ISKUSTVA U KORIŠĆENJU STANDARDA U SISTEMIMA ZA TESTIRANJE PUTEM RAČUNARA

Goran Šimić
E-mail: gshimic@gmail.com
Mladen Veinović
E-mail: mveinovic@singidunum.ac.rs
Aleksandar Jevremović
E-mail: ajevremovic@singidunum.ac.rs

Sažetak: Sistemi za elektronsko testiranje opšte namene predstavljaju efikasno i ekonomično sredstvo za proveru znanja studenata. Rezultati dobijeni ovim sistemima osim informacije o nivou znanja koje su studenti usvojili pružaju i informacije neophodne za unapređivanje nastavnih materijala i procesa. Sa ciljem dobijanja objektivnih rezultata i mogućnosti kolaborativnog rada nastavnika razvijeno je više standarda za elektronsko testiranje. U ovom radu su analizirani popularni standardi za elektronsko testiranje i dati je primer njihovog korišćenja u samostalno razvijenom sistemu za elektronsko testiranje opšte namene.

1. Uvod

Sistemi za testiranje putem računara predstavljaju i dalje aktuelnu problematiku u domenu e-učenja. Iako postoje robusni standardi koji omogućavaju zapisivanje i razmenu materijala provera sa različitim tipovima i formatima problema, malo je uspešnih implementacija. Diferencirale su se dve krajnosti. Jednu grupu čine domenski usko specijalizovani sistemi koji su dizajnirani za naprednije oblike proveravanja. Ovakvi sistemi podsećaju na inteligentne tutorske sisteme, ili predstavljaju njihov deo. Fokusirani su na proveru proceduralnih znanja. Na primer rešavanje zadataka iz medicine [1], ili matematike [2]. Za njih je tipično da se radi o vrlo zatvorenim sistemima koji su uglavnom dizajnirani za samotestiranje studenata u procesu učenja.

Drugu krajnost predstavljaju sistemi za testiranje opšte namene (neki od mnogobrojnih: MCAS [3], FCAT [4], FAIR [5], itd.). Ovi sistemi nemaju generatore problema, rešenja, niti naprednu semantičku obradu materijala. Oni rade po egzaktno postavljenim pravilima, nezavisno od domena primene. Model zadataka je opšti – bez obzira da li se radi o izboru ponuđenih rešenja, unosu kratkih odgovora, numeričkih rezultata, uparivanja reči, ili prevlačenja objekata po radnoj površini, radi smeštanja na određenu lokaciju. U ovim sistemima postoji unapred zadati problem i rešenja, koja mogu biti potpuno, ili delimično tačna. Ocenjivanje se svodi na upoređivanje studentskih odgovora sa sistemskim rešenjima. Sistemi za testiranje opšte namene imaju nekoliko važnih karakteristika: može da se implementira bilo kakav vid testiranja iz bilo koje oblasti; nisu zahtevni po pitanju resursa (memorija i procesorska snaga); zasnivaju se na standardima što im omogućava interoperabilnost.

Fokus ovog rada je na standardima vezanim za elektronsko testiranje i na iskustvima u korišćenju jednog samostalno razvijenog sistema opšte namene.

2. Standardi

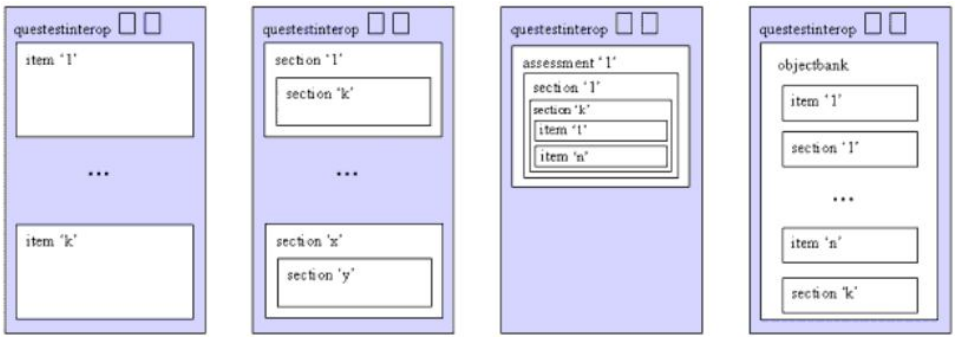
Danas najzančajniji standard za opis podataka elektronskog učenja – SCORM 2004, obuhvata IMS standarda koj se odnosi na proveravanje. To je IMS QTI (Question and Test Interoperability) specifikacija. Ova specifikacija opisuje strukture za razmenu i podataka

