направление 4.6 Информатика и компютърни науки, специалност „Информатика (Изкуствен интелект)”, обявен в ДВ бр. 41/21.05.2019**

**Семантични интернет услуги**

- 1. Agre, G. INFRAWEBs designer - A graphical tool for designing semantic web services. Lecture Notes in Computer Science 4183 LNCS, 2006, pp. 275-289 (SJR = 0.317)**

За да станат нови информационни технологии, базиращи се на семантични интернет услуги, достъпни за крайните потребители, нивото на инструментите, поддържащи такива технологии, трябва да бъде значително подобро. Настоящата статия представя архитектурата, основните принципи и реализационните детайли на един такъв инструмент – INFRAWEBs Designer, разработен в рамките на Европейски проект по 6-та Рамкова програма INFRAWEBs. INFRAWEBs Designer е графичен, основан на онтологии, инструмент за създаване на описания на семантични интернет услуги в съответствие с WSMO методологията. Инструментът е предназначен за крайните потребители – доставчици на семантични интернет услуги, желаещи да превърнат разработените от тях интернет услуги във WSMO ориентирани семантични интернет услуги. Отличителните характеристики на инструмента са интензивното използване на онтологии, автоматичното генериране на логическите описания на семантични услуги от техните графични модели и използване на разсъждения, базирани на прецеденти, за намиране на описания на сходни услуги, които могат да бъдат преизползвани като шаблони при конструиране на нови услуги. Всичко това позволява крайният потребител да създава коректни описания на основани на WSMO семантични интернет услуги без да знае формален логически език за тяхното описание – WSML.

- 2. Agre, G., Marinova, Z., Pariente, T., Micsik, A. Towards semantic Web service engineering. CEUR Workshop Proceedings 243, 2007, pp. 91-105 (SJR = 0.157).**

Статията представя основни резултати от Европейски изследователски проект INFRAWEBs, финансиран по 6-та Рамкова програма. В рамките на проекта е разработен един лесен и ефективен начин за създаване и използване на семантичните описания на съществуващи и нови семантични интернет услуги. Проектът използва WSMO методология за онтологично описание и езика WSML за моделиране на интернет услуги като не налага никакви допълнителни изисквания към тях. По тази причина разработените по проекта нови, високотехнологични софтуерни компоненти представляват интерес за цялата научна общност, използваща WSMO. Проектът предлага интегрираната INFRAWEBs среда - една обща среда, базирана върху ориентираната на услуги архитектура.. Тя се основава върху ESB (Enterprise Service Bus) парадигма за интеграция и може лесно да се използва от различни групи потребители (доставчици на приложения, дизайнери на семантични интернет услуги и др.), като позволява интегриране на компоненти, създадени с различни технологии. Всичко това дава основание интегрираната INFRAWEBs среда да се разглежда като една от първите среди за разработка на семантични интернет услуги, която покрива целия цикъл на живота на такива услуги и позволява създаване на сложни, семантично подпомогнати приложения.

- 3. Agre, G., Dilov, I. How to create a WSMO-based semantic service without knowing WSML. Lecture Notes in Computer Science 4832 LNCS, 2007, pp. 217-235 (SJR = 0.293)**

INFRAWEBBS Designer е графична, базирана на онтологии интегрирана среда за конструиране на описания на семантични интернет услуги и цели съгласно WSMO методологията. Средата е предназначена за доставчици на семантични услуги и приложения, основани на такива услуги, и не изисква от тях никакви предварителни знания на езика за описание на услугите – WSML. Основната цел на настоящата статия е да представи, стъпка по стъпка, процеса на създаване на една семантична интернет услуга с помощта на INFRAWEBBS Designer. Описанието на семантичната услуга се създава по графичен начин от вече съществуващото WSDL описание на интернет услугата и набор от WSML онтологии. За да се улесни този процес, се използва графичното описание на сходни семантични интернет услуги, които се намират автоматично от средата и се използват като шаблони. На края са представени няколко извода за приложимостта на INFRAWEBBS Designer, базирани върху нашия опит по създаване на две пилотни приложения – първото използва сценария на агенцията за пътувания, а второто – сценария на електронното правителство.

- 4. Agre, G; Lobo, TP; Marinova, Z; Nern, J., Micsik, A., Boyanov, A., Atanasova, T., Scicluna, J., Lopez-Cobo, JM , Tzafestas, E. Chapter 12: INFRAWEBBS - A Framework for Semantic Service Engineering. In: DiNitto, E; Sassen, AM; Traverso, P; Zwegers, A (Eds.). AT YOUR SERVICE: SERVICE-ORIENTED COMPUTING FROM AN EU PERSPECTIVE, 2009, 299-325, M I T PRESS, USA (WoS)**

Главата представя основните резултати на Европейски изследователски проект INFRAWEBBS, който се фокусира върху разработка на среда за семантичното инженерство на интернет услуги. Средата позволява създаване, поддръжка и изпълнение на WSMO-базирани семантични интернет услуги, както и поддържа приложения на такива услуги през целия им жизнен цикъл. Разработената в строго съответствие с WSMO методологията, средата осъществява контрол върху сложния процес по създаване на семантичните описания на услугите чрез идентифициране на различните типове потребители на технологии на семантични интернет услуги, уточняване на различни фази на процеса по конструиране на семантични интернет услуги, както и чрез предоставяне на един набор от инструменти, пригодени за идентифицираните типове потребители.

