

ФОНДАЦИЈА
НАРОДНЕ
БИБЛИОТЕКЕ
СРБИЈЕ
Ф



РЕПУБЛИКА СРБИЈА
МИНИСТАРСТВО СПОЉНЕ И
УНУТРАШЊЕ ТРГОВИНЕ И
ТЕЛЕКОМУНИКАЦИЈА



Едукација у Дигитализацији

Народна библиотека Србије





Сврха и процес дигитализације

Фондација Народне библиотеке Србије
у оквиру пројекта „Дигиталне библиотеке – едукација“
под покровитељством Министарства за спољну и
унутрашњу трговину и телекомуникације планом
активности издаје ову брошуру са темом „Сврха и
процес дигитализације“.

Брошура има за циљ да представи процес
дигитализације, а намењена је запосленима у култури
као и будућим младим стручњацима из области
заштите културног наслеђа и библиотекарства.

Дигитализација је процес који доприноси очувању знања
и духовности једне нације за будућа поколења.



- Стратешки принципи дигитализације културног наслеђа
 - Процес дигитализације
 - Општи циљеви дигитализације
 - Дигитални документ
 - Основне врсте дигиталних докумената
 - Дигитални објекат
 - Дигитална збирка
 - Дигитална библиотека
 - Основни критеријуми селекције
- Метаподаци
 - OCR – Optical Character Recognition
 - Дугорочно очување
- Препоручене техничке карактеристике за креирање дигиталних докумената
- Препоручени параметри за дигитализацију библиотечко-информационе грађе и извора
- Правни аспекти дигитализације
 - Ауторска и сродна права
 - Ауторска права и дигитализација
- Дигиталне библиотеке будућности

СТРАТЕШКИ ПРИНЦИПИ ДИГИТАЛИЗАЦИЈЕ КУЛТУРНОГ НАСЛЕЂА

- Дигитализовано културно наслеђе је јавно добро које треба да буде доступно најширем кругу заинтересованих, у мери која не нарушава правила заштите интелектуалне својине и приватности.
- Посебно треба водити рачуна о доступности дигитализоване грађе свим заинтересованим, без обзира на узраст, физичка ограничења, познавање језика и техничку обученост.
- Дигитализовано културно наслеђе, како оно јавно доступно, тако и оно које из неких оправданих разлога није, треба да буде чувано и у архивској форми, у циљу заштите културног наслеђа и употпуњавања фондова институција које нису у могућности да буду у поседу оригиналних културних добара.
- Дигитализовано културно наслеђа треба да буде организовано тако да се омогући његово истраживање и примена у науци, образовању и, уопште, изградњи друштва заснованог на знању, као и у економском развоју земље (на пример: туризам, индустрија дигиталних медија).
- Дугорочно, дигитализовано и адекватно репрезентовано културно наслеђе треба да допринесе очувању националног идентитета и културне разноврсности националних заједница нашег поднебља, као и побољшању позиције наше земље на глобалном плану.
- У поступку дигитализације културног наслеђа потребно је следити принципе и стандарде дефинисане на националном нивоу и усклађене са одговарајућим међународним смерницама које су формулисали UNESCO, Савет Европе, Европска унија итд., чиме би требало да буде постигнута интероперабилност, како у земљи, тако и у окружењу.

Процес дигитализације је претварање аналогног објекта (било да је он састављен од слова, бројева, симбола, да је графички, звучни или видео запис) у дигитални документ/објекат. Обухвата неопходне поступке и активности усмерене ка добијању дигиталног објекта:

1. одређивање приоритета и опсега дигитализације;
2. селекција и припрема библиотечко-информационе грађе и извора за дигитализацију;
3. успостављање система метаподатака;
4. креирање дигиталних докумената;

5. формирање и обрада дигиталне збирке/библиотеке;
6. дуготрајно чување дигиталних објеката;
7. обезбеђивање приступа дигиталној збирци/библиотеци.

Дигитализација представља планску и организовану реализацију активности:

- креирање дигиталних докумената/објеката;
- управљање дигиталним документима/објектима;
- презентовање и унапређивање коришћења дигиталних докумената/објеката.

Општи циљеви дигитализације библиотечко-информационе грађе и извора су:

1. универзална доступност информација;
2. дуготрајно чување и представљање културног и научног наслеђа;
3. промоција и представљање библиотечко-информационе грађе и извора;
4. физичка заштита библиотечко-информационе грађе и извора;
5. ширење корисничке популације.

Дигитални документ је скуп података састављен од битова који представљају слова, бројеве, симболе, графичке, звучне или видео записе.

Основне врсте дигиталних докумената су: текст, слика, звучни запис и видео запис. Врсту одређује садржај дигиталног документа.

Дигитални објекат се састоји од дигиталног документа и одговарајућих метаподатака за тај документ.

Дигитална збирка је скуп дигиталних објеката организованих по одређеном принципу са становишта садржаја и формата.

Дигитална библиотека је скуп дигиталних збирки са алатима и методама за приступ, проналажење, избор, организацију и одржавање збирки.

Основни критеријуми селекције библиотечко-информационе грађе и извора за дигитализацију су:

1. вредност грађе и извора за културно и научно наслеђе;
2. учесталост коришћења;
3. целовитост и обрађеност оригиналне збирке;
4. могућност приступа;
5. физичка обележја (формат, величина, облик, врста материјала) и стање оригинала (ниво очуваности, осетљивост на манипулацију, постојаност медија);
6. и други критеријуми по процени стручњака.

