

UPORABNIKU PRILAGOJENO ISKANJE MEDIJSKIH VSEBIN

Ciril Gale*

UDK: 025.4.03:004

Ciril Gale: Uporabniku prilagojeno iskanje medijskih vsebin. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik referatov z dopolnilnega izobraževanja, Maribor 7/2008, str. 141-148.

Izvirnik v slovenščini, izvleček v slovenščini in angleščini, povzetek v angleščini.

Uporabniku prilagojeno iskanje vsebin ali personalizacija je postopek, s katerim iz množice dostopnih vsebin izberemo tiste, ki so za uporabnika najbolj primerne (uporabniku všečne vsebine). To naredimo na podlagi poznavanja uporabnikovih preferenc, ki so zapisane v uporabniškem modelu. Tega gradimo in posodablamo s pomočjo informacij o v preteklosti izbranih/neizbranih vsebinah, katere dobimo z analizo interakcije uporabnika s sistemom preko uporabniškega vmesnika.

UDC: 025.4.03:004

Ciril Gale: User Adapted Selection of Content. Technical and Field Related Problems of Traditional and Electronic Archiving. Conference Proceedings, Maribor 7/2008, pp. 141-148.

Original in Slovenian, abstract in Slovenian and English, summary in English.

User adapted selection of content, known as personalisation, is a procedure, by which the system for the personalised selection of content, among the available contents chooses those that are most suitable for (or most liked by) the user. This procedure is based on the knowledge of users' preferences, which are recorded in the user model. This model is constructed and updated with the help of information about content chosen or not chosen in the past, which is obtained with the analysis of the user - system interaction through the user interface.

Ključne besede: digitalni mediji, poizvedovanje, digitalna televizija, personalizacija vsebin, interaktivnost, standardi, modeli.

1. STROKOVNA TERMINOLOGIJA IN OKRAJŠAVE

Razumevanje novih tehnologij obravnavanja vsebin s posebnim poudarkom na digitalni televiziji zahteva poznavanje nekaterih novih pojmov, predsem pa okrajšav. V nadaljevanju posredujem izbor najbolj znanih in uporabljenih v tem prispevku¹:

DVB-C - Digital Video Broadcasting - Cable (digitalni način prenosa zvoka in slike po kablu, uporablja sistem kodiranja MPEG-2 ter način modulacije QAM 64, 128, 256, omogoča tudi povratno zvezo - interaktivnost, zato se uporablja za storitve interneta.

VOD - Video On Demand, video na zahtevo.

PPV - Pay Per View, plačilo po ogledu, deluje na MHP-osnovi.

* Mag. Ciril Gale, TV Slovenija, Oddelek za arhiviranje in dokumentacija, Kolodvorska 2, 1000 Ljubljana, Slovenija.

Prispevek je nastal na temelju razmišljanja o ovladovanju množice vsebin, ki jih posredujejo mnogi ponudniki storitev po različnih komunikacijskih kanalih. Prijatelj, univerzitetni profesor dr. Nikola Pavešič, vodja laboratorija za umetno inteligenco na ljubljanski Fakulteti za elektroniko, je pred nastankom tega besedila organiziral pogovor z dr. Matevžem Pogačnikom, ki je pod mentorstvom dr. Jurija Tasiča doktoriral s temo *Uporabniku prilagojeno iskanje multimedijskih vsebin*. Za posredovane usmeritve in podatke se vsem trem iskreno zahvaljujem.

¹ Terminologija povzeta iz: Primož Sernel, *Prihaja digitalna televizija*, Frik, Ljubljana 2007.

DVB-H - Digital Video Broadcasting - Handhelds (digitalni prizemeljski način prenosa zvoka in slike na majhne, ročne, mobilne aparate, imenovan tudi mobilni sprejem, prenos podatkov lahko poteka po istem omrežju kot za DVB-T, sočasno na istem kanalu, vendar je taka uporaba manj verjetna, ker sta sistem obdelave in način prenosa drugačna, uporablja se sistem kodiranja MPEG-4, time-slicing mehanizem na podlagi podatkovnega prenosa IP, pomeni, da prenos podatkov poteka v časovnih svežnjih in ne stalno, potreben je zaradi omejene uporabe mobilnih naprav z vidika energije, zmogljivosti baterije, uporablja se transformacija IFFT nosilcev 2K, 4K ali 8K, modulacija COFDM ter dodatna kodna zaščita MPE-FEC, zmogljivost prenosa na posamezen kanal je odvisna od izbranih parametrov in je med 5 in 25 Mbit/s).