testiranja. Osnovni resursi za testiranje koji se razmatraju su pitanja (items) i provere (assessments). Pored toga, IMS QTI standard opisuje i formate za izveštavanje o rezultatima provera studenata.
U implementaciji QTI specifikacije provera koristi se XML jezik kako bi se obezbedila interoperabilnost različitih e-Learning sistema. Tehnička struktura QTI XML zapisa provere generalno se sastoji od dve komponente:

- ASI komponenta
- RR komponenta koja sadrži objekte izveštavanja o rezultatima

ASI komponenta
ASI komponenta formalno i detaljno opisuje podstrukturu (elemente) provera (assessment), sekcija (section) i pitanja (item). Pored navedenih entiteta, postoje i banke pitanja, koje omogućavaju organizovaniju razmenu i višestruko korišćenje istih kompleta pitanja, bilo da su organizovana za sebe, ili grupisana u sekcije. Svaki QTI XML zapis je prepoznatljiv po korenom elementu – questestinterop. Strukture ASI dela dokumenta su različite. To mogu biti vrlo jednostavne do kompleksne strukture za koje je zajedničko da sve mogu da se razmenjuju između IMS QTI kompatiblinih sistema. Tako može da se obavi razmena jednog pitanja (item), sekcije, jedne provere (Assessment), ili banke pitanja.

U kompleksnijim slučajevima (Slika 1), IMS QTI standard dozvoljava razmenu složenih ASI struktura, kao što su grupe izabranih pitanja, sekcija, provere koje sadrže više nivoa struktuiranih sekcija, ili banki pitanja koje u sebi sadrže različite sturkture (pojedinačna pitanja i pitanja grupisana u sekcije).



Slika 1. Složene ASI strukture

Sve ove varijacije treba da pruže korisnicima mogućnost da pronađu odgovarajući oblik formatiranja podataka provera u skladu sa konkretnim potrebama.

Pitanje (Item)

Pitanje (Item) predstavlja osnovni i najvažniji gradivni element IMS QTI strukture. Istovremeno, pitanje (Item) je najmanja nezavisna celina koja može da se razmenjuje u okviru IMS QTI standarda. Jedna stavka (Item) predstavlja jedno pitanje provere. Osnovni koncept specifikacije je Item. Item odgovara jednom pitanju (zadatku) u testu (proveri). Pored vlastitih atributa (naslov, labela i identifikator), Item je struktuiran i sastoji se od pet delova: zadaci, rubrika, prezentacija, procesiranje odgovora i povratna informacija. IMS QTI specifikacija podržava veliki broj tipova pitanja (zadataka) koji se mogu grupisati po sledećem:

- Logički identifikatori - jednostruka i višestruka selekcija ponuđenih odgovora
- Stringovi - unos jednog, ili više odgovora u slobodnoj formi, kratki odgovori

- Numerička pitanja – unos realnog, ili celog broja tekstualno, ili izborom pomoću klizača na skali
- Logička grupisanja – prevlačenje slika u skupove
- XY koordinate - selekcija regiona (grupe objekata) ili tačke u 2D prostoru (slici), povezivanje tačaka.

```
<questestinterop>
  <item ident="RMr_pitanje_1">
    <presentation label="pitanje_1">
      <material>
      <mattext>Token Ring je najrasprostranjeniji tip RMr</mattext>
      </material>
      <response_lid ident="P01" rcardinality="Single" rtiming="No">
      <render_choice>
        <response_label ident="T">
            <material><mattext>Tacno</mattext></material>
        </response_label>
        <response_label ident="F">
            <material><mattext>Netacno</mattext></material>
        </response_label>
        </render_choice>
      </response_lid>
  </presentation>
    <resprocessing>
      <outcomes><decvar/></outcomes>
      <respcondition title="Correct">
        <conditionvar>
          <varequal respident="P01">T</varequal>
        </conditionvar>
        <setvar action="Set">1</setvar>
        <displayfeedback feedbacktype="Response" linkrefid="Correct"/>
      </respcondition>
    </resprocessing>
    <itemfeedback ident="Correct" view="Candidate">
        <material><mattext>Odgovor je tacan.</mattext></material>
    </itemfeedback>
  </item>
</questestinterop>
```

