

4 festival hrvatskih digitalizacijskih projekata



Korištenje knjižničnih podataka za donošenje odluka u procesu digitalizacije građe

Sofija Klarin Zadravec

Hrvatski zavod za knjižničarstvo

Nacionalna i sveučilišna knjižnica u Zagrebu

Polazište

- Složena materijalna i intelektualna svojstva novina utječu na digitalizaciju i uspostavu sustava digitalne knjižnice
- Nedostatak podataka o svojstvima korpusa starih hrvatskih novina u NSK koji bi mogli poslužiti:
 - za cjelovito planiranje opsega i sadržaja aktivnosti digitalizacije starih hrvatskih novina temeljem relevantnih podataka
 - kao polazište za definiranje stvarnih zahtjeva građe u odnosu na funkcionalnosti sustava digitalne knjižnice starih novina

Knjižnični podaci

- izvori podataka
 - knjižnični informacijski sustavi,
 - podaci o građi/primjercima
 - podaci o knjižničnim korisnicima
 - podaci o korištenju građe
 - podaci o međuknjižničnoj posudbi
 - dnevničke datoteke sustava (log files)
 - podaci o građi
 - podaci o korisnicima
 - podaci o korištenju građe
 - podaci o pretraživanju itd.
 - dokumentacija poslovanja

Bibliomining – rudarenje knjižničnih podataka

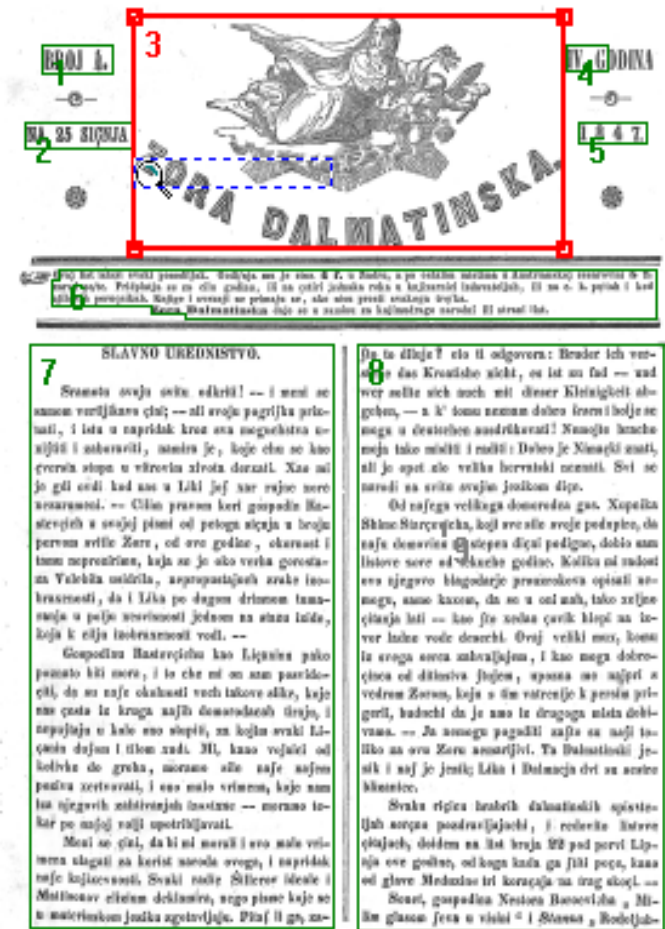
- Korištenje *bibliografskih, administrativnih i transakcijskih podataka* koje posjeduju i proizvode knjižnice može povećati spoznaju o poslovanju knjižnice, olakšati donošenje odluka i prilagodbu/poboljšanje poslovanja
Scott Nicholson (2003.) – nedovoljno korištenje podataka za poboljšanje usluga, bolju raspodjelu sredstava za nabavu, donošenje strateških odluka.

