

ZINĀTNISKIE RAKSTI

DATORTEHNOLOĢIJAS VĒSTURES AVOTU REPREZENTĀCIJĀ, EDĪCIJĀ UN AVOTPĒTNIECISKAJĀ KRITIKĀ

Aleksandrs Ivanovs

Dr. hist., Daugavpils Universitātes profesors, Rēzeknes Augstskolas Reģionālistikas zinātniskā institūta vadošais pētnieks. Zinātniskās intereses: Latvijas vēstures historiogrāfija, arheogrāfija, diplomātika, avotpētniecība.

E-pasts: aleksandrs.ivanovs@du.lv

Raksta mērķis ir parādīt jaunāko datortehnoloģiju potenciālās iespējas vēstures pētniecības tradicionālākajās jomās, vispirms arheogrāfijā, avotpētniecībā, diplomātikā un vēstures palīgzinātnēs. Raksts balstās uz vairākiem pilotprojektiem, kurus kopš 2000. gada īsteno raksta autors sadarbībā ar matemātikas doktoru Alekseju Varfolomejevu (Petrozavodskas Valsts universitāte, Krievija). Rakstā secināts, ka mūsdienās lielākās potenciālās iespējas ir avotu semantiskām publikācijām, kas var kļūt ne tikai par efektīvu informācijas reprezentācijas, meklēšanas un sasaistīšanas līdzekli, bet arī par analītisku instrumentu, kas izmantojams jaunu zināšanu ģenerēšanai un vēsturnieka hipotēžu verificēšanai. Lai gan rakstā aplūkotas datortehnoloģiju izmantošanas iespējas tieši viduslaiku vēstures avotu reprezentācijā, edīcijā un kritikā, taču tās pašas tehnoloģijas var tikt auglīgi izmantotas jauno un jaunāko laiku vēstures avotu, it sevišķi – masveida avotu izpētē.

Atslēgas vārdi: vēstures avoti, datortehnoloģijas, semantiskās publikācijas.

Datortehnoloģiju ieviešana humanitārajās zinātnēs, to skaitā – vēstures pētniecībā,¹ ir diezgan pretrunīgs process. No vienas puses, var apgalvot, ka uz datortehnoloģijām balstītā humanitārā pētniecība (*e-humanities*) ir jau kļuvusi par relatīvi patstāvīgu, ļoti attīstītu un populāru virzienu (nozari) humanitārajās zinātnēs. Par to liecina gan plaša un sazarota infrastruktūra – asociācijas, centri un biedrības, kas veicina datortehnoloģiju ieviešanu humanitārajās zinātnēs,² gan regulāri rīkotas starptautiskās zinātniskās konferences,³ gan arī pētījumu skaita pieaugums un uz humanitāro pētniecību orientēto interneta resursu strauja attīstība kopš 20. gs.

90. gadiem. No otras puses, nepārtraukti norisinās diskusijas par datortehnoloģiju un ar to saistīto metožu pielietošanas sfērām un ierobežojumiem, par šo tehnoloģiju savienošanas iespējām ar humanitāro zinātņu tradicionālām pētniecības metodēm⁴ un izmantošanas perspektīvām vēstures izpētē.⁵ Turklāt dažreiz tiek pat pausta vilšanās informācijas tehnoloģijās, konstatējot, ka 20. gs. 90. gados ieceres cieši sasaistīt vēstures pētniecību un informācijas zinātņi nav pilnīgi piepildījušās.⁶ Šajā sakarā var noteikti runāt par vēstures zinātnes zināmu konservatīvismu, kas neļauj vēsturniekiem sekot jaunākajām tendencēm vēstures metodoloģijā un nesekmē modernāko pieeju un metožu ieviešanu pētniecībā. Tas nenozīmē, ka vēsturniekiem ir viennozīmīgi noraidoša attieksme pret mūsdienu pētniecisko instrumentāriju un kvantitatīvajām metodēm. Bieži vien konservatīvisma izpausme ir utilitārisms: jaunākās pētnieciskās metodes un tehnoloģijas tiek lietotas, ja tās var nodrošināt principiāli nozīmīgus rezultātus pagātnes izziņāšanā. Turpretim vēsturnieku skatījumā, ja “novecojušās” metodes tā vai citādi dod iespēju risināt pētnieciskos uzdevumus, tās nav jānomaina ar modernākām tehnoloģijām un metodēm. Lai gan visumā tāds vēstures zinātnes “minimalisms” ir attaisnojams, tomēr tas vienlaikus apdraud vēstures pētniecības turpmāku attīstību: datortehnoloģiju straujās, lavīnveidīgās attīstības apstākļos var vienkārši nepamanīt kaut kādas jaunas, vēstures pētniecībai noderīgas un auglīgas pieejas un metodes, kas tiek izstrādātas un aprobētas informācijas zinātnē, vēstures informātikā un citās zinātnes nozarēs. Rezultātā uz laiku laikiem tiks palaistas garām potenciālās iespējas, kuras šīs metodes un tehnoloģijas paver.

Iepriekšminēto apsvērumu dēļ šajā rakstā tiek mēģināts parādīt jaunāko datortehnoloģiju potenciālās iespējas vēstures pētniecības pašās tradicionālākajās jomās, galvenokārt arheogrāfijā, avotpētniecībā, diplomātikā un vēstures palīgzinātnēs. Raksts balstās uz vairākiem pilotprojektiem, kurus kopš 2000. gada īsteno šā raksta autors sadarbībā ar speciālistu vēsturiskajā informātikā, matemātikas doktoru, Petrozavodskas Valsts universitātes (Krievija) Matemātikas fakultātes dekānu Alekseju Varfolomejevu. Projektos, kā arī zinātniskajās publikācijās un prezentācijās starptautiskajās konferencēs tika aprobētas dažādas datortehnoloģijas, kas ļāva novērtēt to pielietošanas iespējas, perspektīvas un lietderīgumu viduslaiku vēstures avotu – lielākoties aktu⁷ un lietvedības (kancelejas)

dokumentu – edīcijā un avotpētnieciskajā kritikā, respektīvi, vēstures avotu informācijas reprezentācijā, apstrādē, izpētē un interpretācijā.⁸ Jāatzīmē, ka datortehnoloģiju iespēju aprobācijai ir metodoloģiska nozīme arī jauno un jaunāko laiku vēstures avotu pētniecībā. Ja šīs tehnoloģijas, kuras zināmā mērā strukturē, standartizē un pat unificē unikālo vēstures avotu informāciju, pielietojamas konservatīvākajās pētniecības jomās, tad jauno un jaunāko laiku masveida vēstures avotu izpētē šo metožu un pieeju potenciāls ir patiešām liels.

Tādējādi šis raksts piedāvā specifisku ieskatu mūsdienu avotpētniecībā, kura balstās uz datortehnoloģijām. Šajā sakarā jāatzīmē, ka “datorizētā” avotpētniecība⁹ kā relatīvi patstāvīga joma vēstures avotu mācības ietvaros joprojām ir tapšanas stadijā, jo speciālistu vidū nav panākta vienprātība pat šīs jaunās jomas priekšmeta noteikšanā. Taču ir jau skaidrs, ka “datorizētā” avotpētniecība nevar aprobežoties ar vēstures datu bāzu veidošanas tehnoloģijām un vēstures avotu digitālu publikāciju sagatavošanu (respektīvi, avotu digitālu reprezentāciju internetā¹⁰), jo visspilgtāk “datorizētās” avotpētniecības potenciāls atklājas, datortehnoloģijas izmantojot vēstures avotu analītiskajā un sintētiskajā kritikā,¹¹ kā arī avotu informācijas verifikācijā.¹² Turklāt šīm tehnoloģijām un metodēm jābūt universālām, lai tiktu nodrošināta iespēja tās lietot gan atsevišķu avotu, kuri attiecas uz dažādiem veidiem un paveidiem, dziļā analīzē, gan arī rakstīto avotu reprezentatīvu kopumu sistēmiskā izpētē. Būtībā “datorizētā” avotpētniecība var veikt visus vēstures avotu kritikas galvenos posmus: rakstītā vēstures avota teksta un tā struktūras noteikšanu; vairāku vēstures avotu liecību sistēmisku agregēšanu un sasaistīšanu (*linkage of sources*), informācijas mērķtiecīgu meklēšanu (*open-ended inquiries*), kas palīdz risināt vēstures avotu datēšanas un atribūcijas jautājumus, kā arī interpretēt un verificēt vēstures avotu sniegto informāciju; vēstures faktu noteikšanu, balstoties uz vēstures avotu liecībām, un vēsturiskās realitātes vispusīgu atveidi.¹³

“Datorizētajā” avotpētniecībā iepriekšminētās pētnieciskās operācijas tiek veiktas augstākā “tehnoloģiskā” līmenī nekā tradicionālajā avotu mācībā. Šķiet, pirmo reizi rodas iespēja veikt vēstures avotu kvantitatīvi neierobežotu kopumu pilnvērtīgu, vispusīgu, kompleksu izpēti, kas paaugstina gan avotpētnieciskās kritikas kvalitāti, gan vēstures rekonstrukcijas drošumu. Turklāt jebkuras pētnieciskās operācijas var būt elastīgi modificētas atkarībā no konkrēta pētījuma

uzdevumiem; šajā gadījumā vienīgais ierobežojums ir vēsturnieka spēja izvirzīt oriģinālus pētnieciskus uzdevumus. Perspektīvā mūsdienu “datorizētā” avotpētniecība var radīt vēstures avotu kritikai vairākas iespējas pilnīgi atklāt vēstures avotu informatīvo potenciālu. Pirmā iespēja ir saistīta ar vēstures avotu analītiskajām digitālajām publikācijām, kurās avotu teksti tiek papildināti ar to analīzes instrumentiem.¹⁴ Otrā iespēju nodrošina dalītās piekļuves tīklā iesaistīto pētnieku sadarbība apjomīgu vēstures avotu kompleksu apstrādē tiešsaistes režīmā.¹⁵ Vēstures avotu tekstu apstrādes un izpētes jomā speciālistu – vēsturnieku un filologu – kooperācijas (kopdarba) veiksmīgu priekšzīmi rāda Vācijā izstrādātā sistēma TextGrid.¹⁶ Trešā iespēja ir saistīta ar vēstures avotu izpētes rezultātu atkārtojamību, jo gan pētījumu metodikas, gan arī rezultāti tiek saglabāti koplietojamajās datubāzēs.

