

**ИКТ (ИНФОРМАЦИОНО-КОМУНИКАЦИОНЕ ТЕХНОЛОГИЈЕ) И ЋИРИЛИЦА
ICT (INFORMATION AND COMMUNICATIONS TECHNOLOGY) AND CYRILLIC**

Драгољуб Пилиповић
Факултет за информационе технологије, Слобомир П Универзитет, Бијељина
Жељко Гаврић
Факултет организационих наука, Универзитет у Београду, Београд

РЕЗИМЕ: Овај рад ће почети од настанка ћириличног писма, те ће се одмах прећи на данашње стање. Онда ће се описати који су то електронски уређаји који (не)могу користити споменуто писмо и у ком облику и под којим условима. Услови су дефинисани разним не-хардверским концептима: софтвер, базе података, веб странице, кодовање итд. Посебна пажње ће се посветити кратким порукама на мобилним телефонима.

КЉУЧНЕ РЕЧИ: ИКТ, ћирилица, кодна страница, UTF, смс поруке.

ABSTRACT: Paper will start from the creation of the Cyrillic alphabet, then will immediately move on to the present situation. It will then describe what the electronic devices (not) be used Cyrillic letters and in what form and under what conditions it use. Conditions are defined in a ways of non-hardware concepts: software, databases, web pages, coding, etc. Special attention will be given to short messages on cell phones.

KEY WORDS: ICT, Cyrillic, code page, UTF, SMS.

1. УВОД И ЦИЉ РАДА

Ћирилица је једно од основних особина словенских народа. Јер већина тих народа је користи. Такође српска ћирилица је једна од основних особина српског народа. То писмо у садашњем савременом облику створено Вуковим напорима има скоро јединствену особину на свету – тежи да буде потпуно фонетско. Овим радом желим потакнути размишљање, расправу и делање по овом питању. Тежиште овог рада је примена ћирилице (нарочито српске ћирилице) у информационо-комуникационим технологијама (ИКТ), које се налазе свуда око нас.

Почеће се од настанка ћириличног писма, те ће се одмах прећи на данашње стање. Онда ће се описати који су то електронски уређаји који (не)могу користити споменуто писмо и у ком облику и под којим условима. Услови су дефинисани разним не-хардверским концептима: софтвер, базе података, веб странице, кодовање итд. Посебна пажње ће се посветити кратким порукама на мобилним телефонима. Овај рад је испирисан манифестом Милића од Мачве из фебруара 1962. године и посвећен је њему и српској ћирилици. [1]

2. ИСТОРИЈА ЋИРИЛИЦЕ

Писмо `ћирилица` је добило свој назив по Ћирилу (грчки: Κύριλλος, црквени старословенски: КѢрилль), једном од двојице браће познатих као Солунска браћа и Словенски апостоли. Најмлађи од седморо браће, рођен је у византијском Солуну као Константин 827. године, а замонашио се у Риму мало пре смрти 869. године, добиши у том име Ћирило што изворно значи Господњи. Православна и римокатоличка црква на нашим просторима су браћу Ћирила и Методија уврстила у календаре за спомен и прослављање.

Најпознатији допринос из лингвистичке области Ћирило је дао у формирању првог словенског, па тиме ујед-

но и српског писма – глагољице. Ово писмо се користило самостално прве две-три деценије од настанка у другој половини 9. века, па онда упоредо са ћирилицом, новим писмом које је „савршеније, практичније и за српски језик подесније“ [2]. Глагољица се још дуго користила само на јадранским острвима, у Далмацији и Хрватској, чак до деветнаестог века. Однос глагољице и ћирилице није још увек до краја утврђен. Додатно, питање о постојању и поужење са винчанским писмом овде нисмо изводили.

Претпоставља се да је првобитна ћирилица настала кроз рад ученика Ћирила и Методија. После Методијеве смрти (885.) морали су да напусте западнословенске крајеве и оду на југ, добар део њих под заштиту бугарске власти. Једно од битних места јесте Преславска књижевна школа, где су између осталих радили Наум Охридски и Црноризац Храбар. Друга школа је била у Охриду, где се спомиње Климент Охридски и брат му Наум. На тим просторима су нађена најстарија датирана документна на првобитној ћирилици, па се са великом вероватноћом претпоставља да је на тим просторима и настала.

Ћирилица је настала од грчког унцијалног алфабета. Унцијални алфавет, у нашој терминологији уставни алфавет, је начин писања слова као да су у питању само велика слова (типографски термин за такав изглед и писање слова је верзал или мајускула). Словима из иницијалног грчког алфабета су додата слова за оне гласове које није има грчки језик; она су углавном преузета или прилагођена из глагољице. [3]

Могло би се рећи да су Срби на кратко били у стању двоазбучја (глагољица и ћирилица). Међутим, ћирилица се у српском језику користи више од 10 векова. Ипак, то није потпуно иста ћирилица. Временом је дошло до диференцијације старословенског језика у три редакције: српску, бугарску и руску. Писмо за руску редакцију је промењено одлуком Петра Великог у 18. веку и задњи пут после Октобарске револуције. Бугарско писмо је делимично пратило промене у руској ћирилици, да би 1945. године било реформисано задњи пут.

Тадашња српска ћирилица није била погодна за записивање простог, народног језика, већ за неки од редакција словенског језика који су повремено или упоредо били у употреби (српскословенски, словеносрпски, рускословенски). Поседовала је 42 знака и са њом су се добро служило високо свештенство и образовани појединци.