Метаподаци су структурисане информације које описују, објашњавају, идентификују, лоцирају и на други начин помажу да се информације о изворима могу претраживати, а сами извори користити. Метаподатак је податак који треба посматрати као сет тврдњи од којих се свака односи на извор.

Врсте метаподатака:

1. Описни – намењени проналажењу и идентификацији извора;
2. Структурни – намењени опису структуре вишиделног извора;
3. Технички подаци о грађи и изворима (формати, кодови, лозинке, веб-адресе, услови чувања, миграција података итд.);
4. Административни – намењени за управљање грађом и изворима (набавка, ауторска и сродна права, локација, информације о коришћењу итд.);
5. Метаподаци о интелектуалним правима;
6. Метаподаци о заштити – намењени архивирању и дуготрајној употреби извора.

Генеришу се:

1. Из електронских извора;
2. Харвестујемо их из других система;
3. Креирамо их мануелно.

Сврха метаподатака:

1. Интероперабилност – способност различитих система са различитим хардверским и софтверским платформама, структуром података и интерфејсом да размењују податке са минималним губитком;
2. Идентификација електронских извора;
3. Дуготрајна заштита електронских извора.

Структура метаподатака:

1. Елемент – најмањи елемент који има јасно семантичко значење;
2. Вредност – јединствена информација која је садржај елемента;
3. Запис – структурисана група елемента;
4. Схема – систем правила како се користе елементи и како се креирају вредности унутар елемената.

OCR – Optical Character Recognition подразумева „оптичко препознавање карактера“. OCR програми служе за препознавање и издвајање текста из слика – скенираних докумената. Ови програми имају функцију да „читају“ скенирани

документ (слику), препознају карактере на слици и креирају нови текстуални документ који се може читати или обрађивати.

Дугорочно очување (Long-term preservation) дигиталних објеката обухвата, пре свега, одређивање формата за мастер датотеке, дескриптивне, структуралне и административне метаподатке, као и OCR резултате.

Препоручене техничке карактеристике за креирање дигиталних докумената

Приликом креирања дигиталних докумената потребно је разликовати:

1. мастер (архивску) копију која је намењена за дуготрајно чување,
2. копије за друге потребе.

Препоручује се коришћење отворених стандарда и формата дигиталних докумената који не подлежу власничким лиценцама, када је то могуће, односно оних формата који имају широку подршку у постојећој пракси и литератури.





СТЕФАНИЪ МИХАЙЛОВИЧЪ УРОШИЪ

Князь Сербскій

0,35 1272 - 1521.

Синоскладъ Сръбски изъ Монаховидца.

Вѣстникъ Л.

Препоручени параметри за дигитализацију библиотечко-информационе грађе и извора

ДИГИТАЛИЗАЦИЈА ТЕКСТУАЛНЕ ГРАЂЕ И ИЗВОРА (ТЕКСТ)

МАСТЕР КОПИЈА (ЗА ДУГОРОЧНО ЧУВАЊЕ)

Минимална резолуција скенирања	300 тачака по инчу (dpi)
Формат документа	<ul style="list-style-type: none">• некомпресован – TIFF• компресован – JPEG, JPEG2000• PDF
Дубина боје	<ul style="list-style-type: none">• црно-бели документи – 1 bit• сива скала – 8 бита• пун колор – 24 бита

ОПЕРАТИВНА КОПИЈА СА ТРАНСКРИПЦИЈОМ ПУНОГ ТЕКСТА

Формат документа	<ul style="list-style-type: none">• структурисан текст – XML шема (TEI), PDF/A, HTML, ODF (ODT)• неструктурисани текст – TXT
------------------	---