Područje digitalizacije i digitalnih knjižnica starih novina

- **Stare novine – vrijedan izvor informacija za istraživače i širu javnost**
 - složena materijalna i intelektualna svojstva
- Iznimna povezanost **oblika** (cjelokupni razvoj novinstva), **sadržaja** (razvoj hrvatskog jezika), količine – obuhvat (od kraja 18. st. do prve polovice 20. stoljeća), **postupaka** (skeniranje, OCR i sl.) i **kvalitete** konačnog proizvoda (prevladavanje ograničenja tiskanog oblika, kvalitetno pretraživanje)
- Jedan od zadataka - cjelovita analiza formalnih osobitosti koje mogu utjecati na postupak digitalizacije

Digitalizacija starih novina u NSK

- provodi se od 2005. godine
- dio programa zaštite starih hrvatskih novina
- skeniranje uglavnom s mikrofilma
- obrada preslika i preimenovanje datoteka
- masovni unos datoteka u sustav i OCR
- samostalno upravljanje cjelokupnim procesom od odabira do objave



- 2009. - 2014. – Portal **DaNiCa**
- Sustav DNC izrađen je suradnjom tvrtke Conscius i NSK (sustav za upravljanje digitalnim preslikama, importer za masovni unos datoteka i OCR, portal)
- Sadržaj – 181.505 stranica
 - 128.118 stranica novina
 - 53.387 stranica časopisa
- 9 ustanova suradnika
- 2013. - NSK je postala pridružena članica projekta *Europeana Newspapers*

BETA **Stare hrvatske novine**
portal digitaliziranih novina

Nacionalna i sveučilišna knjižnica u Zagrebu
Hrvatska kulturna baština


naslovnica | pregledavanje | pretraživanje | o projektu | suradnici | o starih novinama | časopisi | english

odaberite naslov

1. Življa Hrvatska (Zagreb...
2. Soborske novine
3. Kvarner (Bakar)
4. Čuk
5. Zatočnik
6. Branik. (Sisak)
7. Obzor. (Zagreb, 1871)
8. Pozor (Zagreb, 1860)
9. Jularnji list (Zagreb, ...
10. Narodne novine (Zagreb...

prikaži sve naslove

Pretraži tekst traži




odaberite datum

SID	VEL	OŽU	TRA
SVI	LIP	SRP	KOL
RUJ	LIS	STU	PRO

N P U S Č P S

odaberite mjesto izdavanja



ULAZ ZA SURADNIKE

Početna stranica | Sadržaj portala | Pomoć | Kontakt | Uvjeti korištenja

© Nacionalna i sveučilišna knjižnica u Zagrebu 2009. Sva prava pridržana.

BETA **Stari hrvatski časopisi**
portal digitaliziranih časopisa

Nacionalna i sveučilišna knjižnica u Zagrebu
Hrvatska kulturna baština

naslovnica | pregledavanje | pretraživanje | o projektu | suradnici | o časopisima | portal novina | english

odaberite naslov

1. Raški rudar
2. Oglasnik Parižke mode
3. Parižka moda
4. Neven
5. Mladi Istranin
6. Mladi Istran
7. Mladi dobrotvor
8. Bosiljak
9. Mladi Hrvat
10. Hrvatica (Zagreb, 1939)

prikaži sve naslove

popis starih hrvatskih znanstvenih i stručnih časopisa

A B C C D D Z D E F
G H I J K L M N N J
O P R S S T U V Z Z

Pretraži tekst traži



Portal digitaliziranih starih hrvatskih časopisa sadrži izbor najstarijih hrvatskih časopisa u digitalnom obliku Nacionalne i sveučilišne knjižnice u Zagrebu te preslike časopisa suradnika na portalu.

Cilj je Portala učiniti sadržaj hrvatskih časopisa što dostupnijim kulturnoj i znanstvenoj javnosti te ih zaštititi i očuvati kao kulturno i znanstveno nasljeđe. Projekt je financijski potpomoglo Ministarstvo kulture Republike Hrvatske i sastavni je dio projekta *Hrvatska kulturna baština*.