DVB-S - Digital Video Broadcasting - Satellite (digitalni način prenosa zvoka in slike po satelitu, najpogosteje uporabljen način digitalnega prenosa, po različnih satelitih Astra, Eutelsat, Helasat, uporablja sistem kodiranja MPEG-2 ter način modulacije QPSK, omogoča zelo veliko zmogljivost prenosa programov, saj lahko zaradi aktivnega in krmilnega sprejemnega dela antene (LNB) sprejemamo signale na štirih različnih ravneh, dveh frekvenčnih in dveh polarizacijskih).

DVB-S2 - Digital Video Broadcasting - Satellite (nadgradnja osnovnega sistema prenosa digitalnih signalov po satelitu, ki ima za 30 % večji izkoristek prenosa zaradi izboljšanih sistemov modulacije in kodiranja).

DVB-T - Digital Video Broadcasting - Terrestrial (digitalni prizemni način prenosa zvoka in slike, imenovan tudi DTT, v Evropi in tudi drugod po svetu opredeljen standard za prenos digitalne televizije in zvoka, ki ga lahko sprejemamo z anteno. Prenos podatkov vsakega programa poteka v sistemu kodiranja MPEG-2, uporablja se enaka širina kanala kot pri analognem prenosu VHF in UHF, 7 MHz, vendar lahko zdaj v istem kanalu).

MPEG-2 (generični standard video kodiranja slike in zvoka s stiskanjem; v resnici ne gre za stiskanje, ampak za omejitev podatkov; uvedli so ga leta 1994 kot naslednika MPEG-1, opredeljen je kot osnovni standard za DVB, omogoča podatkovni prenos do 50 Mbit/s).

MPEG-4 (generični standard video kodiranja slike in zvoka, je naslednik standarda MPEG-2 od leta 1998, namenjen je sistemom z manjšo zmogljivostjo in pasovno širino, kot so prenosni telefoni in video telefoni, vendar ga z razvojem zahtevnejših algoritmov stiskanja in omejevanja podatkov želijo v prihodnosti uporabljati tudi za DVB za večjo zmogljivost prenosa).

HDTV - High Definition Television (televizija visoke razločljivosti). Pri HDTV je slika sestavljena iz veliko več pik kot pri navadni sliki na dvd-ju. HDTV največje ločljivosti je sestavljen iz približno 2.764.800 pik, medtem ko je dvd slika sestavljena iz 414.720 pik. To je ogromna razlika. Pri navadnih televizijah ni nobene razlike, ker so ekrani tako narejeni, da so pike malo večje, in niti ne pomislimo, da bi slika sploh lahko bila bolj ostra.

Pri lcd zaslonih, ki so sestavljeni iz majhnih pikic, pa od blizu hitro opazimo, da slika ni tako čista. Ko pa na lcd ekranu preizkusimo hdtv, opazimo neverjetno razliko. Za primer pri tenisu, ko gre kamera od daleč čez vse gledalce, vidimo na navadni televiziji samo gledalce, vendar jih težko prepoznamo v obraz, ker so obrazi bolj zamegljeni. Pri hdtv pa je slika tako čista, da lahko ločimo vsak obraz posebej.

2. KAKO NAJTI IN IZBRATI USTREZNE TELEVIZIJSKE PROGRAME

Zamislite si naslednje: na razpolago imamo na tisoče televizijskih programov, pa ne vemo, katere naj izberemo in jih gledamo. Vsak dan izgubimo preveč časa s pregledovanjem programov, preden se lahko odločimo za tiste, ki najbolj ustrezajo našim interesom. Nazadnje bomo obupali in gledali televizijo vse manj. Podobno velja za večino televizijskih gledalcev po vsem svetu.