Први покушај реформе писма је у време барока предложио Гаврило Стефановић Венцловић (1680. — 1749.), свештеник и писац, који је писао проповеди на народном језику и превео скоро цело Свето писмо. Предлагао је за употребу три слова у српској ћирилици: ђ, ђ и џ.



Слика 1. Насловна страна Мркаљеве књижице

Учитељ и монах Сава Мркаљ (1783. – 1833.) је значајан по томе што је у својој програмској књижици „Сало дебелог јера или азбукопротрес“ први предложио радикалан раскид са постојећим писмом. Предложио је избацивање већине непотребних слова, те би се тако са првобитних 42 дошло до 26 знакова потребних за српски језик.

На његов рад се директно надовезује Вук Караџић (1787. - 1864.), који користећи крилатицу „пиши као што говориш, а читај као што је написано“ немачког филолога Јохана Аделунга, реформише српски језик али и српско писмо – ћирилицу. Задржао је 24 уобичајена слова (а, б, в, г, д, е, ж, з, и, к, л, м, н, о, п, р, с, т, у, ф, х, ц, ч, ш), а додао још 6. Слово ђ је прихватио од образованог владике Лукијана Мушицког, слово ђ је пронашао у старим српским рукописима, а слово џ у румунским рукописима, док је слово ј преузео из латинице. Осмислио је два нова слова, стапајући н и л са њ (полуглас за умекшавање) и добио слова њ и њ. Његова реформа је била дуго оспоравана због превише радикалног приступа, посебно због слова ј, али је на крају то постало писмо којим пишемо 150 година. Једноставност и економичност правила „један глас, једно слово“ је била неумољива у овом случају.

3. САДАШЊА УПОТРЕБА ЋИРИЛИЦЕ

Данас ћирилицу користе углавном словенски језици и језици разних египћких група на простору бившег Совјетског Савеза.

Користе је руски, белоруску, украјински, русински, бугарски, српски и македонски језик. Такође је користе сви језици бројних несловенских народа у Руској Федера-

цији, који припадају следећим групама језика: индо-аријански, ирански, романски, кавкаски, кинески, монголски, чукотско-камчатски, тунгуски, туркијски, уралски, ескимско-алеутски и други. Државе у којима је ћирилица службено писмо су следеће:

- 1) Русија,
- 2) Украјина,
- 3) Белорусија,
- 4) Киргизија,
- 5) Казахстан,
- 6) Таџикистан,
- 7) Узбекистан,
- 8) Монголија,
- 9) Бугарска,
- 10) Србија,
- 11) Македонија,
- 12) Црна Гора и
- 13) Босна и Херцеговина.



Слика 2. Распрострањеност ћирилице данас [5]

Из свега предходно приказаног може се „извести поуздан закључак да је ћирилица евро-азијско писмо“. [6]

Ко је и када напустио коришћење ћирилице? Румунија је то урадила до 1860. године за румунски језик иако је он од свог првог записаног појављивања у шеснаестом веку био заснован на ћириличном писму. Азерски језик је више пута мењао своје основно писмо. Тако се у Азербејџану ћирилица користила у периоду од 1938. до 1991., када су прешли на (нову) латиницу. Туркменски језик има сличну историју као и азерски: оба припадају породици туркијских језика, прво су користила арапско писмо, па латинично, онда ћирилицу, да би после осамостаљења 1991. године прешле на латиницу која је прављена по угледу на садашње турско писмо. Ипак, „совјетска“ ћирилица се доста користи, нарочито је користи политичка опозиција у својим штампаним и електронским издањима да би нагласила разлику од власти. Узбекистан и узбечки језик који припада ујгурској групи је прошао кроз сличне мене, али је употреба ћирилице још распрострањена иако се латиница уводи на све стране, нарочито кроз школство и медије. Ипак у ове три споменуте земље се због укидања ћирилице одједном губи историјски континуитет, јер се у њиховој укупној историји највише штампаних ствари појавило у прогресивном времену током Совјетског Савеза тј. током ћириличног раздобља.

Број различитих слова (графема) свих постојећих ћирилица је негде око 150 [7]. У табелама испод може се видети скуп слова из најраспрострањенијих (словенских) језика која користе ово писмо:

Руски	А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й	К	Л	М	Н		
Белоруски	А	Б	В	Г	Д	Е	Ё	Ж	З	І	Й	К	Л	М	Н		
Украјински	А	Б	В	Г	Ґ	Д	Е	Є	Ж	З	И	І	Й	К	Л	М	Н
Бугарски	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н			
Српски	А	Б	В	Г	Д	Ђ	Е	Ж	З	И	Ј	К	Л	Љ	М	Н	Њ
Македонски	А	Б	В	Г	Д	Ѓ	Е	Ж	З	И	Ј	К	Л	Љ	М	Н	Њ

Табела 1. Упоредни списак словенских ћириличних писама

Руски	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	33	
Белоруски	О	П	Р	С	Т	У	У	Ф	Х	Ц	Ч	Ш	Щ	Ы	Ь	Э	Ю	Я	32	
Украјински	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	33	
Бугарски	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	30	
Српски	О	П	Р	С	Т	Ђ	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	30
Македонски	О	П	Р	С	Т	Ѓ	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	31

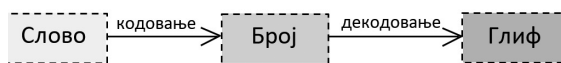
Табела 2. Упоредни списак словенских ћириличних писама, наставак слова

Када је Бугарска држава приступила у пуноправно чланство у Европску Унију 2007. године, ћирилица је постала треће по реду официјелно писмо, поред латинице и грчког алфабета, које се користи као званично за писање докумената. Такође постоји и евро новац са ћириличним натписима.