Slika 4. Pitanje sa ponuđenim odgovorima u QTI XML formatu

Na prethodnoj ilustraciji (Slika 4) dat je primer korišćenja ove specifikacije. Koreni čvor je questestinterop u unutar njega se nalaze pitanja u formi Item-a. Prikazan Item sadrži tri opciona elementa:

- presentation sa pitanjem (crveno uokvireno) i ponuđenim odgovorima (2 odgovora, plavo uokvirena),
- resprocessing sa jednim uslovom (Correct), koji u sebi sadrži uslovnu varijablu (ljubičasto uokvirena) povezanu za response_lid element P01 i koja se postavlja na vrednost 1 u slučaju da je odgovor studenta Tacno (identifikator T za element response_label), i
- jednu povratnu informaciju (itemfeedback) za slučaj korektnog odgovora (zeleno uokvirena).

Sekcija (Section)

Sekcija predstavlja grupu pitanja. Često se, umesto termina sekcija, koristi termin blok pitanja. Korišćenjem koncepta sekcije, moguće je grupisanje pitanja po lekcijama, nedeljama, ili tematskim celinama. Sekcija može da bude heterogena struktura, što zanači da može u sebi, pored pitanja, da sadrži i same sekcije. Pored toga sekcija može da referencira druge sekcije koje nisu definisane u istom XML fajlu.

Smisao sekcija može da bude i drukčiji: ista pitanja mogu da budu različito komponovana u zavisnosti od primenjene edukativne i/ili pedagoške paradigme.

Provera (Assessment)

Assessment objekti su analogni testu – proveri. Namena ovog entiteta nije samo da se izvrši agregacija pitanja, već da se definišu instrukcije sistemu za testiranje, za varijabilno sekvenciranje pitanja i vrednovanje rezultata studenata.

Zaključak o IMS QTI specifikaciji

IMS QTI standard je veoma struktuiran, koncepti su opisani sa velikim brojem atributa, mnoštvom elemenata i podelemenata. Pored hijerarhijske strukture, definisane su i poprečne veze između koncepata. Ove veze se ostvaruju posredstvom referenciranja između elemenata opisa. Splet veza unutar paketa znatno povećava njegovu kompleksnost. Specifikacijom nisu definisani niti granice struktuiranja koncepata (dubina stabla). U praktičnoj primeni, ovaj problem zahteva primenu osetljivih rekurzivnih funkcija za pretragu.

Pored navedenog, nedostatak može da bude i činjenica da ne postoji strategija dodeljivanja identifikatora. Jedino je definisano da unutar paketa sadržaja ne budu koncepti istog tipa i sa istim identifikatorom. Time je odgovornost za rešavanje konflikta imenovanja u slučaju kombinovanja resursa iz različitih paketa na strani korisničkog sistema (aplikacije).

3. Drugi formati za struktuiranje pitanja

Kao što je prethodno opisano, IMS QTI predstavlja najreferentniji (W3C) standard za razmenu pitanja između sistema za elektronsko proveravanje znanja. Složenost ovog formata za implementaciju omogućila je opstanak drugh formata za razmenu podataka provera, vezanih za konkretne sisteme i alate elektronskog učenja.

Moodle format

Iako podržava SCORM 2004 standard, Moodle (Modular Object-Oriented Dynamic Learning Environment) [8] kao jedan od najviše korišćenih LMS, zadržao je vlastiti format za uvoz i izvoz pitanja. Osnovni razlog je velika kompleksnost IMS QTI specifikacije – velika količina neinformativnog sadržaja XML fajlova struktuiranih po QTI formatu.

```
1   <quiz>
2    <question type="multichoice">
3       <name><text>RMr_pitanje_1</text></name>
4       <questiontext>
5         <text>Token Ring je najrasprostranjeniji tip RMr?></text>
6       </questiontext>
7          ..
8       <shuffleanswers>true</shuffleanswers>
9       <answer fraction="100">
10        <text>Tacno</text>
11        <feedback><text>Odgovor je tacan</text></feedback>
12      </answer>
13      <answer fraction="0">
14        <text>Netacno</text>
15        <feedback><text></text></feedback>
16      </answer>
17   </question>
     ...
90  </quiz>
```