NOVOSTI: Na portalu je dostupno 35 naslova časopisa s 53.000 stranica. (2013-10-16). Na zahtjev suradnika portala, na Portalu je započela i objava novih digitaliziranih časopisa.

ULAZ ZA SURADNIKE

Početna stranica | Sadržaj portala | Pomoć | Kontakt | Uvjeti korištenja

© Nacionalna i sveučilišna knjižnica u Zagrebu 2009. Sva prava pridržana.

Temeljni ciljevi Portala

- Osigurati objavu novina na portalu (novinama) primjerenih funkcionalnosti
- Osigurati dinamiku u količini objavljenog sadržaja i broju novih naslova na portalu
- Osigurati ravnomjerni ^Σobuhvat naslova iz svih hrvatskih županija
- Osigurati kvalitetu sadržaja
- Povećati korištenje portala i zadovoljstvo korisnika



Odabir novina za digitalizaciju

Kriteriji za odabir naslova za digitalizaciju:

- objavljeni između 1789. i 1900.
- dostupni na mikrofilmu (=zaštićeni)
- najtraženiji naslovi

Izazovi:

- cjelovitost - obuhvat čitavog naslova :
dinamika objave (naslovi manjeg opsega) :
obuhvat svih povezanih naslova : naslovi iz
različitih županija/gradova
- posebni zahtjevi (npr. obljetnice)



FEST

10. - 11. travnja 2014.
Nacionalna i sveučilišna knjižnica u Zagrebu

Planiranje procesa digitalizacije

Primjeri pitanja koja postavljaju korisnici, novinari, kolege, ankete:

- koliko ste do sada digitalizirali novina (u postotku)?
- kada ćete digitalizirati sve novine?
- koliko je novina izlazilo u razdoblju od 1914. - do 1918. i koliko vam vremena treba da sve digitalizirate?
- koliko je naslova izlazilo u 1914. i koliko vam vremena treba da sve digitalizirate?



- Koliko je moguće skenirati stranica godišnje?
- Koliko je moguće obraditi stranica godišnje?
- Koliko naslova ulazi u taj broj stranica?
- Koliko određeni naslov ima stranica?
- Koliko određeni naslov zauzima diskovnog prostora? Koliko je godišnje potrebno povećati diskovni prostor?
- Koliki je postotak uspješnosti OCR-a? Što je utjecalo na kvalitetu OCR-a?
- Kakve su posljedice na pretraživanje?

Izvori podataka za odabir građe za digitalizaciju/objavu

- Bibliografski podaci iz kataloga NSK/Portala
 - provedena *Analiza formalnih svojstava starih novina u fondu NSK (2012.)*
- Popis mikrofilmiranih naslova novina
 - Podaci o fondu mikrofilmova unose se uz zapise na Portalu novina
- Podaci o korištenju tiskane građe u čitaonici
- Baza podataka o pripremi građe za mikrofilmiranje (Odjel zaštite i pohrane)
- Podaci o raspodjeli naslova novina prema županijama
 - Naslovi novina prikazuju se na interaktivnoj karti Portala novina
- Literatura ; Anketa HKD-a ; Dodatni popisi

Izvori podataka za praćenje procesa i mjerenje kvalitete

- Dokumentacija radnog procesa
- Statistički podaci Portala
- Google Analytics
- Dnevničke datoteke portala
 - Provedena su 2 istraživanja dnevničkih datoteka portala *Stare hrvatske novine* (2012.)
 - *Istraživanje predmeta upita korisnika portala DNC*
 - *Istraživanje obuhvata upita korisnika portala DNC*

