

INFRASTRUKTURA SLOVENISTIČNE LITERARNE VEDE

Miran Hladnik

Filozofska fakulteta, Ljubljana

UDK 821.163.6.09:004.738.5

Slovenistična literarnovedna infrastruktura je predmet nujne posodobitve, ki zajema digitalizacijo njenega predmeta slovenske književnosti od začetkov do danes, preselitev kompletne strokovne komunikacije, raziskovalnega ter pedagoškega dela na internet ter digitalizacijo preteklih literarnovednih objav. Članek predstavlja trenutno stanje teh procesov in njihovo perspektivo.

slovenistika, digitalizacija dokumentov, podatkovne zbirke, Digitalna knjižnica Slovenije, Videolectures, Wikivir, Wikipedija

Slovene literary studies is being rapidly brought up to date through the digitalisation of Slovene literature from its beginnings to the present, the relocation of academic communication, research and teaching work to the Internet, and the digitalisation of past published works on literary studies. The article describes the current situation and takes a look into the future.

Slovene Studies, digitalisation of documents, databases, Slovene Digital Library, video lectures, Wikisource, Wikipedia

Infrastruktura je družboslovni pojem in pomeni temeljne tehnološke in organizacijske pogoje za funkcioniranje socialnega sistema oz. podjetja. Pojavil se je v 20. letih 20. stoletja, popularen pa je postal v 80. letih, zadnje čase tudi v humanistiki. Pred leti bi na vprašanje, kaj si predstavljamo pod infrastrukturo literarne vede, najprej pomislili na raziskovalne in izobraževalne inštitucije s knjižnicami, po vsebinski strani pa na tista strokovna področja s pripadajočimi inštitucijami, ki smo jim dodelili status t. i. pomožnih ved: biografiko, bibliografijo in tekstologijo. Danes pomislimo na prost dostop do podatkov in na komunikacijske kanale, ki lajšajo dostop do

njih in omogočajo distribucijo raziskovalnih rezultatov. Pod infrastrukturo razumemo v prvi vrsti internetne vsebine in orodja za udobno delo v stroki.¹ Neobhodno in samoumevno se nam zdi, da so na klik dosegljivi vsi mogoči bibliografski in biografski podatki pa tudi besedila sama. Novo je, da med osnovne pogoje za funkcioniranje literarnovednega sistema štejemo možnost nemudnega spletnega objavljanja, hrambe študij ter strokovnega diskutiiranja.

Obsežnost pojma infrastruktura je odvisna od dožemanja naše lastne dejavnosti. Svojtčas je bila običajna identifikacija raziskovalcev in univerzitetnih učiteljev s stroko, kar bi lahko

1 Christine L. Borgman, *Scholarship in the Digital Age: Information, Infrastructure and the Internet*, MIT Press. Ker je navajanje spletnih lokacij v tisku nesmiselno, saj jih nikoli ne pretipkavamo v brskalnikovo okno, nanje opozarjam samo s podčrtavo. Tudi datuma ogleda ne navajam, ker je samoumevno, da se povezave preverijo tik pred objavo.

povzeli z izjavo *Literarna veda, to smo mi* ali v ekstremnih primerih celo *Literarna veda, to sem jaz*. S tega stališča spada v infrastrukturo vse, kar omogoča naše delo. Bolj je v skladu z današnjim časom in zdi se bolj perspektivno dojemanje lastnega dela kot javnega servisa. S tega stališča javni interes oz. uporabniki določajo, kaj je potrebno za zadovoljevanje njihovih potreb v literarnovednem sistemu. Infrastruktura tedaj zajema tudi fakultetne oddelke in inštitute, ki se posvečajo literarnemu študiju in raziskovanju, skupaj z zaposlenimi učitelji, asistenti in raziskovalci. Te službe naj bi servisirale področja, kjer je literatura najbolj v obtoku, to pa so, navajam iz svojega *Praktičnega spisovnika*, »šolstvo s predmetom slovenščina, založništvo s potrebo po spremnih besedah, recenzijah in redakciji besedil, mediji s potrebo po literarnih kritikah, lokalne in državne oblasti ter javnost s potrebo po kulturni in nacionalni reprezentanci, ki se kaže v praznovanju jubilejev literarnih osebnosti ter podeljevanju literarnih nagrad«. ² Vloga servisa je daleč manj ugledna kot tradicionalno samorazumevanje, ki so ga legitimirale posvečenost, predanost in odgovornost zgolj vzvišenemu predmetu samemu, to je slovenski književnosti, ki je bila zaslužna za vzpostavitev nacionalnega sistema in zavesti.