Pašlaik “datorizētajā” avotpētniecībā pati attīstītākā joma ir vēstures avotu un vēstures faktogrāfijas reprezentācija datubāzēs. Jau 20. gs. 90. gados datubāzu veidošanā ir nostiprinājušās divas pamatpieejas.¹⁷ Pirmo pieeju var nosacīti saukt par problēmorientēto pieeju.¹⁸ Problēmorientēto datubāzu mērķis ir reprezentēt internetā maksimāli pilnu vēstures faktu un datu sakopojumu par kādu tēmu vai problēmu. Otrā pieeja ir orientēta tieši uz vēstures avotu, respektīvi, to tekstu, liecību un digitālu attēlu reprezentāciju datubāzēs (*source-oriented approach*); parasti šī pieeja neparedz avotu informācijas iepriekšējo interpretāciju, novērtēšanu un analīzi. No praktiskā skatpunkta problēmorientēto relāciju datubāzu veidošana nav sevišķi sarežģīta un darbietilpīga, bet to izmantošanas rezultāti ir labi pārskatāmi: lietotājam tiek piedāvāta plaša un detalizēta informācija par konkrētiem vēstures jautājumiem un aspektiem; šo informāciju var izmantot gan speciālisti, gan arī amatieri, studenti, skolēni u.c.

Tomēr uzkrātā pieredze ļauj apgalvot, ka problēmorientētajām datubāzēm ir būtiski trūkumi. Šajā sakarā var minēt komplekso datubāzi “Latgales Dati”,¹⁹ kura reprezentē daudzveidīgo informāciju par Latgales novada vēsturi.²⁰ Datubāzes veidošana tika iesākta vēl 1994. gadā;²¹ tās pilnveide un papildināšana turpinās arī mūsdienās. Diemžēl šī elektroniskā resursa izmantošanas prakse liecina, ka tāda datubāze nav sevišķi piemērota nopietnu, oriģinālu pētījumu realizācijai. Datubāze piedāvā jau mērķtiecīgi atlasīto, izvērtēto un interpretēto vēstures faktu sakopojumu (bieži vien – atrauti

no vēstures avotiem, kuri sniedz liecības par šiem faktiem); šīs faktogrāfijas atkārtota analīze un interpretācija praktiski nav iespējama. Tāpēc datubāze “Latgales Dati” var būt noderīga tikai nedaudziem vēsturniekiem, kuriem ir ļoti konkrēti pētnieciski uzdevumi, kā arī sabiedrības pārstāvjiem, pasniedzējiem un studentiem, kuriem ir vajadzīga retrospektīvā un operatīvā informācija par Latgales novadu. Tas vēlreiz apstiprina vēstures informātikā izplatīto viedokli, ka tikai uz vēstures avotiem orientētās datubāzes var kļūt par vēstures izpētes efektīvu instrumentu un par tradicionālās avotu mācības turpmākās attīstības “virzītājspēku”.²²

Tādējādi datortehnoloģiju un “datorizētās” avotpētniecības metožu lietošanas priekšnoteikums ir vēstures avotu reprezentācija tieši uz avotiem orientētajās datubāzēs; tādu avotu reprezentāciju internetā var uzskatīt par vēstures avotu publikāciju jaunajā tehnoloģiskajā līmenī. Jāatzīmē, ka pēdējos gados datortehnoloģijas arvien biežāk tiek izmantotas, lai zinātnes aprītē ievestu ne tikai jaunajos un jaunākajos laikos tapušos dokumentus, bet arī viduslaiku vēstures avotus. Par to liecina vairāki avotu edīcijas un vienlaikus izpētes projekti un vēstures mantojuma publikācijas internetā.²³ Lai gan daudzām vēstures avotu publikācijām internetā nav pievienots adekvāts pētniecisks un datu meklēšanas un sasaistīšanas instrumentārijs, tomēr, bez šaubām, tās popularizē vēstures avotu kopumus, nodrošinot attālināto piekļuvi dokumentiem un to informācijai. Šajā sakarā var minēt vairākas datubāzes, kas tiek veidotas Latvijā: Latvijas Nacionālā arhīva (turpmāk – LNA) Latvijas Valsts vēstures arhīva (turpmāk – LVVA) projekts un datubāze “Raduraksti”,²⁴ datubāze “Latvijas kultūras vēsture attēlos. Sērija A: Portreti”,²⁵ LNA Latvijas Valsts arhīva datubāze “Deportēto Latvijas iedzīvotāju saraksti. 1941., 1949., 1945.–1953.”.²⁶

Bieži vien vēstures avotu elektroniskās publikācijas imitē tradicionālās (“papīra”) publikācijas. Nevar noliegt, ka šajā gadījumā digitālajām publikācijām ir dažas dabiskas priekšrocības: dokumentu attēlu kvalitāte kļūst augstāka, publicēto avotu skaits var būt ļoti liels (teorētiski – pat neierobežots) u.c. Taču orientēties iespēstajā avotu krājumā ir daudz vieglāk nekā dokumentu elektroniskajā publikācijā. Tas nozīmē, ka pārliecinoši demonstrēt savas priekšrocības var tikai digitālās publikācijas, kas risina ne vien arheogrāfiskus

(avotu ieviešana zinātnes aprītē), bet arī analītiskus uzdevumus, piedāvājot lietotājiem avotu kritikas modernus līdzekļus. Digitālo publikāciju vidū sevišķi izceļas semantiskās publikācijas, kurām ir lielas (kaut gan pašlaik – tikai potenciālas) iespējas rakstīto vēstures avotu un to informācijas reprezentācijā, sasaistīšanā, meklēšanā un izpētē.²⁷

Pēdējos gados termins “semantiskās publikācijas” kļuvis ļoti populārs, pateicoties semantiskā tīmekļa (*Semantic Web*)²⁸ tehnoloģiju straujai attīstībai. Parasti ar “semantisko publikāciju” saprot teksta digitālo publikāciju internetā, šim tekstam pievienojot papildu informācijas slāņus (*information layers*), kuri formalizētā veidā atklāj teksta jēgu un saturu, kā arī reprezentē esošās zināšanas par tekstu, tādējādi nodrošinot šīs informācijas automātisko apstrādi. Mūsdienās semantiskās publikācijas bieži parādās elektroniskajos zinātniskajos žurnālos un digitālajās bibliotēkās.²⁹ Turpretim vēstures avotu semantiskās publikācijas joprojām ir aprobācijas stadijā, tāpēc arī literatūrā tām veltīta nepietiekama uzmanība.³⁰

Semantisko publikāciju koncepcija diezgan sīki pamatota vairākos pētījumos. Piemēram, Deivids Šotons uzmanību veltījis tā saucamajiem semantiskās pastiprināšanas līdzekļiem (*semantic enhancements*): tabulām, kas satur sākotnējos pētījuma datus; interaktīvajām diagrammām un grafikiem (grafiem), kuri veidoti, balstoties uz šiem datiem; atsaucēm uz terminoloģiskām vārdnīcām u.tml.³¹ Acīmredzot šajā gadījumā koncepts “semantisks” izmantots, lai apzīmētu noteiktu līdzekļu kopumu, kuri palīdz tieši *cilvēkam* (nevis datoram) uztvert raksta jēgu. Detalizētāk semantisko publikāciju ideju raksturo Anita de Vārda. Viņa izdala trīs semantiskos slāņus, kuri pievienojami zinātniskā raksta tradicionālajai elektroniskajai publikācijai.³² Pirmais slānis ietver tekstos iezīmētos terminus un konceptus, kuri sasaistīti ar personu, ģeogrāfisko nosaukumu un terminu vārdnīcām, tēzauriem un ontoloģijām. Otrajā slānī ir dažādi ar publicēto tekstu saistīti fakti, citiem vārdiem sakot, metainformācija par tekstu, kura reprezentēta triplete formā (subjekts–predikāts–objekts), jo semantiskajā tīmeklī tādi tripleti ir universāls datu reprezentācijas veids. Trešais slānis satur informāciju par tā saucamo zinātnisko diskursu: par pētāmo jautājumu nostādni, izvirzītajām hipotēzēm, atzinumu pamatošanu, diskusijām, eksperimentu gaitu utt. Šie slāņi dod iespēju izdalīt tekstā nozīmju bloku hierarhiju, kā arī

parādīt attiecības starp šiem semantiskajiem blokiem. Sevišķi svarīgi, ka šos semantiskos slāņus uztver ne tikai cilvēks, bet arī dators.

Var apgalvot, ka semantisko publikāciju ideja pilnīgi atbilst vēstures avotu digitālās reprezentācijas (respektīvi, publikācijas) uzdevumiem. Jebkurā avotu publikācijā piedāvā savstarpēji saistītu dokumentu sistēmu,³³ kuras ietvaros starp tiem ir daudzveidīgas saites un sakari. Dažus sakarus veido pats vēsturnieks (arheogrāfs) publikācijas sagatavošanas gaitā, piemēram, sasaistot dokumentu tekstus un to fragmentus – lappuses, rindas, klauzulas, teikumus, vārdkopas, leksēmas, terminus, personvārdus, vietvārdus u.c. ar attiecīgiem zinātniskiem komentāriem, bibliogrāfiskām norādēm, attēliem, vienībām rādītājos u.tml. Semantiskajā publikācijā visus šos sakarus starp avotu tekstiem un metainformāciju³⁴ ir iespējams precīzi reprezentēt un pat vizualizēt, atvieglinot digitālās publikācijas izmantošanu vēstures avotu pētniecības nolūkos.

Taču jebkurā vēstures avotu publikācijā (un it sevišķi – semantiskajā publikācijā) lielāka nozīme ir sakariem starp vēstures avotu, kuri ietilpst noteiktā “dabiskā” avotu kompleksā,³⁵ tekstiem (precīzāk – starp dažādu avotu sniegtām liecībām). Var izcelt pašas svarīgākās iezīmes, kas piemīt dabiskajam avotu kompleksam. Pirmkārt, tā ir kompleksu veidojošo vēstures avotu kopīga izcelsme: bieži vien tādi avoti tapuši vai nu kādas vienas institūcijas (iestādes, organizācijas, biedrības utt.), vai savstarpēji saistītu institūciju, vai arī atsevišķas personas darbības gaitā. Otrkārt, vēstures avoti, kuri ietilpst vienā kompleksā, ir savstarpēji saistīti. Saiknes starp avotiem veidojas spontāni; tās var būt gan tiešas (piemēram, lietvedības dokumentu kompleksā viens dokuments atsaucas uz citu un otrādi), gan arī netiešas (piemēram, avoti, kas tapuši dažādu cilvēku un institūciju darbības gaitā, sniedz informāciju par vienu un to pašu vēstures faktu). Visas tādas saiknes izveidojušās un nostiprinājušās pagātnē, vēstures avotu tapšanas un “aktīvās” funkcionēšanas posmā, bet vēstures avotu kompleksos, kas ir nonākuši līdz mums, konstatējamas tikai līdzšinējo saišu “pēdas”. Treškārt, viena kompleksa avoti parasti veido hierarhisku, visai noturīgu sistēmu, kurā katram avotam ir sava “niša”. Bieži vien šajās sistēmās ir daudz “robu”, tukšu vietu, jo dažādu iemeslu dēļ vairāki vēstures avoti gājuši bojā. Jāievēro, ka kompleksa ietvaros avotu hierarhiskā

sakārtotība atspoguļo nevis vēsturnieku subjektīvos priekšstatus par dažādu vēstures avotu nozīmi konkrētu vēstures jautājumu izpētei, bet avotu relatīvo svarīgumu savā laikmetā, kad tie veica savas sākotnējās funkcijas. Tādējādi vēstures avotu komplekss ir dabiska vēsturiska parādība; tā ir saistīta ar noteiktu laikmetu, ar sociālo, ekonomisko un kultūras vidi, kurā šis komplekss izveidojies. Tāpēc vēsturnieka un avotpētņieka uzdevums ir rekonstruēt vēstures avotu kompleksu un konstatēt reālas saiknes starp avotiem. Turpretim avotu kompleksa mākslīga izveide, patvaļīgu saikņu uzspiešana avotiem, kas vēsturiski nav savstarpēji saistīti, avotu sakārtošana pēc “nozīmīguma” utt. galu galā sagroza un modernizē priekšstatus par pagātņi. Tādējādi semantiskās publikācijas sagatavošanas priekšnoteikums ir vēstures avotu dabiskā kompleksa rekonstrukcija (noteikšana).