Валидността на разработения подход е проверена върху две пилотни приложения с високата степен на сложност. Първото е системата STREAM Flows!, в която потребителят може да създава и да преизползва пакети за пътувания. Второто приложение се базира на сценария от областта на електронното правителство. То илюстрира работата на семантични интернет услуги при осъществяване на взаимодействия между публична администрация, граждани и фирми.

Наборът от софтуерни модули и инструменти, разработени по проекта, предоставя иновативните решения за създаване на описания, откриване и динамично композиране на семантични интернет услуги. По този начин, процесът по създаване на приложения на семантични интернет услуги, който обикновено е доста сложен и изисква експертните знания както на WSMO модела, така и на логическия език WSML, се превръща в една задача, достъпна за решаване от обикновени разработчици на интернет услуги. Улесняване на задачата по създаване на WSML обекти е постижението на проекта, което има потенциал за по-лесно адаптиране на технологиите

за семантични интернет услуги в глобален мащаб. Като цяло, разработената среда INFRAWEBBS може да се разглежда като първата среда за инженерството върху семантични интернет услуги, която покрива целия им жизнен цикъл и позволява създаване на сложни приложения, базирани на семантични интернет услуги.

- 5. G. Agre. Chapter 5: Engineering Semantic Web Services. In: Gabriel Fung (Ed.). Engineering Semantic Web Services. 2010, iConcept Press, USA, pp. 75-98, ISBN: 978-0-9807330-1-3 (Hard Cover) / 978-1-4536364-0-4 (Paperback)**

Изследванията в областта на семантични интернет услуги (СИУ) са свързани с автоматизацията на процеса по създаване, поддръжка и изпълнение на базирани на интернет услуги приложения чрез прилагане на семантичните технологии. Чрез предоставяне на формалните описания с добре дефинирана семантика, СИУ представляват една нова стъпка в посока решаване на такива фундаментални проблеми, свързани с конструиране на услуги, като взаимодействие между услуги, откриване, хореография и оркестрация на услуги. Европейският изследователски проект INFRAWEBBS, който е успешно завършен през 2007 г., предлага една технологично-ориентирана стъпка за решаване на някои от указаните по-горе проблеми. Той се фокусира върху разработка на среда за семантичното инженерство на интернет услуги, позволяваща създаване, поддръжка и изпълнение на WSMO-базирани СИУ, както и да поддържа приложения, използващи такива услуги, през целия им жизнен цикъл. Настоящата глава се акцентира основно върху две „горещи“ задачи от областта на СИУ инженерството – създаване на описания и динамично композиране на СИУ. Описанията се създават чрез INFRAWEBBS Designer - графична, базирана на онтологии интегрирана среда за конструиране на описания на WSMO-базирани СИУ и цели, която е предназначена за доставчици на семантични услуги и приложения и не изисква от тях никакви предварителни знания на езика WSML за описание на СИУ. Динамичното композиране на СИУ се базира върху оригиналния, воден от данни подход, в който процесът за намиране на подходяща комбинация от услуги е воден от изпълняваната в режим на реалното време декомпозиция на зададената цел върху под-цели и откриване на съществуващи услуги, способни да удовлетворят тези под-цели. Съвместимостта на услуги, участващи в композицията, се постига чрез използване на съвместимото описание на шаблона, представящ сложна съставна цел, подготвена офлайн от разработчика на приложения. Използвайки неявно предоставена информация за желаня ред на изпълнение на услугите в композицията, предложеният алгоритъм е способен да намери подходяща оркестрация на услугите в композицията, както и да открие подходящите замени, ако някои от услугите в композицията не могат да бъдат използвани по различни физически причини.

#### **Технологично поддържано обучение**

- 6. Dochev, D., Agre, G. Towards semantic web enhanced learning. KMIS 2009 - 1st International Conference on Knowledge Management and Information Sharing, Proceedings, 2009, pp. 212-217 (Scopus)**

Качеството на преподавания материал винаги се разглеждало като краеъгълен камък във всички ситуации, възникващи както при класическото, така и при електронното обучение. По тази причина подготовката на учебния материал е една от най-важните и трудоемки дейности в практиката на съвременното технологично поддържано обучение

(ТПО). В настоящата статия ние предлагаме подход за разработка на нова семантична среда с архитектура, базирана на услуги, която е ориентирана към ТПО приложения. Подходът се базира на анализа и използването на предимствата на технологиите на семантични интернет услуги за автоматизацията на процеси по откриване, избор и композиция на учебни обекти в една разпределена, ориентирана на услуги среда, интегрирана чрез онтологии. Подходът отразява текущите изследвания, извършвани в рамките на българския изследователски проект „СИНУС: Семантични технологии за интернет услуги и технологично поддържано обучение”.