4. КОДОВАЊЕ

Кодовање (кодирање, encoding) представља скуп правила који успостављају једнозначну везу између података и њиховог приказа у дискретном облику, попут знакова или симбола [8]. Циљ кодовања је представљање података у погодном облику за запис, преношење, обраду и приказ. Кодовање преводи неки податак (нпр. слово) у неки други облик – код (нпр. бројчану вредност) који представља његову замену. Декодирање изводи обрнут поступак. Обично се кодови и припадајући подаци дају у облику речника односно табеле са два ступца. Да би се учесници у комуникацији разумели међусобно, потребно је да имају исти речник са кодовима. Пример опште познатог кода је Морзеова азбука, где су слова замењена са цртама и тачкама.

За потребе овог рада даље ћемо разматрати само део кодовања које се односи на кодовање слова. Такво кодовање преводи слово у број, број се записује у меморију или се преноси преко комуникационог медијума, те се по декодовању слово приказује на екрану (тзв. глиф) или се оно даље штампа на папир.



Слика 3. Дијаграм словног кодовања

Прва шема за кодовање слова на електронским комуникационим уређајима попут телепринтера је био ASCII (American Standard Code for Information Interchange), који користи 7 бита (нула и јединица) за представљање свих слова енглеског/латинског алфабета, цифара, знакова за интерпункцију и специјалних карактера. Са 7 бита се може представити 27 комбинација односно 128 слова и других знакова, нпр. слово 'А' је представљено са нумеричком вредношћу 65 (бинарни еквивалент је 100 0001), слово 'В' је 66, слово 'С' је 67 итд.

Ипак за оперисање на рачунарима је погодније користити бајт, који се састоји од 8 бита. Тако је ASCII код доживео проширење за један бит, те је сад могао представити 28 комбинација. У тих 256 комбинација могла су стати сва западноевропска латинична слова. Услед ширења броја личних рачунара на послу и по домовима, настала је потреба да се са овим проширеним ASCII кодом представе друга писма. За ту потребу су осмишљене кодне странице, које су постојале за разне језике и писма. Ћирилична слова су била покривена у кодној страници 855 (друга имена за ову страницу су CP 855, IBM 00855, OEM 855, MS-DOS Cyrillic). Ту кодну страницу нису користили Источни Словени, мало више су је користили Јужни Словени за своје ћирилице. Сличну судбину је имала Мајкрософтова кодна страница 1251, која се користила под њиховим оперативним системом Windows. У Русији се у то време користио највише KOI8-R код, али је он био прилагођен само руским словима па зато није био раширен другде.

Није постојао ауторитет у овом домену рачунарства да би стандардизовао кодне странице, те су се оне недоследно користиле и документи са ћирилицом су се тешко преносили између рачунара. Ипак највећи проблем је било ограничење да се у једном документу може користити само једна кодна страница. Није могуће било написати (или није било лако и општег решења) документ на два различита писма. Зато се приступило изради универзалног кодног стандарда.

Уникод (Unicode) тежи да буде кодни стандард за све знакове, свеједно да ли су некад били у употреби, тренутно су у употреби или ће тек бити у употреби. Има потенцијал за милион знакова, од којих је тренутно искоришћено око 22% односно око 250 хиљада позиција, где је за око 100 светских писама постављено око 110 хиљада слова. [9] Стандард прописује непрофитна организација Unicode Consortium (unicode.org), а 2012. године је издата последња верзија 6.1 овог стандарда. [10] Сличан стандард само мањег обима је Универзални карактер скуп (UCS) Међународне организације за стандардизацију ISO 10646.

Уникод уводи један додатни слој апстракције у кодовање, тако што слова кодује у кодне тачке. Он не кодује слово у конкретан број записан у меморији, већ повезује слово са апстрактном вредношћу – кодном тачком. Кодне тачке се даље кодују у неки конкретан запис на меморијском медијуму, што одређује заузеће меморијског простора. (Кодне равни нећемо овде спомињати због једноставности излагања.)



Слика 4. Дијаграм словног кодовања са Уникод и UTF кодовањем

Да би се до краја завршило кодовање потребно је кодну тачку превести у број који ће се записати на меморијски медијум или ће бити прослеђен кроз комуникациону мрежу. То другостепено кодовање изводи се са UTF (Unicode Transformation Format) превођењем. UTF је стандард за

превођење свих Уникод знакова, постоје варијанте UTF-8, UTF-16 и UTF-32. Најзаступљенија варијанта је UTF-8 са променљивом дужином записа у меморији од једног до четири бајта за једну кôдну тачку. Ако је у питању текст са знаковима само из ASCII табеле, онда се користи само један бајт по знаку/кôдној тачки. У случају додатних слова из Гајеве латинице (č, ć, š, ž, đ) или било ког слова српске ћирилице, користе се два бајта. Кинеско (али и јапанско, корејско) писмо захтева бар три а некад и четири бајта. [11]

Можемо рећи да стања пре и после Уникод кодовања (слово, кôдна тачка) су апстрактни делови овог система, они могу бити записани на обичан папир или замишљени као идеје. UTF кодовање има за резултат конкретан запис на физички медијум и то је оно што одређује величину записаног текста. Графички приказ слова се добија после Уникод и UTF-8 подфаза декодовања и на екрану се види као глиф. На пример, слово 'Буки' (старословенско име) је има U+0411 кôдну тачку, његов UTF-8 запис има два бајта 0xD0 0x91, а у неком програму за приказ текста то слово овако изгледа: Б.