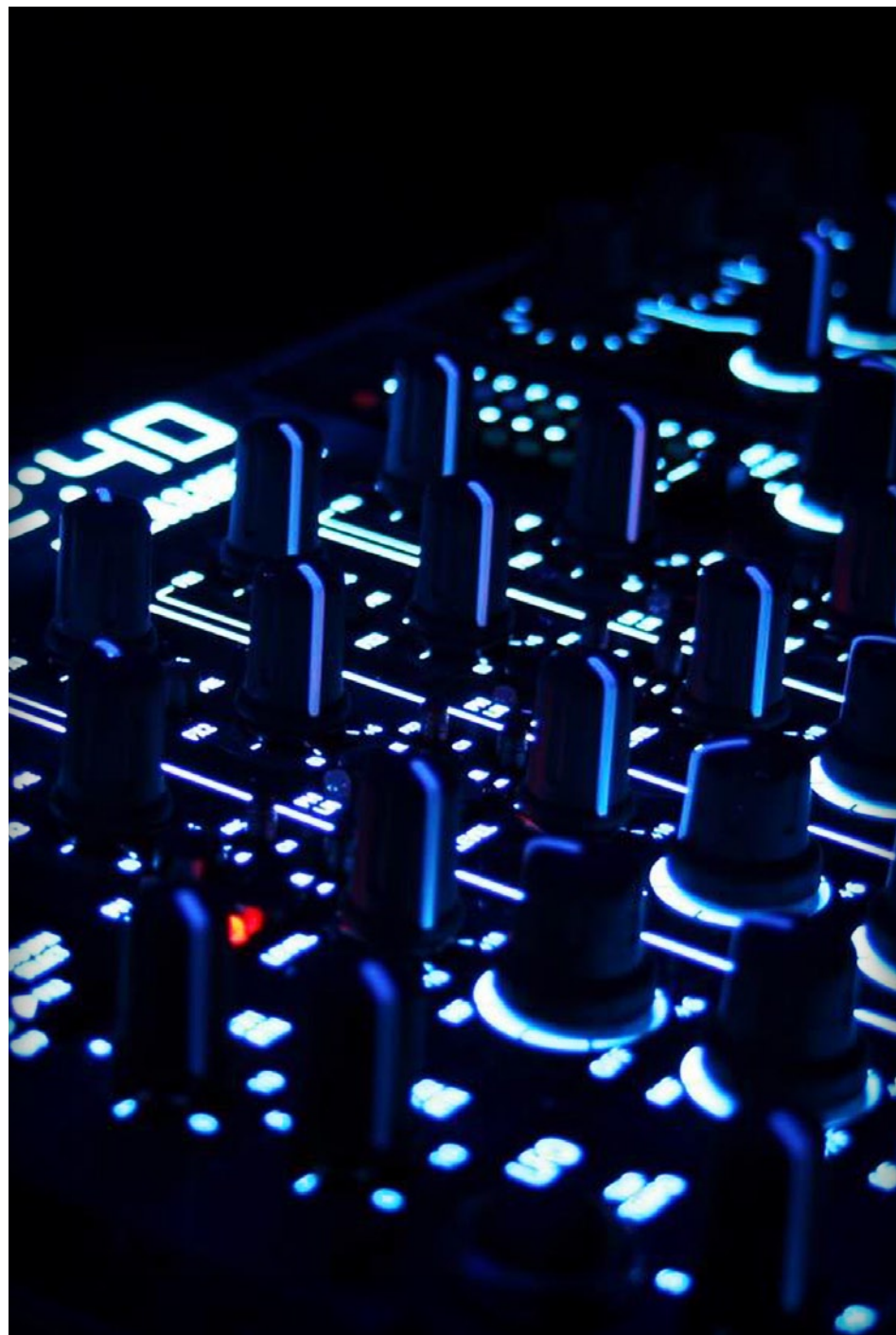
ДИГИТАЛИЗАЦИЈА СЛИКОВНЕ ГРАЂЕ И ИЗВОРА (СЛИКА)

МАСТЕР КОПИЈА (ЗА ДУГОРОЧНО ЧУВАЊЕ)

Минимална резолуција скенирања	• 300 тачака по инчу (dpi)
Формат документа	<ul style="list-style-type: none">• некомпресован – TIFF• компресован – JPEG, JPEG2000
Дубина боје	<ul style="list-style-type: none">• црно-бели документи – 1 bit• сива скала – 8 бита• пун колор – 24 бита

ОПЕРАТИВНА КОПИЈА (ЗА РАЗМЕНУ, ОБЈАВЉИВАЊЕ НА ВЕБУ)

Резолуција документа	• 72–150 тачака по инчу (dpi)
Формат документа	• компресован – JPEG, JPEG2000, PNG
Дубина боје	<ul style="list-style-type: none">• црно-бели документи – 1 bit• сива скала – 8 бита• пун колор – 24 бита



ДИГИТАЛИЗАЦИЈА ГРАЂЕ И ИЗВОРА (АУДИО)

МАСТЕР КОПИЈА (ЗА ДУГОРОЧНО ЧУВАЊЕ)

Минималан квалитет семплованог (узоркованог) звука	16 Bit, 48 KHz 24 Bit, 96 KHz (препорука)
--	--

Формат документа	<ul style="list-style-type: none"> • некомпресован – WAV, AIFF, BWF • компресован компресијом недеструктивног карактера – FLAC, APE, WV, TAK • компресован деструктивном компресијом – OGG
------------------	---

ОПЕРАТИВНА КОПИЈА (ЗА РАЗМЕНУ, ОБЈАВЉИВАЊЕ НА ВЕБУ, „СТРИМОВАЊЕ”)

Формат документа	<ul style="list-style-type: none"> • компресован (отвореног кода) – OGG • компресован (лиценцирани) – WMA, MP3, AAC <p>напомена : све наведене компресије намењене оперативним копијама компресоване су компресијама деструктивног карактера</p>
------------------	---

ДИГИТАЛИЗАЦИЈА ВИДЕО ГРАЂЕ И ИЗВОРА (ВИДЕО)

МАСТЕР КОПИЈА (ЗА ДУГОРОЧНО ЧУВАЊЕ)

Резолуција слике	<ul style="list-style-type: none"> • 640 x 480 пиксела (минимално за снимке са односом страна слике 4:3, „дигиталних по настанку”) • 720 x 576 пиксела (телевизија стандардне резолуције, PAL система за 25 слика у секунди)
------------------	--

Формат документа	<ul style="list-style-type: none"> • контејнери: AVI, MOV, MPEG, MPEG-2, MP4 MKV (отвореног кода) • кодекси: H.262/MPEG-2 codecs H.263/MPEG-4 Part 2 codecs H.264/MPEG-4 AVC codecs WMV, MS MPEG-4v3 = DivX :-)
------------------	---

