

## СЕМАНТИЧКО ОБОГАЋЕЊЕ МУЛТИМЕДИЈАЛНИХ ИНФОРМАЦИЈА И ПРОЦЕСА

Синиша Рудан

Електротехнички факултет у Београду, Србија и Црна Гора

### I. УВОД

Више није довољно пласирати сирове информације на тржиште информација. Корисници очекују додатну вредност, семантички обогатење информације. Разумевање информација од стране мултимедијалних система заснива се на овим обогатењима. Намећу се потребе за развијањем метода за премошћавање семантичког јаза који се појављује између захтева корисника и могућих операција од стране данашњих система. Управо су мултимедијалне информације те које су и продубиле и прошириле овај јаз.

### II. КОРЕНИ СЕМАНТИЧКИХ ОБОГАЋЕЊА

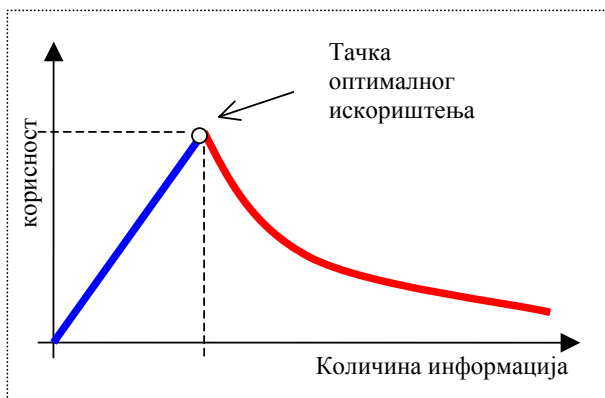
Револуција концепата мултимедије десила се оног момента када су мултимедијални подаци, место исечака забаве, постали информационо добро.

Паралелно овоме, мултимедијални садржај, који већ заузима незаобилазно место у подршци електронског пословања, постаје и сам предмет (електронског) пословања.

Од овог тренутка **мултимедијални садржај представља информационо и тржишно добро.**

Оба атрибута померају акценат са пуког емитовања, на потражњу и захтев мултимедијалних добара по садржају. Ради испуњења ових, новонасталих потреба, постао је приоритетан развој техника за каталогизацију, претрагу и семантички опис мултимедијалних садржаја.

У готово свим доменима употребе мултимедијалних података ушли смо или улазимо у критичну зону десно од тачке оптималног искориштења (слика 1.).



Слика 1. Крива корисности

График нам приказује зависност корисности/употребљивости информација од количине информација у неком просторно/експертском домену. Повећањем количине информација, пропорционално се повећава и употребљивост, при чему се новопридошла информација настоји уградити у већ постојећи систем информација, систем знања. Ипак, у једном тренутку долази до преоптерећености услед ограничења система или немогућности обраде информација од стране људског фактора. Као последица, употребљивост више не успева истим темпом испратити повећање новоукљученог садржаја. Шта више, нови садржај додатно повећава неуређеност досадашњег скупа, отежавајући приступ и проналажење већ постојећих информација. Услед овога имамо негативну реакцију на увећање количине (мултимедијалних) информација. Наравно, ова појава у многим практичним случајевима нема оволико изражене ефекте те се само квари линеарност уз одржање позитивне реакције.

У свим системима описаним овим графиком, циљ нам је поправити радну карактеристику повећањем X или/и Y координате тачке оптималног искориштења.

У циљу тога уводимо **мета-податке (metadata)**.

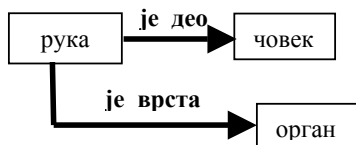
Мета-податак је "податак о податку".

У случају слике као основног податка, "податак о слици" (metadata) могао би да буде опис, семантички приказ слике. При овоме уочавамо мета-податке високог нивоа апстракције, попут управо наведеног, и мета-податке ниског нивоа: хистограм слике, резолуција, аутор, итд. Између ове две крајности налази се више нивоа апстракције са мета-подацима попут текстуре, уочених региона на слици, итд. Кречући се ка вишим слојевима апстракције, мета-подаци садрже већу семантичку вредност.

