

KOMPIUTERINĖS DUOMENŲ BAZĖS PANAUDOJIMAS LIETUVOS PARLAMENTARIZMO ISTORIJS TYRIMAMS

VIGINTAS STANCELIS

*Vilniaus pedagoginis universitetas
T. Ševčenkos g. 31, LT-2009 Vilnius
El. paštas: stancelis@vpu.lt*

Santrauka

Lietuvos parlamentarizmo istorijos duomenų bazės projektas prieš pradėdant parlamentarų žodyno rašymą leidžia preliminariai išanalizuoti tyrinėtojų jau turimą medžiagą ir atmesti kriterijus, pagal kuriuos toliau reikės ieškoti ir kaupti trūkstamą informaciją. Kompiuterinės duomenų bazės (toliau – DB) panaudojimas gali žymiai palengvinti informacijos apsikėitimą tarp tyrinėtojų, o taip pat sąlygoti efektyvesnę pasiskirstymą vaidmenimis (informacijos rinkėjas, įvedimo operatorius, galutinių tekstų rašytojas) tarp bendradarbiaujančių kolegų. Į kompiuterinės DB parengimo darbus galima įtraukti tam tikrą vyresnių kursų studentų skaičių, tai turės nemažą edukacinį aspektą – kaip supažindinimo su reikšmingo Lietuvos istorijos aspekto problematika, taip ir kompiuterinių technologijų istorijos mokslė propagavimo prasme. Duomenų prieinamumas ne tik tiesioginiams projekto dalyviams, bet ir jų publikacija platesniam tyrinėtojų ratui prieinama forma sudarytų galimybes panaudoti surinktas žinias kitų Lietuvos istorijos tyrimų kontekste.

Raktažodžiai: parlamentarizmas, informacinės technologijos, duomenų bazės, programinė įranga, lentelės, užklauso.

Išvadas

Naujų informacinių technologijų panaudojimas istorijos tiriamajame darbe kol kas tebėra nepakankamai perprasta sfera. Dažniausiai kompiuteris išnaudojamas kaip tekstų suvedimo ir spausdinimo priemonė, ir tik labai retai jis dirba pagal tikrąją paskirtį – atlieka sukauptos medžiagos analizę. Dirbant su dideliais, iš esmės vieningos struktūros duomenų srautais, kaip kad parlamentarų biografiniai duomenys, kompiuterinių technologijų panaudojimo tikslingumas tampa ypač akivaizdus.

Viena vertus, iš anksto nustatčius vieningus reikalavimus duomenims, kurie vėliau bus kaupiami DB, gali būti žymiai palengvintas ir pats žinių rinkimo procesas – visi tyrinėtojai naudotųsi tuo pačiu klausimynu, dėdami surinktus faktus į gatavai paruoštas „lentynėles“. Tokia iš anksto struktūrizuota paieška paprastai būna kur kas greitesnė ir tikslesnė už laisvą faktų „medžioklę“.

Iš kitos pusės, jau pačioje DB tai leistų lengvai atrinkti rūpimus faktus iš didelio pirminės informacijos kiekio, o taip pat atlikti statistinės analizės ir kitus apibendrinančius veiksmus.

DB kūrimo bandymas neabejotinai būtų naudingas ir kaip naujos metodikos kūrimo ir bandymo dirva, nes skaitmeninių technologijos panaudojimas istoriko darbe ateityje neišvengiamai plėsis ir tam reikia ruošti tiek pačius istorikus, tiek ir metodinę medžiagą.

Tikslai: sudaryti kompiuterinę duomenų bazę, apimančią informaciją apie visus XV-XXI a. Lietuvos parlamentarus. Pritaikyti ją istorija besidominčiam vartotojui, neturinčiam specializuoto kompiuterinio pasirengimo. Išleisti kompaktinę plokštelę ir padaryti prieinamą interneto vartotojams.