- 7. Dochev, D., Agre, G., Pavlov, R. An approach to learning-by doing through user creation of learning content. EUROMEDIA 2011 - 16th Annual Scientific Conference on Web Technology, New Media Communications and Telematics Theory Methods, Tools and Applications, 2011, pp. 9-12 (Scopus)**

Статията изследва организацията на дейности по „обучение чрез правене” в контекста на някои специфики на процеса на обучение в хуманитарните области на знания, които се реализират чрез подготовка на аналитични материали от самите обучаеми. Семантичните технологии се прилагат за подпомагане на действията на обучаемия при създаване на специализирани колекции от мултимедийни обекти с ограничен обем – първо за намиране на обектите, отговарящи на предварително зададени цели, а след това - за сравнение на избраните обекти с цел извършване на анализа. Обсъжда се разработваната от авторите среда за технологично поддържано обучение с приложение в една конкретна проблемна област – българската иконография, която се използва за обучение в такива дисциплини като иконография, изкуство, история, теология и др. Статията съдържа и пример, представящ структурираното представяне на една задача за обучение, както и нейната формализация във вид на заявки към средата, целящи подпомагане създаването на колекция от обекти, удовлетворяващи заявката, както и за оценка на адекватността на избраните от обучаемия обекти, съставляващи колекцията.

- 8. Dochev, D., Agre, G., Pavlov, R. User authoring in learning-by-doing situations. ACM International Conference Proceeding Series 578, 2011, pp. 577-582 (SJR = 0.184)**

В статията се разглеждат дейности от типа „обучение чрез правене”, изпълнявани от обучаемия при разработка на аналитични материали в добре определени сценарии от процеса на обучение. Изпълнението на тези дейности се подпомага от технологиите за семантични интернет услуги, които се използват за достъп и филтриране на информационни обекти, необходими за извършване на самостоятелен анализ, както и за оценяване (в определена степен) на избраните материали. В качеството на пример се обсъжда експерименталното приложение на една среда за подпомогнато от технологии обучение, разработена в рамките на текущия български изследователски проект „СИНУС: Семантични технологии за интернет услуги и технологично подпомогнато обучение”. Разглежда се задачата за създаване на колекция от български икони, удовлетворяваща определени, предварително зададени изисквания.

- 9. Hristov, I., Agre, G., Dochev, D. Measuring the learning progress in a "learning by authoring" semantic web services based ecosystem. CSEDU 2011 - Proceedings of the 3rd International Conference on Computer Supported Education 2. 2011. pp. 413-418 (Scopus)**

Статията представя един подход за подпомагане на процеса на оценяване на някои специфични дейности от типа на „обучение чрез правене”, които се свеждат до написване на аналитични есета/проекти върху ограничени по размер тематични колекции от мултимедийни обекти, създадени от самия обучаем в съответствие с предварително дефинирани учебни задачи. Поддръжката на обучаемия се прави чрез непрекъснатото оценяване на три типа учебни дейности. Първият тип са дейностите, свързани със създаване на колекции, изпълнявани чрез сравнение на резултатите от заявките, изпращани от обучаемия, с формализираното представяне на учебната задача, подготвено от учителя. Вторият тип дейностите е оценка на качеството на анализа – тя се прави от човека-оценител, който е подпомогнат от системата чрез предоставяне на информация за броя на понятията от предметната област, използвани и изпуснати от обучаемия, както и за процента на използваните понятия, съвпадащи с тези, използвани от учителя. Третият тип дейностите е оценяване на визуалния характер на колекцията, което, в нашия случай, не е предмет на обучението и се изпълнява от други участници в проекта.

**10. Dochev, D., Agre, G. Supporting learning-by-doing situations by semantic technologies. EUROMEDIA 2012 - 17th Annual Scientific Conference on Web Technology, New Media Communications and Telematics Theory Methods, Tools and Applications, 2012, pp. 49-53 (Scopus)**

В настоящата статия се разглежда един подход към реализация на дейности от типа на „учене чрез правене”, които се състоят в създаване на аналитични материали в определени учебни ситуации. Представени са основни характеристики на експерименталната обучаваща среда, базирана на технологията за семантични интернет услуги, която реализира този подход. Кратко е обсъдено как средата може да води и да подпомага обучаваните при решаване на поставени учебни задачи при предположения за неточност и непълнота на притежаваните от обучаваните знанията за предметната област. В качеството на илюстративен пример е описано едно пилотно приложение на средата в конкретната предметна област – българската иконография, както и реализацията на предложения подход за изпълнение на две конкретни учебни задачи – създаването на тематична мултимедийна колекция и нейния аналитичен анализ.