Мета-подаци свакако увећавају употребљивост мултимедијалних података, поправљајући поменути карактеристику; подижући и истовремено померајући удесно тачку оптималног искориштења. Ипак, оно чему тежимо при уређењу информација/знања, је уређење у хијерархијске структуре. У случају мултимедијалних информација, ми ћемо мета-податке придружене овим информацијама увезивати у хијерархијске структуре, градећи тиме (инстанце) онтологије над посматраним доменом.

**Онтологија** је хијерархијски опис посматраног домена, концептуална шема домена. Прецизније,

онтологија дефинише појмове и правила која постоје унутар посматраног домена. Ова хијерархијска структура сачињена је од ентитета повезаних семантичким релацијама попут: **је\_део** (*is part of*), која говори да је неки концепт (појам) део другог концепта и **је\_врста** (*is kind of*), која говори да је неки концепт специјализација другог, генералнијег концепта.



Слика 2. Основне релације у концептуалним шемама

Структуре у које узимамо податке могу такође исказивати временске (темпоралне) и просторне (спацијалне) односе, тако да имамо темпоралне, спацијалне и темпорално-спацијалне мреже.

Поред многих добробита које проистичу из стварања онтологијских структура над информацијама у нашој (мултимедијалној) бази знања, од круцијалног интереса је и омогућавање **семантичког претраживања**.

Семантичко претраживање нам омогућава извлачење свих релевантних информација, семантички везаних за пронађену информацију. Овиме као резултат претраге добијамо један од чворова на некој од **семантичких нити** коју хватамо и њоме повлачимо све информације из свеукупног структурног графа, градећи подграф који опусује посматрани проблем.

Системи за семантичку претрагу су свесни семантике (значења) и упита и информација које се претражују, семантичких веза међу информацијама.

У глави АПЛИКАТИВНИ ДОМЕН дат је приказ таквог система развијаног од стране аутора овог рада.

### III. СЕМАНТИЧКИ ЈАЗ

Семантички јаз (*Semantic Gap*) генерално посматрано представља семантичку разлику два погледа на исти домен. У нашем разматрању дефинисаћемо **семантички јаз** као семантичку разлику у представљању (схватању) или нивоу операција над мултимедијалним садржајима која се појављује између човека и аутоматизованих (рачунарских) система. Семантички јаз се савлађује сваки пут када се концепти вишег нивоа апстракције преводу у појмове нижег нивоа апстракције и обратно. Семантички јаз је корен неуспешности упита за претрагу над мултимедијалним садржајима. Корисник, желећи да пронађе у мултимедијалној бази података или на Интернету мултимедијални садржај који представља тражени појам, биће разочаран са непрецизношћу, тј. богатством нерелевантних резултата претраге за унесени појам. Узрок томе је традиционално, општекориштено претраживање, претраживање на нивоу кључних речи, без уграђеног схватања појма који се тражи, без икаквог познавања домена претраге, другим речима без семантичког обогаћења процеса претраге. Зато је савлађивање семантичког јаз предмет интезивног научног истраживања. Претрага мултимедијалних података се врши традиционалним техникама, развијеним за текстуална претраживања, упаривањем предикатског израза састављеног од кључних речи, унесених од стране корисника, са именом под којим је меморисан мултимедијални

податак или са речима које се налазе у окружењу коме овај податак припада. Али док су речи биле интегрални део садржаја текстуалних информација, ово не важи у општем случају за мултимедијалне информације.

Оно чему тежимо у овом истраживању је семантичко претраживање на основу садржаја мултимедијалних података. Циљ нам је задовољити следеће захтеве:

- претраживање на основу семантичке сличности са траженим мултимедијалним садржајем . (Корисник тражи слике које задовољавају одређени ниво сличности са сликом коју корисник прилаже, или скицом коју је нацртао у алату за претрагу.)
- проналажење података који садрже, или су садржане у траженом мултимедијалном податку. (Претрага за, рецимо, прозором, треба да пронађе и слике зграда са прозорима. Корисник може да тражи и видео материјал који у некој од секвенци садржи ту слику/скицу)

Овакви семантички упити нису у релацији са подацима ниског нивоа апстракције као што су: боја, хистограм, итд. Стога наше истраживање концентришемо на савлађивање овог семантичког јаз, на генерисање података високог нивоа апстракције, богате семантичке вредности, на основу екстракованих података ниског нивоа апстракције.