Vēstures avotu semantisko publikāciju koncepcijas izstrādāšanai un aprobācijai tika speciāli veidots viduslaiku aktu avotu semantiskās publikācijas vairākfunkciju prototips.³⁶ Tajā tika iekļauti pieci savstarpēji saistīti dokumenti senkrievu valodā, kuri veido subkompleksu viduslaiku un agro jauno laiku dokumentu kompleksa “Moscovitica–Ruthenica”³⁷ ietvaros. Prototipā izmantotie dokumenti atspoguļo attiecības starp Rīgu un Smoļensku 13. gadsimtā.³⁸

Lai pilnīgi atklātu vēstures avotu semantiskās publikācijas priekšrocības avotiskās informācijas, kā arī metainformācijas reprezentācijā, sasaistīšanā un meklēšanā, prioritārā uzmanība jāveltī semantiskajam slānim, kurā reprezentēta informācija par kādā konkrētā avota tekstā tieši vai netieši minētajiem tā saucamajiem vēstures objektiem³⁹ – personām, vietām, organizācijām un iestādēm, vēstures notikumiem un atgadījumiem, citiem dokumentiem (atsauces uz citiem dokumentiem), terminiem u.tml. Digitālajā vidē tādu objektu izdalīšana tekstā tiek veikta, izmantojot tekstu iezīmēšanas (*markup*) shēmas un standartus.⁴⁰ Nav šaubu, ka vēstures avotu edīcijas un izpētes nolūkiem visvairāk atbilst tieši tehnoloģija XML (*Extensible Markup Language*).⁴¹ Tā dod iespēju jebkurā tekstā iezīmēt semantiskās vienības, plašāku vienību ietvaros izdalot sīkākas vienības utt., tādējādi stadiāli un secīgi detalizējot vēstures avota struktūru no tā tekstu kopuma līdz pat (nepieciešamības gadījumā) atsevišķām leksēmām vai simboliem. Semantiskajām vienībām var pievienot atribūtus, kuri sasaista šīs vienības ar noteiktām

objektu klasēm vai arī piešķir objektiem standartizētas nozīmes (piemēram, līdzšinējam vietvārdam var pievienot pašlaik lietotu vietvārdu). Vienību iezīmēšanai izmanto speciālus marķierus – tā saucamos tagus (*tag*): pāra tagi – “atverošais” tags un “aizverošais” tags izdala katru atsevišķu semantisku vienību.⁴² Pateicoties tagiem, kuri ir lielākoties standartizēti un unificēti, vairāki avoti tiek sasaistīti savā starpā,⁴³ bet atsevišķu avotu tekstiem un to fragmentiem tiek pievienoti tulkojumi citās valodās, komentāri, avotu oriģinālu digitāli attēli utt. Būtībā vēstures avots, kas iezīmēts, izmantojot tehnoloģiju XML, veido vienu failu.⁴⁴ Tas nozīmē, ka tehnoloģija XML ļauj saglabāt katra vēstures avota integritāti, nesadalot tā tekstu, bet integrējot vienā failā vairāku avotu tekstus un attiecīgu metainformāciju (atšķirībā no avotu reprezentācijas relāciju datubāzēs, kad avotu tekstu fragmenti tiek ielikti relāciju tabulu šūnās; rezultātā avoti tiek faktiski “dezintegrēti”). Tāpēc var apgalvot, ka šī tehnoloģija ir uz vēstures avotiem orientēto datubāzu tehnoloģija.⁴⁵

Balstoties uz tehnoloģiju XML, ir jau izstrādātas vairākas iezīmēšanas shēmas. Vēstures avotu semantisko publikāciju sagatavošanā, kā arī vēstures avotu kritikā būtu lietderīgi izmantot labi izstrādātās divas iezīmēšanas shēmas TEI⁴⁶ un CEI,⁴⁷ jo tajās ir gandrīz visi vēstures avotu tekstu iezīmēšanai vajadzīgie elementi (tagi). Tajā pašā laikā semantiskās publikācijas prototipā tika izmantoti arī papildu elementi, kuru iepriekšminētajos standartos nav. Turpmāk tiek piedāvāts vienas akts – Rīgas arhibīskapa Jāņa II raksta Smoļenskas kņazam Fjodoram Rostislavičam, kurā Vitebskas pilsētas iedzīvotāji apsūdzēti nepamatotas sūdzības iesniegšanā pret rīdziniekiem,⁴⁸ – iezīmēšanas piemērs, kas uzskatāmi parāda iezīmēšanas operācijas specifiku un XML-faila “ārējo” izskatu:

```

<charter id="6">
<chDesc> ... </chDesc>
<tenor>
<protocol>
<invocatio> † </invocatio> <salutatio> Svētība </salutatio> no
<intitulatio> <issuer> Rīgas <roleName id="Archbishop"> metropolīta
</roleName> </issuer> </intitulatio> <lb n="1"/> <inscriptio> savam
mīļajam dēlam, <addressee> <persName id="Fedor Rostislavich Prince
of Smolensk"> <roleName id="Grand Prince"> lielkņazam </roleName>
<lb n="2"/> Fjodoram </persName> </addressee>, un viņa bērniem un
<supplied> Smoļenskas </supplied> <roleName id="Bishop"> valdniekam
</roleName> un <lb n="3"/> <recipient> <supplied> Smoļenskas
</supplied> <roleName id="governor of Smolensk"> vietvaldim
</roleName> </recipient>, un visiem bajāriem. </inscriptio>
</protocol>
<context>
Lai <lb n="4"/> būs tev zināms par to sūdzību, kurā <placeName
id="vitebsk"> vitebskas </placeName> iedzīvotāji <lb n="5"/>
sūdzējās par rīdziniekiem, ar ko viņi gribēja <lb n="6"/>
attaisnoties pret <persName id="Helmich merchant from Munster">
Gelmiku </persName>. Un viņu vārds ir šāds - <lb n="7"/> gribēja ar
to vārdu attaisnoties - un <supplied> viņi </supplied> runāja <lb
n="8"/> šādi <roleName id="Prince of Briansk"> <placeName
id="Briansk"> Brjanskas </placeName> kņaza </roleName> priekšā:
izbrauca <lb n="9"/> 50 vīru no <placeName id="Riga"> Rīgas
</placeName> un nogalināja cilvēku <lb n="10"/> un paņēma 10 <term
id="unit of weight"> berkovecu </term> vaska. <lb n="11"/> Un
pašlaik es, <roleName id="Archbishop"> metropolīts </roleName>, tā
teicu, ka tie <lb n="12"/> <placeName id="vitebsk"> vitebskas
</placeName> iedzīvotāji netaisnīgi sūdzējās par rīdziniekiem. <lb
n="13"/> Un tagad tas ir man zināms, ka rīdzinieki <lb n="14"/> tajā
<supplied> lietā </supplied> būtībā nav vainīgi. Un tagad es tam
brīnos, <lb n="15"/> ka tavš <roleName id="governor of Smolensk">
vietvaldis </roleName> klausās <lb n="16"/> katra cilvēka vārdus,
bet tā <refToDoc> tiesa ir starp <lb n="17"/> jums un mums, kad
strīds radīsies, <supplied> tad </supplied> tas <lb n="18"/> jāizšķir
</refToDoc>. Tagad es ļoti lūdzu jūs, lai varētu <lb n="19"/> palikt
pie tās <refToDoc> tiesas un pie krusta <lb n="20"/> skūpstīšanas
</refToDoc>. Ja jums kāds sūdzēsies <lb n="21"/> par rīdziniekiem -
vai nu <persName id="Helmich merchant from Munster"> Gelmiks
</persName>, vai arī kāds <lb n="22"/> cits -, tad jūs sūtiet pie
mums, bet mēs dosim tiesu <lb n="23"/> pēc Dieva tiesas. <lb
n="24"/>
</context>
</tenor>
</charter>

```

Šajā piemērā dokumenta iezīmēšanas shēma ir būtiski vienkāršota. Tajā ar pāra tagiem iezīmētas tikai atsevišķas dokumenta iekšējās struktūras (formulāra) daļas: piemēram, <protocol> ... </protocol>, <context> ... </context>, bet protokola ietvaros – tās sastāvdaļas: <invocatio> ... </invocatio>, <salutatio> ... </salutatio>, <intitulatio> ... </intitulatio> u.c. Dokumenta pamatteksta (tā saucamā – konteksta) struktūras sastāvdaļas nav iezīmētas, lai nepārlīvētu iezīmēšanas piemēru ar tagiem. Turklāt šajā piemērā iezīmēti atsevišķi vēstures objekti, tiem pievienojot atribūtus,⁴⁹ kuri dod iespēju semantiskajā tīklā (publikācijā) tos identificēt: personu vārdi (<persName> ... </persName>), vietvārdi (<placeName> ... </placeName>), personu amati vai tituli (<roleName> ... </roleName>). Atsevišķi elementi parāda teksta izkārtojumu avota oriģinālā – tie ir tagi, kas apzīmē rindu beigas (piemēram, <lb n="18"/>, respektīvi, "18. rindas beigas"). Atšķirībā no citiem tagiem tie nav pāra tagi. Ar atsevišķiem tagiem dokumenta tekstam tiek pievienota metainformācija par publicējamo avotu (<chDesc> ... </chDesc>); taču šajā piemērā metainformācija kā tāda tika izlaista, uz tās vietu publikācijā norāda šie tagi. Iezīmēšanas piemērā izmantoti arī papildu elementi (tagi), kuru nav CEI vai TEI standartos:⁵⁰ piemēram, pāra tagi <refToDoc> ... </refToDoc> tiek izmantoti, lai avota tekstā apzīmētu esošās norādes uz citiem dokumentiem, kuri iekļauti semantiskās publikācijas prototipā. Protams, iezīmēšanai jābūt daudz detalizētākai, lai avotu tekstos tiktu izdalītas arī vēstures reālijas, vēstures faktu atsevišķi aspekti, par kuriem sniedz liecības avotu teksts. Taču šim nolūkam ir nepieciešams izveidot vairākus jaunus elementus (tagus), kuri tiks speciāli orientēti uz vēstures notikumu un atgadījumu unificētu reprezentāciju digitālajā vidē.