Infrastruktura stroke je iz treh segmentov. Na tisti segment, ki v skladu z etimologijo izraza (*infra*- lat. 'spodaj') čisto na dnu omogoča naše delo, to je na računalniško tehnologijo, se humanisti povečini ne spoznamo, zato o njej ne bom izgubljal besed. Naše kompetence se začenjajo pri drugem segmentu, to je pri vsebinah elektronske infrastrukture, kjer naj bi določali, katere informacije naj bodo vsak čas dostopne in v kakšnem preferenčnem zaporedju naj bodo postavljene, ter bili pri pripravi in

postavljanju informacij sami dejavni. Mislimo na gradnjo elektronske knjižnice in podatkovnih zbirk s področja literarne vede. Tretji segment infrastrukture, ki ga konstituira razumevanje inštitucij kot javnega servisa, so t. i. »človeški viri« oz. po starem kadrovska zasedba; o tem sem nekaj že pisal, ³ podrobnejšo analizo pa bo treba prihraniti za drugo priložnost.

Posvetimo se torej samo digitalizaciji strokovnih vsebin. Ne gre samo za selitev obstoječih informacij s težko dostopnih arhivskih in knjižničnih lokacij na prosto dostopna internetna mesta, ampak tudi za spremembo razmerij med udeleženci v tem interesnem krogu. Včasih je javnost iz zaupanja v t. i. stroko prepuščala pobudo in selekcijo informacij literarnovednim avtoritetam. Ker ni več prepričljivih avtoritativnih arbitrov in splošno sprejetega kriterija za selekcijo in ker so bile literarnovedne inštitucije v razmerju do tehnoloških možnosti prej zadržane kot ne, so pobudo prevzemali posamezniki z zavestjo o pomembnosti nove raziskovalne in kulturne paradigme, ki jo definira možnost slehernikove aktivne in kreativne participacije. Slehernik v tej kulturi ni dojet samo kot uporabnik, ki posega po dostopnih informacijah, ampak kot človek, ki želi sodelovati tudi pri ponudbi informacij. Nov koncept lepo dokumentira ključni pojem *sharing*, katerega pomen 'kooperativna udeležba pri ponujanju informacij ali izkušenj' ali 'soudeležba pri uporabi, izkušnji ali veselju', ⁴ v slovenščini nima utrjenega prevoda. Slovenci pod *sharing* vztrajno razumemo samo udeležbo pri delitvi dela in dobička, ne uzaveščamo pa dovolj tistega pomena besede, ki izhaja iz koncepta *sharing economy* (obdarovalna ekonomija) in prihaja do izraza, npr. pri pojmu *shareware* (programska oprema, ki jo zastonj snamemo z interneta in jo plačamo

² *Praktični spisovnik* (2002, 1. izdaja 1990) seznanja študenta s slovenistično literarnovedno infrastrukturo: z inštitucijami, publikacijami in orodji za ukvarjanje s književnostjo.

³ Slovenska književnost na Filozofski fakulteti 1919–1979–2009, *Slovenistika: Pogledi na eno od ustanovnih strok Univerze v Ljubljani ob njeni 90-letnici*, Ljubljana: Znanstvena založba FF, 2009, 30–49.

⁴ Prosto po *The Free Dictionary*.

samo v primeru, če jo zares in dalj časa uporabljamo) ter pri praksi licenciranja avtorskih del pod znamko Creative Commons. S tem izkazujemo svoje počasno sprejemanje in nezaupanje do nove kulturne paradigme.

Slovenske univerze so zadržane pri nujnem prehodu s tiskanih na digitalne objave. Stremljive univerze po svetu ena za drugo prehajajo s tiska na digitalno objavljanje. Tiska s tem seveda ni konec, ker bralec natisnjeno varianto publikacije vedno lahko naroči (*print on demand*, 'knjige na zahtevo'), čemur znajo ustreči tudi naše tiskarne,⁵ ali pa publikacijo, ki jo dobi na zaslon, iztisne sam. Primarna oblika objave je zdaj elektronska. Novo dejstvo z eno potezo pomete s tradicionalnim »argumentom«⁵ proti objavljanju na spletu, češ da je *knjiga le knjiga*. Verodostojnost spletnih objav ni nič manjša, če za njimi stoji enak recenzijski mehanizem kot pri tiskanih znanstvenih revijah. Z ustreznim vmesnikom se da zavarovati tudi komercialna dimenzija znanstvenih objav, čeprav je pritisk na inštitucije, da rezultate raziskav, ki so financirane iz davkoplačevalskega denarja, objavljajo zastoj, vedno večji. Hramba in distribucija sta pri elektronskih publikacijah neprimerno učinkovitejši kot pri tiskanih publikacijah.

Digitalizacija zadeva vsa področja literarne vede (literarno zgodovino in teorijo, metodologijo in zgodovino stroke). Digitalizirati je treba obstoječe informacije leksikonske in enciklopedične narave, bibliografije, kataloge, članke, zbornike in monografske razprave. Digitalizirati se mora aktualno znanstveno publiciranje, komunikacija ob raziskovalnih projektih, konferencah in predavanjih ter merjenje znanstvene odmevnosti. Posneta javno dostopna predavanja bodo do neke mere nadomestila fizično obiskovanje predavanj.