### IV. САВЛАЂИВАЊЕ СЕМАНТИЧКОГ ЈАЗА. СИСТЕМИ ЗА РЕКОГНИЦИЈУ

Изнећемо основне принципе на којима је заснован пројекат под радним називом **IHLERS** (*Interactive Human Leded Environment Recognition System*) који развија аутор.

Систем се користи за препознавање окружења, тј. за упоређивање скупа наших опажаја са низом познатих правила и хипотеза похрањених у систему.

Ипак, приказ ћемо почети сликовитијим приступом, кроз једну од могућих примена коришћеног принципа и то применом на семантичко упоређивање мултимедијалних садржаја.

Претпоставимо да је корисник задао упит за налажење слике попут оне на слици 3.а.



Слика 3.

а) корисников упит.

б) слика из базе.

Користићемо **семантичку сличност**, која дефинише метрику за **удаљеност** слика на основу препознавања онога што се налази на слици коришћеној за упит, а потом истим поступком препознавања свих слика из базе, које су семантички, по смислу, везане са упитом, тј. оне које би биле сличне по људском опису. У посматраном примеру, препознајемо у упиту 3.а кућу и препознајемо је и у слици 3.б из базе коју претражујемо. Стога, иако би слика брда са дрвећем у позадини по другим критеријума (нпр. грубо препознавање региона унутар слике и упаривање боја) била сличнија упиту

него слика 3.б, у нашем систему је удаљеност такве две слике веома велика.

Овако дефинисана метрика је од већег практичног значаја, што ће бити илустровано у наставку рада.

Процес препознавања почећемо екстраковањем особина ниског нивоа семантике помоћу већ развијених алгоритама који стога нису предмет нашег истраживања. Евидентираћемо све регионе ограничене уоченим контурама, и они ће бити практично једини улазни подаци целог процеса. Од сваког региона креирамо ентит обогашен мета-подацима: *положај* и *боја*.

Закључак мојих претходних изучавања рекогнитивних процеса у људи је да у развоју будућих рацунарских система рекогниције треба мењати парадигму:

уместо досадашњег приступа са нагласком на монотоним резоновању, базираном на исцрпној и "скупој" анализи по микро-плановима, треба применити холистички приступ, кроз рад у неизвесном окружењу подржаном од стране ТМСа (*Truth Maintenance System / Систем за одржавање истинитости*) и итеративним подизањем нивоа апстракције и извесности релација међу препознатим објектима у циљу **постепеног** изоштравања виђења посматраног света.

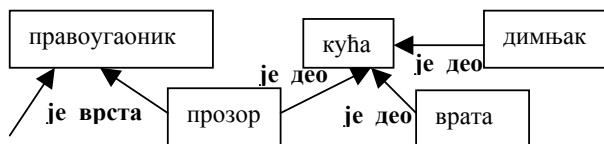
Објекте посматрања, који окидају закључивања, сортирамо по величини и геометријско-текстурној једноставности.

У првом кораку површно сагледавамо "свет опажених ентитета" и изграђујемо почетни скуп претпоставки, од којих ћемо у предстојећим итереацијама неке оповргнути а у неке ћемо повећати поверење, ширећи скуп новим претпоставкама.

У нашем приступу увек ћемо тежити, да сваким опажањем, дедукцијом, утврдимо или учврстимо веровање о контексту света који рекогнујемо.

Илустроваћемо овај принцип једним од правила које се користи при самом почетку рекогновања: ако у посматраном приказу уочимо троуглове, четвороуглове јасних ивица, веома је извесно да је приказ у контексту урбане средине. Стога ћемо у наставку значајно сузити област претраживања.

При уочавању облика, узевши за пример слику 3. и на њој уочени правоугаоник који смо уокрили контуром 1, залазимо у онтологију и уочавамо прозор као једну од специјализација правоугаоника. Упоредивањем са придруженим визуелним патернима уочава се одређени ниво сличности који још увек нема довољну тежину. Али не губимо време на исцрпној анализи овог детаља, већ настављамо претрагу. Проналажењем још једног прозора, једна претпоставка ојачава другу.



Слика 4. приказ дела кориштене онтологије

Поновним уласком у онтологију, сазнајемо да је прозор део (*part\_of*) куће, што постаје нова хипотеза. Откривањем нечега што личи на димњак, извесност хипотезе «видимо кућу» расте. А сазнање да смо уочили прозоре, додатно повећава извесност уочавања димњака.