**11. Staykova, K., Agre, G. Use of Ontology-to-Text Relation for creating semantic annotation. ACM International Conference Proceeding Series, 2012, pp. 64-71 (SJR = 0.181)**

Основният акцент в настоящата статия е върху задачата за създаване на семантични анотации на мултимедийни обекти, съхранявани в една дигитална библиотека. В повечето от случаите създаването на такива анотации е много трудоемък и отнемаш значително време процес, когато той се изпълнява ръчно от експерти в дадена проблемна област. Този тромав процес може да бъде значително облекчен, ако някоя част от информацията, използвана за аотиране, е налична във вид на естествени езикови текстове. В тези случаи е възможно към тях да бъдат приложени определени техники от областта на обработка на естествен език (ОЕЕ), за да бъде извлечена от тях необходима за аотиране информация, която след това да бъде превърната в семантични анотации. Настоящата статия описва един подход за решаване на задача по разпознаване на онтологични термини в текстове на български език чрез използване на една специфична ОЕЕ техника, наречена „релация онтология-към-текст” (Ontology-to-

Text relation). Дискутирани са всички предварителните стъпки, за да бъдат подготвени всички необходими ресурси за реализация на тази техника.

**12. Agre, G., Dochev, D., Slavkova, L. Technology Enhanced Learning for humanities by active learning - The SINUS project approach. Cybernetics and Information Technologies. 2012, 12(4), pp. 25-42 (SJR = 0.111)**

Настоящата статия анализира възможностите на технологично поддържано обучение (ТПО) за подпомагане на някои специфични дейности, възникващи при активно обучение в някои хуманитарни области на знанието – написване на аналитични материали чрез активно използване на мултимедийното съдържание на дигитални библиотеки (ДБ). За да реализира това подпомагане, са били идентифицирани и формализирани типове и съдържание на основни знания от проблемната област. Предлаганият подход за моделиране на тези знания се състои в комбиниране на базовата онтология, описваща само основните признаци, които явно или неявно са заложени в структурата на използваните ДБ, с допълнителни специализирани онтологии. Подходът позволява да се запази непокътнати съдържанието, анотациите и методът за достъп към използваните библиотеки и в същото време да се добави възможността за по-богат семантичен достъп към съхраняваните в ДБ информационни обекти чрез използване на допълнителни описателни (онтологични) признаци, дефинирани в допълнителните специализирани онтологии. Фокусът на анализа е върху дейностите на обучаемия, изпълнявани по време на първата фаза на дефинираната задача за обучение – създаване на тематични колекции от мултимедийни обекти, съхраняване в ДБ. Представени са модели на правилните решения и възможните грешки на обучаемия, които са използвани за създаване на алгоритъма, който проверява коректността и пълнотата на предложеното от обучавания решение, независимо от това, по какъв начин то е било получено. Тези конструктивни решения формират функционалността на експерименталната ТПО среда, която е реализирана чрез използване на най-съвременните семантични технологии.

**13. Agre, G. SINUS - A Semantic Technology Enhanced Environment for Learning in Humanities. Cybernetics and Information Technologies. 2012, 12(4), pp. 5-24 (SJR = 0.111)**

Статията представя програмната среда СИНУС – една базирана на семантични технологии среда за разработка на приложения за технологично поддържано обучение (ТПО) в хуманитарните области на знанието. Средата се състои от три слоя – слой за съхранение, съдържащ хетерогенни хранилища за съхраняване на знания за предметната област и педагогическите знания, слой на инструментите, съдържащ набор от инструменти за обработка на различни видове знания, и междинен слой, реализиран като разширена машина за търсене, изпълняваща всички необходими комуникации между инструментите и хранилищата. Средата има голяма степен на хетерогенност – в момента тя включва три вида хранилищата: една релационна база от данни, съхраняваща модели на учебните задачи, една SQL-базирана дигитална библиотека за съхраняване на мултимедийни обекти и техните анотации, реализирана като WSDL интернет услуга, и едно RDF-базирано семантично хранилище, съхраняващо OWL онтологии и семантични анотации на обекти, описани в дигиталната библиотека. Компонентите на средата са написани на различни програмни езици - C# (Разширена машина за търсене) и Java (Редактор за семантични анотации и Редактор на учебни задачи) и са реализирани като WSDL или RESTful интернет услуги. В статията са

обсъдени и някои реализационни детайли, свързани с успешна разработка на две ТПО приложения – разширената дигитална библиотека с архитектура, ориентирана на услуги, и системата за обучение в областта на Източноправославната иконография. Представена е и една първоначална оценка на разработената среда, която доказва правилността на избрания подход

**14. Dichev, C., Dicheva, D., Agre, G., Angelova, G. Current practices, trends and challenges in k-12 online learning. *Cybernetics and Information Technologies*. 2013, 13(3), pp. 91-110 (SJR = 0.173)**

Онлайн обучението, един общ систематичен подход за образование, който използва интернет, е една от най-бързо развиващите се тенденции в технологично поддържаното обучение. Технологията в комбинация с предписания, която е насочена към когнитивни и социални процеси по създаване на знания, може да предложи по-разнообразни и по-ефективни възможности за онлайн обучение, отколкото традиционното „лице в лице” обучение. В този обзор ние опитахме да обобщим разнообразни форми и практики на основното и средното (K 12) образование във вида на онлайн или смесено обучение, както те се прилагат в различни регионални, национални и културни контексти. Статията започва с въвеждане на някои основни понятия и терминология, обобщава състоянието на K12 онлайн обучение по света и завършва с обсъждане на забелязаните тенденции и предизвикателства в съществуващите практики на K12 онлайн обучението.