Za katere vsebine konkretno gre v literarni vedi? Teoretiki razlikujejo med dvema vrstama informacij, med dokumenti in podatki.

Dokument je fizična reprezentacija informacije s komunikacijsko intenco in funkcijo. V literarni vedi je dokument drugo ime za besedilo. Tradicionalno so besedila na papirju, neobjavljena kot rokopis, objavljena kot knjiga ali članek, po novem pa v obliki računalniške datoteke različnih formatov: besedilnih (txt, doc, rtf, pdf, html), slikovnih (jpg, gif, png), filmskih (avi, wmv), zvočnih (mp3, wav), kot blogi, e-pošta ali wikiji. Podatki so osnovni elementi informacije, pridobljeni večinoma z meritvami in prikazani v formatiranih zapisih v podatkovni zbirki ali v obliki tabel in grafov. Idealno literarnovedno spletišče naj bi vsebovalo tako podatke kot dokumente. Tako je zgrajena Digitalna knjižnica Slovenije [dLib](#). Iskalcu ponudi najprej vrsto podatkov, ki ustrezajo njegovi poizvedbi, ponudi pa mu tudi, da s klikom na izbrani podatek prikljče na zaslon celotni dokument. Do neke mere gre v to smer tudi zbirka [Nova beseda](#) Primoža Jakopina. Uporabniku, ki išče določeno besedo, ponudi rezultat v obliki serije konkordanc (vsak zapis je svoj podatek), v posameznih primerih pa preko povezave na ustrezna spletišča omogoči tudi dostop do celotnega teksta, v katerem se iskana beseda pojavi.

Običajno je bilo težko priti do podatkov. Če jih nismo generirali sami pri svojih raziskavah, se je bilo treba potruditi in jih poiskati na mestih, kjer so bili dostopni za javno uporabo: besede v slovarjih ali kataložnih zbirkah, publikacije v objavljenih bibliografijah ali listkovnih katalogih knjižnic, avtorje in druga dejstva v leksikonih in enciklopedijah, bodisi natisnjenih bodisi v arhivirani obliki. Specialni podatki so bili v lasti in hrambi strokovnih kolegov in je bilo treba zanje prositi osebno, če avtorji niso poskrbeli za knjižno objavo, kot npr. Gregor Kocijan v bibliografijah slovenske kratke proze. Neredko je raziskovalec čepel nad svojimi podatki in jih ljubosumno čuval pred strokovnimi kolegi v

5 [Storitve digitalnega tiska: Birografika Bori](#).

strahu, da jih bodo uporabili pred njim. Število javno dostopnih podatkovnih zbirk, iz katerih je mogoče poljubno črpati, včasih pa jim tudi kaj dodajati,⁶ narašča. Za slovensko literarno vedo pridejo v poštev slovarji (*SSKJ*, *Besede slovenskega jezika*), besedilna korpusa *Nova beseda* in *Fidaplus*, bibliografije (slovenski vzajemni katalog *Cobiss*, zbirka *Diplomske naloge iz slovenske književnosti na FF 1950–2008*, zbirka *Slovenski zgodovinski roman in Slovenska kmečka povest*), leksikonske informacije biografskega značaja (Wikipedija, *Slovenski biografski leksikon (SBL)*, *Primorski slovenski biografski leksikon (PSBL)*). Pri konkretnem delu pridejo prav tudi druge informacije javnega značaja (*Telefonski imenik Slovenije* in slovenski zemljevidi na *Geopediji*).

Literarnemu zgodovinarju prideta prav dve vrsti digitalnih dokumentov, literarnovedni in literarni. Literarnovedna vsebina prihaja do uporabnika na zaslon z digitalizacijo starih in svežih letnikov strokovnih in znanstvenih revij. *Slavistična revija* od 1997 dalje sproti digitalizira svoje izdaje, za nazaj pa še leta 1993–1995, revija *Jezik in slovstvo* pa je dostopna v celoti tudi preko portala dLib. *Primerjalna književnost* se z enoletnim časovnim zamikom pojavlja na spletu od 2005 dalje, *Slovene Studies* napoveduje skorajšnji dostop do arhivskih izvodov od začetka izhajanja (1979–), *Acta neophilologica* pa do vzporedne digitalne oblike še ni prišla. Med zborniki z literarnovedno vsebino so od 2002 dalje dostopna *Obdobja*, od 2004 pa zborniki *Seminarja slovenskega jezika, literature in kulture*. Oddelek za slovenistiko je l. 2009 na splet postavil dva zbornika (*Zbornik 4. slovensko-hrvaškega slavističnega srečanja in Slovenistika: Pogledi na eno od ustanovnih strok Univerze v Ljubljani ob njeni 90-letnici*), knjižno izdane študijske programe pa postavlja na splet od leta 1997. Oddelčna kronika (*Novice Oddelka za slovanske književnosti na Filozofski fakulteti*