Vairāku vēstures avotu informācijas, kā arī metainformācijas par šiem avotiem sasaistīšanai tiek veidots viens XML-fails, kurā tiek iekļauti vairāki pēc iepriekšminētā parauga iezīmētie avoti. Vispārīgos vilcienos šis fails izskatīsies šādi:

```

<docGroup>
  <charter id="1">
    <chDesc> ... </chDesc>
    <tenor> ... </tenor>
  </charter>
  <charter id="3a">
    <chDesc> ... </chDesc>
    <tenor> ... </tenor>
  </charter>
  ... ..
  <charter id="6">
    <chDesc> ... </chDesc>
    <tenor> ... </tenor>
  </charter>
</docGroup>

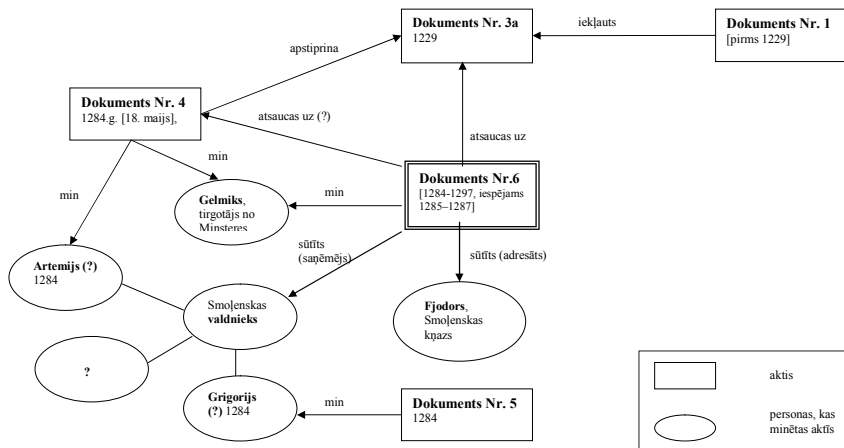
```

Tāda faila ietvaros kļūst iespējama iezīmēto informatīvo aspektu automātiskā sasaistīšana, meklēšana, klasificēšana un salīdzināšana (piemēram, dažādu dokumentu formulāru sastāvdaļu salīdzināšana savā starpā, kura ir tā saucamās formulāra analīzes pamatā⁵¹), kā arī šo pētniecisko operāciju rezultātu vizualizācija.

Šajā sakarā jāatzīmē, ka semantisko publikāciju pamats ir dažādas ontoloģijas, kuras variējas plašās robežās – no vienkāršām terminoloģiskām vārdnīcām un klašu sistēmām objektu aprakstīšanai līdz sarežģītām struktūrām, kas klašu, objektu, attiecību, loģisku kārtulu utt. formā reprezentē daudzveidīgās zināšanas par kaut kādu priekšmetisku jomu. Veidojot semantisku publikāciju, vēstures avotā iezīmētie objekti jāsaista ne tikai ar analogiskiem objektiem, kuri iezīmēti citos avotos, bet arī ar attiecīgiem datiem, kuri reprezentēti specializētās (vēstures) ontoloģijās. Te var minēt, piemēram, ontoloģiju CIDOC (*le Comité International pour la Documentation*) CRM (*Conceptual Reference Model*).⁵² Tā tika speciāli izstrādāta, lai vispusīgi reprezentētu informāciju par muzeja priekšmetiem. CIDOC CRM pamatā tika veidotas vairākas ontoloģijas, kas izmantojamas vēstures aspektu aprakstīšanai,⁵³ to skaitā *Pearl Harbor Project*⁵⁴ (ASV) un *CultureSampo Project*⁵⁵

(Somija). Diemžēl visas šīs ontoloģijas tikai daļēji atbilst vēstures avotu publikāciju sagatavošanas uzdevumiem, jo tajās rakstīto vēstures avotu specifika netiek pietiekami ievērota.

Tieši tāpēc ir nepieciešams veidot speciālu ontoloģiju aktu materiāla semantiskās publicēšanas nolūkiem. Šīs ontoloģijas pamatā ir uz dokumentiem orientētā pieeja (*document-oriented approach*) vēstures zināšanu reprezentācijā, kas atšķir šo ontoloģiju no esošajām ontoloģijām.⁵⁶ Ontoloģija atspoguļo divu veidu semantiskās saites, kas tiek konstatētas semantiskajā publikācijā. Pirmkārt, tās ir saites starp tematiski un/vai vēsturiski savstarpēji saistītajām aktīm, kuras izveidojās šo dokumentu “aktīvās” funkcionēšanas posmā 13. gs., attiecību starp Smoļensku un Rīgu dokumentēšanas gaitā. Otrkārt, tās ir saites ar citiem vēstures avotiem, kuri netiek iekļauti semantiskajā publikācijā reprezentētajā dokumentu subkompleksā. Tas nozīmē, ka publicējamo dokumentu informācija tiek sasaistīta ar “ārējo” informāciju, kuru sniedz vai nu citi vēstures avoti,⁵⁷ vai arī pētījumi, citas ontoloģijas u.tml. Šīs saites veido noteiktu semantisku tīklu, kas reprezentē zināšanas par publicējamajām aktīm. Būvējot šīs semantiskais tīkls ir datu tveršanas (*data capture*) modelis, kas var veidot pamatu avotu semantiskajai publikācijai. 1. attēls parāda šā semantiskā tīkla vienkāršotu segmentu, kura centrā ir iepriekšminētais Rīgas arhibīskapa Jāņa II raksts Smoļenskas kņazam Fjedoram (dokuments Nr. 6). Tematiski un daļēji arī tekstuāli tas ir saistīts ar četriem citiem dokumentiem.⁵⁸ Saites starp dokumentiem tika konstatētas, balstoties uz semantiskās publikācijas sniegtās “iekšējās” informācijas interpretāciju. Turpretim “ārējā” informācija tika izmantota aktis minēto personu identificēšanai, kā arī citu vēstures objektu pilnvērtīgai aprakstīšanai. Speciāli būtu jāatzīmē, ka semantiskā tīkla ietvaros saites starp objektiem – personām, dokumentiem u.c. – aprakstītas, izmantojot tripletus:⁵⁹ “*charter written by a person*” (dokumentu uzrakstījusi persona), “*charter sent to a person*” (dokuments nosūtīts personai), “*charter mentions a person*” (dokuments min personu) u.tml. Hipotētiska (nedroša) informācija par saitēm tiek apzīmēta, izmantojot jautājuma zīmi.



1. att. Dokumentu subkompleksa reprezentācija semantiskā tīkla formā (fragments)

Veidojot ontoloģiju šī semantiskā tīkla aprakstīšanai, var izmantot plaši izplatītās tehnoloģijas – RDF (*Resource Description Framework*) un OWL (*Ontological Web Language*).⁶⁰ Taču šajā gadījumā semantisko publikāciju veidošana būs ļoti darbietilpīga, tāpēc ir vērts meklēt citus instrumentus, kas atvieglina semantiskā tīmekļa tehnoloģiju izmantošanu. Mūsdienās tādus instrumentus var piedāvāt semantiskās *Wiki*-sistēmas.⁶¹ Tās dod iespēju, vairākiem speciālistiem sadarbojoties, vienkāršotas iezīmēšanas shēmas pamatā veidot un rediģēt tīmekļa lappuses, kā arī pievienot tekstiem papildu semantiskos slāņus. Viens no *Wiki*-sistēmu moduļiem ir *Semantic MediaWiki* (SMW),⁶² kas ļauj vēstures avotu tekstos izdalīt terminus, personu vārdus utt., kā arī sasaistīt avotu tekstus un metainformāciju.⁶³

SMW principiāli svarīga īpatnība ir tipizētās hipersaites starp tīmekļa lappusēm (*Wiki*-lappuses). Šīs lappuses reprezentē semantiskā tīmekļa objektus, turklāt katram objektam – katra vēstures avota tekstam, katrai avotos minētajai personai, vietai, institūcijai u.tml. – ir sava *Wiki*-lappuse. Pateicoties hipersaitēm, vēstures avota noteiktus fragmentus var sasaistīt ar citām tīmekļa lappusēm, kas satur citu avotu tekstus, kā arī informāciju par personām, vietām, notikumiem.⁶⁴ SMW arī dod iespēju semantiskā tīkla ietvaros vizualizēt tīmekļa hipersaites. Turklāt šīs hipersaites automātiski ģenerē faktu kopumus, kuri attiecas uz publicējamo vēstures avotu.

Tomēr *Wiki*-sistēmām ir zināmi trūkumi un ierobežojumi, it sevišķi vēstures avotu semantisko publikāciju sagatavošanā: avotu tekstus (tekstu fragmentus) nav iespējams sasaistīt ar vēstures avotu oriģinālu digitāliem attēliem; nestandarta šriftu izmantošanas iespējas ir visai ierobežotas;⁶⁵ avotu teksti nevar būt iezīmēti, balstoties uz TEI vai CEI iezīmēšanas shēmām; pašas *Wiki*-lappuses neļauj izsmeltoši reprezentēt zināšanas par vēstures avotiem. Neskatoties uz iepriekšminētajiem trūkumiem, SMW būtiski paātrina semantiskās publikācijas sagatavošanu.

Visas iepriekš raksturotās tehnoloģijas, kuras var izmantot semantisko publikāciju sagatavošanā, prasa publicējamo avotu un ar tiem saistītās metainformācijas rūpīgu un sīku iezīmēšanu, jo semantiskā tīkla ietvaros ir iespējams atrast, sasaistīt, apkopot, sistematizēt un galu galā analizēt tikai mērķtiecīgi iezīmētos informatīvos aspektus. Pašus avotu tekstus – to saturu un jēgu dators “nesaprot”. Tas nozīmē, ka semantiskās publikācijas potenciālās iespējās ir pilnībā atkarīgas no iezīmēšanas algoritma, kuru nosaka vēsturnieks (arheogrāfs). Tomēr jau pastāv reāla iespēja “izvairīties” no vēsturnieka iejaukšanās vēstures avotu tekstos, lai viņš ar noteiktu iezīmēšanas shēmu neuzspiestu savu subjektīvo vēstures avota informācijas un semantisko saišu vīziju citiem semantiskās publikācijas lietotājiem.