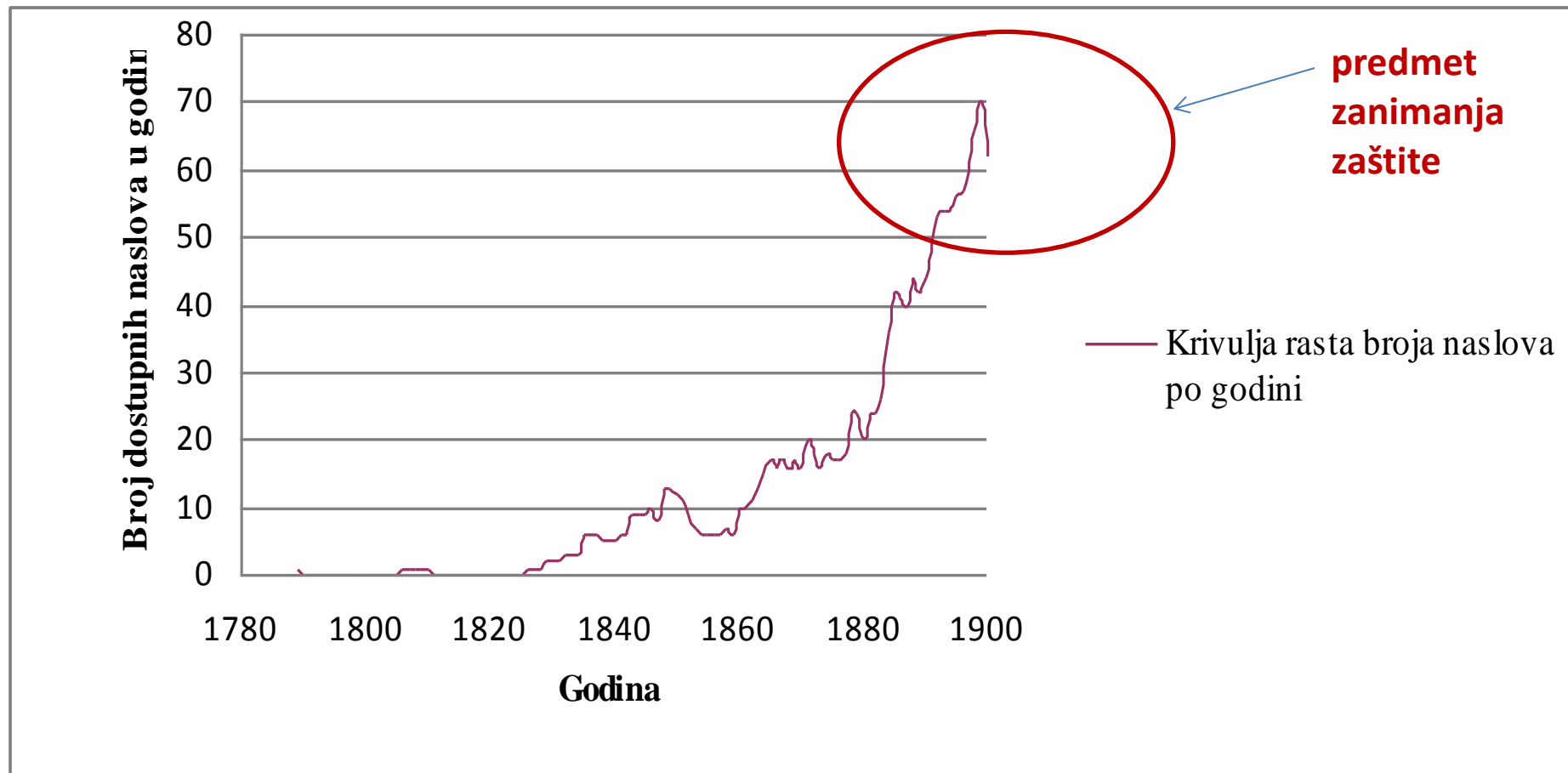
Primjer 1: *Analiza bibliografskih podataka zbirke starih hrvatskih novina (1789.-1900.)*

- Analiza provedena 2012. godine
- Uzorak - bibliografski zapisi starih novina (izvornika) na portalu *Stare hrvatske novine*
- Cilj: utvrđivanje formalnih svojstava zbirke starih novina iz fonda NSK radi preciznijeg sagledavanja svojstava zbirke i pomoći pri donošenju odluka u procesu digitalizacije i objave preslika