Univerze v Ljubljani ali Kronika ljubljanske slavistike 1996–1998) je izšla vzporedno na spletu leta 1998. Publikacije Slavističnega društva Slovenije so pospletene selektivno, npr. Kozma Ahačič, *Izvirne slovenske pesmi Jovana Vesela Koseskega 2006* (Slavistična knjižnica, 9), *Zborniki SDS 16–18 za leta 2005–2007*. Založbe za digitalizacijo niso zainteresirane, pa naj gre za založbe akademskega ali komercialnega tipa. Na svojih spletnih straneh objavljajo samo naslovnice in kazala, v najboljšem primeru kakšen odlomek iz knjige, cele knjige pa zgolj na papirju. To velja za Znanstveno založbo FF, za zbirko *Zora* Filozofske fakultete v Mariboru in za zbirko *Studia litteraria* pri ZRC SAZU. Tudi med učbeniki so prosto dostopne spletne knjige redkost in posledica avtorjeve osebne iniciative, npr. Hladnikova *Praktični spisovnik* (2002) in *Slovenski zgodovinski roman* (2009).

Spremembe nakazuje nova serija *Digitalne objave* pri ZRC SAZU, kamor so vključene tudi *Elektronske znanstvenokritične izdaje slovenskega slovstva (e-ZISS)*, ki jih pripravlja Inštitut za slovensko literaturo in literarne vede in prinaša sedem znanstveno kritičnih izdaj. Gre v glavnem za starejša besedila (Brižinski spomeniki, Škofjeloški pasijon, pridige Antona Martina Slomška, korespondenca Žiga Zoisa, mestne prisege), bližje sodobnosti pa so pesmi Antona Podbevška in roman *S poti* Izidorja Cankarja.

Filozofska fakulteta UL je pred leti poskusila s sprotnim objavljanjem diplomskih del na svojih spletnih straneh, vendar zadeva ni stekla in fakulteta je celo ukinila povezavo do njih, tako da je deset slovenističnih literarnovednih diplomskih del dostopno le z internega *Seznama avtorjev diplomskih nalog iz slovenske književnosti*. Posamično objavljene študije najdemo v osebnih spletnih bibliografijah nekaterih literarnih zgodovinarjev, npr. pri *Miranu Hladniku* in *Marku Juvanu*.

⁶ *Razvezani jezik: Prosti slovar žive slovenščine*.

Bolj kot aktualni literarnovedni razpravni dosežki so digitalizaciji naklonjene starejše publikacije, ki so 70 let po smrti avtorjev postale javna last. Rokopisno *Slovensko slovstvo* (1871/1872) Frana Levca smo digitalizirali in prepisali sami,⁷ druge so prinesli digitalizacijski projekti pri Googlu ali dLibu in je tako na naših ramenih samo njihovo korigiranje in predstavljanje na Wikivir; strojno prepoznavanje je namreč pri starejših slovenskih tekstih pogosto problematično in je korektorjeva ročna intervencija nujna. DLibovo pospletanje edicije Knezova knjižnica, ki jo je izdajala prva slovenska založba znanstvenega tiska Slovenska matica, in pospletanje Slovenskih večernic je naplavilo tudi nekaj zgodnjih literarnovednih študij. To so Frana Levca *Dr. Janez Bleiweis* (SV 35, 1880), *Matija Valjavec: Životopis*, 1895 (KK 2, 1895) in *Dr. Fr. Prešeren* (KK 7, 1900), Josipa Cimpermana *Življenje in pesni Franje Ser. Cimpermana* (1874), Gregorja Kreka *Einleitung in die slavische Literatur-geschichte; akademische Vorlesungen, Studien und kritische Streifzüge* (1887) in *Anton Aškerc, Studie mit Übersetzungsproben* (1899), Ivana Prijatelja *Drama Prešernovega duševnega življenja (Naši zapiski* 1905) in Dragotina Lončarja ocena Prijateljeve študije *Janko Kersnik, njega delo in doba*, ki je izšla v goriški reviji *Veda* 1914.

Največja pridobitev digitalizacij zunaj Slovenije je serija literarnovednih knjig. Najprej sta tu dve knjigi Ivana Macuna, *Cvetje slovenskiga pesništva* (Trst, 1850) in *Kratak pregled slovenske literature* (1863) – poetika z antologijo in literarnozgodovinski pregled. V več kopijah so na voljo štiri knjige *Zgodovine slovenskega slovstva* Karla Glaserja (1894–98)⁸ in deloma se slovenski književnosti posveča tudi Matije Murka *Geschichte der älteren*

südslawischen Litteraturen (1908). Dokler ne bodo pretvorjene v obliko podatkovne zbirke, bomo med dokumenti naštevati tudi digitalizirane bibliografske monografije: Franc Simonič, *Slovenska bibliografija (1550–1900)*, Janko Šlebinger, *Slovenska bibliografija za l. 1907–1912* in Niko Kuret, *Slovenska knjiga: Seznam po stanju v prodaji dne 30. junija 1939*.