Z vključevanjem Slovenije v večje svetovne povezave, s prehodom na digitalno televizijo, razširjenost osebnih računalnikov tako pri nas kot v tujini, dostopnostjo vse večje količine množičnih medijev, se je enormno povečala količina informacij, ki so dostopne vsakomur.

Prek satelitov v naše domove prihaja na stotine različnih televizijskih programov, z uvedbo digitalne televizije pa jih bo na tisoče. Zato pravica do obveščenosti ni nikakršen problem več, pač pa prihaja do nove nevarnosti. Zaradi preobilice informacij prihaja do prezasičenosti uporabnikov z njimi. Zato se postavlja vprašanje, kako izbrati iz nepregledne množice tiste vsebine, ki posamezniku najbolj ustrezajo, upoštevajoč pri tem njegove individualne interese.

3. KAJ JE TO PERSONALIZACIJA ISKANJA?

Rešitev omenjenega gledalčevega problema je personalizacija iskanja vsebin. Uporabniku prilagojeno iskanje vsebin ali personalizacija je postopek, s katerim iz množice dostopnih vsebin izberemo tiste, ki so za uporabnika najbolj primerne. To naredimo na podlagi poznavanja uporabnikovih interesov. Te so zapisane v uporabniškem modelu. Zgradimo in posodobljamo ga s pomočjo informacij o izbranih in s strani uporabnika ocenjenih vsebinah.

Pri tem se lahko vprašamo, za katera področja izbiranja vsebin je ta model ustrezen, in ugotovimo, da so področja uporabe izbire vsebin številna; iskanje spletnih dokumentov, filtriranje pošte, iskanje novic v novičarskih skupinah, izbira televizijskih programov in glasbe, iskanje strokovnih člankov in celo iskanje šal.

Široka uporaba tovrstnih sistemov je povsem razumljiva, saj je število uporabniku dostopnih informacij in vsebin v zadnjih letih tako silovito naraslo, da si učinkovite izbire vsebin brez postopkov personalizacije ne moremo več predstavljati.

Ob tem je treba dodati dejstvo, da postajajo digitalni televizorji in spremljajoče naprave v nekem pogledu podobni računalnikom in bodo v prihodnje imeli možnost izvajanja programov za personalizacijo iskanja medijskih vsebin. To pomeni, da gledalcu ne bo treba več razmišljati o reševanju tovrstnih težav, pač pa jim bo rešitev dana že z nakupom aparata samega.

Trenutne rešitve že omogočajo personalizacijo, vendar se ta izvaja na strežnikih ponudnika storitve. Glavna naloga tovrstnih sistemov je predvsem razbremenitev uporabnika, ki se v množici informacij ne znajde več.

Poleg tega v času, ki mu je na voljo, ne more lastnoročno pregledati vsega. Avtomatizirani sistemi neprimerno hitreje pregledajo sezname opisov dostopnih vsebin in uporabniku ponudijo tiste vsebine, ki so mu všeč.

4. RAZBREMENITEV IN PRIHRANEK DRAGOCENEGA ČASA

Poznamo nekaj osnovnih metod personalizacije izbire vsebin, med katerimi sta najpogostejša dva: vsebinsko ter skupinsko izbiranje vsebin. Pri vsebinskem filtriranju vsebin se primernost posamezne vsebine določa na podlagi primerjave podatkovnega opisa vsebine in uporabniškega modela.

Pri skupinskem filtriranju se ugotavlja stopnjo podobnosti med različnimi uporabniki, nato pa izbere tiste vsebine, ki so bile vseh podobnim uporabnikom. Drugi pristopi delujejo na podlagi demografskih podatkov, z njimi povezanih stereotipov, pa tudi kombinirani (hibridni) pristopi so vse pogostejši. Pri oblikovanju tovrstnih sistemov se postavljajo vsakršne dileme. Vedno je prisotna dilema ob odločitvi med uporabo implicitne ali eksplicitne povratne informacije, na podlagi katere sistem posodablja in prilagaja uporabniški model.