Elementi analize

- raspodjela broja objavljenih naslova po godinama
- učestalost izlaženja
- format
- odnos učestalosti izlaženja i formata
- količina građe (procjena)
- mjesta izdavanja
- veze među bibliografskim zapisima
- jezik i pismo
- zastupljenost ilustracija
- postojanje preslika (mikrofilm, digitalna preslika)

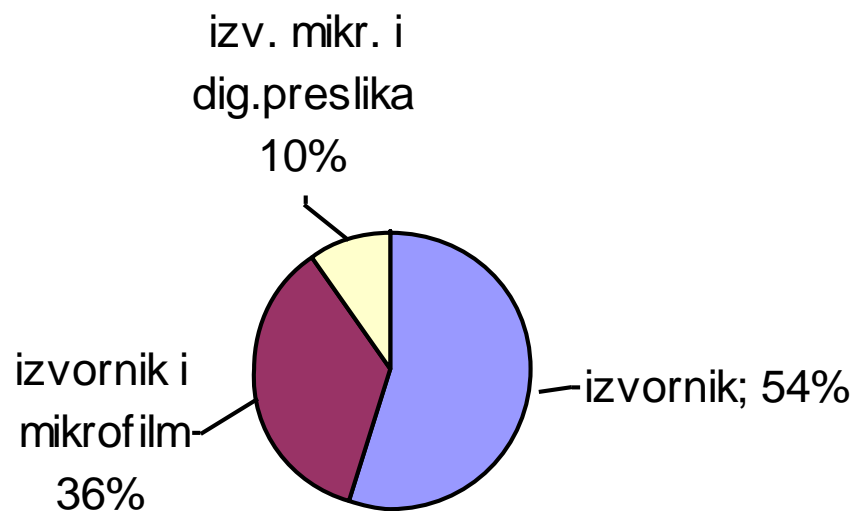
Broj objavljenih naslova u godini (1789.-1900.)



Veze među zapisima

R.broj	Zastupljene veze u bibliografskim zapisima	Broj veza	Broj zapisa
1.	Prethodni naslov	1	13
2.	Prethodni naslov, Ima prilog	2	1
3.	Prethodni naslov, Prethodni naslov	2	1
4.	Prethodni naslov, Sljedeći naslov	2	7
5.	Prethodni naslov, Sljedeći naslov, Ima prilog	3	4
6.	Prethodni naslov, Sljedeći naslov, Ima prilog, Ima prilog	4	2
7.	Prethodni naslov, Sljedeći naslov, Usporedno izdanje na drugom jeziku	3	1
8.	Prethodni naslov, Usporedno izdanje na drugom jeziku	2	2
9.	Sljedeći naslov	1	13
10.	Sljedeći naslov, Ima prilog	2	3
11.	Sljedeći naslov, Ima prilog, Ima prilog	3	1
12.	Sljedeći naslov, Sljedeći naslov	2	1
13.	Sljedeći naslov, Usporedno izdanje na drugom jeziku	2	1
14.	Usporedno izdanje na drugom jeziku	1	3
15.	Ima prilog	1	4
16.	Prilog uz	1	7
17.	Prilog uz, Sljedeći naslov	2	2
18.	Prilog uz, Prethodni naslov	2	3
19.	Prilog uz, Prethodni naslov, Sljedeći naslov	3	4
			Ukupno: 73