**15. Dichev, C., Dicheva, D., Angelova, G., Agre, G. From gamification to gameful design and gameful experience in learning. *Cybernetics and Information Technologies*. 2014, 14(4), pp. 80-100 (SJR = 0.138)**

Обучението представлява една целенасочена социална дейност, която се определя от емоционални фактори. За да може обучението да бъде ефективно геймифицирано (т.е. да включва игрови елементи) с цел да повиши мотивацията и ангажираността на обучаемите, преподавателите трябва да разбират както специфичните аспекти, свързани с игрите, така и мотивационната психология и педагогика. Това ще им помогне за идентификация на факторите, които определят и обясняват желаното поведение на обучаемите. Настоящата статия представлява един обзор на основните подходи, използвани при геймификацията, както и на възникващите нови направления в контекста на релевантна мотивационна психология и педагогика. Фокусът е върху мотивационните фактори, които влияят върху обучението и разбиране на промените в поведението. Целта на статията е да направи анализ на релевантните работи, свързани с геймификацията заедно с отбелязване на нововъзникващите тенденции, както и да предостави основи за оценяване и идентифициране на области за възможни подобрения.

**16. Dicheva, D., Dichev, C., Agre, G., Angelova, G. Gamification in education: A systematic mapping study. *Educational Technology and Society*, 2015, 18(3), pp. 75-88 (IF – Q2)**

Докато геймификацията с успех се прилага в такива области като бизнес, маркетинг и корпоративно управление, нейното приложение в областта на обучението все още се разглежда като една нововъзникваща тенденция. Настоящата статия представя едно проучване на публикуваните емпирични изследвания на приложенията на геймификация в образованието. Проучването е направено само върху статиите, които явно дискутират ефектите от използване на игровите елементи в специфични

образователни контексти. Тя прилага подхода за систематичното отобразяване. Предложена е една структура от категории, направена от анализа на темите, дискутирани в прегледаните статии, която се прилага за класифициране на резултати от изследванията. Тези категории включват принципи на дизайн на геймификацията, механика на игрите, контекст на прилагането на геймификацията (тип на приложение, образователното ниво и предмет на обучението), реализация и оценяване. Чрез отобразяване на публикуваните работи върху класификационните критерии и техния анализ, проучването изяснява основните направления на изпълняваните в момента емпирични изследвания по прилагане на геймификацията към образованието. То също така показва някои основни пречки и нужди, като, например, необходимостта от подходяща технологична поддръжка, от контролирани надеждни изследвания, убедително демонстриращи положителни или отрицателни резултати от използването на специфични игрови елементи в конкретни образователни контексти, и т.н. Макар че повечето от изследваните статии докладват обещаващи резултати, необходимо е наличието на повече и на по-задълбочени емпирични изследвания, за да се определи, дали вътрешни или външни мотивации на обучаемите могат да бъдат повлияни от геймификацията,

**17. Dichev, C., Dicheva, D., Agre, G., Angelova, G. Trends and opportunities in computer science OER development. Cybernetics and Information Technologies, 2015, 15(3), pp. 114-126 (SJR = 0.158)**

В момента светът прегръща модела на отвореното образование. Успехът на този процес предполага наличието на адекватно запознанство с този модел – едно предположение, което не намира потвърждение в съвременните отчети и статистически факти. Не зависимо от големия напредък в последните години, отворените учебни ресурси (ОУР) все още не са в центъра на процеса по създаване на курсове в областта на компютърните науки. Мотивирани от необходимостта да бъде запълнена тази празнота, настоящата статия анализира еволюцията на създаването на ОУР и на набелязаните тенденции, релевантни за обучение в областта на компютърните науки. Целта е да бъде повишено нивото на информираността и да бъде подпомогнат един практически преход към подходящи модели, които са еднакво интересни за всички участници в процеса на обучение по компютърни науки.