Rezultati raziskav so pokazali, da večina uporabnikov želi sisteme, ki od njih ne zahtevajo odvečne interakcije, bistvena je le funkcionalnost. Po drugi strani pa vemo, da sistemi, kjer uporabniki eksplicitno izrazijo všečnost posameznih vsebin, delujejo precej bolje.

Zanimiva in aktualna postaja interaktivna televizija in k temu je botroval predvsem razvoj digitalnih naprav, ki postajajo vse bolj zmogljive v smislu procesorske moči in pomnilniških kapacitet. Digitalni videorekorderji, pa tudi mobilni telefoni, dlančniki in druge naprave so vse bolj podobni osebnim računalnikom, kar v praksi pomeni, da lahko dostopajo, prikazujejo ter shranjujejo slike in avdio/video zapise. Pomembno vlogo ima tudi razvoj komunikacijskih povezav, ki omogoča vse hitrejši dostop do vsebin in spremljajočih podatkov, potrebnih za izvajanje storitev interaktivne digitalne televizije. Uvajanje standardiziranih tehničnih rešitev za izvajanje aplikacij v digitalnih napravah, kot je Multimedijška domača platforma (ang. Multimedia Home Platform - MHP), omogoča poenotenje načrtovanja in uporabe storitev interaktivne televizije.

5. DVA OSNOVNA PRISTOPA

Seveda pa brez težav tudi tu ne gre. Eden od problemov, povezanih z interaktivno digitalno televizijo, je opisovanje vsebin. Primernost vsebin za uporabnika se, kot smo že omenili, izračunava na osnovi primerjave med uporabniškim modelom in metapodatkovnimi opisi vsebin. Zato je pomembno, da so ti opisi kvalitetni, predvsem pa standardizirani. Na področju multimedijev se je v zadnjih letih pojavilo nekaj standardov, med katerimi so najpomembnejši MPEG-7, TV Anytime ter MPEG-21. MPEG-7 je izredno splošen standard, ki omogoča opisovanje najrazličnejših multimedijjskih vsebin (avdio, video, slike, 3D modeli itd.) in je zaradi tega izredno kompleksen.

Preprostejši standard, usmerjen v opisovanje televizijskih programov, je TV Anytime. Osnovan je na standardu MPEG-7 in predpisuje mehanizme tako za opisovanje, kakor tudi za iskanje vsebin. Oba standarda se lotevata tudi opisovanja uporabnikov. MPEG-21 je standard, ki je namenjen zagotavljanju kvalitete storitev, nadzoru dostopa do vsebin ter v zagotavljanju spoštovanja avtorskih pravic. Poleg izbire standarda za opisovanje vsebin pa je pomembna tudi kvaliteta opisovanja vsebin. Veliko vsebin ima namreč zelo skromne metapodatkovne opise, pri čemer so problemi tako v skoposti samih opisov, kakor tudi (pre)splošnosti pri kategorizaciji vsebin (razvrščanju glede na žanr).

Razvili so nov pristop uporabniškega modeliranja, ki temelji na personalizirani izbiri vsebin, opisanih po standardu TV Anytime. Pri razvoju sistema za personalizirano izbiro vsebin so izhajali iz izhodišča, da mora ta delovati, ne glede na to ali sistem podpira izbiro vsebin za večje število uporabnikov ali le za enega samega. To jih je vodilo k razvoju algoritma na osnovi vsebinskega filtriranja. Pri tem mora znati ocenjevati tudi primernost vsebin s slabimi opisi, predvsem tistih s presplošno kategorizacijo.

Želeli so razviti tak algoritem, ki ne bi bil prezahteven za izvajanje na procesorsko manj zmogljivih napravah, kot je npr. digitalni-videorekorder. Odločili so se za sistem, ki bo množico dostopnih vsebin razvrstil v dva razreda in sicer v razred vsebin, primernih za uporabnika, in razred vsebin, neprimernih za uporabnika.