Slika 8. Prikaz Moodle formata zapisa provere namenjen razmeni između Moolde platform

U prethodna dva primera, u kojima je predstavljena konstrukcija najjednostavnijeg pitanja i izveštavanje o rezultatu studenta na tom pitanju, preko 80% sadržaja (računajući u bajtovima) su tagovi XML elemenata koji moraju da se parsiraju, ali najveći deo njih ne mora da se koristi. U sledećem primeru (Slika 8) predstavljen je XML zapis pitanja koji se koristi u razmeni (izvoz/uvoz) između Moodle platformi. Primetna je mnogo manja kompleksnost u odnosu na IMS QTI format. Deskriptivnost je znatno manja, provera sadrži kolekciju pitanja sa odgovorima u okviru kojih su metapodaci o tipu pitanja, o primenjenim pedagoškim

taktikama (npr, nasumičan redosled pitanja i odgovora, način ocenjivanja/bodovanja, ponderisanje i sl.).
Nedostatak Moodle formata je da nema referenciranja, nema posebno izdvojenh elemenata za izveštavanje, a mogućnost da se pedagoška taktika definiše na nivou pitanja predstavlja potencijalni problem kod velikog broja korisnika koji istovremeno pristupaju rešavanju provere. Procesiranje i priprema pitanja pre isporuke postaje kritični faktor za brzinu odziva sistema.

Hot Potatoes
Jedn od popularnijih alata koji omogućava izgradnju provera putem računara je HotPotatoes [7]. Za razliku od Moodle-a, koji predstavlja integrisanu platformu, HotPotatoes je specijalizovan alat Interesantna je za razmatranje jer se i dalje koristi, iz razloga što je intuitivna, besplatna i zato što je SCORM 1.2 kompatibilna. HotPotatoes je namenjen da se kreira provera, a zatim zapamti kao XML fajl(ovi) u nekom od formata (JQuiz, ili SCORM1.2). Iako JQuiz nije zvanično prihvaćen standard, veliki broj LMS sistema ga podržava. Sledeća ilustracija predstavlja već obrađivani primer pitanja, zapamćenog u HotPotatoes JQuiz formatu (Ilustracija 9).

```
1  <?xml version="1.0" encoding="ISO-8859-1"?>
   ...
3  <hotpot-jquiz-file>
4    <data>
5      <title>RMr_pitanje_1</title>
6      <questions>
7        <question-record>
8          <question>Token Ring je najrasprostranjeniji tip RMr?</question>
9          <weighting>100</weighting>
10         <question-type>1</question-type>
11         <answers>
12           <answer>
13             <text>Tacno</text>
14             <feedback>Odgovor je tacan</feedback>
15             <correct>1</correct>
16             <percent-correct>100</percent-correct>
17           </answer>
18           <answer>
19             <text>Netacno</text>
20             <feedback></feedback>
21             <correct>0</correct>
22             <percent-correct>0</percent-correct>
23           </answer>
24         </answers>
25       </question-record>
26     </questions>
27   </data>
   ...
90 </hotpot-jquiz-file>
```

Slika 9. Deo Hot Potatoe JQuiz fajla u kom se definiše isto pitanje kao u prethodnim primerima

Predstavljeni kod predstavlja jedan deo JQuiz fajla, deo u kome su definisana pitanja. Drugi deo fajla isključivo je namenjen opisu parametara provere (grafička reprezentacija, pedagoške strategije, povratne poruke koje se odnose na proveru u celini i sl.). XML struktura omogućava preglednost fragmenta, na kome se jasno ističe pitanje i odgovori sa bodovima i povratnim porukama. Jednostavnost zapisa u odnosu na IMS QTI specifikaciju unosi mnoga ograničenja: Nemogućno je spoljnje referenciranje pitanja, jer ne postoje identifikatori. Sadržaj pitanja kao i odgovora je isključivo tekstualan, izuzev ako se referencira slika koja je dostupna posredstvom lokalne, ili globalne mreže. Format koji je predstavljen čitljiv je za Moodle platformu (uvoz pitanja iz fajla, sa konverzijom pre upotrebe).