По Уникод стандарду постоји укупно 434 кôдних тачака за знакове који се могу повезати са свим историјским и садашњим ћирилицама. То укључује мала и велика слова, специфичне знакове за интерпункцију, ознаке за бројеве и друге ознаке за модификацију слова. [12]

По ISO 639 међународном стандарду српски језик има двословну ознаку sr и трословну ознаку srp. [13] Стандард ISO 15924 за ћирилично писмо одређује четворословну ознаку Cyrl и троцифрени број 220. [14]

5. ХАРДВЕР И СОФТВЕР

Подршка ћирилице у хардверу се огледа у различитим хардверским уређајима. Првенствено је ту у питању лични рачунар, али и преносиви рачунари, таблети и мобилни телефони. Лични и преносиви (лаптоп) рачунари скоро да немају функционална разлике; таблети су на пола пута између првих и мобилних телефона а имају функционално доста сличне делове али обично немају физичку тастатуру. Тај недостатак све чешће прате и мобилни телефони.

Код личних рачунара ћемо прво раздвојити битне хардверске делове за подршку ћирилице. У излазне уређаје спадају монитори и штампачи, који могу да прикажу било које слово јер формирају слова са тачка-по-тачка техником (растерска графика). Код обрадног подсистема, сваки данашњи процесор је (бар) 32-битни па може да обрађује Уникод слова. Меморије, у које спадају радна меморија, тврди дискови, оптички медијуми и флеш меморије, имају довољно капацитета да похране било који ћирилични документ. Ипак треба узети у обзир да при коришћењу Уникод записа, документи са ћириличним словима заузимају двоструко више простора него документи са истим бројем слова из енглеске латинице. Код излазних уређаја имамо тастатуре са налепљеним словима које се код нас продају у три варијанте: са енглеском латиницом, Гајевом латиницом и српском ћирилицом. Друга варијанта је ви-

шеструко заступљенија од осталих. Гајева латиница лажно задовољава услове локализације на српски језик јер не подржава ћирилицу, већ се куцају еквивалентна слова уз додатак да слово „q“ замењује „љ“, слово „w“ замењује „њ“ и слово „x“ замењује „џ“. Не постоји обавеза увозника да локализују тастатуре на ћирилично писмо, иако таква операција је веома јефтина (десетак ситних налепница).



Слика 5. Тастатура са српском азбуком

Скоро све споменуто важи за таблети и мобилне уређаје. За уређаје који немају физичку тастатуру, подршка ћирилице зависи од оперативног система (као што ћемо видети, софтверске тастатуре имају одличну подршку независно од врсте и верзије оперативног система).

Основни софтвер без кога не може функционисати нити један хардверски уређај је оперативни систем. Овде ће бити споменути Мајкрософтови Windows 7 и Windows 10, Linux Ubuntu 20.04, Android и фабрички оперативни системи за мобилне телефоне. Тиме ће бити покривени скоро сви случајеви коришћења хардверских уређаја код нас. Код оба Мајкрософтова оперативна система постоји локализација у три нивоа. Први ниво је подешавање културе, која се подешава кроз бирање језика (понекад државе) плус додатно прецизирање. За ћириличне културе у верзији 7 и 10 се спомињу: Serbian (Cyrillic, Serbia), Serbian (Cyrillic, Montenegro), Serbian (Cyrillic, Bosnia and Herzegovina) и Bosnian (Cyrillic, Bosnia and Herzegovina). Подешавање мења начин исписа и обраде бројчаних вредности, новчаних износа, времена и датума за цели оперативни систем. Други ниво локализације је избор језика и припадајуће тастатуре за унос текста. За верзију 7 и 10 постоје исте оне опције као код првог нивоа локализације. Трећи ниво је превод корисничког интерфејса. Верзија XP има потпуни превод на српски језик исписан ћирилицом, док верзија 7 и 10 се може прво купити на Гајевој латиници и српском језику (или на неком другом општем језику, нпр. енглеском) па онда на њу додатно инсталирати језички интерфејсни пакет (LIP) са ћириличним писмом који садржи превод најфреквентнијих делова корисничког интерфејса. Фонтови служе за тачан приказ слова на екрану или папиру. Ипак, неки фонтови немају подршку за српску ћирилицу.

Оперативни систем Linux Ubuntu је бесплатан и слободан за употребу, јер свако може да има у виду шта изнутра чини њега (open-source). У њему, слично као у Мајкрософтовом систему, се одвојено подешавају тастатура (постоје за Serbian језик варијанте Serbia и Montenegro Cyrillic), формати за испис бројева, датума и новчаних износа (постоје `српски – Србија` и `српски – Црна Гора`) и на крају превод свих натписа (постоје `српски`, `српски – Србија` и `српски – Црна Гора`). Превод је потпун за оперативни систем али чак и за накнадно инсталиране програме попут

OpenOffice-a/LibreOffice-a, јер сви они деле базу података са преводом истих ставки у менијима а остатак превода се посебно инсталира. Пакет са преводом на српски језик се скида са Интернета, док се језик бира при пријави на систем из листе инсталираних језика којих може бити већи број.