У исто време градимо и просторне релације међу уоченим и дедукованим концептима и кроз сваку итерацију их дижемо на виши семантички ниво.

Дакле, нису нам довољне само онтолошке везе **је део**, **је врста** те уводимо мреже које описују просторне (спацијалне) односе међу објектима и уочене релације упоређујемо са њима. Исто тако, уводимо мреже за опис темпоралних релација. Рецимо да смо оучили светлост и проласком кроз више сцена уочили да се пали и гаси, велика је извесност да је у питању светионик, па посматрамо контекст мора и при претрази користимо релевантне онтологије. Ако смо, рецимо, уочили жуто светло, па у неком од следећих фрејмова да се угасило, а испод њега појавило зелено светло, онда бисмо увидом у теморално-спацијалну мрежу «уочили» семафор и створили нови јак контекст улице, за наставак анализе.

## V. СЕМАНТИЧКО ПРЕТРАЖИВАЊЕ. БОРБА ПРЕТРАЖИВАЧА

У нашем истраживању разматрамо могуће токове борбе Веб претраживача у светлу наступајућих захтева за претрагама вишег нивоа апстракције, семантички обогашеним претрагама.

**Google** је несумњиво добио битку у брзини претраге и веродостојности резултата (комплексном **PageRank** технологијом). Ипак узимајући у обзир чињенице изнесене у овом раду, уочљиво је да већ код веродостојности резултата предност имају претраживачи који полуаутоматизовано, у процесу потпомогнутом/ревидираном од стране људи, класификују Веб странице у стабловске каталоге. За очекивати је, и то је сенка која се надвија над **Google**-ово царство, да ће се корисници све више оријентисати ка порталима који имају уграђене **семантичке везе** и одређени ниво онтологије изграђен над терминима. По овоме сценарију, први у ланцу семантичке претраге Интернета постају Онлајн енциклопедије попут портала **Wikipedia** или **Directory Service портали** попут **ODP-a** (*Open Directory Project*) – који се мануелно одржавају и у којима се мануелно граде семантичке везе међу појмовима, већином од стране волонтера. Након упознавања структуре проблема тј. појмова у вези са траженим појмом, корисник се спушта на нижи ниво апстракције преласком на следећи корак; на употребу **Google**-ликих портала ради квантитавног проширења претраге и ради потраге за свежијим резултатима.

Поставило се питање: «Енциклопедије, па и оне онлајн, већ дужи временски период постоје, у чему је еволутивни помак?» Из одговора ће се уочити да је овај еволутивни помак толики да ће присилити класичне претраживаче на убрзано увођење револуционарних промена у технологијама претраге. Већ сада се мочни **Google** не преза да користи услуге ODPa, а то чине и многи други. Еволутивни помак је просто отварање онтолозијских структура ових енциклопедија-портала за јавну изградњу, чиме је омогућено нудиоцима/власницима Веб страница и осталих ресурса, да на овим порталима самостално додају странице које их описују и реферишу на њих и да их уграђују у структуру портала, кроз семантичке међувезе. Ово је такозвани **The Wiki Way**. Импликације су заиста револуционарне: енциклопедије од архивских контејнера постају **Up-To-Date Digests of Web and Real World**, тј. *Увек-свежи сажеци Веба и реалног света*, и

све то са одређеним нивоом семантиких међувеза, лаких за праћење, које доводе до нових релевантих података. Власнички претраживачи не могу мануелном класификацијом од стране свога особља подржати исти ниво услуге. Опасан недостатак који уочавам у тренутној реализацији *Wiki* приступа је некориштење онтолојских језика који би омогућили аутоматизовану претрагу ових садржаја. Доласком семантичког Веба овај недостатак ће учинити цео пројекат технолошки застарелим. Али бројна армија волонтера ће и тај јаз премостити.

Ако је потреба за семантичким претраживањима постала неопходна код претраживања текстуалних садржаја, јасно је колики су тек притисак на до сада коришћене методе изазвали мултимедијани садржаји.

*Google*-у преостаје да на исти начин како се попео на трон, остане на њему: научним достигнућима. У овом случају, за очекивати је да ће снажно подржати развој семантичког веба који ће омогућити *Google*-овој мочној машинерији да изврши семантичке упите, вероватно много квалитетније него што то сада раде волонтери у опонентским отвореним пројектима.