4. Iskustva

U nastavku rada dati su neku od detalja implementacije i evaluacija nezavisno razvijenog sistema za elektronsko testiranje MTutor. Kod realizacije pomenutog sistema ispitivao se

odziv sistema pri korišćenju različitih standarda, na različite načine, kao bi se odabrao najoptimalniji pristup.

Osnovni kriterijumi pri dizajniranju sistema zahtevali su visok stepen pouzdanosti sistema, visoke performanse u pogledu broja paralelnih korisnika, mogućnost uključivanja grafičkih elemenata i jednostavnost upotrebe. Pri izboru standarda za organizovanje i kodovanje baze pitanja analizirani su IMS QTI i Moodle XML standardi. U fazi istraživanja sadržaj pitanja je kodovan u ovim formatima i smeštan u bazu podataka. Za oba formata su merene performanse sistema pri radu sa nekoliko hiljada pitanja u čiji su sadržaj bile uključene grafičke ilustracije.

Na osnovu rezultata merenja izabran je podskup QTI srukture u koju su uključeni elementi koji su se pokazali kao najčešće korišćeni u razvoju baza pitanja za testiranje studenata. Struktura je implementirana direktno u bazu podataka a zadržana kompatibilnost sa IMS QTI i Moodle XML standardima omogućila je ugrađivanje funkcija za izvoz i uvoz baza pitanja u ovim formatima. Za potrebe ostvarivanja željenih performansi sistema isključeni su određeni elementi. Zahtevani nivo performansi je postignut a rezultati evaluacije sistema dati su u nastavku.

Rezultati

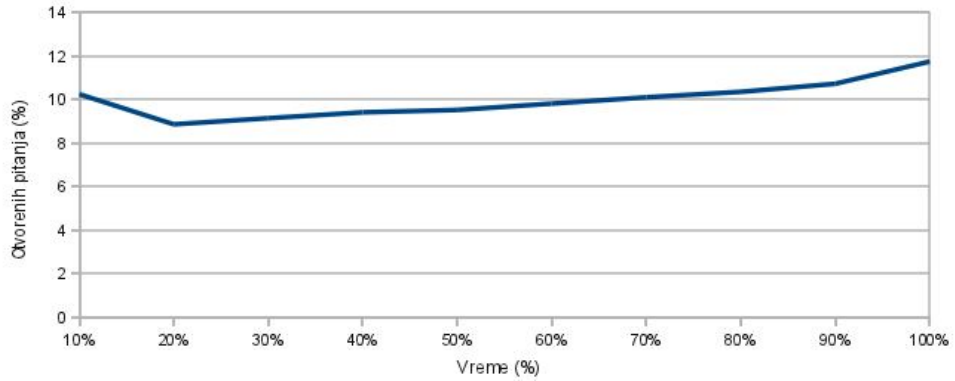
U školskoj 2009/2010 godini sistem za elektronsko testiranje korišćen je prema sledecem prikazu (tabela 1) na Fakultetu za Informatiku i Računarstvo.

Tabela 1. Podaci u bazi sistema za testiranje za školsku 2009/2010 godinu

Broj predmeta	33
Broj oblasti	222
Broj pitanja	7.015
Ukupan broj odgovora	27.197
Broj provera znanja	367
Broj testiranja studenata	6.381
Broj testiranih studenata	10.96
Broj otvorenih pitanja	245.939
Broj datih odgovora	364.872