Vēstures avotu tekstu jēgu un saturu var samērā vispusīgi reprezentēt, izmantojot vienu no tā saucamajām kontrolējamajām dabiskajām valodām (*controlled natural languages*):⁶⁶ avotu tekstus, kuri ierakstīti šajās valodās, saprot ne tikai cilvēks, bet arī dators. Mūsdienās kontrolējamās dabiskās valodas zināšanu reprezentācijai tiek intensīvi attīstītas. Šīm valodām ir raksturīga vienkāršota sintakse un stingri ierobežota semantika. Var minēt šādas kontrolējamās dabiskās valodas: *Controlled English to Logic Translation* (CELT),⁶⁷ *Processable English* (PENG),⁶⁸ *Computer-Processable Language* (CPL)⁶⁹ un *Attempto Controlled English* (ACE).⁷⁰ Šķiet, ACE ir visplašāk izplatītā kontrolējamā dabiskā valoda; turklāt tai ir labāki oriģinālu nestrukturētu avotu satura izteiksmes līdzekļi,⁷¹ tāpēc tieši ACE tika izmantota semantiskās publikācijas prototipa veidošanā un aprobācijā.

Vēstures avotu publikāciju sagatavošanā ACE priekšrocības ir acīmredzamas. Tekstus, kuri ierakstīti šajā valodā, var pārveidot tā saucamajās diskursa reprezentācijas struktūrās (*Discourse Representation Structures* – DRS) un tādējādi tieši transformēt loģiskajās formulās.⁷² Pastāv arī specializēti instrumenti,⁷³ kuri

ļauj veikt visādas analītiskas operācijas, to skaitā – transformēt ACE tekstus diskursa reprezentācijas struktūrās – DRS (*ACE parser* – APE), veikt loģiskus izvedumus un pat formulēt hipotēzes (*ACE reasoner* – RACE), koriģēt tekstus, kuri ierakstīti šajā valodā (*ACE editor*), u.c. Visumā ACE virsuzdevums ir ne tikai sasaistīt informāciju, bet arī automātiski “ģenerēt” jaunas zināšanas. Vienīgais trūkums saistīts ar to, ka vēstures avotu teksti jāpārtulko ACE; tas nozīmē, ka vēstures avoti tiek interpretēti, bet avotu saturs tiek būtiski vienkāršots.⁷⁴ Turklāt ACE “konstruēšanas” likumi ir diezgan sarežģīti un neelastīgi; tie izskatās diezgan savādi un bieži vien ir pretrunā ar angļu valodas pareizrakstības likumiem. Kā piemēru var piedāvāt dažus teikumus no iepriekšminētā dokumenta Nr. 6 latviešu valodā un tulkojumā ACE:⁷⁵ “Un pašlaik es, metropolīts, tā teicu, ka tie Vitebskas iedzīvotāji netaisnīgi sūdzējās par rīdziniekiem” – “*The metropolitan of Riga says that the inhabitants of Vitebsk complain on the inhabitants of Riga unjustly*”; “Un tagad tas ir man zināms, ka rīdzinieki tajā [lietā] būtībā nav vainīgi” – “*The metropolitan of Riga knows that the inhabitants of Riga are not guilty*”; “Un tagad es tam brīnos, ka tavs vietvaldis klausās katra cilvēka vārdus” – “*The metropolitan of Riga is surprised by the credulity of the governor of Smolensk*”. Protams, tādi teksti (jāteic – tekstu surogāti) nevar aizvietot tekstus oriģinālvalodās; tāpēc semantiskajā publikācijā teksti kontrolējamajā dabiskajā valodā jāsaista gan ar oriģināliem tekstiem, gan arī ar avotu digitāliem attēliem. Tādējādi ACE var uzskatīt tikai par palīglīdzekli, kas dod iespēju sasaistīt avotisku informāciju un metainformāciju un rezultātā ģenerēt semantisko tīklu kā tādu. Visumā ACE ļauj iekļaut semantiskajā tīklā daudz vairāk informācijas aspektu nekā iezīmēšanas shēmas TEI, CEI un SMW.

Vēstures avotu semantiskās publikācijas prototipa aprobācija apstiprina tēzi, ka tieši šis vēstures avotu reprezentācijas veids semantiskajā tīmeklī var kļūt ne tikai par efektīvu avotiskās informācijas, kā arī metainformācijas reprezentācijas, meklēšanas, semantiskās sasaistīšanas līdzekli, bet arī par analītisku instrumentu, kas perspektīvā izmantojams jaunu zināšanu ģenerēšanai automātiskā loģiskā izveduma pamatā un vēsturnieka (avotpētņieka) hipotēžu verificēšanai, veicot formalizētās, uz pašām jaunākajām dator tehnoloģijām balstītās operācijas. Tas nozīmē, ka semantiskās

publikācijas faktiski ievieš jaunu paradigmu tradicionālajā avotpētniecībā, arheogrāfijā un vēstures palīgzinātnēs. Turklāt datortehnoloģijas (to skaitā tās, kas izmantotas vēstures avotu semantisko publikāciju veidošanā) var būtiski paplašināt vēstures pētījumu avotu bāzi un nodrošināt attālinātu piekļuvi unikāliem vēstures avotiem. Tajā pašā laikā nevar neatzīmēt, ka datortehnoloģijas nemaina ne vēstures avotu izpētes loģiku, ne avotpētniecisko metožu būtību; jaunajā tehnoloģiskajā līmenī tās faktiski atkārtoti vēstures avotu ārējās un iekšējās kritikas, kā arī sintētisko operāciju gaitu.⁷⁶ Tāpēc par datortehnoloģiju izmantošanas “dabiskām” robežām var kļūt tieši vēstures avotu teorijas (un pat filozofijas) attīstības līmenis, kā arī vēsturnieka (avotpētnieka) radošā iztēle.

ATSAUCES UN PIEZĪMES

- ¹ Par datortehnoloģiju un kvantitatīvo metožu izmantošanas vēsturi humanitārajās zinātnēs sīkāk sk.: Susan Hockey (2004). *The History of Humanities Computing*. In: Susan Schreibman, Ray Siemens, John Unsworth (eds.). *A Companion to Digital Humanities*. Malden (MA); Oxford; Carlton: Blackwell Publishing, pp. 3–19.
- ² Mūsdienās pētniecības koordinācijā un ideju apmaiņas nodrošināšanā sevišķi liela nozīme ir šādām organizācijām: *The Alliance of Digital Humanities Organizations (ADHO)*, sk. šīs organizācijas mājaslapu: <http://adho.org/> (skatīts 26.12.2012.); *The European Association for Digital Humanities*, <http://www.allc.org/> (skatīts 26.12.2012.); *The Association for Computers and Humanities (ACH)*, <http://www.ach.org/> (skatīts 06.08.2012.); *Digital Medievalist*, <http://www.digitalmedievalist.org/> (skatīts 06.08.2012.); *The Australasian Association for Digital Humanities (aadH)*, <http://aa-dh.org/> (skatīts 06.08.2012.); un daudzas citas. Ilgu laiku, no 20. gs. 80. gadu vidus, pastāvēja arī starptautiska organizācija, kuras uzdevums bija sekmēt informācijas tehnoloģiju ieviešanu tieši vēstures pētniecībā (*Association for History and Computing – AHC*), taču pēdējos gados aktīvi darbojas tikai šīs asociācijas reģionālā nodaļa Krievijā (kopš 1992. g.) un NVS valstīs – *Assotsiatsiia “Istoriia i komp’iuter” (AIK)*, <http://aik-sng.ru/> (skatīts 06.08.2012.).
- ³ Par reprezentatīvākām var uzskatīt starptautiskās konferences “*Digital Humanities*”, kuras katru gadu tiek rīkotas Eiropā un ASV (pamīšus). Pēdējā konference notikusi nesen – 2012. gada 16.–22. jūlijā Hamburgā. Sk. konferences mājaslapu: <http://www.dh2012.uni-hamburg.de/> (skatīts 06.08.2012.). Var arī minēt specializētas konferences, piemēram, “*Digital Diplomats*”, kas veltītas diplomātikas metodēm un aktu avotu izpētes un reprezentācijas problēmām digitālajā vidē. Sk. informāciju par konferenci “*Digital Diplomats 2011: Tools for the Digital Diplomatist*”, kura notika Neapolē 2011. gada 29. septembrī–1. oktobrī: <http://www.cei.lmu.de/digidip11/> (skatīts 06.08.2012.).

- ⁴ Manfred Thaller (2012). Controversies around the Digital Humanities: An Agenda. *Historical Social Research = Historische Sozialforschung*, 37 (3) (A Special Issue: Controversies around the Digital Humanities), pp. 7–23.
- ⁵ Piemēram, sk.: Leonid I. Borodkin (1998). Kvantitatīvnaia istoriia v sisteme koordinat modernizma i postmodernizma. *Novaia i noveishaia istoriia*, 5, s. 3–16; William G. Thomas (2004). Computing and the Historical Imagination. In: Schreibman et al. *A Companion to Digital Humanities*, pp. 56–68; Lorna Hughes (2008, reprinted in 2010). Conclusion: Virtual Representations of the Past – New Research Methods, Tools and Communities of Practice. In: Mark Greengrass, Lorna Hughes (eds.). *The Virtual Representation of the Past*. [S.l.]: Ashgate, pp. 191–201. (Digital Research in the Arts and Humanities).
- ⁶ Onno Boonstra, Leen Breure, Peter Doorn (2004). *Past, Present and Future of Historical Information Science*. Amsterdam: NIWI – KNAW, pp. 88–89.
- ⁷ Pētījumos diplomātikā latviešu valodā terminoloģija pārņemta no latīņu un vācu valodas. Latīņu termini (piemēram, dokumenta formulāra sastāvdaļu apzīmējumi) akceptējami bez ierunām. Taču par citu terminu lietošanu var rasties diskusijas. Piemēram, tāds ir termins, ar kuru apzīmēts diplomātikas izpētes objekts, vācu valodā – *Urkunde*, kas latviešu valodā parasti atveidots kā “dokuments”, to pretstatot lietvedības/kancelejas dokumentam, vācu valodā – *Akte*. Taču termins “dokuments” ir ļoti plašs, un tādējādi (šajā kontekstā) tam trūkst terminoloģiskās noteiktības. Turklāt nevar noliegt, ka diplomātikas starptautiskā valoda ir franču valoda. Franču valodā diplomātikas objekts ir *acte*. Sk.: Maria M. C. Ortí (ed.) (1997). *Vocabulaire International de la Diplomatique* [Diplomātikas starptautiskā vārdnīca], 2nd ed. València: Universitat de València; vārdnīcas modernizētā versija pieejama arī: <http://www.cei.lmu.de/VID/> (skatīts 06.08.2012.). Pēc raksta autora domām, latviešu valodā diplomātikas objektu būtu lietderīgāk saukt par “akti”, lai to norobežotu no pārējiem dokumentu veidiem. Tieši šajā nozīmē vārds “akts” turpmāk lietots rakstā.
- ⁸ Piemēram, sk.: Aleksandrs Ivanovs, Aleksei Varfolomejev (2005). Editing and Exploratory Analysis of Medieval Documents by Means of XML Technologies. In: *Humanities, Computers and Cultural Heritage: Proceeding of the XVI International Conference of the Association for History and Computing*. Amsterdam: Royal Netherlands Academy of Arts and Sciences, pp. 155–160; Aleksandr Ivanov, Aleksei Varfolomeev (2004). Ispol'zovanie tekhnologii XML dlia vvedeniia v nauchnyi oborot kompleksa dokumentov “Moscovitica–Ruthenica”. V kn.: *RCDL'2004. Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kolleksii. Trudy shestoi Vserossiiskoi nauchnoi konferentsii*. Pushchino, s. 285–289; Aleksandr Ivanov, Aleksei Varfolomeev (2006). Proekt izdaniia kompleksa dokumentov “Moscovitica–Ruthenica”. V kn.: *Sovremennye informatsionnye tekhnologii i pis'mennoe nasledie: ot drevnikh rukopisei k elektronnym tekstam*. Izhevsk: Izdatel'stvo IzhGTU, s. 65–70; u.c. Sk. arī raksta autora publikācijas, kas minētas atsaucēs turpmāk.
- ⁹ Par “datorizēto” avotpētniecību kā vispārīgās avotmācības virzienu var runāt, sākot tikai ar 20. gs. 90. gadu pirmo pusi. Sikāk sk.: Aleksandr