Počasi narašča število posnetih predavanj. Na portalu *Videlectures* je v kategoriji Humanities: Literature osem predavanj (Milena Blažič, Miran Hladnik, Gregor Pompe, Alojzija Zupan Sosič, Andrej E. Skubic, Borut Telban, Jana Kolar), v drugih kategorijah pa še predavanja Tarasa Kermaunerja, Petra Svetine in Toneta Smoleja. Še druge najdemo na portalu *Youtube* (Helgo Glušič, Igorja Krambergerja) ali na seznamu v *Slovenistični literarnovedni kašči* (Gregorja Kocijana, Janeza Vrečka, Marijo Mitrovič, Matjaža Zupančiča).⁹ Popularnejše oblike literarnovednega diskurza v zvočnem ali filmskem zapisu se vsaj za nekaj časa pojavljajo tudi na radijskih in televizijskih spletnih straneh.

Na Wikipediji izvajajo infrastrukturne projekte študentje: *Slovenski literarni zgodovinarji*, *Slovenska mladinska književnost*, *Slovenska literarna veda* itn. V treh letih smo v okviru diplomskega seminarja postavili 160 gesel o slovenskih literarnih zgodovinarjih. Na seznamu jih je skoraj še enkrat toliko, torej za nadaljnja tri seminarska leta, vendar bomo v prihodnje raje dali prednost drugim enciklopedičnim informacijam (literarnovednim terminom, inštitucijam, tekstom, avtorjem), kar bo osnova za literarni portal na domači Wikipediji.¹⁰

Enako pomembna kot digitalizacija literarnovednih virov je digitalizacija glavnega predmeta slovenistične literarne vede, to je

7 Za natis se pripravlja kritična izdaja tega rokopisa s spremno študijo.

8 Včasih je publikacije težje najti, ker je jezik pri marsikateri označen kot hrvaški ali slovaški.

9 Predavanja je snemal Matjaž Rebolj ali pa kar sami s funkcijo video na fotoaparatu.

10 Angleški *Literature: Portal* navaja blizu 40 jezikov z Wikipedijinim literarnim portalom. Pomenljivo je, da Slovenci, ki se sicer radi postavljamo kot »literarna nacija«, moči za kaj takega do zdaj nismo zbrali.

slovenske književnosti. Infrastruktura literarne vede je v preteklosti nastajala ob dejavni participaciji literarnih zgodovinarjev in naša udeležba je nujna tudi pri digitalizaciji. Prvo slovenistično podjetje tega tipa, Prešernove konkordance, se je rodilo v tujini.¹¹ Tam se je dogajala tudi prva obširnejša digitalizacija klasikov v začetku 90. let.¹² Javno dostopni so ti teksti postali leta 1995, ko je v okviru literarne spletišča na strežniku Instituta Jožef Stefan zaživela stran z naslovom Zbirka slovenskih leposlovnih besedil. Zadnja leta sveže digitaliziranih besedil ne postavljamo več sem, ampak na Wikivir, ki omogoča širšo udeležbo strokovnjakov in laikov pri tem početju.

Naštejmo še druge udeležence pri gradnji digitalnega skladišča slovenske književnosti. Za 300 knjig obsežni kontingent je poskrbel Franko Luin v zbirki Beseda,¹³ ki pa po urednikovi smrti leta 2005 ni redno vzdrževana in vedno dosegljiva. Ker je bil Luin tipograf, so knjige lepo oblikovane, vendar smo jih, ker so nepriročno zaklenjene v pdf-format, konvertirali in predstavili na Wikivir. Skeniranje in prepoznavanje več kot 200 knjig, prevedenih iz nemščine, je organiziral Erich Prunč z graške univerze.¹⁴ Delo za njegov projekt so opravili pri dLibu, zato so knjige dostopne tudi preko portala Digitalne knjižnice Slovenije. Projekt Rastko Slovenija je vseboval prevode slovenske književnosti v srbsčino. Nekaj dragocenih slovenističnih publikacij se je znašlo na zaslونih po zaslugi Googleove digitalizacije štirih ameriških univerzitetnih knjižnic (Google knjige). Od tod imamo prve izdaje Valentina Vodnika, Franceta Prešerna in Antona Martina Slomška. Googleove knjige prezentira portal

Internet Archive, kjer v rubriki Texts med 1,7 milijona knjigami najdemo tudi letnike Stritarjevega *Zvona* (1877–1879), Janeza Trdino, Antona Aškerca, Slovenske večernice in Knezovo knjižnico. Vsi drugi ponudniki slovenske književnosti iz naštetih virov večinoma le zajemajo: vatikanski Intratext, Logos Library, zvočne knjige, elektronske knjige za čitalnike in za mobilne telefone, Virtualna knjižnica s teksti, primernimi za branje z Braillovo vrstico.