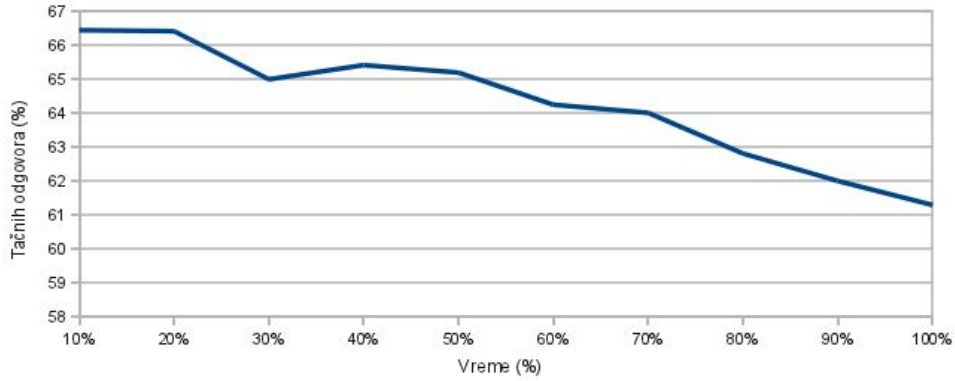
Prosečno trajanje provere znanja je nešto manje od 90 minuta. Prosečno trajanje pojedinačnih studentskih radova je oko 11,5 minuta (696,29 sekundi), prosečan broj otvorenih pitanja po radu je oko 39 a prosečan broj datih odgovora oko 57. Prosečno vreme za davanje odgovora na pitanje je približno 18 sekundi.

Intenzitet eksploatacije sistema je neravnomeran usled periodične prirode proveravanja (kolokvijalne nedelje i ispitni rokovi): skoro 50% ukupnog vremena (oko 555 sati) - utrošeno je za svega 7 sati testiranja. To znači da je u tom periodu sistem intenzivno korišćen sa oko 80 paralelnih pristupa studenata. U to vreme je centralni server isporučivao oko 267 pitanja u minutu (oko 4,5 pitanja u sekundi) i za njih prihvatao odgovore koje su studenti izabrali. Prosečna raspodela rešavanja pitanja u toku izrade testa je skoro uniformna. (slika 10). Nakon prve petine vremena, postoji ravnomerni priraštaj broja otvaranja do zaključivanja provere. To znači da studenti prilagođavaju brzinu izrade testa preostalom vremenu.



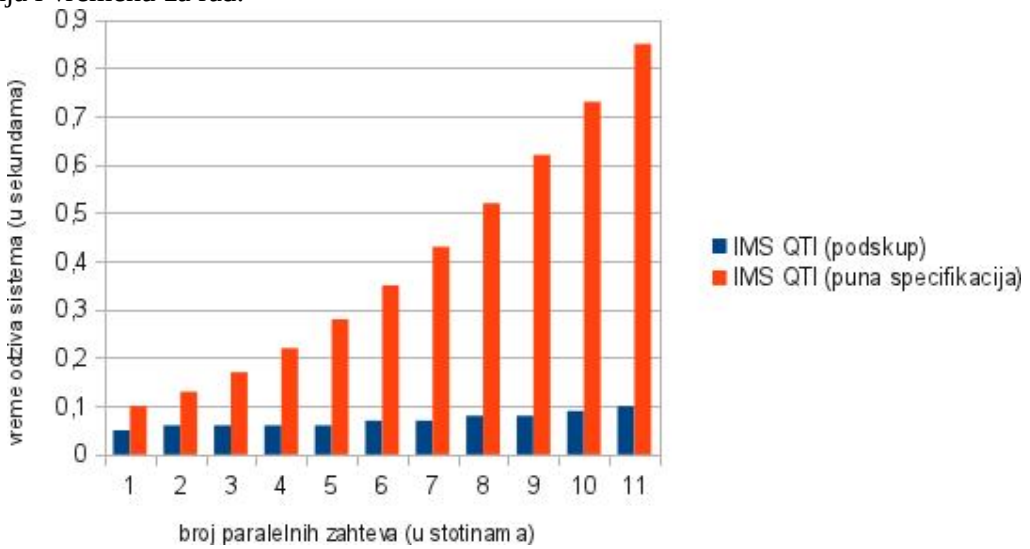
Slika 10. Raspodela otvaranja pitanja na trajanje rada

Procenat davanja tačnih odgovora blago opada tokom izrade testa (slika 11). Najviši procenat tačnosti beleži se u prvih 20 procenata vremena (66,45%), dok najnižu tačnost čini 61,29 procenata u poslednjih 10 procenata vremena izrade testa.