Pojavni oblici*

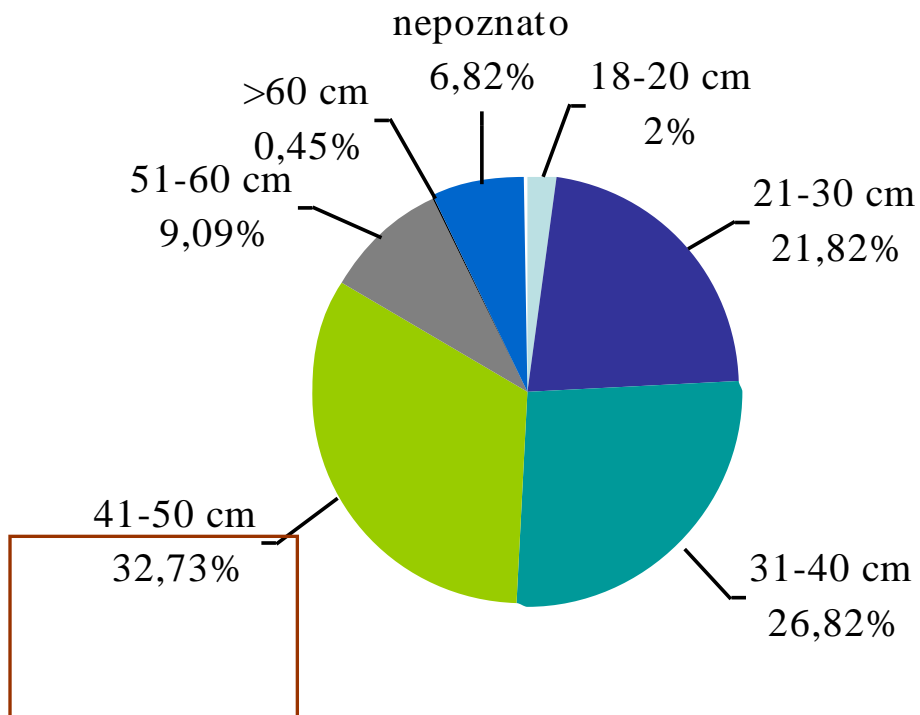


* uz bibliografske zapise korišteni su i dodatni izvori podataka

Učestalost izlaženja

Učestalost izlaženja	Broj naslova	Postotak udjela naslova
dnevno	30	13,39%
tri puta tjedno	9	4,02
polutjedno	25	11,16%
tjedno	64	28,57%
dvotjedno	12	5,36%
polumjesečno	29	12,95%
<i>tri puta mjesečno</i>	4	1,79%
<i>mjesečno</i>	2	0,89%
<i>tromjesečno</i>	1	0,44%
nepoznato, neredovito i ostalo	49	21,88%
	225	100 %