Pri ocenjevanju vsebin so uporabljali več atributov. Vsebine so bile opisane z od dvema do šestimi opisnimi atributi. Opise s šestimi atributi so imeli igrani filmi (žanr, ključne besede, igralci, režiser, država ter leto nastanka). Programi ostalih žanrov (televizijske serije, športni programi itd.), imenujmo jih nefilmi, so bili večinoma opisani z dvema atributoma (žanr ter ključne besede). Nekateri opisi nefilmov so vsebovali tudi informacijo o pripadnosti vsebine skupini podobnih programov.

Na dva koraka so razdelili postopek ocenjevanja vsebin. V prvem koraku za vsakega od opisnih atributov vsebine posebej izračunajo "podobnost" z uporabniškim modelom in ga izrazijo z realnim številom.

Za izračun podobnosti na nivoju opisnih atributov uporabljajo dva matematična postopka: modificiran naivni Bayesov klasifikator (za žanr, igralce, režiserje ter pripadnost skupini programov) in modifikacijo postopka imenovanega TF-IDF (ang. Term Frequency - Inverse Document Frequency), ki temelji na pogostosti posameznih besed v opisnih atributih (za ključne besede, državo ter leto nastanka).

To je le prva stopnja. Na podlagi teh izračunanih vrednosti podobnosti v drugem koraku vsebino razvrstijo v razred primernih ali neprimernih vsebin z metodo podpornih vektorjev (ang. Support Vector Machines - SVM). Sistem naučijo razvrščanja vsebin z metodo SVM s pomočjo učne množice, ki vsebuje že izračunane vrednosti podobnosti na nivoju posameznih atributov ter informacijo o razredu, v katerega je uporabnik razvrstil to vsebino.

V praksi to pomeni, da so testne uporabnike prosili, da vsebine iz učne množice ocenijo kot primerne ali neprimerne, nato pa so za te vsebine izračunali podobnost z uporabniškim modelom na nivoju posameznih atributov. Rezultat tega učenja je model SVM, s pomočjo katerega lahko nove in še neocenjene vsebine, za katere so že izračunali podobnost na nivoju posameznih atributov, razvrstijo v enega izmed obeh razredov.

Z uporabo modelov SVM za razvrščanje vsebin je povezan tudi njihov (t. j. skupine, ki deluje na ljubljanski Fakulteti za elektroniko) izvorni prispevek. Njihov pristop omogoča razvrščanje vsebin na podlagi takoimenovanega univerzalnega modela SVM, ki je izdelan z učenjem na podlagi ocen vsebin velike množice uporabnikov. Univerzalni model SVM je tako uporaben v začetku delovanja sistema, ko le-ta še nima dovolj (zanesljivih) informacij o uporabnikovih izbirah vsebin in s tem možnosti za priučitev prilagojenega modela SVM.

Tega sistem zgradi kasneje, ko ima na voljo dovolj informacij o uporabnikovih izbirah vsebin. Testiranja kažejo, da je zanesljivost delovanja z uporabo univerzalnega modela SVM dovolj dobra, z uvedbo prilagojenega modela SVM pa se še izboljša.

Da bi uporabniki televizijskega gradiva čim hitreje prišli do ustreznih vsebin, pa to še ni dovolj. Naslednja stopnja možnosti ocenjevanja vsebin s presplošno kategorizacijo atributa žanr omogoča strukturo njihovega uporabniškega modela in izvorni način njegovega posodabljanja. Uporabniški model je tako sestavljen iz množice delnih uporabniških modelov (DUM). Vsak DUM vsebuje informacije o uporabnikovih preferencah glede posameznih atributov (ključnih besed, igralcev, režiserjev itd.), vezanih na vsak posamezen žanr.

Tako govorimo o DUM-ih, ki pripadajo posameznim žanrom, kar daje atributu žanr posebno vlogo. Poleg tega DUM vsebuje tudi informacijo o priljubljenosti samega žanra, kateremu pripada. DUM-i so povezani v hierarhično drevesno strukturo (taksonomijo), ki je povzeta po hierarhični taksonomiji žanrov TV Anytime, oziroma modelu hierarhije TV Anytime uporabljane v okviru mednarodnega projekta MyTV.