Tyrimo metodai: sukurti standartizuotą klausimyną, pagal kurį tyrinėtojai rinks žinias archyvuose, spausdintuose šaltiniuose ir mokslinėse publikacijose. Sudaryti kompiuterinę reliacinę duomenų bazę ir į ją suvesti rastas žinias. Sukauptos duomenų bazės pagrindu atlikti statistinę informacijos analizę, apibendrinti gautus rezultatus.

Dėstymas

Viena pagrindinių Vilniaus pedagoginio universiteto Istorijos fakulteto Lietuvos istorijos katedros darbo krypčių artimiausiems keliems metams – keturių tomų Lietuvos parlamentarų žodyno sudarymas. Numatoma surinkti kiek galima išsamesnę medžiagą apie visus XV-XXI a. aukščiausios atstovaujamosios valdžios asmenis, sudaryti jų biogramas, pateikti svarbiausius apibendrinimus. Į šį darbą bus įtraukti ne vien tik katedros darbuotojai, bet ir pasitelkta kitur dirbančių kolegų. Didelė darbo apimtis, būtinybė dalintis darbais ir keisti informacija tarpusavyje, standartizuoti galutinio produkto reikalavimai, kai reikės sukurti tūkstančius tokios pat struktūros biogramų – gimimo data, partinė priklausomybė, išrinkimas į Seimą, politinė veikla jame – savaiame kelia mintį apie kompiuterinės duomenų bazės panaudojimo tikslumą.

Kuriama duomenų bazė (DB) – sudėtinė didesnio viso Lietuvos istorijos katedros darbo proceso dalis, tad jos įgyvendinimas planuotinas derinant su būsimu parlamentarų žodyno rašymo procesu, siekiant, kad kompiuteriniai duomenys būtų surinkti ir pateikti kitiems kolegoms prieinamu formatu kiek įmanoma anksčiau, tam, kad remiantis turimais išrašais būtų galima sudarinėti galutines tekstines parlamentarų biogramas. Skaitmeniniu formatu užrašyta istorinė informacija turi potencialiai didelę ir ilgalaikę vertę tyrinėtojų ir pedagogų bendruomenei, tačiau racionaliai ji bus panaudota tuo atveju, jeigu priėjimą prie jos turės ne tik tiesioginis duomenų kaupėjas, bet ir kiti suinteresuoti vartotojai.

Tradicinės gairės tokio darbo pradžioje – pasitarti su kolegomis, kurie jau praėję panašaus darbo kelią, susipažinti su šiai temai skirta literatūra, įvertinti galimos naudoti programinės įrangos galimybes ir pasirinkti iš jų tinkamiausią. Deja, kol kas neteko girdėti apie politinei istorijai skirtos DB panaudojimą Lietuvos istorijos tyrimams, todėl tenka remtis tik užsienio kolegų patirtimi ir rekomendacijomis. Šio straipsnio tikslas – ne tiek atsiskaityti apie pirmąjį parlamentarizmo istorijos DB projektavimo etapą, kadangi projektas dar tik įsibėgėja, bet ir plačiau pristatyti Lietuvos kolegoms kompiuterinės DB koncepciją, atskleisti šio metodo esminius pranašumus ir trūkumus, pabandyti savotiškai paskatinti efektyvių technologijų naudojimą istoriko darbe.

Užsienio patirtis. Vakarų Europoje ir Jungtinėse Amerikos Valstijose DB panaudojimas istoriniams tyrinėjimams jau turi nemažas tradicijas. Tačiau pati šiose šalyse naudojama *duomenų bazės* sąvoka yra ganėtinai neapibrėžta, dažnai tuo pačiu terminu vadinami tiek paprasčiausi kokių nors pareigūnų ar valdų sąrašai, tiek ir griežtai standartizuotos, pritaikytos kompiuterinei analizei elektroninės lentelės – tai, ką duomenų bazėmis laiko informatikos specialistai. Toliau šiame straipsnyje bus nagrinėjamas tik antras, „tikrasis“ DB variantas.