- Ivanov (2011). Komp'uternoe istochnikovedenie. V kn.: *Informatsionnyi biulleten' Assotsiatsii "Istoriia i komp'iuter"* (turpmāk: IB-AIK). № 37. Petrozavodsk, s. 56–62.
- ¹⁰ Aleksandr Ivanov (2012). Arkheografiia i aktovoe istochnikovedenie v epokhu komp'uternykh tehnologii. V kn.: *Problemy diplomatiki, kodikologii i aktovoi arkheografii: Materialy XXIV Mezhdunarodnoi nauchnoi konferentsii. Moskva 2–3 fevralia 2012*. Moskva: RGGU, s. 321–323.
- ¹¹ Datortehnolģiju atbilstība vēstures avotu ārējās un iekšējās kritikas, kā arī sintētiskās kritikas uzdevumiem tika speciāli aplūkota vairākās publikācijās un referātos konferencēs. Sk.: Ivanovs, Varfolomejev. Editing and Exploratory Analysis; Aleksandr Ivanov, Aleksei Varfolomeev (2005). Tekhnologiia XML kak instrument komp'uternogo istochnikovedeniia (na primere formuliarnogo analiza dokumentov prikaznogo deloproizvodstva). V kn.: Leonid Borodkin, Vladimir Vladimirov (red.). *Krug idei: Algoritmy i tekhnologii istoricheskoi informatiki*. Moskva; Barnaul: Izdatel'stvo Altaiskogo universiteta, s. 241–281; Aleksei Varfolomeev, Aleksandr Ivanov (2006). Tekhnologiia XML: Sovremennaia realizatsiia istochniko-orientirovannogo podkhoda v rabote s kompleksami istoricheskikh dokumentov. V kn.: IB-AIK. № 34. Moskva; Tambov: Izdatel'stvo Tambovskogo universiteta, s. 61–62; Aleksei Varfolomejev, Henrihs Soms, Aleksandrs Ivanovs (2008). Knowledge-Based Information Systems in Research of Regional History. In: Lisa L. Opaš-Hänninen, Mikko Jokelainen, Ilkka Juuso, Tapio Seppänen (eds.). *Digital Humanities 2008. Conference Abstracts*. [Oulu]: University of Oulu, pp. 210–211; u.c.
- ¹² Aleksei Varfolomeev, Aleksandr Ivanov (2010). Modeli dlia adekvatnoi reprezentatsii nepolnoi i/ili protivorechivoi istoricheskoi informatsii v istoricheskoi semanticheskoj seti: K postanovke problemy. No: Irēna Saleniece (atb. red.). *Vēsture: Avoti un cilvēki. Humanitārās fakultātes XIX starptautisko zinātnisko lasījumu materiāli. Vēsture XIII*. Daugavpils: Saule, 210.–216. lpp.
- ¹³ Sīkāk sk.: Aleksandr Ivanov, Aleksei Varfolomeev (2009). Idei I. D. Koval'chenko ob informatsionnoi tsennosti i neischerpaemosti istochnika v sovremennom komp'uternom istochnikovedenii. V kn.: S. Karpov (otv. red.). *Idei akademika I. D. Koval'chenko v XXI veke. Materialy IV nauchnykh chtenii pamiati akademika I. D. Koval'chenko*. Moskva: Izdatel'stvo Moskovskogo universiteta, s. 154–164, ņeit s. 158–159.
- ¹⁴ Aleksei Varfolomeev, Aleksandr Ivanov (2008). Printsipy elektronnykh publikatsii kompleksov istoricheskikh dokumentov so sredstvami paleograficheskogo, tekstologicheskogo i diplomaticheskogo analiza. V kn.: *Sovremennye informatsionnye tekhnologii i pis'mennoe nasledie: ot drevnykh tekstov k elektronnyim bibliotekam. El'Manuscript-08*. Kazan': Izdatel'stvo Kazanskogo universiteta, s. 60–63.
- ¹⁵ Varfolomejev et al. Knowledge-Based Information Systems, pp. 210–211; Aleksei Varfolomeev, Aleksandr Ivanov (2010). Virtual'nye sredy dlia nauchnykh issledovanii v arsenale istorika: tendentsii i perspektivy. V kn.: IB-AIK. № 36. Moskva: Izdatel'stvo Moskovskogo universiteta, s. 5–7. Sk. arī: Aleksei Varfolomeev, Aleksandr Ivanov, Henrihs Soms (2008). Semanticheskaia set' kak model' predstavleniia znanii po regional'noi istorii. V kn.: IB-AIK. № 35. Moskva; Barnaul: Azbuka, s. 158–160; Aleksei

- Varfolomejev, Aleksandrs Ivanovs (2009). Knowledge-Based Scholarly Environment Project for Regional Historical Studies. In: *Interactive Systems and Technologies: The Problem of Human-Computer Interaction*. Vol. III: *Collection of Scientific Papers*. Ulyanovsk: Ulyanovsk State Technical University, pp. 273–276.
- ¹⁶ Andreas Aschenbrenner, Peter Gietz, Marc W. Küster, Christoph Ludwig, Heike Neuroth (2006). TextGrid – a Modular Platform for Collaborative Textual Editing. In: *Proceedings of the International Workshop on Digital Library Goes e-Science (DLSci06), September 21, 2006*. Alicante, pp. 27–36; Marc W. Küster, Christoph Ludwig, Andreas Aschenbrenner (2007). TextGrid as a Digital Ecosystem. In: *2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference*. [S.l.], pp. 506–511. Ieee. doi: 10.1109/DEST.2007.372029.
- ¹⁷ Peter Doorn (1997). Eshche raz o metodologii. Staroe i prekrasnoe: “myl’naia opera” o neponimanii mezhdou istorikami i modeliami. *Novaia i noveishaia istoriia*, 3, s. 87–98, šeit s. 88.
- ¹⁸ Citi varianti: *data-oriented* (sk.: Donald Spaeth (2008, reprinted in 2010). Representation of Sources and Data: Working with Exceptions to Hierarchy in Historical Documents. In: Greengrass, Hughes. *The Virtual Representation*, pp. 49–61, šeit p. 50), vai arī *model-oriented approach* (Boonstra et al. *Past, Present and Future*, pp. 47–49).
- ¹⁹ <http://latgalesdati.lv/> (skatīts 06.08.2012.).
- ²⁰ Aleksandrs Ivanovs, Henrihs Soms (1999). Latgales vēstures izpētes kompleksā programma: historiogrāfisks aspekts, metodoloģija, saturs. *Latvijas Vēsture*, 1, 96.–110. lpp.
- ²¹ Henrihs Soms (1997). Vēstures informātika: saturs, struktūra un datu bāze “Latgales dati”. *Latvijas Vēsture*, 4, 136.–139. lpp.
- ²² Sīkāk sk.: Aleksandr Ivanov (2000). Metodologicheskie problemy sozdaniia bazy dannykh dlia regional’nykh issledovaniia: (Primer Latgal’skogo kraia). V kn.: *Novye informatsionnye resursy i tekhnologii v istoricheskikh issledovaniiax i obrazovanii: Sbornik tezisov dokladov i soobshchenii Vserossiiskoi konferentsii*. Moskva, s. 71–72; Aleksandr Ivanov (2002). Baza dannykh kak dinamicheskaiia model’ istorii regiona. V kn.: *IB-AIK*. № 30. Moskva, s. 170–172.
- ²³ Piemēram, sk.: Monasterium.Net: Europe’s virtual documents online [Monasterium Project]. Pieejams: <http://monasterium.net/> (skatīts 06.08.2012.); Codices Electronici Ecclesiae Coloniensis (CEEC). Pieejams: <http://www.ceec.uni-koeln.de/> (skatīts 06.08.2012.); Manuskript. Slavianskoe pis’mennoe nasledie. Pieejams: <http://manuscripts.ru/> (skatīts 06.08.2012.); un daudzi citi projekti.
- ²⁴ <http://www.lvva-raduraksti.lv/> (skatīts 07.08.2012.).
- ²⁵ <http://data.lnb.lv/retumiprojekts/main.htm> (skatīts 07.08.2012.).
- ²⁶ <http://www.itl.rtu.lv/LVA/dep1941/meklesana41p.php?id=> (skatīts 07.08.2012.). Šī datubāze būtībā ir dokumentu regestu publikācija, jo lietotājiem tā piedāvā speciāli atlasītu un strukturētu avotisku informāciju.
- ²⁷ Semantiskajām publikācijām veltīti vairāki šā raksta autora darbi, kuri sagatavoti līdzautorībā un ciešā sadarbībā ar Dr. A. Varfolomejevu. Piemēram, sk.: Aleksei Varfolomeev, Aleksandr Ivanov (2010). Semanticheskie