Sam uporabniški model je zasnovan splošno in kot tak lahko posnema poljubno hierarhično ali nehierarhično strukturo žanrov. V primeru hierarhične strukture žanrov je število hierarhičnih nivojev lahko poljubno. Prav tako lahko DUM-i vsebujejo informacije o uporabnikovih preferencah glede poljubnih atributov. Izviren je tudi pristop posodabljanja uporabniškega modela oziroma DUM-ov, ki ob posodabljanju DUM-a, vezanega na nek konkreten žanr (npr. DUM("Šport:Košarka")) posodobi tudi DUM na višjem nivoju v hierarhični strukturi (v tem primeru DUM("Šport")). Skušali so tudi izboljšati delovanje sistema za personalizirano izbiro vsebin z upoštevanjem leksikalnih povezav med ključnimi besedami z iskanjem sinonimov, hiponimov ter hipernimov. V ta namen so uporabili leksikalno bazo WordNet.

6. BISTVENA JE FUNKCIONALNOST

Kakšni pa so prvi rezultati? Rezultate so ovrednotili z znanimi kriteriji za ocenjevanje uspešnosti delovanja tovrstnih sistemov. V literaturi obstaja več kriterijev, med njimi pa sta najbolj znana in po našem mnenju tudi najbolj uporabna delež najdenih ter natančnost (ang. recall in precision). Med obema se zdi bolj pomemben kriterij natančnosti, saj je celotno število primernih vsebin ponavadi tako veliko, da si nihče ne more ogledati vseh. Tako manjši delež najdenih vsebin v realnosti ne vpliva preveč na uspešnost delovanja sistema.

V kolikšni meri pa ta projekt deluje v praksi? Testni rezultati so pokazali, da sistem deluje dobro. Natančnost razvrščanja vsebin je bila preko 80 %, ne glede na uporabljeni model SVM. Delež najdenih vsebin je dosegal približno 45 % pri uporabi prilagojenega modela SVM pri razvrščanju vsebin, oziroma približno 30 % pri uporabi univerzalnega modela SVM.

Rezultati pri razvrščanju vsebin s presplošno kategorizacijo niso bistveno slabši. Veliko informacije vsebujejo namreč ključne besede v opisih vsebin (predvsem pri športnih programih), hierarhična struktura uporabniškega modela pa omogoča uporabo teh informacij tudi v primeru slabše kategoriziranih vsebin.

Upoštevanje leksikalnih povezav med ključnimi besedami ne prinese izboljšanja rezultatov. To lahko najverjetneje pripišemo že sicer dobremu delovanju sistema za personalizirano izbiro vsebin, ki je rezultat uporabe hierarhične strukture delnih uporabniških modelov, standardizirani terminologiji v opisih vsebin ter (pre)splošnosti leksikalne baze WordNet.

Eden od hitov CEBITA je DVD recorder, ki ne snema propagandnih sporočil, ki jih je vsak dan več. Uporabniki so nejevoljni zaradi številnih prekinitev oddaj z reklamnimi

sporočili, zato ni čudno, da je eden od hitov prav tovrstni DVD recorder. To je le eden od vidikov uporabe omenjenih programov, ki so vgrajeni v tem primeru v najnovejših zapisovalcih televizijskih vsebin. Preprosto povedano: zapisovalec najrazličnejših sporočil je, glasov, glasbe, slike, barvnih odtenkov ipd., vse to »registrira« kot dele neke celote. Ko se ti signali zaradi reklam radikalno spremenijo, pa je program pripravljen tako, da preneha s snemanjem, saj v tem trenutku pridejo v sistem signali, ki so opazno različni od prejšnjih. Ko se ti signali ob koncu reklam zopet opazno spremenijo, rekorder spet preklopi na prvotni ukaz in se snemanje nadaljuje.