Natūralu, kad į tokius ankštus rėmus nėra lengva įsprausti bet kokius duomenis, todėl dažniausia kompiuterinių DB panaudojimo sfera – ekonominė ir žemėvaldos istorija, demografiniai procesai, visa tai, kas natūraliai reikalauja statistinio apibendrinimo. Visgi nuolatos buvo stengiamasi kuo labiau išplėsti tyrimų problematiką, aprėpiant karinę, religinę, net kultūrinę sferas.

Peržiūrint užsienio šalių istorikų nuveiktą darbą, daugiausia dėmesio teko skirti anglų kalba publikuotai medžiagai. Nors JAV yra nuveiktas grandiozinis darbas pritaikant DB istoriniams tyrimams, pavyzdžiui, vien *Historical Data Systems* tvarkomai Amerikos Pilietinio karo duomenų bazei sudaryti skirta per 60 000 darbo valandų¹, vis dėlto kaip naudingiausias parlamentarizmo istorijos prasme tenka įvardyti Didžiojoje Britanijoje atliktus tyrimus ir paskelbtas metodikas. Tai apibrėžia pati JAV istorijos tyrimų specifika, kur praktiškai visas dėmesys sutelktas naujesiems ir naujausiems laikams, tuo tarpu britų mokslininkams tenka aprėpti kur kas platesnę chronologinę ir probleminę erdvę.

Išskirtinai Didžiojoje Britanijoje reikėtų paminėti Esekso universitetą, kuris yra savotiškas kompiuterinių inovacijų istorijos moksle centras. Jame veikia Didžiosios Britanijos Duomenų Bazė², skirta XIX ir XX a. statistiniams tyrimams, kurioje apjungtos ne tik Esekso, bet ir Aberdyno, Kembridžo, Stafordšyro universitetų pastangos. Kitas padalinys – Istorinių duomenų tarnyba (History Data Service) dirba leisdama rekomendacijas ir metodinius patarimus kaip projektuoti ir eksploatuoti naujas DB. Ją įsteigęs Jungtinis

¹ *Civil War Research Database*, <http://www.civilwardata.com>

² *The Great Britain Historical Database*, <http://hds.essex.ac.uk/gbh.asp>

informacinių sistemų komitetas³ apibrėžė gana tikslus ir patogius naudojimo kriterijus, kuriais remiantis būtų renkami, kataloguojami, valdomi, saugomi ir palaikomi skaitmeniniai istorijos mokslo resursai. Jie gali būti pritaikyti daugeliui skirtingų duomenų apdorojimo lygmenų, nuo studentiškų tyrimų iki plataus masto tiriamųjų projektų. Pagrindinis reikalavimas – duomenis kaupti standartizuotų lentelių pavidalu, kurias vėliau bus galima panaudoti kompiuterinėse duomenų bazėse, skaičiuoklėse ar statistinėse programose.

Daug padaryta verčiant istorinę informaciją elektroniniu formatu ir pritaikant sistemingam kompiuteriniam apdorojimui Skandinavijos šalyse. Stokholmo miesto archyvas dalį savo medžiagos lentelių pavidalu skelbia internete⁴, Danijos Duomenų archyvas (Dansk Data Arkiv⁵, specialus Danijos valstybinio archyvo padalinys) kaupia skaitmeninių duomenų banką, apimančią XVI iki XX a. Ta pačia kryptimi dirbama ir Norvegijos aukštosiose mokyklose – Oslo⁶, Bergeno universitetuose⁷. Pradeda aktyviau dirbti šioje srityje ir Rusijos istorikai, ir nebūtinai vien didžiuosiuose mokslo centruose – paimekime kad ir Tverės universitete ruošiamas DB: „Тверская усадьба. Век за веком“, „Тверское дворянство. Гербы“⁸.