Infrastrukturalna podjetja so na vseh področjih (promet, komunala, komunikacija) obsežna in finančno zahtevna, zato se jih lotevajo velike inštitucije z državno podporo. Zelo nenavadno bi bilo, če bi pri naštevanju akterjev, ki vzpostavljajo infrastrukturne pogoje za funkcioniranje literarne vede, ostali le pri tujcih in domačih individualnih volonterjih. Domače inštitucionalne inštanace, od katerih smo pričakovali akcijo (univerze, inštituti, nacionalna knjižnica, akademija in ustrezna ministrstva), so na civilizacijski izziv odreagirale sprva nekam zaspano in bile zato deležne mojih kritičnih pripomb, vendar je v zadnjih letih delo steklo in vzbuja zaupanje, da bo digitalni dostop do vsebin postopoma nadomestil tradicionalnega.¹⁵ Od leta 2005 največ k rasti digitalnih vsebin prispeva Digitalna knjižnica Slovenije v okviru Nacionalne in univerzitetne knjižnice. Prednostno digitalizira starejšo fizično občutljivo in težko dostopno periodiko (210.000 enot v obliki člankov) in neknjižno gradivo (stare fotografije, plakate, notne zapise, rokopise, zemljevide, slike) ter disertacije; knjige (trenutno je prebranih samo 450) pridejo na vrsto pozneje. Literarni zgodovinarji

11 Peter Scherber, 1977: *Slovar Prešernovega pesniškega jezika*. Maribor: Obzorja.

12 Skeniral sem jih v prostorih Gesellschaft für wissenschaftliche Datenverarbeitung v Göttingenu (GWDG), kjer je bil tedaj zaposlen moj neformalni mentor v času Humboldtove štipendije Peter Scherber. Digitalizacija angleške književnosti je 20 let starejša: Michael Hart je zastavil z digitalizacijo literarne dediščine pod naslovom Gutenberg Project že leta 1971.

13 Tudi ta zbirka je locirana v tujini, na Švedskem, kjer je živel in delal Tržačan Luin.

14 TraDok - Translation-Dokumentation, Die Datenbank für "Literarisches Übersetzen" aus dem Projekt "Deutsch-slowenische/kroatische Übersetzung 1848-1918".

15 Leta 2009 je NUK za nakup elektronskih publikacij namenil že 45 % vseh sredstev.

smo dLibu hvaležni za *Kmetijske in rokodelske novice*, za nemško časopisje na Slovenskem v 19. stoletju in za časopise *Ljubljanski zvon* (1881–1941), *Dom in svet* (1888–1944), *Sodobnost*, *Novi svet*, *Ženski svet*, *Zora* itn.

Digitalizirana besedila so v formatih pdf in html in so različne kvalitete. Preference so bile primerno določene: digitalizacija prvih natisov namesto poznejših redakcij, redkih ali slabo ohranjenih tiskov in tistega gradiva, po katerem obstaja povpraševanje. 51 knjig Ivana Cankarja imamo tako zdaj pred sabo v neredigirani, avtentični obliki. Malo nerodno je, da so bile zaradi principa povpraševanja med starimi knjigami najprej skenirane prevedene, ne pa izvirne. Teksti niso šli skozi ročno (človeško) korekturo in zato pri začetnicah, akcentih, delitvah in pri slabši podlagi tudi na drugih mestih vsebujejo napake. Skenirani so bili skupaj s paginacijo, ilustracijami in slučajnimi bralčevimi rokopisnimi posegi, pri revijalnih objavah so tekstu priključeni tudi koski drugih besedil na isti strani, pri knjižnih izdajah pa kolofon, reklame ipd. Vse to je pomembno za poznavanje konteksta objave, vendar bi ga bilo treba v html-obliki izločiti, pri pdf-obliki pa dodati listanje naprej in nazaj in tako omogočiti ogled širšega konteksta objave. Ker so napake tudi v meta-podatkih (imenih avtorjev, naslovih, izpuščene so strani in poglavja), nekaterih posameznih objav ne moremo najti. Najdene napake sproti sporočamo in smo z odzivom zadovoljni. Največja hiba tega sicer zelo hvalevrednega podjetja je izoliranost podatkov na spletu. Napovedano vključitev dLibovih enot v Cobiss pozdravljamo, nujno pa bo treba tudi poskrbeti, da se spletišče odpre indeksiranju spletnih iskalnikov in iskanje tako postane hitrejše in udobnejše.