7. TREBA BO ZAMENJATI TELEVIZIJSKE SPREJEMNIKE ALI DOKUPITI PRETVORNIKE

Če za trenutek ostanemo na področju individualnega uporabnika, potem se postavlja vprašanje uporabnosti aparatov za širšo uporabo tudi v prihodnje. Vsekakor tehnološki razvoj zahteva svoje in tudi prehod na digitalno televizijo bo od vsakega posameznega naročnika zahteval nakup novih naprav oz. sprejemnikov televizijskih signalov (ang. settop-box). Ti bodo morali biti namenjeni sposobnosti sprejemanja digitalnih signalov (DVB-T) s podporo kodeku MPEG4.

Torej: treba se je osredotočiti na nakup digitalnih sprejemnikov televizijskih signalov. V kolikšni meri bodo uporabni digitalni sprejemniki, ki so naprodaj v tem trenutku, je težko reči, menda pravih še ni na trgu. Na vsak način pa je treba biti pri nakupih sprejemnikov televizijskih signalov previden, saj se je Slovenija odločila za digitalizacijo televizije z uporabo novejšega kodeka MPEG-4. Tovrstne naprave, ki so že v prodaji in uporabi v drugih državah EU, niso kompatibilne z bodočo digitalno televizijo v Sloveniji, saj je v teh državah digitalna televizija zasnovana na kodeku MPEG-2.

Postavlja se vprašanje, ali lahko digitalne sisteme uporabljamo uspešno tudi pri arhiviranju televizijskega gradiva. Nekje v prihodnosti bodo na voljo sistemi, ki bodo znali analizirati televizijske programe. Tako bodo lahko že sami »razpoznavali« vsebinske sklope in jih opisovali z opisnimi atributi, ki bodo potem na razpolago ustvarjalcem novih televizijskih oddaj. Tako bodo računalniki že sami ustvarjali arhive, čeprav bo tudi v prihodnje potrebna pomoč dokumentalista.

Z arhivskim gradivom v digitalni obliki pa bo izpolnjen glavni pogoj za njegovo najširšo uporabo in sicer takoj po nastanku. Vsekakor bo gradivo dostopno vsem v digitalni obliki in tako novinarjem in režiserjem ne bo potrebno vedno znova iskati trakov v arhivu, jih pregledovati in šele nato med njimi izbirati inserte. V prihodnje bodo povezani neposredno z digitalnimi arhivi in jih bodo lahko pregledovali ter uporabljali glede na svoje trenutne potrebe.

V kolikšni meri pa je obstojnost arhivskega gradiva primerljiva z dosedanjimi nosilci podatkov na področju avdiovizualnih medijev, je tudi pomembno vprašanje. Analogni signali oziroma zapisi na kasetah z večkratno uporabo vse bolj izgubljajo na kvaliteti. Z leti se tako ali tako kvaliteta manjša, zlasti pa če kasete ali trakove uporabljamo pogosto. Digitalni zapisi ohranjajo svojo kvaliteto ne glede na število uporab in slabšanje samega fizičnega zapisa na traku, disku ali DVD-ju. Slaba stran tega je, da se po več desetletjih digitalni zapis v nekem trenutku poslabša do te mere, da ga sploh ni več mogoče prebrati. Rešitev je seveda v sprotnem prepisovanju gradiva v celoti. Sicer pa imajo televizijski arhivi gradivo skoraj povsod tudi kot t. i. varnostne kopije, ki pa jih ne uporabljajo.

Do leta 2012 naj bi vse evropske nacionalne televizije oddajale televizijski signal izključno v digitalni obliki. V kolikšni meri bomo ta cilj dosegli, je odvisno predvsem od denarja, ki so ga institucije pripravljene vložiti. Po besedah slovenskega ministra naj bi Slovenija ta cilj dosegla že leta 2010.

SUMMARY

USER ADAPTED SELECTION OF CONTENT

User adapted selection of content, known as personalisation, is a procedure, by which the system for the personalised selection of content, among the available contents chooses those that are most suitable for (or most liked by) the user. This procedure is based on the knowledge of users' preferences, which are recorded in the user model. This model is constructed and updated with the help of information about content chosen or not chosen in the past, which is obtained with the analysis of the user - system interaction through the user interface. There are numerous areas of user-adapted search. Among them are searching for documents on the web, filtering of e-mail, searching for news in newsgroups, selection of television programs, searching for articles and even searching for jokes.