Programinės ir techninės įrangos pasirinkimo klausimas. Visos populiariausios šiuolaikinės DB iš esmės veikia vienu ir tuo pačiu principu: tarpusavyje susietų *lentelių* visuma, todėl vadinamos reliacinėmis (nuo angl. žodžio *relation*). Lentelės savo ruožtu yra sudarytos iš stulpelių, kitaip vadinamų *laukų*, kuriuose apibrėžta, kokie duomenys bus renkami, pvz., „Vardas“, „Pavardė“, „Tėvavardis“, ir eilučių, kuriuose saugoma konkretūs duomenys, pvz., Petras, Jonaitis, Antano. Visa vienoje eilutėje saugoma žinių visuma vadinama *įrašu*, ir atitinka visą turimą DB informaciją apie atskirą

³ Joint Information Systems Committee, <http://www.jisc.ac.uk>

⁴ Stockholms stadsarkiv, <http://www.ssa.stockholm.se>

⁵ Dansk Data Arkiv, <http://www.dda.dk>

⁶ Enhet for digital dokumentasjon ved HF, <http://www.dokpro.uio.no/engelsk/index.html>

⁷ Digitalisering av eldre materiale, <http://www.ub.uib.no/avdeling/fdok/digi/index.htm>

⁸ Генскалогия и история тверских дворян (до 1890 года). <http://library.tversu.ru/usadyba/dvoryane/poisk.htm>

asmenį, įvykį ir pan. Tokia DB struktūra buvo specialiai sukurta prekių apyvarta paremtam verslo pasauliui, todėl natūralu, kad jos taikymas istorinio tyrimo tikslams kelia nemažų sunkumų. Vis dėlto mūsų sąlygomis nėra kito pasirinkimo, kaip tik naudoti kurią nors iš jau patikrintų, populiarią ir lengvai prieinamą reliacinę DBVS (duomenų bazių valdymo sistemą).

Buvo galimi šie komercinės programinės įrangos pasirinkimai – Oracle, Microsoft FoxPro ir Microsoft Access. Tačiau pirmieji du variantai negalėjo būti traktuojami kaip rimta alternatyva dėl daugelio priežasčių – visų pirma, tai iš tiesų brangios ir profesionaliam vartotojui skirtos programos, be to, jos kelia aukštesnius reikalavimus techninei įrangai. Tuo tarpu Microsoft Access įeina į standartinį Microsoft Office paketą, taigi juo jau iš karto buvo aprūpinta dauguma VPU naudojamų kompiuterių, jį nesunku gauti tiek katedros darbuotojams, tiek talkinantiems studentams. Sistemai keliami tokie pat reikalavimai, kaip ir kitoms Microsoft Office programoms. Pastaruoju metu šiai programai išleista nemažai mokomosios literatūros lietuvių kalba⁹.

Galimas ir kitas kelias, būdingas daugeliui aukštųjų mokyklų – naudoti nemokamą Atviro kodo (Open Source) programinę įrangą, kuri suteikia galimybę pačiam vartotojui keisti, tobulinti sistemą, papildyti ją savo sukurtais moduliais. Tačiau ši laisvė, brangi profesionaliam informatikui, gresia dideliais rūpesčiais tam, kuriam kompiuterija – tik papildoma disciplina, bet ne gyvenimo pašaukimas. Nemokamos ir neužbaigtos sistemos pritaikymas konkreitiems vartotojo poreikiams gali pareikalauti kvalifikuotų specialistų pramos ir didelių darbo laiko sąnaudų. Todėl bent jau pirmuoju etapu buvo nutarta apsistoti ties komercine, bet jau iš anksto turima Microsoft Access programa, turint omenyje galimybę susidarius palankioms aplinkybėms pereiti prie nemokamos Atviro kodo programinės įrangos. Materialinė techninė bazė – VPU IF turimi kompiuteriai, tarp kurių dominuoja toli gražu jau nebe naujos Pentium II klasės sistemos, kurios tačiau iš esmės patenkina šiam darbui keliamus reikalavimus.