Brez dLiba je trenutna zaloga elektronsko dostopnih knjig v Sloveniji, kakor jo prešteje Cobiss, 43.000 enot, ko omejimo iskanje na

slovenščino, pa 23.000. Številka je v največji meri posledica vpisov v zadnjih letih, pri čemer je opazno veliko diplomskih del – univerze so očitno pobudnik nove oblike knjige.¹⁶ Delež elektronskih izdaj je med 3,5 milijona vpisov v Cobissu v najboljšem primeru komaj dober odstotek. Ko se pridruži še 300.000 dLibovih enot, se bo povzpел na 10 %.

S podporo Ministrstva za kulturo na Filozofski fakulteti Univerze v Ljubljani že drugo leto v okviru projekta Slovenska leposlovna klasika digitaliziramo leposlovje v skupnem obsegu osem milijonov besed. Študenti svoje digitalizacije ali popravilne strojno prebranih del uveljavljajo kot del študijskih obveznosti, pozneje za višjo oceno pri izpitih in najbolj zavzeti, med katerimi je marsikdo z drugih univerz, tudi za honorar. Avtorji z obsežnejšim ali celo skoraj popolnim opusom so France Prešeren (vključno z nemško poezijo), Stanko Vraz, Janez Trdina, Fran Erjavec, Simon Jenko, Josip Stritar, Janez Mencinger, Jakob Alešovec, Josip Jurčič, Simon Gregorčič, Josip Cimperman, Fran Detela, Anton Koder, Jakob Sket, Ivan Tavčar, Janko Kersnik, Fran Maselj Podlimbarski, Pavlina Pajk, Fran Zbašnik, Anton Aškerc, Miroslav Malovrh, Peter Bohinjec, Fran Milčinski, Fran Jaklič, Rado Mur-nik, Ksaver Meško, Ivan Cankar, Vladimir Levstik, Milan Lipovec, Sonja Koranter.

Koliko dela je že opravljenega in koliko ga še čaka? Globalne količine so takele: na svetu je med 65 in 75 milijoni knjig, od česar jih je okrog 25 milijonov v javni lasti – kar pomeni, da so njihovi avtorji pokojni že več kot 70 let – in torej brez vprašanja na voljo digitalizaciji. Približno 10 milijonov knjig je že prosto dostopnih na internetu: dobrih osem milijonov jih ponuja Google (Google Book Search), 1,3 milijona jih dobimo na portalu Internet Archive, 0,5 milijona v World Public Library in 0,1 milijona na straneh projekta Gutenberg.¹⁷

¹⁶ Strožje iskanje samo v omrežju dostopnih monografij vrne precej manj zadetkov (4300 vseh oz. 2200 slovenskih).

¹⁷ Michael Hart, Re: [Humanist] 22.639 looking back, Humanist Discussion Group 22/641 (2009).

Slovenske količine so ustrezno manjše. Zelo preprost ukaz v Cobissu pokaže okrog 600.000 vseh monografskih publikacij v slovenščini.¹⁸ Leposlovja je skupaj s ponatisi 24.500 knjig.¹⁹ O skupnem deležu leposlovja je iz te številke težko sklepati, ker se je leposlovje v 19. stoletju in prvi polovici 20. stoletja dogajalo v veliki meri na straneh literarnih revij in v podlistkih dnevnega časopisja. Poznani so podatki o količini dveh obsežnih pripovednih žanrov, kmečke povesti (7,7 mio. besed do 1945) in zgodovinskega romana (26 mio. besed do danes), poznani sta tudi količini digitaliziranega leposlovja v okviru obeh slovenskih korpusov: v Novi besedi 12 mio besed (5 % od skupaj 240 mio. besed), v Fidiplus 21,5 mio. besed (3,5 % od skupaj 621 mio. besed). Po zelo grobi oceni bi bilo vsega izvirnega slovenskega leposlovja 300–600 milijonov besed (brez ponatisov pa ustrezno manj). Ker se digitalizirane enote iz različnih virov deloma prekrivajo, jih ni mogoče kar preprosto sešteti in lahko le zelo približno ocenimo, da imamo v digitalni obliki za 30–60 mio. besed leposlovja, kar je nekje med 5 in 20 %. Dveh tretjin celotne mase leposlovja žal zaradi avtorske zakonodaje, ki preprečuje prosto uporabo del tistih avtorjev, ki so mrtvi manj kot 70 let (zadnjih 70 let pa je, če sklepamo po zgodovinskem romanu, prispevalo dve tretjini leposlovne zaloge), ne moremo digitalizirati. Domnevamo, da smo nekje na polovici tega, kar je mogoče postaviti na splet brez spraševanja avtorjev ali avtorske agencije oz. brez dodatnih stroškov, kar je podobno globalnim razmeram.