Android је сличан Linux Ubuntu оперативном систему јер има исту основу. Налази се на већини таблет рачунара и тзв. паметних мобилних телефона. Они углавном немају физичку тастатуру већ се по потреби приказује софтверска тастатура на екрану осетљивом на додир. Постоје бројне верзије (4.*, 5.*, 6.*, 7.*, 8.*, 9.*, 10.*) које се брзо испоручују на тржиште и упоредо постоје и продају. Код свих верзија постоји могућност коришћења већ уграђене или накнадно инсталиране ћириличне тастатуре. За неке уређаје постоји локализација на српски језик потпуно, али има доста и случајева када у питању српски језик исписан Гајевом латиницом. Мобилни оператери у Србији и већи увозници мобилних уређаја мисле да је довољна локализација на овај начин, јер их иначе закон на то не обавезује а корисници ретко када траже превод са ћириличним словима.

Код обичних мобилних телефона са фабричким оперативним системом у најбољем случају локализација је урађена са Гајевом латиницом и могу се укуцавати та слова у порукама. Иначе сви савремени мобилни телефони могу примити и приказати слова српске ћирилице.

Документи попут формата .doc, .docx, .xls, .xlsx, .ppt, .pptx, .mdb, .acscdb, .pdf итд. већ дуже од две деценије могу садржати српску ћирилицу. За неке од њих постоје алати за пресловљавање тј. пребацивање текста из српске ћирилице у Гајеву латиницу и обрнуто. При томе, популарни Мајкрософт канцеларијски алат Офис у верзијама 2007, 2010 и 2016 има ћириличну локализацију.

6. ИНТЕРНЕТ

Интернет представља рачунарску мрежу података светских размера, која је свима јавно доступна. На њој се пружају разне услуге, попут електронске поште и хипертекстуалних докумената. HTML (Hyper Text Markup Language) је формат у коме се описују и записују такви документи, HTTP (Hyper Text Transfer Protocol) представља протокол за размену а сами документи се идентификују са URI (Unified Resource Locator) именом. У HTML описном језику се ћирилични текст може записати са разним врстама кодовања (већ спомињаним раније). Софтвер за приказ HTML докумената (web browser) понекад не прикаже исправно слова па уместо њих исписује упитнике, празне квадратиће и сл. Разлог је непрепознавање врсте кодовања, па се она у опцијама мора подесити на одговарајућу. Ипак најчешће се користи UTF-8, што се поставља додавањем ознаке `charset="utf-8"`. Сама слова се записују у форматама `�` и `�`, где је `aaaa` број кодне тачке по Уникод стандарду а `oooo` исто то записано хексадекадним бројним системом. То повлачи за собом чињеницу да су HTML странице са ћириличним текстом веће од оних са западноевропским латиницама, али је то занемариво

повећане јер су капацитети размене преко интернета далеко већи од тог увећања. Пример: слово `ћ` је записано као `Ћ`. Платформа Flash фирме Adobe намењена интерактивним и мултимедијалним презентацијама унутар HTML страница подржава такође UTF-8 запис.

Колико је подржана српска ћирилица код најпопуларнијих и најфреквентнијих сајтова на интернету, можемо да видимо да је гигант у тој области – Гугл – користи подразумевано односно без икакве додатне акције од стране корисника када приступате из Србије или као јасно уочљива опција када приступате из БиХ. То се односи на бројне сервисе које нуди Гугл (поред оног основног за претрагу на Google.com страници): за електронску пошту Џимејл, за документе и колаборацију Документи, интерактивне дигиталне Гугл мапе, претраживање научно-истраживачких и стручних радова Гугл академик, вођење јавних дневника Блогер и доста других. Од њихових познатих подружница једино YouTube сервис за видео садржај нема србизовано име. Један од највећих сајтова за комуникацију и дружење – Фејсбук – нуди такође добро локализован превод на српску ћирилицу.

Сајтови државне управе у земљама у којима се претежно говори српски језик морали би имати испис ћирилицом али ту постоје велике разлике:

- i. Подразумевани испис ћирилицом: Република Србија – председник, скупштина, влада, бројна министарства и војска; Република Српска – председник, скупштина, влада и унутар ње бројна министарства; АП Војводина – скупштина и влада.
- ii. Мора се изабрати писмо на почетку: Босна и Херцеговина – председништво и оружане снаге; Федерација Босне и Херцеговине – ентитетска влада;
- iii. Ћирилица је једна од додатних опција: Република Црна Гора – влада и унутар ње министарства; Босна и Херцеговина – парламентарна скупштина и савет министара; Дистрикт Брчко – влада и скупштина; влада Тузланског и Средњобосанског кантона Федерације БиХ.

Сајтови делова државне управе који нису поменути (попут владе и скупштине привремених органа АП Косова и Метохије, осталих кантона у БиХ или председника Црне Горе) немају подршку ћирилице.

Системи за управљање садржајем (CMS, Content Management System) на Интернету постоје да би се на лакши, аутоматизован начин постављали садржаји попут текста и слика, прилагођено знању и вештинама просечног корисника рачунара. Постоје за разне области па и њихове карактеристике од тога зависе. Најпопуларнији систем опште намене је Joomla, која има превод на српски ћирилицом. Исто је и са најпопуларнијим системом за вођење блогова Вордпрес, системом за дискусионе форуме - Simple Machines Forum (SMF), системом за сакупљање знања и колаборациони рад – MediaWiki (на којем ради и чувена Википедија), као и системом за електронско удаљено учење – Мул.

Базе података које се израђују у производима фирми Мајкрософт (SQL Server 2019 и раније), Oracle (Database 19c и бесплатни MySQL 8.0), IBM (DB2) и Postgres (бесплатни PostgreSQL 13.1), имају могућност смештања података са српским ћириличним писмом одавно, најчешће по UTF-8 стандарду али и по другим. Подаци који се смештају у XML датотеке као вид база података чак ако се не нагласи подразумевано користе UTF-8.