**18. Dobрева, M., Angelova, G., Agre, G. Bridging the gap between digital libraries and e-learning. Cybernetics and Information Technologies, 2015, 15(4), pp. 92-110 (SJR = 0.158)**

Дигиталните библиотеки предлагат достъп до огромните обеми дигитално съдържание, което на практика е свързано с всички области на човешкото знание. Това ги прави много подходящи за усъвършенстване на процесите по преподаване и обучение. Настоящата статия представя един обзор на съществуващите подходи за премахване на съществуващата пропаст между потенциалната полезност на дигиталните библиотеки за целите на обучение и съвременните системи за електронно обучение. Основните изводи от обзора са, че от най-голямата важност и спешност е да бъде постигнато ясно разбиране на важността от използването на качествено съдържание при електронното обучение, както и че е необходимо създаване на по-голямо количество отворени учебни ресурси. Предлагането на по-голямо количество отворени учебни ресурси ще представлява една важна стратегическа стъпка за засилване на ангажираността на



дигиталните библиотеки в процеса на обучение и ще е от полза както за преподаватели, така и за обучаеми.

### **Машинно самообучение и извличане на закономерности от данни**

- 19. Hristov, V., Agre, G. A software system for classification of archaeological artefacts represented by 2D plans. *Cybernetics and Information Technologies*, 2013, 13(2), pp. 82-96 (SJR = 0.172)**

Една важна част от археологическия анализ е задачата за класифициране на керамика на базата на формата на съдовете, тяхната украса, материал и т.н. Формата е едно от основните свойства на един съд, която се използва за неговата идентификация. Обаче, традиционните описания на формата се базират върху интуитивни, често неясни характеристики, които е сложно да бъдат оценени количествено, или използват няколко лесно измерими атрибути (от типа на височина, диаметър на гърлото на съда и т.н.), които описват формата на съда само частично. Настоящата статия представя предварителните резултати, свързани с текущ изследователски проект, целящ създаване на една програмната система, облекчаваща археологическия анализ на артефакти на базата на тяхното представяне чрез двумерни археологически чертежи. Акцентът е върху задачата за идентификация (класификация) на керамични съдове (или техните фрагменти), които принадлежат към определено дискретно множество от предварително дефинирани типове (класове). Входната информация за решаване на тази задача са сканирани рисунки на съдове (от книги или археологически отчети) и интернет базирана база данни, съдържаща предварително класифицирани примери на съдове от различни типове. Описан е метод за представяне на цели и фрагментирани съдове, както и процес по създаване на това представяне. Обсъдени са и резултати от началните експерименти, използващи разработеното представяне, за решаване на задача по идентификация на древногръцките амфори, намерени в затворени археологически комплекси от територия на черноморското крайбрежие.

- 20. Strandjev, B., Agre, G. On applicability of Principal Component Analysis to concept learning from images. 2013 IEEE International Symposium on Innovations in Intelligent Systems and Applications, INISTA 2013, IEEE Xplore, August, 2013, DOI: 10.1109/INISTA.2013.6577623 (Scopus)**

Настоящата статия представя набор от експерименти, целящи изследването на приложимостта на метода за анализ на основни компоненти за решаване на задачи по научаване на различни абстрактни понятия от изображения на лица. Резултатите показват, че в повечето от случаи прилаганата трансформация води до подобряване на класификационната точност на използваните алгоритми за научаване на понятия. Като едно допълнение, експериментите потвърдиха съществуването на възможна връзка между качеството на получените подобрения и сложността на понятията, подлежащи за научаване. Тази връзка може да се използва като една обективна мярка за оценяване на сложността на научаваните понятия.

- 21. Strandjev, B., Agre, G. On impact of PCA for solving classification tasks defined on facial images. *International Journal of Reasoning-based Intelligent Systems*. 2014, 6(3-4), pp. 85-92 (SJR = 0.104)**

В настоящата статия се дискутира въпросът за влияние на метода на анализа на основни компоненти (АОК) за решаване на задачи, свързани с класификация на различни абстрактни понятия, научавани от изображенията на човешки лица. Проведените експерименти показваха, че в повечето случаи, когато научаваното понятие има смисъл (т.е. не е дефинирано по изкуствен начин), прилагането на АОК трансформацията води до повишаване на класификационната точност на алгоритъма на най-близкия съсед, използван за разпознаване на такова понятие. Като допълнение, накратко е дискутиран един възможен подход за количествено измерване на сложността на понятие, подлежащо на научаване.

**22. Marinchev, I., Agre, G. A customised metric for foods categorization. ACM International Conference Proceeding Series 883, 2014, pp. 234-239 (SJR = 0.25)**

Настоящата статия представя една евристична мярка, пригодена за категоризация на хранителни продукти. Нейната разработка е част от текущия проект с Националния иновационен фонд за създаване на експертно-консултационната система за здравословно и диетично хранене. Мярката се базира върху използване на две функции – първата оценява разстоянието между продукт, който трябва да бъде категоризиран, и еталонни продукти – примери на различни хранителни категории, представени чрез продуктова онтология; втората функция оценява степента на доверие в намерената категория, изведена след прилагането на метода за най-близкия съсед. Всички хранителни продукти (еталонни и подлежащи на категоризирането) се описват чрез свои нутриенти. Този начин е много подходящ за описание на продукти, тъй като лежи в основата на съставяне на диети, различни фитнес програми и т.н. Проведените първоначални експерименти показваха, че използваната метрика води до добри резултати при категоризиране на продукта към по-общи категории (т.е. категории от по-горните нива на продуктовата онтология), както и при решаването на задача за търсене на подходяща замяна на дадения хранителен продукт.