ОПЕРАТИВНА КОПИЈА (ЗА РАЗМЕНУ, ОБЈАВЉИВАЊЕ НА ИНТЕРНЕТУ, „СТРИМОВАЊЕ”)

Формат документа	<ul style="list-style-type: none"> • контејнери: AVI, WMV, FLV
------------------	---

ПРАВНИ АСПЕКТИ ДИГИТАЛИЗАЦИЈЕ

Закон о култури („Службени гласник РС”, бр. 72/09)

Члан 6 (Општи интерес у култури)

– подстицање примене нових технологија у култури, нарочито информационих технологија и дигитализације;

Члан 8 (Културна делатност)

– дигитално стваралаштво и мултимедији;

Закон о библиотечко-информационој делатности („Службени гласник РС” бр. 52/11)

Члан 44 (Дигитализација библиотечко-информационе грађе и извора)

1. Библиотека има обавезу да континуирано ради на дигитализацији библиотечко-информационе грађе и извора као дела културног наслеђа Србије.
2. Народна библиотека Србије спроводи и координира послове на дигитализацији библиотечко-информационе грађе и извора у Републици Србији.
3. Ближе услове за дигитализацију библиотечко-информационе грађе и извора прописује министар надлежан за културу.

Ауторска и сродна права

При креирању дигиталних докумената морају се поштовати ауторска и сродна права у складу са позитивним правним прописима.

Да би библиотеке дигитализовану библиотечко-информациону грађу и изворе заштићене ауторским и сродним правима учиниле доступним корисницима, неопходно је да обезбеде ауторска и сродна права.

Библиотечко-информациона грађа и извори који су у јавном домену (над којима су истекла ауторска права) могу се дигитализовати без дозволе аутора.

Стара и ретка библиотечка грађа се може дигитализовати без дозволе аутора, под условом да су истекла ауторска права.

У складу са позитивним правним прописима, библиотекама је дозвољено да без дозволе аутора и без плаћања ауторске накнаде дигитализују дела само за сопствене потребе и то из сопствених примерака и без намере да се оствари непосредна или посредна имовинска корист.

Са аспекта Закона о ауторском и сродним правима, у процесу дигитализације постоје три категорије грађе:

1. дела заштићена ауторским и сродним правима;
2. дела чији се носилац ауторског и сродног права не може идентификовати („ауторски сирочићи”);
3. дела из јавног домена.

У случају коришћења дигитализоване грађе и извора библиотеке су у обавези да искористе правна средства ради проналажења носиоца ауторских права и тек у случају да су носиоци и даље непознати, може дела користити као дела из јавног домена.

Библиотеке су дужне да за дигитализовану библиотечко-информациону грађу и изворе у јавном домену омогуће отворен приступ.

Ауторска права и дигитализација

Дозвољено је без дозволе аутора и плаћања ауторске накнаде дигитализовање и стављање ауторског дела у промет за потребе особа са инвалидитетом и то у случају ако то дело не постоји у траженом облику, ако је његова употреба у директној вези са инвалидитетом тих особа и у обиму који захтева одређена врста инвалидитета и ако то дигитализовање и стављање у промет није учињено ради остваривања посредне или непосредне имовинске користи.

Тим стручњака за дигитализацију



ДИГИТАЛНЕ БИБЛИОТЕКЕ БУДУЋНОСТИ

Хуманистичке науке — и само човечанство — су се одувек налазили на ивици провалије, у стању трајног пропадања. Катастрофа — као и грчка реч катастрофѣ која означава окрет или обрт, али и свршетак, пад, смрт или пропаст — је у западној традицији уписана, с једне стране, у предпросветитељски начин мишљења, по коме је историја исто што и каталог опадања вредности: од истеривања из Раја до класицистичког идеализовања и извитоперавања слике древне Грчке; док се и само Просветитељство, са својом вером у прогрес и доминацијом редукованог разума, испоставило, како је то Адорно показао, као зачетник тријумфа великог зла. Катастрофа је зато, хтели ми то или не, неизбежна.

Ако је класични модел библиотеке био у исто време и архив и храм, то значи да је библиотека била одбрана против катастрофе; институција која је чувала не само систем вредности него и одређене ритуале коришћења знања: књигу као физички објекат и из саме природе физичког објекта модел текста као неповредивог и коначног. Усред вечите катастрофе, библиотека је била ван катастрофе, као простор сећања на имагинарно време пре катастрофе.

Данас, међутим, живимо у свету у коме влада екстаза комуникације, у коме има све више информација, а све мање значења. У бесконачном клонирању слика и бесконачној пролиферацији знакова, сам знак је, ако је веровати бодријаровској хиперболи, постао невидљив. После смрти аутора и превазилажења аристотеловске линеарности кроз хипертекст, могућност смисаоне текстуалности није више дата сама по себи. Као добро распеване нарикаче у великом, познокапиталистичком, постиндустријском хору светске мреже, унисоно постављамо питање: коме су још потребне библиотеке — и какве — на заласку хуманизма, у друштву презасићености и обесначености, у време перманентне катастрофе?