Format novina

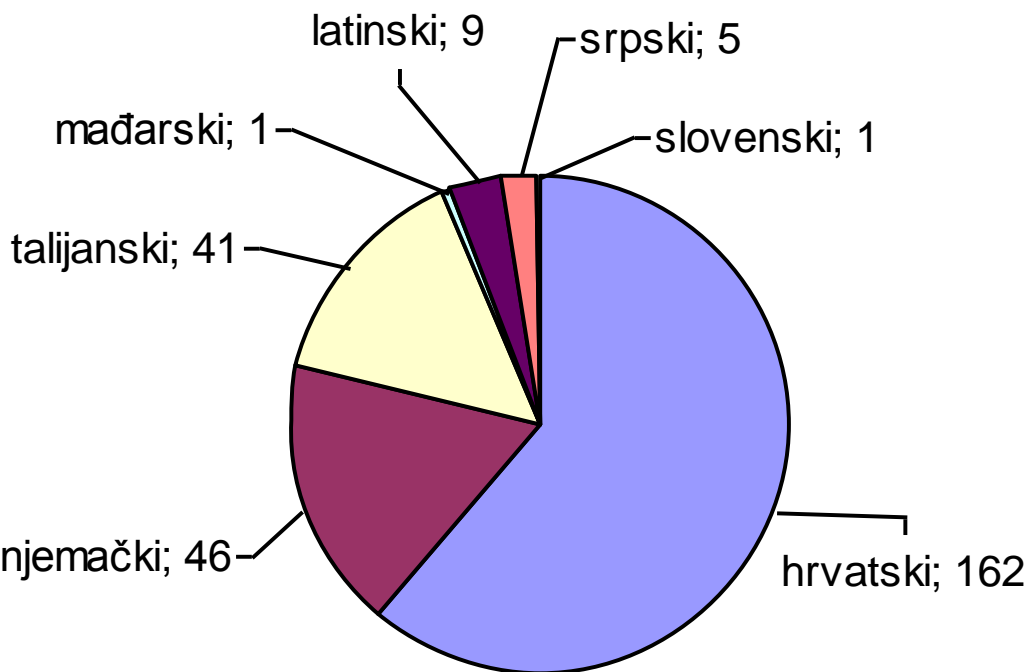


Mjesta izdavanja*

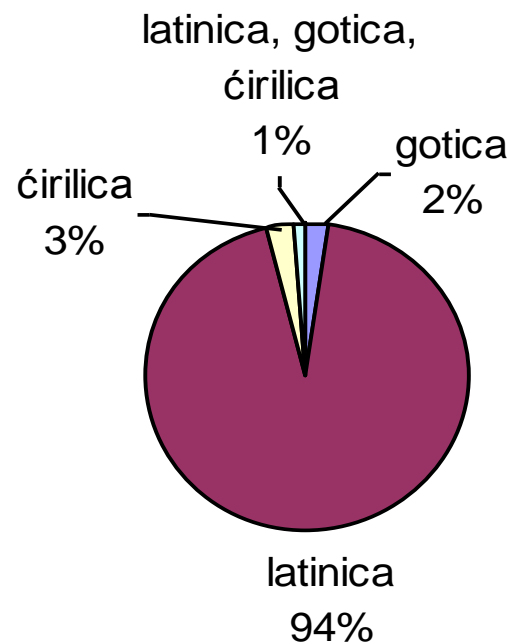
R. br.	Mjesto izdavanja	Broj naslova
1.	Zagreb	103
2.	Zadar	36
3.	Osijek	11
4.	Rijeka	11
5.	Dubrovnik	7
6.	Varaždin	6
7.	Karlovac	5
8.	Split	5
9.	Vukovar	5
10.	Sisak	4
11.	Gospić	3
12.	Vinkovci	3
13.	Bakar	2
14.	Ogulin	2

* Prikazan dio podataka

Jezici



Pismo



Ilustrirane novine

R.br.	Zastupljenost podatka o ilustracijama	Mjesto izdavanja	Razdoblje izlaženja	Učestalost izlaženja
1.	Giornale di Fiume	Rijeka	1865.	tjedno
2.	Esseker allgemeine illustrierte Zeitung	Osijek	1869.	tjedno
3.	Illustrierte, belletristische Sonntags-Beilage der esseker Zeitung	Osijek	1876.	tjedno
4.	Slovinac	Dubrovnik	1878-1884	dvotjedno
5.	Hrvatska vila	Sušak	1882-1885	tjedno
6.	Svjetlo	Karlovac	1884-	polutjedno
7.	Arkiv za šalu i satiru	Zagreb	1885-1886	polumjesečno
8.	Bič	Zagreb	1885-1885	polumjesečno
9.	Ćuk	Zagreb	1888.	polumjesečno
10.	Dom i svijet	Zagreb	1888-1923	polumjesečno
11.	Stršen	Zagreb	1897-1901	polumjesečno
12.	Agramer Lloyd	Zagreb	1898-1908	tri puta mjesečno

Zaključno o provedenoj analizi

- Dobiveni podaci daju okvir za donošenje odluka o odabiru građe, odabiru opreme, praćenju i analizi rezultata OCR-a te za ukupno planiranje i praćenje procesa digitalizacije
- Dostupnost cjelovitih bibliografskih podataka starih hrvatskih novina u katalogu NSK odnosno postojanje skupnog kataloga osigurali bi i zanimljiv uvid u povijest hrvatskog novinskog nakladništva i tiskarstva

Zaključno o korištenju knjižničnih podataka

- Dostupno je više izvora vrijednih podataka u knjižnicama koji, nakon okupljanja i primjene odgovarajućih metoda analize, mogu poslužiti za donošenje strateških odluka i poboljšanje poslovanja
- Razmisliti o načinu njihova povezivanja, analize i korištenja
- Voditi brigu o zaštiti privatnosti korisnika