- publikatsii kompleksov istoricheskikh istochnikov. V kn.: *Informatsionnyye tekhnologii i pis'mennoe nasledie. El'Manuscript-10*. Ufa; Izhevsk: Vagant, s. 42–46; Aleksei Varfolomeev, Aleksandr Ivanov (2011). Semantic publications of Old Russian Charters (The Case of a Diplomatic Edition of the Complex of Old Russian Charters “Moscowitica–Ruthenica”). In: *International Conference “Digital Diplomats: Tools for the Digital Diplomatist”*. Napoli, 29 settembre – 1 ottobre 2011. Naples, pp. 36–40; Aleksey Varfolomeyev, Aleksandr Ivanov (2012). Wiki Technologies for Semantic Publication of Old Russian Charters. In: J. Ch. Meister (ed.). *Digital Humanities 2012: Conference Abstracts*. Hamburg: Hamburg University Press, pp. 405–407; u.c.
- ²⁸ Tim Berners-Lee, James Hendler, Ora Lassila (2001). The Semantic Web. *Scientific American*, 284 (5), pp. 34–43; Nigel Shadbolt, Tim Berners-Lee, Wendy Hall (2006). The Semantic Web Revisited. *IEEE Intelligent Systems*, 21 (3), pp. 96–101.
- ²⁹ David Shotton (2009). Semantic Publishing. The Coming Revolution in Scientific Journal Publishing. *Learned Publishing*, 22 (2), pp. 85–94; Andrea Baruzzo, Paolo Casoto, Prasad Challapalli, Antonina Dattolo, Nirmala Pudota, Carlo Tasso (2009). Toward Semantic Digital Libraries: Exploiting Web2.0 and Semantic Services in Cultural Heritage. *Journal of Digital Information*, 10 (6). Pieejams: <http://journals.tdl.org/jodi/article/viewArticle/688/576> (skatīts 07.08.2012.).
- ³⁰ Piemēram, sk.: Vanesa Mirzaee, Lee Iverson, Babak Hamidzadeh (2005). Computational Representation of Semantics in Historical Documents. In: *Humanities, Computers and Cultural Heritage*, pp. 199–206; Eeva Ahonen, Eero Hyvönen (2009). Publishing Historical Texts on the Semantic Web – A Case Study. In: *Proceedings of the Third IEEE International Conference on Semantic Computing (ICSC2009)*. Berkeley (CA), pp. 167–173.
- ³¹ David Shotton, Katie Portwin, Graham Klyne, Alistair Miles (2009). Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article. *PLoS. Computational Biology* 5 (4). e1000361. Pieejams: <http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000361> (skatīts 07.08.2012.).
- ³² Anita de Waard (2010). From Proteins to Fairytales: Directions in Semantic Publishing. *IEEE Intelligent Systems*, 25 (2), pp. 83–88.
- ³³ Piemēram, sk.: Vladimir Kozlov (2008). *Osnovy teoreticheskoj i prikladnoj arheografii*. Moskva: ROSSPEN, s. 91–102.
- ³⁴ Arheogrāfijas skatījumā metainformācija (jeb metadati) ir vēsturnieka (arheogrāfa) sniegtā papildu informācija par publicējamo vēstures avotu, proti, informācija, kura pausta virsrakstā (avota satura iss izklāsts) un leģendā (dati par avota glabāšanas vietu, autentiskumu, iepriekšējām publikācijām utt.), tekstuālas piezīmes un vēsturiski komentāri, nepieciešamības gadījumā avota datējuma un atribūcijas pamatošana, informācija

- par vēstures avotos minētajām personām, organizācijām un vietām utt. Tādējādi metainformācija ir informācija, kura pastāv it kā “ārpus” vēstures avota teksta un kura tiek sasaistīta ar avota “iekšējo” informāciju.
- ³⁵ Aleksandrs Ivanovs (2006). Dokumentu kompleksa apzināšana, rekonstrukcija un izpēte: arheogrāfiskie un avotpētnieciskie aspekti. *Humanitāro Zinātņu Vēstnesis*, 10, 76.–88. lpp.; Aleksandr Ivanov (2010). Istochnikovvedcheskie i arheograficheskie aspekty rekonstruktsii istoricheskikh kompleksov istochnikov (otdel “Moscovitica–Ruthenica” v byvsheem arkhive Rizhskogo magistrata). V kn.: Valentin Ianin, Vladimir Nazarov (red.). *Soslovitā, instituty i gosudarstvennaia vlast’ v Rossii. (Srednie veka i rannee Novoe vremia): Sbornik statei pamiati akademika L. V. Cherepnina*. Moskva: Iazyki slavianskikh kul’tur, s. 97–105. (*Studia philologica*).
- ³⁶ Ivanovs, Varfolomejev. *Semantic Publications of Charter Corpora*. Sk. arī 27. atsaucē minētās publikācijas.
- ³⁷ Par šo dokumentu kompleksu, kas glabājas LNA LVVA, sīkāk sk.: Aleksandrs Ivanovs (2004). Kompleksa “Moscovitica–Ruthenica” ieviešana zinātnes aprītē: arheogrāfisks apskats. *Latvijas Arhīvi*, 2, 47.–85. lpp.; Aleksandr Ivanov (2004). “Moscovitica–Ruthenica” v Latviiskom gosudarstvennom istoricheskom arkhive: istoriia formirovaniia kompleksa, sostav i vvedenie v nauchnyi oborot. *Drevniaia Rus’: Voprosy medievistiki*, 3, s. 47–54; 4, s. 94–106.
- ³⁸ Aleksandrs Ivanovs, Anatolijs Kuzņecovs (2009). *Smolenskas–Rīgas aktis: 13. gs.–14. gs. pirmā puse: Kompleksa Moscovitica–Ruthenica dokumenti par Smolenskas un Rīgas attiecībām / Smolensko-rizhskie akty: XIII v. – pervaiā polovina XIV v.: Dokumenty kompleksa Moscovitica–Ruthenica ob otnosheniiakh Smolenska i Rigi*. Rīga: LVVA; Latvijas Arhīvistu biedrība; DU LPI. (Vēstures avoti, 6). Sk. dokumentus Nr. 1, 3a, 4, 5, 6. Semantiskās publikācijas prototipā dokumentiem tiek konvencionāli piešķirti numuri atbilstoši šim izdevumam.
- ³⁹ Varfolomejev et al. *Knowledge-Based Information Systems*, pp. 210–211.
- ⁴⁰ Sīkāk sk.: Desmond Schmidt (2012). The Role of Markup in the Digital Humanities. *Historical Social Research*, 37 (3), pp. 125–146.
- ⁴¹ <http://www.w3.org/XML/> (skatīts 13.08.2012.).
- ⁴² Piemēram, tags <charter> apzīmē publicējama dokumenta (viduslaiku akts) sākumu, bet tags </charter> – beigas. Lauka starp šiem tagiem iezīmēšana tiek detalizēta, izmantojot citus pāra tagus (sk. dokumenta iezīmēšanas piemēru turpmāk).
- ⁴³ Jāatzīmē, ka šajā gadījumā vēstures avotu tekstus var publicēt jebkurā valodā, to skaitā – oriģinālvalodā. Vienīgais ierobežojums ir prasība izmantot standartizētus simbolus (*Unicode’s characters*). Veidojot vēstures avotu diplomātisko publikāciju, nestandarta simbolu atveidei jāizmanto speciāli elementi – <glyph>, kuri dod iespēju aprakstīt simbolus (piemēram, senkrievu burtus), kuriem ir īpatnējas grafiskas formas. Šis paņēmieni izmantots semantiskās publikācijas prototipā.
- ⁴⁴ Jāteic, ka ar tehnoloģijas XML palīdzību var veidot ļoti lielus failus, kuros ir iespējams secīgi iekļaut daudzus vēstures avotus.
- ⁴⁵ Par XML izmantošanu vēstures avotu edīcijā un izpētē sk. arī: Ivanov, Varfolomejev. *Tehnoloģiā XML kak instrument*. Sk. arī: Aleksandr Ivanov,

- Aleksei Varfolomeev (2009). Publikatsiia i analiz rukopisnykh istoricheskikh dokumentov s pomoshch'iu tekhnologii XML. *Humanitāro Zinātņu Vēstnesis*, 16, 69.–84. lpp.; Aleksei Varfolomeev, Aleksandr Ivanov (2004). Tekhnoloģiia XML kak instrument komp'uternogo istochnikovedeniia: (na primere izucheniia kompleksa srednevekovykh dokumentov po istorii Dinaburga). V kn.: *IB-AIK*. № 32. Moskva; Tomsk: Izdatel'stvo Tomskogo universiteta, s. 137–139.
- ⁴⁶ Text Encoding Initiative, sk.: <http://www.tei-c.org/index.xml> (skatīts 07.08.2012.). Šis iezīmēšanas standarts tiek plaši izmantots tekstu digitālu kolekciju veidošanas projektos humanitāro zinātņu vajadzībām.
- ⁴⁷ Charters Encoding Initiative, sk.: <http://www.cei.lmu.de/> (skatīts 07.08.2012.). Šis standarts tieši paredzēts viduslaiku aktu publicēšanai, to struktūras (formulāra) izpētei un aktu digitālu kolekciju veidošanai.
- ⁴⁸ [1284.–1297. g., iespējams, tuvāk 1285.–1287. g.]. Pergaments. 16,5 × 23,3 cm. Noraksts (koncepts). Tulkojums. *LNA LVVA*, 673–4–18(kast.)–8, 3. lp. Sk.: Ivanovs, Kuzņecovs. *Smolenskas–Rīgas aktis*, Nr. 6. Tieši šis dokuments veido semantiskās publikācijas prototipa “centru”.
- ⁴⁹ Tagu ietvaros atribūti apzīmēti šādi: id="...". Atribūti sniegti angļu valodā – tajā pašā valodā, kurā veidota iezīmēšanas shēma un notiek informācijas sasaistīšana semantiskā tīkla ietvaros.
- ⁵⁰ Jāatzīmē, ka gan CEI, gan TEI standarts ļauj papildināt elementu klāstu.
- ⁵¹ Ivanovs, Varfolomeyev. Editing and Exploratory Analysis, pp. 157–159.
- ⁵² The CIDOC Conceptual Reference Model. Pieejams: <http://www.cidoc-crm.org/> (skatīts 07.08.2012.). Sk. arī: Martin Doerr (2003). The CIDOC CRM. An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine*, 24 (3), pp. 75–92.
- ⁵³ Piemēram, sk.: Michele Pasin, Enrico Motta, Milton Keynes (2008). PhiloSURFical: Browse Wittgenstein's Tractatus with the Semantic Web. In: Alois Pichler, Herbert Hrachovec (eds.). *Wittgenstein and the Philosophy of Information. Proceedings of the 30th International Ludwig Wittgenstein Symposium*. Vol. 1. [S.l.]: Ontos Verlag, pp. 319–333.
- ⁵⁴ Nancy Ide, David Woolner (2007). Historical Ontologies. In: Khurshid Ahmad, Christopher Brewster, Mark Stevenson (eds.). *Words and Intelligence*. Vol. II: *Essays in Honor of Yorick Wilks*. Dordrecht: Springer, pp. 137–152.
- ⁵⁵ Ahonen, Hyvönen. Publishing Historical Texts.
- ⁵⁶ Varfolomeyev, Ivanovs. Knowledge-Based Scholarly Environment Project, pp. 273–276.
- ⁵⁷ Piemēram, sk.: Karl E. Napierskii (sost.). (1857). *Gramoty, kasaiushchiesia do snoshenii severo-zapadnoi Rossii s Rigoiu i ganzeiskimi gorodami v XII, XIII i XIV veke*. Sankt-Peterburg: Tipografiia imperatorskoi Akademii nauk. № III, u.c.
- ⁵⁸ Nr. 4: Smolenskas kņaza Fjodora Rostislaviča vārdā sastādīts raksts Rīgas arhibīskapam Jānim I, Livonijas ordeņa mestram Villekinam un Rīgas rātskungiem par brīva ceļa piešķiršanu vācu un krievu tirgotājiem (1284. g. [18. maijs]); Nr. 5: Smolenskas kņaza Fjodora Rostislaviča raksts sakarā ar tiesas prāvu par vācu “zvanu” – atsvaŗu (1284. g.); Nr. 3a: 1229. gadā Smolenskas kņaza Mstislava Davidoviča noslēgtā līguma ar Rīgu, Gotlandi un vācu pilsētām Gotlandes recenzijas noraksts A; Nr. 1: Noteikumi

Smoļenskas līgumam ar Rīgu un Gotlandi ([1219.–1228./9. g.]). Sk.: Ivanovs, Kuzņecovs. *Smoļenskas–Rīgas aktis*.