Тренутак у коме живимо — тренутак у коме се дигитализација призива као нека врста технолошких додола, које ће својом кишом освежити и уместо народа позлатити библиотекарску делатност, занимљив је и изазован из више разлога: не само зато што се преношењем рукописне и штампане грађе у машински читљив облик радикално мења политика приступа библиотечким фондовима, а самим тим и избор доступних материјала; и не само зато што библиотека, да би уопште преживела, постаје поприште технолошких експеримената у којима почиње да се непосредно надмеће са моћним комерцијалним подухватима; него и зато што библиотеке, захваљујући развоју технологије и дигиталних хуманистичких наука, најзад имају могућност да почну да иду корак у корак са новијим теоретским моделима текстуалности.

Уосталом, грчка реч катастрофа означава још и прекретницу или обрт у драмској радњи, са којом почиње отплитање замршеног драмског чвора. Ја верујем да нам не преостаје ништа друго него да пробамо да спасимо катастрофу од лошег гласа који је бије и да пробамо да у данашњем технолошком изазову видимо прилику за преосмишљавање улоге које библиотеке могу да имају у даљем развоју хуманистичких наука.



Под појмом дигиталних библиотека се подразумевају системи различитих обима и различите функционалности — од репозиторија дигиталних објеката и метаподатака, референсних система и система за управљање садржајем до сложених организација које се баве пружањем дигиталних библиотечарских услуга. Мене, као некога чији је основни предмет истраживања — и даље, упркос катастрофи — текст, зато највише занима статус текстуалног садржаја у дигиталним библиотекама и развој функција које библиотека нуди у односу на текстуални садржај. Како је текст сачуван? Како се испоручује кориснику? Шта корисник може да ради са текстом? И како му у свему томе библиотека стоји на располагању?

И управу ту — у домену употребе, односно испуњавања текстуалног потенцијала — ја видим проблем са основном идејом и методологијом многих пројеката дигитализације. Данашње дигиталне библиотеке су махом засноване на моделима штампаног (линеарног) текста и затвореном просторном (и економском) моделу који подразумева приступ библиотечкој грађи преко само једног одређеног интернет портала. Уместо да разложи текст на његове конститутивне елементе и омогуће претраживање целог садржаја — када то још увек важећи, мастодонтски закони о ауторским правима дозволе — неке дигиталне библиотеке још увек испоручују текст као слику, која, иако дигитална, није ни рачунарски разумљива, нити обрадива као текст. У исто време, изолација дигиталних библиотека у индивидуалним институционалним оквирима је у супротности са могућношћу које интернет пружа на стварању глобалног хипертекста. Мој циљ није да стичем иконоборачки стаж тако што ћу — ко о чему — заменити пословичне уштипке за дигиталне слике, већ само да укажем на нелогичност приступа дигитализацији који тексту одузима могућност да буде машински обрађен као текст.

Приликом дигитализације већ постојећег штампаног текста, начелно говорећи, постоје два могућа приступа: скенирање текста, односно стварање дигиталне слике, или прекуцавање истог у текстуалну датотеку. Оба приступа стварају дигитални садржај који је машински испоручив, али само текстуалне датотеке (или дигиталне слике подвргнуте режиму оптичког препознавања карактера — OCR) чине садржај директно машински читљивим. Нема сумње да сваки текст — било да је он писан руком, штампан или састављен уз помоћ софтвера за обраду текста — пре него што се чита, разуме и тумачи, прво доживљава као визуелни предмет. Дигитална фототипска издања су зато веома важна за изучавање физичког изгледа текста, његових димензија, типографских одлика, маргиналних забелешки итд.

Али дигитална слика је ипак само — слика, а рачунарска обрада текста и могућност да се са текстом опходимо као са текстуалном датотеком — неопходност, и на практичном и на теоретском нивоу. Дигитализовани отисак штампаног документа је „замрзнути” текст; текст, који се панично крије од машине и чезне за величањем сопствене неповредивости. Ако желимо да остваримо помак на пољу дигиталних библиотека, онда не смемо да заборавимо да је дигитални текст флуидан и да се од дигиталних библиотека мора захтевати флексибилност које традиционална библиотечка грађа никада није могла да пружи. У односу на текст, слика може бити вредна хиљаду речи, али често је то хиљаду погрешних речи.

Текст, као темељни појам и књижевне теорије и теорије кодирања текста, налази се у процепу између две методолошке праксе. Али тај усек између две литице,