**23. Marinchev, I., Agre, G. On speeding up the implementation of nearest neighbour search and classification. ACM International Conference Proceeding Series 1008, 2015, pp. 207-213 (SJR = 0.253)**

Настоящата статия представя някои практически техники за ускоряване на имплементации на алгоритъма за най-близкия съсед за случаи на бази данни с голям брой примери и/или примери с голяма размерност. За да бъде предварително ограничен броят кандидат класове при търсене на отговор на подадената заявка, се използва декомпозиционната матрицата, изчислена в резултат от прилагане на една итеративна форма на полярна ортогонализация. Показано е, че при наличието на голям брой обучаващи примери за всеки клас, допълнителното ускоряване може да бъде постигнато чрез разделянето на класовете на под-класове чрез прилагане на бързите версии на клъстерните алгоритми и използване на получените клъстери за построяване на декомпозиционната матрица. Предложената стъпка по предварителната обработка, която линейно или почти линейно зависи от броя на примерите и тяхната размерност, както и стъпката по предварителната селекция, зависеща линейно от броя на класовете, може да се използва с всеки от добре известните индексни методи (kd-trees, r-trees, metric trees и т.н.). Накрая ние предлагаме една нова индексна структура, наречена „клъстерен индекс”, която позволява прилагане на по-сложни индексни структури към по-големи бази данни.

**24. Marinchev, I., Agre, G. An expert system for healthful and dietary nutrition. ACM International Conference Proceeding Series 1164, 2016, pp. 229-236 (Scopus)**

Статията представя една експертно-консултационна система за диетично и здравословно хранене, разработена по проект, финансиран от Национален иновационен фонд. Основната цел на системата е да облекчи класификацията на хранителни продукти, продавани в магазините, като принадлежащи към една от предварително дефинирани категории на потребителски групи (класове) от типа на „подходящи или не за деца”, „подходящи или не за възрастни под 60 години с висока степен на физическа активност” и т.н. Експертните знания са представени във вид на онтология от стандартни категории на хранителни продукти, описани чрез свои нутриенти, и чрез система от продукционни правила, свързващи различни категории от продукти със съответните категории потребители. Процесът на класификацията се състои от две стъпки – на първата стъпка се осъществява идентификация на конкретен продукт чрез намиране в онтологията на стандартен продукт с достатъчна степен на доверие, който има най-голямото сходство с проверяемия продукт. На втората стъпка се прилагат продукционните правила, анализиращи нутриентния състав на продукта. Статията представя детайлното описание на архитектурата на системата, на процеса на класификация, на механизма за разпространение на доверието, на възможните потребителски роли, достъпни в системата, както и на нейната имплементация.

**25. Nikolova, I., Dicheva, D., Agre, G., Angelova, G., Dichev, C., Madzharov, D. Emerging applications of educational data mining in Bulgaria: The case of UCHA.SE. Studies in Computational Intelligence 648, 2016, pp. 113-131 (SJR = 0.219)**

Настоящата статия представя едно изследване на учебния интернет портал UCHA.SE, което цели подобряване на качеството на предоставяните от портала учебни услуги. Алгоритмите за машинно самообучение са използвани за извличане на учебни данни от записи за работата на потребители със системата, които след това са превърнати в обучаващи примери, използвани за оценка на портала и за идентификацията на недостатъци, които биха могли да бъдат оправени. Чрез използване на методи за манипулиране на обучаващи данни със силен дисбаланс на класове, е получено множество от примери, използвани за научаване на един достатъчно точен и разбираем за потребителя класификационен модел, базиран на правила. Моделът предсказва дали потребителят на портала ще поднови своя абонамент или не. Наученият модел беше анализиран, за да бъдат определени важните фактори, влияещи на решението на потребителя за продължаване на своя абонамент. Анализът показва, че такива фактори са периодът на първия абонамент, скоростта, с която потребителят губи интерес към портала и степента на завършване на упражнения след разглеждане на съответните видеоматериали. Основните рекомендации към разработчиците на портала са да разширят системата с повече възможности за научаване на предлагания учебен материал, които да включват по-разнообразни упражнения и тестове с различна степен на сложност, съчетани с персонализирано, водено от необходимост подпомагане на ученика при тяхното извършване. Предлаганите подобрения ще повишат степента на персонализация, рекомендация на ресурси, проследяване и визуализация на напредъка на обучаемия и на неговите постижения.