Slika 11. Raspodela tačnih odgovora na trajanje rada

Opisani intenzitet korišćenja zahteva precizno projektovanje sistema za elektronsku proveru znanja u cilju postizanja neophodnog niova performansi i najvišeg nivoa pouzdanosti. Na osnovu ovog podatka, u korisnički interfejs su uključene informacije o preostalom broju pitanja i vremenu za rad.



Slika 12. Implementacija pune IMS QTI specifikacije negativno se odrazila na performanse

S obzirom da je vreme za rad studenata ograničeno, nastojalo se minimizovati vreme odziva. Korićenjem navedenig podskupa IMS QTI specifikacije postignuto je vreme manje od 0,1 sekunde za celokupan proces prelaska na sledeće pitanje (slanje odgovora putem mreže, prihvatanje odgovora, kreiranje reprezentacije sledećeg pitanja, slanje reprezentacije putem mreže) za 100 paralelnih korisnika koji koriste komunikacione kanale sa propusnom moći od 100Mb/s (Slika 12). Upotrebom hardverske platforme poslednje generacije opisani nivo performansi postignut je sa oko 30 procenata zauzeća sistemskih resursa, prvenstveno centralnog procesora.
U cilju onemogućavanja zloupotreba i manipulacije rezultatima, sistem koristi centralizovanu arhitekturu (aplikacija i skladište na istoj serverskoj platformi) sa minimalnim delegiranjem odgovornosti na Web klijente tako da je njihova jedina uloga da renderuju korisnički interfejs i prosleđuju podatke serveru. Sve aktivnosti studenata beleže se tako da je moguća detaljna postprocesna analiza.

Zaključak

Zahvaljujući ugrađenosti u SCORM 2004, IMS QTI predstavlja referentni standard za struktuiranje i predstavljanje podataka za testiranje i rezultata testiranja. Njegova sveobuhvatnost (podršaka za različite vrste pitanja) u velikoj meri je uticala na kompleksnost standarda za implementaciju. Veliki broj meta – podataka (tagova) i mogućnost neograničene dubine strukturiranja, najveći su problemi sa kojima se susreću implementatori standarda. Stablasta struktuiranost sekcija zahteva implementaciju korišćenjem rekurzivnih funkcija, što može dovesti do prekoriačenja veličine podataka u RAM (na *stack*-u). Kao posledica, većina aplikacija za kreiranje testova i njihovu realizaciju, pored podrške za QTI standard, razvila je vlasite interne formate, davajući mogućnost uvoza, izvoza i konverzije u sve formate.

Na isti način pristupilo se realizaciji sistema za elektronsko testiranje MTutor. Za potrebe ovog sistema razvijen je podskup originalne IMS QTI strukture za organizovanje i kodovanje baze pitanja, kompatibilan sa popularnim javnim standardima, prvenstveno sa IMS QTI i Moodle XML standardima. U cilju postizanja odgovarajućeg nivoa performansi i jednostavnosti korščenja u samostalno razvijenu specifikaciju nisu uključene neke od funkcija referentnih standarda. Željeni nivo performansi je postignut a sveukupna efikasnost razvijenog sistema i izabranih pristupa pozitivno je ocenjena na osnovu realnih podataka dobijenih u toku jednogodišnje praktične primene sistema na Fakultetu za Informatiku i Računarstvo.

Literatura

- O'Callaghan, T., Popple, J., McCreath, E.: *SHYSTER-MYCIN: A Hybrid Legal Expert System,* dostupan preko http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1335659_code814811.pdf
- Aleven, V., McLaren, B., Roll, I., Koedinger, K. *Toward Meta-cognitive Tutoring: A Model of Help-. International*, Journal of Artificial Intelligence in Education 16 (2006). 2006. str. 101 - 130
- Sistem za elektronsko testiranje MCAS, dostupan preko http://www.doe.mass.edu/mcas/
- Sistem za elektronsko testiranje FCAT, dostupan preko http://fcat.fldoe.org/default.asp
- Sistem za elektronsko testiranje FAIR, dostupan preko http://www.justreadflorida.com/instrreading.asp
- IMS Global Learning Consortium, link http://www.imsglobal.org
- Hot Potatos, http://hotpot.uvic.ca/
- Sharable Content Object Reference Model, http://www.adlnet.gov/