11. ¹ ² ³ ⁴ ⁵ ⁶ ⁷ ⁸ ⁹ ¹⁰ ¹¹ ¹² ¹³ ¹⁴ ¹⁵ ¹⁶ ¹⁷ ¹⁸ ¹⁹ ²⁰ ²¹ ²² ²³ ²⁴ ²⁵ ²⁶ ²⁷ ²⁸ ²⁹ ³⁰ ³¹ ³² ³³ ³⁴ ³⁵ ³⁶ ³⁷ ³⁸ ³⁹ ⁴⁰ ⁴¹ ⁴² ⁴³ ⁴⁴ ⁴⁵ ⁴⁶ ⁴⁷ ⁴⁸ ⁴⁹ ⁵⁰ ⁵¹ ⁵² ⁵³ ⁵⁴ ⁵⁵ ⁵⁶ ⁵⁷ ⁵⁸ ⁵⁹ ⁶⁰ ⁶¹ ⁶² ⁶³ ⁶⁴ ⁶⁵ ⁶⁶ ⁶⁷ ⁶⁸ ⁶⁹ ⁷⁰ ⁷¹ ⁷² ⁷³ ⁷⁴ ⁷⁵ ⁷⁶ ⁷⁷ ⁷⁸ ⁷⁹ ⁸⁰ ⁸¹ ⁸² ⁸³ ⁸⁴ ⁸⁵ ⁸⁶ ⁸⁷ ⁸⁸ ⁸⁹ ⁹⁰ ⁹¹ ⁹² ⁹³ ⁹⁴ ⁹⁵ ⁹⁶ ⁹⁷ ⁹⁸ ⁹⁹ ¹⁰⁰ ¹⁰¹ ¹⁰² ¹⁰³ ¹⁰⁴ ¹⁰⁵ ¹⁰⁶ ¹⁰⁷ ¹⁰⁸ ¹⁰⁹ ¹¹⁰ ¹¹¹ ¹¹² ¹¹³ ¹¹⁴ ¹¹⁵ ¹¹⁶ ¹¹⁷ ¹¹⁸ ¹¹⁹ ¹²⁰ ¹²¹ ¹²² ¹²³ ¹²⁴ ¹²⁵ ¹²⁶ ¹²⁷ ¹²⁸ ¹²⁹ ¹³⁰ ¹³¹ ¹³² ¹³³ ¹³⁴ ¹³⁵ ¹³⁶ ¹³⁷ ¹³⁸ ¹³⁹ ¹⁴⁰ ¹⁴¹ ¹⁴² ¹⁴³ ¹⁴⁴ ¹⁴⁵ ¹⁴⁶ ¹⁴⁷ ¹⁴⁸ ¹⁴⁹ ¹⁵⁰ ¹⁵¹ ¹⁵² ¹⁵³ ¹⁵⁴ ¹⁵⁵ ¹⁵⁶ ¹⁵⁷ ¹⁵⁸ ¹⁵⁹ ¹⁶⁰ ¹⁶¹ ¹⁶² ¹⁶³ ¹⁶⁴ ¹⁶⁵ ¹⁶⁶ ¹⁶⁷ ¹⁶⁸ ¹⁶⁹ ¹⁷⁰ ¹⁷¹ ¹⁷² ¹⁷³ ¹⁷⁴ ¹⁷⁵ ¹⁷⁶ ¹⁷⁷ ¹⁷⁸ ¹⁷⁹ ¹⁸⁰ ¹⁸¹ ¹⁸² ¹⁸³ ¹⁸⁴ ¹⁸⁵ ¹⁸⁶ ¹⁸⁷ ¹⁸⁸ ¹⁸⁹ ¹⁹⁰ ¹⁹¹ ¹⁹² ¹⁹³ ¹⁹⁴ ¹⁹⁵ ¹⁹⁶ ¹⁹⁷ ¹⁹⁸ ¹⁹⁹ ²⁰⁰ ²⁰¹ ²⁰² ²⁰³ ²⁰⁴ ²⁰⁵ ²⁰⁶ ²⁰⁷ ²⁰⁸ ²⁰⁹ ²¹⁰ ²¹¹ ²¹² ²¹³ ²¹⁴ ²¹⁵ ²¹⁶ ²¹⁷ ²¹⁸ ²¹⁹ ²²⁰ ²²¹ ²²² ²²³ ²²⁴ ²²⁵ ²²⁶ ²²⁷ ²²⁸ ²²⁹ ²³⁰ ²³¹ ²³² ²³³ ²³⁴ ²³⁵ ²³⁶ ²³⁷ ²³⁸ ²³⁹ ²⁴⁰ ²⁴¹ ²⁴² ²⁴³ ²⁴⁴ ²⁴⁵ ²⁴⁶ ²⁴⁷ ²⁴⁸ ²⁴⁹ ²⁵⁰ ²⁵¹ ²⁵² ²⁵³ ²⁵⁴ ²⁵⁵ ²⁵⁶ ²⁵⁷ ²⁵⁸ ²⁵⁹ ²⁶⁰ ²⁶¹ ²⁶² ²⁶³ ²⁶⁴ ²⁶⁵ ²⁶⁶ ²⁶⁷ ²⁶⁸ ²⁶⁹ ²⁷⁰ ²⁷¹ ²⁷² ²⁷³ ²⁷⁴ ²⁷⁵ ²⁷⁶ ²⁷⁷ ²⁷⁸ ²⁷⁹ ²⁸⁰ ²⁸¹ ²⁸² ²⁸³ ²⁸⁴ ²⁸⁵ ²⁸⁶ ²⁸⁷ ²⁸⁸ ²⁸⁹ ²⁹⁰ ²⁹¹ ²⁹² ²⁹³ ²⁹⁴ ²⁹⁵ ²⁹⁶ ²⁹⁷ ²⁹⁸ ²⁹⁹ ³⁰⁰ ³⁰¹ ³⁰² ³⁰³ ³⁰⁴ ³⁰⁵ ³⁰⁶ ³⁰⁷ ³⁰⁸ ³⁰⁹ ³¹⁰ ³¹¹ ³¹² ³¹³ ³¹⁴ ³¹⁵ ³¹⁶ ³¹⁷ ³¹⁸ ³¹⁹ ³²⁰ ³²¹ ³²² ³²³ ³²⁴ ³²⁵ ³²⁶ ³²⁷ ³²⁸ ³²⁹ ³³⁰ ³³¹ ³³² ³³³ ³³⁴ ³³⁵ ³³⁶ ³³⁷ ³³⁸ ³³⁹ ³⁴⁰ ³⁴¹ ³⁴² ³⁴³ ³⁴⁴ ³⁴⁵ ³⁴⁶ ³⁴⁷ ³⁴⁸ ³⁴⁹ ³⁵⁰ ³⁵¹ ³⁵² ³⁵³ ³⁵⁴ ³⁵⁵ ³⁵⁶ ³⁵⁷ ³⁵⁸ ³⁵⁹ ³⁶⁰ ³⁶¹ ³⁶² ³⁶³ ³⁶⁴ ³⁶⁵ ³⁶⁶ ³⁶⁷ ³⁶⁸ ³⁶⁹ ³⁷⁰ ³⁷¹ ³⁷² ³⁷³ ³⁷⁴ ³⁷⁵ ³⁷⁶ ³⁷⁷ ³⁷⁸ ³⁷⁹ ³⁸⁰ ³⁸¹ ³⁸² ³⁸³ ³⁸⁴ ³⁸⁵ ³⁸⁶ ³⁸⁷ ³⁸⁸ ³⁸⁹ ³⁹⁰ ³⁹¹ ³⁹² ³⁹³ ³⁹⁴ ³⁹⁵ ³⁹⁶ ³⁹⁷ ³⁹⁸ ³⁹⁹ ⁴⁰⁰ ⁴⁰¹ ⁴⁰² ⁴⁰³ ⁴⁰⁴ ⁴⁰⁵ ⁴⁰⁶ ⁴⁰⁷ ⁴⁰⁸ ⁴⁰⁹ ⁴¹⁰ ⁴¹¹ ⁴¹² ⁴¹³ ⁴¹⁴ ⁴¹⁵ ⁴¹⁶ ⁴¹⁷ ⁴¹⁸ ⁴¹⁹ ⁴²⁰ ⁴²¹ ⁴²² ⁴²³ ⁴²⁴ ⁴²⁵ ⁴²⁶ ⁴²⁷ ⁴²⁸ ⁴²⁹ ⁴³⁰ ⁴³¹ ⁴³² ⁴³³ ⁴³⁴ ⁴³⁵ ⁴³⁶ ⁴³⁷ ⁴³⁸ ⁴³⁹ ⁴⁴⁰ ⁴⁴¹ ⁴⁴² ⁴⁴³ ⁴⁴⁴ ⁴⁴⁵ ⁴⁴⁶ ⁴⁴⁷ ⁴⁴⁸ ⁴⁴⁹ ⁴⁵⁰ ⁴⁵¹ ⁴⁵² ⁴⁵³ ⁴⁵⁴ ⁴⁵⁵ ⁴⁵⁶ ⁴⁵⁷ ⁴⁵⁸ ⁴⁵⁹ ⁴⁶⁰ ⁴⁶¹ ⁴⁶² ⁴⁶³ ⁴⁶⁴ ⁴⁶⁵ ⁴⁶⁶