**26. Nikolova, I., Dicheva, D., Agre, G., Angelova, G., Dichev, C., Madzharov, D. Exploring the use of resources in the educational site ucha.SE. Lecture Notes in Computer Science 9883 LNAI, 2016, pp. 347-351 (SJR = 0.339)**

В настоящата статия се обсъждат някои резултати от текущия изследователски проект, целящ подобряване на качеството на учебните услуги, предоставяни от учебен интернет портал UCHA.SE. Използваните за оценяване данни са извлечени от 3000000 записа за достъп на потребителите на портала към 3897 учебни ресурса, отнасящи се за 65 учебни категории. Всяка категория се описва чрез набор от признаци, конструирани на базата на записи от системния журнал за взаимодействия на обучаемия с портала, съхранявани в базата данни на портала. Признаците включват 1) общия брой на видеоматериали в категорията, 2) среден брой на достъпи до категория за период от 3 месеца, 3) процент на обучаемите, осъществили достъп до конкретна категория, 4) брой забележки на преподаватели за конкретна категория. На база на тези признаци е извършен статистически анализ на учебните ресурси на портала, както и на интереса на обучаемите към всичките 65 предлагани учебни категории. Резултатите от анализа са използвани за формулиране на рекомендации за подобряване на учебното съдържание и за идентифициране на групите обучаеми, за взаимодействие с които е необходим персонализиран подход. Освен това бяха идентифицирани категории изключения, които се нуждаят от по-голямо внимание.

**27. Agre, G., Dzhondzhorov, A. A weighted feature selection method for instance-based classification. Lecture Notes in Computer Science 9883 LNAI, 2016, pp. 14-25 (SJR = 0.339)**

Статията представя един нов метод за избор на признаци, който е подходящ за методите за класификация, базирани върху запомняне. Изборът се основава върху оценка на качеството на признаци, получавана чрез алгоритъма ReliefF, който се прилага към ортогонализираното пространство на признаците, получено след прилагане на метода за анализ на основните компоненти. Получените тегла се интерпретират като стойности, пропорционални на размера на обяснени промени в описания на понятията. Чрез установяване на прага се определя, какъв процент от цялата вариабилност на тази стойност трябва да обясняват избраните признаци. Избират се само първите „най-важни“ признаци, чиято комбинация от теглата заедно надвишава установения праг. По време на класификацията избраните признаци се използват заедно с теглата си. Проведените експерименти върху 12 еталонни бази данни показаха предимства на предложения метод в сравнение с традиционния алгоритъм ReliefF.

**28. Agre, G., Petrov, D., Keskinova, S. A new approach to the supervised word sense disambiguation. Lecture Notes in Computer Science 11089 LNAI, 2018, pp. 3-15 (SJR = 0.283)**

Статията представя един нов подход към решаване на пълната задача за снемане на многозначността на думи (all-words sense disambiguation), който позволява да бъде избегната необходимостта от построяване на множество от класификатори, специализирани в снемане на многозначността само на една единствена дума. В основата на подхода лежи една нова интерпретация на понятието „клас“, която свързва всяко възможно значение на многозначната дума с честота, с която то се среща в някой текстов корпус или аотиран речник. По този начин всички възможни значения на

различни многозначни думи могат да бъдат отобразени по унифициран начин в едно ограничено множество от класове, започвайки от най-често срещано значение и завършвайки с най-рядко срещаното. За представяне на думи подходът използва вектори от вграждания (embedding) и информацията, към коя част на речта те принадлежат. Проведените експерименти показват, че точността на класификатори, обучени върху примери, създадени чрез прилагане на този подход, надхвърля стандартните базови точности, използвани за измерване на поведението на класификатори, решаващи пълната задача за снемане на многозначността на думи.

**29. Agre, G., Petrov, D., Keskinova, S. Word Sense Disambiguation Studio: A Flexible System for WSD Feature Extraction. Information, 10(3), 97, Multidisciplinary Digital Publishing Institute, 2019, ISSN:2078-2489Close, DOI:10.3390/info10030097, (SJR = 0.222)**

Настоящата статия представя WSD Studio – гъвкава система за избор на признаци и конструиране на обучаващи и тестови примери за решаване на пълната задача за снемане на многозначността на думи (all-words sense disambiguation). Системата позволява използване на вектори на вграждания (embeddings) на думи и техните значения като част от описание на примерите. WSD Studio има две основни характеристики, различаващи я от всички подобни системи: възможността да се конструира специално компресирано представяне на вграждания и възможността да се генерират примери с различни нива на грануларност. Първата характеристика позволява генерирането на множества от данни с ниска размерност, които могат да бъдат използвани за бързо обучение на различни типове класификатори с висока класификационна точност. Втората характеристика позволява генерирането на множества от обучаващите примери, които могат да бъдат използвани за научаване на класификатори, специализирани в снемане на многозначността на отделни думи, на думи, принадлежащи на една и съща част на речта или на всички многозначни думи. Проведените интензивни експерименти показаха, че класификатори, обучени върху примери, създадени с помощта на системата, имат по-голяма класификационна точност от стандартните базови точности, използвани за измерване на поведението на класификатори, решаващи пълната задача за снемане на многозначността на думи.