Сајтови на Интернету имају своја лако препознатљива имена, тзв. доменска имена, скраћено домен (нпр. председник.срб). Део доменског имена на самом десном крају преко кога се одређује земља се зове домен највишег нивоа. Те домене додељује организација IANA (Internet Assigned Numbers Authority). Од пре неколико година је могуће регистровати доменска имена са ћириличним писмом [15]. Тако постоје домени највишег нивоа .рф (Руска Федерација, од 2010. године), .срб (Србија, од 2011. године), .каз (Казахстан, од 2012. године), онда за Украјину .укр (2013.-те), Монголију .мон, Белорусију .бел, Северну Македонију .мкд (2014.) године, док је предложени домен за Бугарску .бг био дуго у процесу разматрања до одобрења 2016.-те године. Поред ових постоје и .ею (бугарски Европјски сојуз), .католик (власништво Ватикана), .ком, .орг и сл. Од почетка постојања .срб домена је регистровано око 3 хиљаде пуних доменских имена [16], а иначе постоје поддомени: .пр.срб (привреда/предузећа), .орг.срб (организације), .обр.срб (образовање), .од.срб (лични домени), .упр.срб (управа) и .ак.срб (академска мрежа).

Слање порука преко сервиса електронске поште подржава ћирилична слова, док крајња подршка зависи од програма за е-пошту и његових подешавања.

7. СМС ПОРУКЕ

Кратке поруке са мобилног уређаја на други се могу послати посредством посебне услуге мобилних оператора по имену SMS (Short Message Service). Овај сервис је направљен као део GSM (Global System for Mobile Communications) стандарда из 1985. године. Данас је прилагођен разним новијим мобилним технологијама, нпр. за пренос преко сателитских веза или чак фиксне телефоније. Прва порука за потребе тестирања је испоручена у децембру 1992. године у виду честитке за Божић. Смс поруке са мобилног уређаја се шаљу на предефинисани смс центар (SMSC), који користећи принцип похрани-и-испоручи, покушава да на најбржи начин испоручи поруку другом мобилном уређају без гаранције испоруке. Порука има одређени рок трајности, обично седам дана, после чега се брише. Могуће је примити поруку у току разговора јер смс поруке иду преко канала за сигнализацију, док говор има свој одвојени канал. (Овде користимо `смс порука` из разлога велике учесталости употребе ове речи у српском језику. Додатно у прилог томе може се сматрати да је `смс` скраћеница од `сервис малих саопштења`.)

Дужина било које смс поруке је 140 бајтова плус заглавље у коме се налазе информације ко шаље, коме се

шаље, смс центар преко кога се шаље, која врста кодовања се користи итд. Врста кодовања утиче на број слова и других знакова који могу стати у једно поруку. Постоје два основна кодовања у смс порукама: подразумевано 7-битно (GSM-ASCII) и 16-битно (UTF-16). [17] Дужина од 140 бајтова односно 1120 битова (140 бајтова * 8 је 1120 битова) у првом случају дају дужину од 160 могућих знакова (1120 / 7 = 160) а у другом случају дужину од 70 знакова (1120 / 16 = 70). Подразумевано 7-битно кодовање је модификација ASCII табеле, па она покрива енглески, италијански, француски, немачки језик и језике из скандинавске групе језика; додатно садржи велика слова из модерног грчког језика. [18] Друга врста је Уникод кодовање фиксне 16-битне дужине, те се са њом кодују ћирилична слова али и латинична слова ван латинске основе (č, ć, š, ž, đ). Када се куца порука на мобилном телефону довољно је откуцати бар једно овакво слово и цела порука ће бити кодована по Уникод стандарду а њена дужина ће бити смањена на 70 знакова.

Вишеструке смс поруке су оне поруке чији број знакова прекорачује највећи дозвољени број знакова, било 160 или 70 знакова. Оне се називају спојене поруке (техничким речником PDU Mode SMS) а информација о тој особини троши одређени простор садржаја поруке. У овом случају један део вишеструке поруке сада може имати највише 153 знака. Тај део поруке се код мобилних оператора појављује као једна основна смс порука и тај начин се тарифирају односно наплаћују. У случају вишеструке смс поруке по Уникод стандарду имамо смањену дужину од 67 знакова. Свака појединачна порука се посебно наплаћује.

Можемо да видимо да су креатори техничких стандарда за смс поруке као полазну основу узели своје алфавете. Број узтеих алфавета је био минималан, те се за друге алфавете могла наћи подршка кроз надоградњу постојећег техничког решења. Та надоградња је `плаћена` кроз краће поруке писане ван слова (неких) латинских алфавета. Ако се жели написати порука са истим бројем знакова (првенствено овде мислимо на латиничне и ћириличне), онда је однос 2 пута (и мало више) неповољнији за не-латиничне поруке. Поставља се питање равноправности писама јер писати неким писмом кошта више него неким другим, као и поштовања правописних правила, закона и устава у неким земљама. Ово укључује и потребу за локализацијом физичких тастатура на телефонима, где се нпр. у Русији поред латинских слова налазе и слова руске ћирилице. Овде разматрамо и предлажемо више решења за равноправно тарифирање латиничних и Уникод смс порука (ово задње укључује и српску ћирилицу):

А) Тржишна утакмица. Данас је српска ћирилица у дефанзиви и у некој врсти подређеног положају у односу на латиницу из разлога у које неће се овде улазити. Иако имамо велики број корисника мобилне телефоније, тек мањи број зна да је уопште могуће слати ћириличне смс поруке. Модели телефона који се сада продају углавном имају Уникод подршку. Остаје да се шири популација информ-

ше и да у неком следећем кораку потражња за ћиричним смс порукама и телефонима који је подржавају прерасте у свеопшту потражњу коју неће моћи игнорисати произвођачи телефона, оперативних система за њих, мобилних оператора, увозника и свих других који су укључени у ово коло. На крају под притиском тржишта и удружења грађана или потрошача мобилни оператори изједначају цену за латиничне и ћириличне смс поруке.