$\hat{C}^{\alpha\beta\gamma\delta} = \frac{1}{2}(\hat{C}^{\alpha\beta\gamma\delta} + \hat{C}^{\delta\gamma\alpha\beta})$
 $\hat{C}^{\alpha\beta\gamma\delta} = \frac{1}{2}(\hat{C}^{\alpha\beta\gamma\delta} - \hat{C}^{\delta\gamma\alpha\beta})$

тај узани, дубоки теснац, је — без обзира на то што се теоријски појмови исувише ретко преводе из једног научног дискурса у други — ипак и место додира. Када речником књижевне теорије говоримо о деконструкцији текста, говоримо заправо о његовом децентрисању, измештању из јединствене идеолошке матрице. Када говоримо о смрти аутора, говоримо о кризи ауторитета. А када говоримо о интертекстуалности, говоримо о тексту који није само оно што можемо из њега да прочитамо и који није идентичан са ликом/машином која га је написала. Другим речима, говоримо о тексту као „разоривом” предмету: предмету који више није условљен својим конвенционалним почетком и крајем; предмету чије се флуидне границе топе у општем феномену текстуалности и предмету чије тумачење, само по себи, укључује и могућност трансформације, превода, дописа.

Електронски текст је зато, између осталог, технолошка реализација „разоривог” текста: не само на динамичком нивоу софтверског исецања и преношења делова текста, него и на нивоу метаподатака који се захваљујући проширивом језику за обележавање (XML) могу кодирати, условно говорећи, унутар граница самог текста. Електронски текст који је обележен у XML-у је текст који садржи сопствене метаподатке, и који је бесконачно надоградив: од корисника, или кодера, зависи колика количина интерпретативног материјала ће бити маркирана у самом тексту.