Pregled nad kompletnim slovenskim digitaliziranim leposlovjem skuša vzdrževati seznam [Slovenska literatura na internetu](#), ki ga s študenti ročno dopolnjujemo. Popisuje vsa

obsežnejša digitalizirana leposlovna dela (knjige, zbirke kratke proze, pesniške zbirke, daljšo in srednjo dolgo prozo v revijah, pa tudi starejša krajša dela) in ponuja povezavo nanja, če so le javno dostopna. Trenutno vsebuje okrog 1800 zapisov, takih, ki pripeljejo na dokument, pa je 1142: 406-krat k Luinu, 400-krat na Wikivir, 200-krat na dLib in 70-krat na Zbirko slovenskih leposlovnih besedil.

Sodobno leposlovje tistih avtorjev, ki ne računajo več na to, da bi živeli od avtorskega honorarja, skladišči online literarna revija [Locutio](#), zelo popularen je tudi pesniški portal [Pesnik](#), česar vse se literarna veda še ni naučila izkoriščati.

Poleg podatkov in dokumentov spadajo v kibernetško infrastrukturo še orodja za obvladovanje podatkov in dokumentov. Najpogosteje so vgrajena v ustrezna spletišča kot iskalniki in konkordančniki. Orodje, ki ga slovenistična literarna veda skupaj s celo neangleško humanistiko močno pogreša, je specializirani citatni sistem. Nadomešča ga [Googlov Učenjak](#), v sili pa kar običajni Googlov iskalnik, s katerim najdemo omembe ali ocene določenega dela v spletnih objavah. Če gre za spletno objavo, v Googlu z ukazom *link* zlahka preverimo, kje vse jo na spletu citirajo.²⁰ Sicer je literarna veda glede računalniških orodij nezahtevna; v sili si jih sposodi od jezikoslovcev ali družboslovcev, svojih pa vsaj na Slovenskem ne razvija.²¹

Mislím, da je treba med infrastrukturne pripomočke šteti tudi orodja za strokovno komunikacijo. Tak je moderirani diskusijski forum [SlovLit](#), ki letos obhaja desetletnico obstoja in 1200 naročnikom postreže s povprečno dvema strokovnima obvestiloma dnevno.

Potrdilo za to, da moje delo spada med infrastrukturna početja, mi je priskrbel Google.

¹⁸ Brez diplom in separatov je enot nekaj manj, če odštejemo prevode, pa precej manj.

¹⁹ Številko dobimo z ukazom `uc=821.163.6*` and `py=1550:1989` and `la=slv`; vrsta gradiva: monografske publikacije, potem ko od zadetkov odštejemo literarnovedne knjige (`uc=821.163.6.09`).

²⁰ Za *Praktični spisovnik* npr. `link:http://lit.ijs.si/spisovn.html`.

²¹ Česa vse se poslužujejo za analizo besedil kolegi v tujini, si je mogoče ogledati na spletišču [TAPoR](#) (Text Analysis Portal for Research).

Med zadetki pri iskanju infrastrukturnih zadev, ki zadevajo slovensko književnost, se je znašel neki dokument pri evropskem projektu HERA,²² kjer sta med 32 infrastrukturnimi referencami navedeni dve podatkovni zbirki v mojem uredništvu: Zbirka diplomskih nalog iz slovenske književnosti in Zbirka slovenskih leposlovnih besedil. Kot odgovorni zanj je naveden Oddelek za etnologijo in kulturno antropologijo, kar je seveda narobe. Etnologi so verjetno zaslužni le za vpis obeh zbirk v kakšno od svojih raziskovalnih poročil, mi pa go-di, da se sklicujejo prav na naše slovenistično literarnovedno delo.

Digitalizacijo infrastrukturnih vsebin in menjavo komunikacijskih kanalov v stroki bi lahko primerjali s selitvijo. Selitve so naporna reč, toda vsaka selitev je priložnost za vzpostavitvevi boljšega reda in za blagodejno prevetritev. Vzemimo tako tudi digitalizacijo in pospletanje v literarni vedi. V novem mediju se nam razkrivajo nove razsežnosti predmetov, za katere smo mislili, da jih v celoti in dovolj poznamo, in to je čisto zadosten iztržek.

Literatura

- HLADNIK, Miran, 2006: Sodobna slovenska literatura in elektronski viri / Contemporary Slovene Literature and Electronic Sources. Zupan Sosič, Alojzija, Nidorfer Šiškovič, Mojca (ur.): *Almanah Svetovni dnevi slovenske literature / Almanac World Festival of Slovene Literature*: Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko. 20–23.
- HLADNIK, Miran, 2008: Slovene student projects in Wikipedia and Wikisource. Referat na 3. regionalni konferenci Wikimedije v Beogradu 20. decembra 2008.
- HLADNIK, Miran, 2009: Za elektronsko knjigo. *Šolska knjižnica* 19/2–3. 126–33.

22 Slovenia: Academic field and financial aspects of the existing research infrastructure.