У класичним ревоуцијама долазило је до смене власти, тако и у овој семантичкој револуцији могу да се појаве неки нови момци који ће «двојицу пребогатих момака» новом револуционарном технологијом збацити са трона.

Можемо ли очекивати и (семантичко) сједињење ODP-а и Wikipedia-е?

Још једно поље на коме се води трка је «**персонализовано и осетљиво-на-контекст претраживање**» које даје резултате различите за сваког корисника на основу његових подешавања и историје претраживања, предвиђајући значај одређених информација за свакога понаособ. Куповином компаније *Kaltix*, *Google* је потврдио опште предвиђање великих финансијских добитака од оваквих система, Водећи системи за е-Пословање и е-Продавнице улажу приметан истраживачки рад у развијање личних система персонализоване претраге.

«*Персонализација је један од светих зралова претраживања*», Крис Шерман, едитор индустријског билтена *Search Day*.

## VI. АПЛИКАТИВНИ ДОМЕН

*IHLERS* систем се развија са циљем помагања слепим особама, које уз помоћ осталих чула и својим резоновањем или свесношћу о месту боравка могу наводити системе рекогновоња, помажући им у декодовању контекста. У овом процесу два субјекта ограничених мочи опажања испомажу се кроз синергистичку интеракцију. Систем користи камере постављене на наочаре корисника, тако да увек прати правац његовог погледа. Сцену детектовану камерама препознаје и препричава је кориснику. Овде је од изузетног значаја персонализација и семантичка компресија.

**Семантичком компресија** је још једна од *State of The Art* примењених техника, која је у интензивном развоју. Ауторова истраживања у овој области нуде да се оваквом компресијом, компресијом заснованом на схватању објекта компресије, издвајају семантички битни детаљи слике. Као и персонализацију, користимо је ради изостављања кориснику небитних детаља да би се

цела сцена могла препричати у реалном времену. Персонализацијом изостављамо делове описа на основу подешавања/избора корисника и сортирамо изнесене чињенице по приоритету у контексту активне функције, изабране од стране корисника.

Аутор је у периоду од 1999.г. до 2000.г. развијао семантичку базу информација за интернет новинску агенцију са графичким приказом релација. Концепт се заснивао на креирању **виртуелних соба** у које је корисник улазио преко неке о вести, анализе или слике и у њој би проналазио све информације семантички везане за иницијалну информацију. Систем је развијан да омогући навигацију како кроз хронолошке, тако и узрочно-последичне везе.

Овакв систем нуди кориснику *added-value*, смањујући време које он улаже у претрагу релевантних информација и омогућава му откривање нових знања.

Корисник пошавши од жеље за задовољењем почетног питања, проналази њему дотада непознате информације, на посредан начин везане па самим тим и тешко уочљиве претрагама заснованим на кључним речима.

Оно што је од још већег значаја је и структурализација приказа информација на начин једноставан за брзо сагледавање оквира проблема.

Овај принцип се данас користи у оквиру пројекта CNET News.com под именом *The Big Picture*, али са нешто слабијим карактеристикама. Скоро је развијен и систем [www.grokker.com](http://www.grokker.com) који ради ово на глобалном нивоу, а визуелно је сличан ауторовом решењу.

## ЛИТЕРАТУРА

- [1] Sinisa Rudan, *Multimedia – Riding Multi-headed Dragon*
- [2] Рудолф Карнап, *Филозофија и логичка синтакса*
- [3] João Miguel da Costa Magalhães, *Semantic Multimedia: Mining, Fusion and Extraction*
- [4] <http://www.w3.org/2001/sw/>
- [5] Marc Spaniol, Ralf Klamma, Matthias Jarke, *Semantic processing of multimedia data by MPEG-7 for capacious knowledge management.*

**Abstract:** This paper explains emerging needs for semantic enrichments of multimedia content and gives the methods for fulfilling them from the point of the information management. It also presents author's research targeted to reveal the fruits coming with these enrichments. Author is trying to predict the flows of upcoming struggle between the Web's largest search engines and portals in their aspiration to satisfy users' need for searches of higher semantic level, especially emphasized by multimedia content. Work is part of the more comprehensive work on the upcoming book on the theme of *Multimedia and digital television*, and is also result of author's research and development efforts in this area, especially in the recognition process.

## SEMANTIC ENRICHMENT OF MULTIMEDIA INFORMATION AND PROCESSES

Sinisa Rudan