- ⁵⁹ 1. attēlā šīs saites aprakstītas latviešu valodā, taču pašā tīklā apraksts veidots, protams, angļu valodā.
- ⁶⁰ Piemēram, sk.: Fabio Ciravegna, Mark Greengrass, Tim Hitchcock, Sam Chapman, Jamie McLaughlin, Ravish Bhagdev (2008, reprinted in 2010). Finding Needles in Haystacks: Data-mining in Distributed Historical Datasets. In: Greengrass, Hughes. *The Virtual Representation*, pp. 65–79.
- ⁶¹ Adam Souzis (2005). Building a Semantic Wiki. *IEEE Intelligent Systems*, 20 (5), pp. 87–91; Sören Auer, Sebastian Dietzold, Thomas Riechert (2006). OntoWiki – A Tool for Social, Semantic Collaboration. In: Isabel Cruz et al. (eds.). *The Semantic Web - ISWC 2006: 5th International Semantic Web Conference. ISWC 2006. Athens, GA, USA, November 5–9, 2006. Proceedings*. Berlin; Heidelberg: Springer-Verlag, pp. 736–749. (Lecture Notes in Computer Science, 4273).
- ⁶² Semantic MediaWiki (SWM), <http://www.semantic-mediawiki.org/> (skatīts 07.08.2012.).
- ⁶³ Markus Krötzsch, Denny Vrandečić, Max Völkel, Heiko Haller, Rudi Studer (2007). Semantic Wikipedia. *Journal of Web Semantics*, 5 (4), pp. 251–261.
- ⁶⁴ Internetā ir pieejama dokumenta Nr. 6 un attiecīgās informācijas publikācija, kura veidota, izmantojot tehnoloģiju SMW. Sk.: Historical Documents Semantic Publishing: http://histdocs.referata.com/wiki/Main_Page (skatīts 07.08.2012.).
- ⁶⁵ Pašlaik izstrādā SMW paplašinājumus (piemēram, XML2Wiki vai Web-Fonts), kuri perspektīvā dos iespēju šo problēmu risināt.
- ⁶⁶ Norbert E. Fuchs (ed.) (2010). *Controlled Natural Language: Workshop on Controlled Natural Language, CNL 2009. Marettimo Island, Italy, June 2009. Revised Papers*. Berlin; Heidelberg: Springer-Verlag. (Lecture Notes in Computer Science, 5972). Sk. arī: Michael Rosner, Norbert E. Fuchs (eds.) (2010). *CNL 2010: Second Workshop on Controlled Natural Languages. Pre-Proceedings... Marettimo Island, Sicily, Italy, September 13–15, 2010*. Pieejams: <http://ceur-ws.org/Vol-622/> (skatīts 07.08.2012.).
- ⁶⁷ Adam Pease, William Murray (2003). An English to Logic Translator for Ontology-Based Knowledge Representation Languages. In: *Proceedings of the 2003 IEEE International Conference on Natural Language Processing and Knowledge Engineering*. Beijing, pp. 777–783.
- ⁶⁸ Rolf Schwitter (2011). Processing Coordinated Structures in PENG Light. In: Dianhui Wang, Mark Reynolds (eds.). *AI 2011: Advances in Artificial Intelligence. 24th Australasian Joint Conference. Perth, Australia, December 5–8, 2011. Proceedings*. Berlin; Heidelberg: Springer-Verlag, pp. 658–667. (Lecture Notes in Computer Science, 7106).
- ⁶⁹ Peter Clark, Phil Harrison, Tom Jenkins, John Thompson, Rick Wojcik (2005). Acquiring and Using World Knowledge Using a Restricted Subset of English. In: *The 18th International FLAIRS Conference (FLAIRS'05)*. Pieejams: <http://www.cs.utexas.edu/users/pclark/papers/flairs.pdf> (skatīts 07.08.2012.).
- ⁷⁰ Juri L. de Coi, Norbert E. Fuchs, Kaarel Kaljurand, Tobias Kuhn (2009). Controlled English for Reasoning on the Semantic Web. In: François Bry, Jan Małuszynski (eds.). *Semantic Techniques for the Web: The Reverse*

Perspective. Berlin; Heidelberg: Springer-Verlag, pp. 276–308. (Lecture Notes in Computer Science, 5500).

- ⁷¹ Jāatzīmē arī, ka viena no *Wiki*-sistēmām (*AceWiki*) tekstu ierakstīšanai izmanto tieši ACE.
- ⁷² Sk.: Hans Kamp, Uwe Reyle (1993). *From Discourse to Logic: Introduction to Mode-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers.
- ⁷³ Attempto Tools. <http://attempto.ifi.uzh.ch/site/tools/> (skatīts 13.08.2012.).
- ⁷⁴ Tas attiecas pat uz tekstiem angļu valodā, ja tie tiek atveidoti ACE.
- ⁷⁵ Teikumi, kuri turpmāk sniegti kontrolējamajā dabiskajā valodā, tulkoti nevis no šā avota tulkojuma latviešu valodā, bet no oriģinālvalodas.
- ⁷⁶ Sk. arī: Aleksandr Ivanov (2006). Rabota s XML-dokumentom kak vosproizvedenie osnovnykh etapov istochnikovedcheskoi kritiki: novye tekhnologii i vozmozhnost' korreksii traditsionnykh podkhodov. V kn.: *IB-AIK*. № 34. Moskva; Tambov: Izdatel'stvo Tambovskogo universiteta, s. 66–67; Aleksandr Ivanov, Aleksei Varfolomeev (2005). O perspektivakh i predelakh primeneniia metodov komp'uternogo istochnikovedeniia v medievistike. *Drevniaia Rus': Voprosy medievistiki*, 3, s. 34–35.

COMPUTER-BASED TECHNOLOGIES IN REPRESENTATION OF HISTORICAL RECORDS, EDITING, AND SOURCE CRITICISM

Aleksandrs Ivanovs

Dr. hist., professor at Daugavpils University, leading researcher at the Scientific Institute of Regionalism at Rēzekne Higher School. Scientific interests: historiography of the history of Latvia, historical records editing, diplomatics, source studies.

E-mail: aleksandrs.ivanovs@du.lv

The aim of this paper is to demonstrate potentialities of computer-based technologies in traditional spheres of historical research, mainly in historical records editing, criticism, diplomatics, and auxiliary historical disciplines. The paper is based on a number of pilot projects that since 2000 have been implemented by the author in cooperation with Dr. math. Aleksey Varfolomeyev (Petrozavodsk State University, Russia). The following conclusion has been drawn: nowadays, semantic publications of historical records are of utmost importance, since they afford effective tools for data-mining, representation and linkage of sources, as well as analytical tools in order to generate new knowledge and to verify hypotheses advanced by historians. The paper reveals potentialities of computer-based technologies, first and foremost, in the field of representation, editing, and criticism of medieval historical records; however, the same technologies can be efficiently used in in-depth research of vast complexes of modern history records.

Key words: historical records, computer-based technologies, semantic publications.

Summary

The paper deals with the problems of historical records studies in the age of e-science. The author focuses his attention on the potentialities of computer-based technologies in traditional spheres of historical research, first and foremost in historical records editing, external and internal source criticism, historical diplomatics, and auxiliary historical disciplines. The impact of computer-based technologies on the above-mentioned traditional and fairly conservative fields of historical research poses a number of theoretical and methodological problems concerning interconnection of traditional historical source studies and computer-based source studies, as well as applicability and advantages of computer-based technologies in publication (representation on the web) and exploratory analysis of the written historical records. In order to solve these problems, since 2000, a number of pilot projects have been implemented by the author of the paper in close cooperation with Dr. math. Aleksey Varfolomeyev (Petrozavodsk State University, Russia). The results have been presented at different international conferences and in a number of published research papers.

It can be stated that nowadays the progress in computer-based studies of historical records is determined by the development of the semantic web technologies. Therefore, semantic publications of historical records assume utmost importance, since they afford effective tools for data-mining, representation and linkage of sources, as well as analytical tools in order to generate new knowledge on the basis of logical inference and to verify hypotheses advanced by historians.

In order to show how the basic principles of semantic publications can be applied to electronic scholarly editing and analysis of the written historical records, the paper presents a multifunctional prototype of the semantic publication of a sub-complex of medieval historical records. It is obvious that semantic publications should be based on “natural” (historical) complexes of interconnected documentary records. Actually, such complexes represent integral parts of definite systems that documents form in the course of performing their initial functions. Taking into account the above-mentioned considerations, in the prototype of the semantic publication, five interconnected charters that reflect the course of relations between Riga and Smolensk in the 13th century have been used. This sub-complex of the documents forms a constituent part of the vast collection of medieval and early modern records “Moscowitica–Ruthenica” kept at the State Historical Archives of Latvia, the structural department of the National Archives of Latvia (Riga).

In the paper, different approaches to the creation of additional semantic layers in digital publications of documentary records have been

tested. Since the advantages of semantic publications are determined by a possibility to capture, link, and aggregate historical information, which is provided by different documentary records, paramount importance is attached to the semantic layer that embraces information about persons, place-names, institutions, terms, events, documents (references to documents), etc., which are directly or indirectly mentioned in the charters' texts. These objects have been marked up using CEI (Charters Encoding Initiative) markup scheme. In the prototype, this semantic markup has been supplemented with the structural markup of the charters' texts, which forms a basis for further pattern analysis.

Since marked out objects should be linked either with analogous objects marked up in other documents, or with corresponding data, which are represented in specialized ontologies, the ontology has been specially designed for the purposes of the prototype of the semantic publication. Unfortunately, production of semantic publications based on ontologies, which are directly recorded using RDF or OWL, is time-consuming. Therefore, the paper also proposes an alternative modification of the prototype of the semantic publication based on Semantic MediaWiki tools.

However, the semantic layers that capture marked up objects do not represent the tenor of historical records exhaustively. Nowadays, the possibility to represent fully the contents of historical records is afforded by controlled natural languages, which are understood not only by a human reader, but also by a computer. The paper shows that one of these languages – Attempto Controlled English (ACE) – seems to be a very promising tool for the purposes of the semantic publication and analysis of the complexes of historical records.

FIGURE CAPTION

Fig. 1. Representation of the records sub-complex that forms a semantic network (fragment)

Iesniegts 13.08.2012.