Б) Техничко решење 1. Користећи постојећу телекомуникациону структуру и стандарде, променити подразумевано 7-битно кодовање и у њега укључити српску ћирилицу. Или укључити све важне светске ћирилице (можда уз додатак малих грчки слова, те тако добити кодовање за православне земље/писма), што доводи до већих промена у кодовању али захтева и већу међудржавну иницијативу. Међутим постоје два значајна проблема. Први је компатибилност са постојећим врстама кодовања, јер нова врста кодовања не мора, уопште неће или неће ускоро бити читљива за друге ван новог кодовања. Техничка изводљивост је такође дискутабилна, јер је потребно осмислити нови софтвер за слање и примање порука на мобилним телефонима. Ако су у питању Android и слични тзв. паметни оперативни системи, онда то није велика препрека, док би у преосталим требало писати Java ME софтвер што не значи да ће добро радити на свим моделима телефона. Ако би ово све успешно било завршено, једна смс порука са ћирилицом би могла имати уобичајених 160 знакова.

В) Техничко решење 2. Код мобилних оператора у смс центрима се провера заглавље поруке и испитује се да ли је оно UTF-16 кодовано, та ако јесте онда се пар таквих порука (укупно 134 знака) наплаћује као једна порука. Овим се не постиже потпуна недискриминација ћирилице (134 знака наспрам 160), али се она бар умањује. Потребно је тражити од мобилних оператора да имплементирају ово техничко решење у своје billing подсистеме, што није превише технички захтевно али ће свакако довести до пада перформанси читавог обрадног система. Међутим, то је једнократан трошак куповине јачих рачунарских система што се обично периодично и чини, углавном због повећања број корисника. Поред додатног трошка, мобилни оператори могу тврдити да се овако нарушава приватност преписке али то није тачно јер се чита само заглавље поруке односно само првих неколико бајтова поруке које се и иначе читају ради техничке манипулације поруком.

Г) Техничко решење 3. По стандардима описаним у [17] и [18], предвиђене су `National Language Shift Table`. Ове кодне табеле пружају могућност кодовања у писмима разних језика (од ћирилица спомиње се руски језик). То 7-битно кодовање због активирања UDH (User Data Header) секције у оквиру поруке доноси највећу дужину од 155 знакова по једној поруци. Потребно је створити српски кодни распоред који би 3GPP организација усвојила као стандард. Ово је слично решење као што је предложено под Б. Ипак ова кодна проширења немају подршку ни код најновијих мобилних уређаја.

Д) Политичко решење. Донесе се закон о равноправности ћирилице којим се тражи да сваки технички уређај (рачунар, мобилни телефон, телевизор итд.) мора да има подршку за ћирилицу или неће добити атест за продају на локалном тржишту. У оквиру њега се обавезу мобилни оператори да ћириличне поруке третирају на исти начин као и оне друге. Оператори могу да примене неко од технички решења предложених овде или чак да креирају неко сопствено решење.

Ђ) Јапанско или интернет решење. Мобилни оператори за поруку дужине 140 бајтова наплаћују екстремно високу цену ако се она упореди са ценом за пренос података кроз мобилну мрежу. При том, нуди се и неограничени интернет проток по ниској цени, додуше и при ниској брзини преноса али више него довољно за размену порука. Поруке више не би биле смс типа, већ би се користили програми за инстант поруке (IM, instant messaging) којих има много и добар део њих је бесплатан за тзв. паметне телефоне. На тај начин је могуће за цену једне смс поруке дописивати се непрекидно дужи период времена у току дана. Такође то отвара могућност дописивања са корисницима на рачунарима јер се користи исти сервери за поруке или исти протоколи за инстант поруке. Употреба е-порука је друга опција код овог решења, мало је спорија што се тиче испоруке поште и потребно је наместити програм за скидање поште да често проверава да ли је дошло нешто ново. Овај тип комуникације је претежан у Јапану, јер слање смс порука са њиховим традиционалним словима је скупо и непрактично, при том користе посебан тип е-поште тзв. push e-mail који се одмах приказује на екрану чим пристигне, по угледу на уобичајено понашање смс порука.

Предложена решења су дата на различитим нивоима апстракције, а нека се међусобно не искључују односно могу се комбиновати. Имена решења су радног типа. Редослед навођења није дат по неком приоритету или пожељности.

8. ЗАКЉУЧАК

Трећа обимна технолошка револуција - информатичка, после неолитске и индустријске револуције, је омогућила пробој ћирилице у све поре постиндустријског друштва. На интернету се може наћи највећи број ћириличних записа него икада до сада у историји. За веровати је да ће тај тренд да се настави. Поређење језика и писама се симболично може видети на почетној страни Википедије, где се најзначајнији ћирилични језик – руски, налази у друштву енглеског, немачког, француског, шпанског, италијанског и других језика који имају највећи број одредница на овој енциклопедији. Видели смо да је ћирилица у ИКТ врло добро подржана и треба је само користити. Ипак неки делови подручја ИКТ се често мењају те је потребно увек тражити се подржи ћирилично писмо у новим верзијама, било хардвера, софтвера или стандарда. То је напор који треба да покрене држава, удружења или појединци а сви они заједно да испоштују. Тај напор и настојање треба да буде константно као и наше коришћење ћирилице.