Дигитална библиотека би зато морала да третира текст као разорив: не само као слику (која открива једну димензију текста) него као вишеслојни, машински обрадив, феномен, и као отворено поље за интеракцију између података и метаподатака. Текст мора бити „отворена књига” за анализу, трансформацију и уклапање у друге колекције. Ако дигиталне библиотеке желе да прате развој теоретског дискурса — и да се тиме и саме трансформишу из архива у лабораторије знања — онда је неопходно почети са размишљањем о томе како ће библиотеке интегрисати аналитичке алатке у систем испоруке, како ће градити мрежне заједнице за своје кориснике, стварати виртуелне просторе за колаборативно анализирање и аотирање материјала, и како ће се уклопити у шире оквире интернета као платформе знања. Кодирање текстуалног садржаја би свакако представљало изазов за традиционално поимање улога библиотека. Али ако се временом мења наше поимање текстуалности, зашто се не би мењало и схватање улоге коју библиотекар има у односу на текст?

У мојој визији будућности, дигиталне библиотеке би својим корисницима морале да пруже текстуалне корпусе који могу да се алгоритамски обрађују у потрази за обрасцима употребе, скривеним цитатима, идентификацијом ауторства итд; паралелне текстуалне корпусе — нпр. оригиналних књижевних дела и њихових превода на стране језике; или тематске корпусе, који би се градили на лицу места на основу претраге метаподатака али и семантички проширених претрага самог текста. Слободно доступни обележени текстови би тада могли да се подвргну статистичком тумачењу, рударењу или ископавању података (data mining) и стварању визуелних модела, укључујући ту презентацију невизуелних података и развој онога што Арнхајм назива „визуелним мишљењем”. Нама који се бавимо књижевношћу или дигиталним хуманистичким наукама, потребне су библиотеке у којима текст није само објекат, већ врста услуге; није статички елемент, него интерактивни метод.

Нема сумње да овде говорим о тектонским променама које за институције

које су по својој природи конзервативне, али дигитализација, отворени приступ и отворена компутација, односно рачунарско збрајање података, са собом носе интелектуалне, политичке, правне и етичке изазове за све институције знања, а пре свега библиотеке и универзитете. Веб је почео као издавачка платформа и извор ауторитета, док се сада увелико развија кроз „архитектуру учешћа” и „искориштавање колективне интелигенције” такозваног веба 2.0. Али веб 2.0 је и даље статички оквир у коме се подаци налазе на одређеним тачкама мреже, институционално или правно везани за појединачне изворе. Захваљујући претраживачима као што је Гугл могуће је релативно ефикасно пронаћи велики број извора и података, али архитектура мреже је таква да од корисника и даље захтева „путовање” од сајта до сајта.

Следећа генерација интернета — тзв. веб 3.0 или семантички веб — је утопија података који слободно плутају. У тој проширеној и надограђеној верзији светске мреже, циљ је развој универзалних софтверских агената који су у стању да интегришу и анализирају податке са различитих извора. Семантичка чворишта различитих ресурса (докумената, слика, људи, концепата) уз помоћ технологија као што су XML, RDF и OWL имају потенцијал да интернет учине паметнијом и амбициознијом платформом за размену информација. Библиотекама зато тек предстоје велики изазови укључивања сопствених колекција у семантички веб као мрежу знања.

Ако кренемо у правцу примене „разоривог” текста, можемо да замислимо, на пример, речник (или било које друго референсно дело) као умрежени сервис са сопственим API-ем (интерфејсом за програмирање апликација), који је интегрисан са садржајем дигиталног текстуалног спремишта. То значи да би свака реч сваког текста који је објављен у дигиталној библиотеци могла да буде повезана са електронским речницима. Електронски речници, пак, могли би да представљају један од могућих динамичких, дискурзивних полазних тачака за истраживање садржаја дигиталне библиотеке: од одреднице ка корпусу и назад. Дигиталне библиотеке будућности нису само виртуелна и дистрибуирана места у којима живе дигитални текстови, него места у којима они смислено и продуктивно интерагирају једни с другим.

Разориви текст није само теоретски концепт, већ физичка могућност дигиталног медија. Дигиталне библиотеке не треба да буду галерије слика, него корпуси машински обрадивих текстова који могу да се семплиују, упоређују један с другим и користе као сервис. За сада је тешко и претпоставити где леже — и да ли уопште постоје — границе примене слободног протока текста. Оно што је далеко извесније тиче се злослутног страха од краја хуманизма. Расплет се, како је то случај у многим наративима, садржи у самом заплету, а крај назире већ близу самог почетка: катастрофа никада неће бити коначна, док год постоји језик.

Тома Тасовац
Центар за дигиталне
хуманистичке науке, Београд

СВРХА И ПРОЦЕС ДИГИТАЛИЗАЦИЈЕ

Издавач
Фондација Народне библиотеке Србије

За издавача
Александар Гаталица

Координатор пројекта
Мирољуб Стојановић

Лектура и коректура
Ана Килибарда

Штампа
БИНДЕР ДИГИТАЛ, Београд

Дизајн и припрема за штампу
Тамара Ивљанин

Тираж 100 комада
Београд 23. април 2014.

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

025.85:004.91

СВРХА и процес дигитализације / [Мирољуб
Стојановић]. - Београд : Фондација Народне
библиотеке Србије, 2014 (Београд : Биндер
дигитал). - 24 стр. : илустр. ; 21 cm

Тираж 100.

ISBN 978-86-89921-00-7

а) Дигитализација
COBISS.SR-ID 206881036

ISBN 978-86-89921-00-7



9 788689 921007