9. ЛИТЕРАТУРА

- [1] Милић Станковић од Мачве: „Манифест – нови реализам или национална уметност“, фебруар 1962. (http://cirilica-beograd.bestfromserbia.com/Други-пишу/111120_Milic_od_Macve/), прочитано 01.11.2020.)
- [2] Драгољуб Збиљић: „(Не)решено питање српског писма у српском правопису“, Скуп „Савремене информационе технологије – Интернет и ћирилица“, Бијељина, 2003.
- [3] Предраг Пипер: „Увод у славистику I“, Завод за уџбенике и наставна средства, Београд, 1998.
- [4] „Кирилица“, Википедија (<http://ru.wikipedia.org/wiki/Кирилица>), прочитано 01.11.2020.)
- [5] „Distribution of the Cyrillic script worldwide“, Википедија (http://en.wikipedia.org/wiki/File:Cyrillic_alphabet_distribution_map.png), прочитано 01.11.2020.)
- [6] Зоран Милошевић: „Геополитика и ћирилица“, Политичка ревија, Београд, 2011. (<https://doi.org/10.22182/pr.3042011.7>), прочитано 01.11.2020.)
- [7] „Кирилица в Юникоде“, Википедија ([http://en.wikipedia.org/wiki/Сyрилlic_\(Unicode_block\)](http://en.wikipedia.org/wiki/Сyрилlic_(Unicode_block))), прочитано 01.11.2020.)
- [8] Петар Бошњаковић: „Кодовање и квантовање“, Висока школа електротехнике и рачунарства струковних студија, Београд
- [9] „Mapping_of_Unicode_characters“, Википедија (http://en.wikipedia.org/wiki/Mapping_of_Unicode_characters), прочитано 01.11.2020.)
- [10] „Unicode“, Википедија (<http://en.wikipedia.org/wiki/Unicode>), прочитано 01.11.2020.)
- [11] „Request for Comments 2044 - UTF-8, a transformation format of Unicode and ISO 10646“, Internet Engineering Task Force, 1996. (<http://www.ietf.org/rfc/rfc2044.txt>), прочитано 01.11.2020.)
- [12] „Кирилица в Юникоде“, Википедија (http://ru.wikipedia.org/wiki/Кирилица_в_Юникоде), прочитано 01.11.2020.)
- [13] „Codes for the Representation of Names of Languages, arranged alphabetically by alpha-3/ISO 639-2 Code“, (http://www.loc.gov/standards/iso639-2/php/code_list.php), прочитано 01.11.2020.)
- [14] „Codes for the representation of names of scripts“, ISO/Unicode, 2004 (<http://www.unicode.org/iso15924/codelists.html>), прочитано 01.11.2020.)
- [15] „срб“, Википедија (<http://sr.wikipedia.org/wiki/срб>), прочитано 01.11.2020.)
- [16] Регистар националног интернет домена Србије (РНИДС), (ruidc.rs), прочитано 01.11.2020.)
- [17] „Alphabets and language-specific information 3GPP TS 03.38“, 3rd Generation Partnership Project - 3GPP (<http://www.3gpp.org/ftp/Specs/html-info/0338.htm>), прочитано 01.11.2020.)
- [18] „GSM 03.38“, Википедија (http://en.wikipedia.org/wiki/GSM_03.38), прочитано 01.11.2020.)



Др Драгољуб Пилиповић, доцент,
Слобомир П универзитет, Бијељина,
РС-БиХ

Имејл: dragoljub.pilipovic@gmail.com
Области интересовања: методологије
развоја софтвера, електронско
образовање, сигурност рачунарских
система, електронско гласање, ИТ
сервисни менаџмент



Мр Жељко Гаврић, истраживач
сарадник, Факултет организационих
наука, Универзитет у Београду
Имејл: gavric.zeljko@yahoo.com
Области интересовања: Мултимедија,
НСИ, Интернет технологије



info m

UPUTSTVO ZA PRIPREMU RADA

1. Tekst pripremiti kao Word dokument, A4, u kodnom rasporedu 1250 latinica ili 1251 ćirilica, na srpskom jeziku, bez slika. Preporučeni obim – oko 10 strana, single prored, font 11.
2. Naslov, abstrakt (100-250 reči) i ključne reči (3-10) dati na srpskom i engleskom jeziku.
3. Jedino formatiranje teksta je normal, bold, italic i bolditalic, VELIKA i mala slova (tekst se naknadno prelama).
4. Mesta gde treba ubaciti slike, naglasiti u tekstu (Slika1...)
5. Slike pripremiti odvojeno, VAN teksta, imenovati ih kao u tekstu, radi identifikacije, u sledećim formatima: rasterske slike: jpg, tif, psd, u rezoluciji 300 dpi 1:1 (fotografije, ekranski prikazi i sl.), vektorske slike – cdr, ai, fh,eps (šeme i grafikoni).
6. Autor(i) treba da obavezno priloži svoju fotografiju (jpg oko 50 Kb), navede instituciju u kojoj radi, kontakt i 2-4 oblasti kojima se bavi.
7. Maksimalni broj autora po jednom radu je 5.

Redakcija časopisa Info M