9 Microsoft Access 2000. Žingsnis po žingsnio: priemonė mokytis pačiam. Kaunas, 2001; Minkevičius S. Microsoft Access XP pradžiamokslis: 8 pamokos. Vilnius, 2003; Vidžiūnas A., Marčiulytė R. Access XP: taikomųjų duomenų bazių projektavimo pagrindai. Kaunas, 2003; Gulbinienė E., Petrauskienė J. Duomenų bazių valdymo sistema MS Access: praktiniai darbai: mokomoji knyga. Šiauliai, 2004.

Parengiamieji darbai. Skaitmeninių resursų kūrimas remiantis istoriniais dokumentais ar tiriamąja medžiaga kaip savarankiškas, pilnavertis mokslinis projektas kol kas dar sutinkamas retai, dažniausiai šio darbo imamas kaip sudėtinės kokios nors didesnės tiriamosios programos dalies, suteikiant jam pagalbinį, papildomą vaidmenį. Kita vertus, resursų kaupimas kaip atskiras, su tiriamuoju, apibendrinančiu darbu nesusietas darbas gali tapti ambicingu, bet neatsiperkančiu, nepasiteisinančiu užsiėmimu. Todėl vadovaujantis vientiso ryšio tarp informacijos kaupimo ir apdorojimo principu buvo pradėti ir Lietuvos parlamentarizmo istorijos duomenų bazės kūrimo darbai, iš anksto ketinant ją integruoti į VPU IF vykdomą Lietuvos parlamentarų žodyno kūrimo programą.

Iš karto pradėti kurti galutinai išbaigtą, „tobulą“ DB iš esmės remiantis vien tik Istorijos fakulteto personalo pajėgomis – skambėtų kaip nelabai apgalvota mintis. Dėl to buvo imtasi nemažai parengiamųjų darbų, kuriais siekta patikrinti technines ir žmoniškąsias galimybes, supažindinti kolegas su pagrindiniais darbo metodais, atlikti pirmuosius bandymus, taip išsiaiškinant problemines vietas ir tolimesnio darbo perspektyvas.

2003-2004 m. antrame semestre istorijos specialybės magistrantams buvo pradėtas dėstyti specialus kursas „Darbas su istorinių duomenų bazėmis“, sudarytojai – dr. doc. A. Ragauskas, asist. V. Stancelis. Viena iš kurso paskirčių – apmokyti studentus dirbti su atitinkama programine įranga, supažindinti su pagrindine kompiuterinio duomenų kaupimo problematika, iš kitos pusės – išbandyti pirmąjį, „pilotinį“ parlamentarizmo istorijos DB variantą. Ketvertą mėnesių 10 studentų grupė svarstė ir rinkosi informacijos atrankos kriterijus, dalyvavo sudarant darbinį DB modelį, suvedinėjo sukauptą informaciją. Šios praktikos rezultatai leido patikrinti pirminės koncepcijos teisingumą, apytiksliai įvertinti būsimo darbo apimtis laike, rasti silpnąsias projekto vietas.

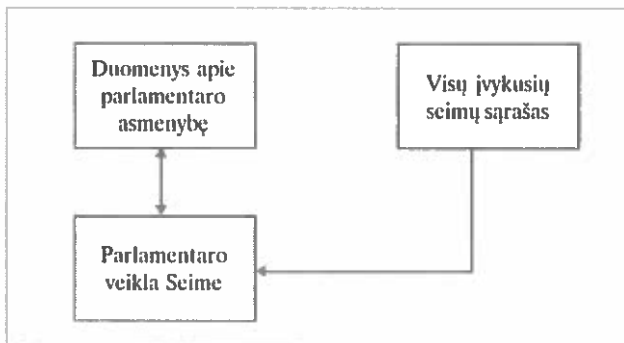
Išankstinis planavimas yra kritiškai svarbus duomenų bazės kūrimo etapas. Bet kuris kompiuterinių duomenų bazių vadovėlis pabrėžia, kad prieš pirmą kartą sėdant prie kompiuterio būtina pasėdėti bent jau keletą valandų su paprastu pieštuku ir popieriumi. Reikėtų aiškiai apibrėžti, kokie duomenys būtini, kokie antraciliai, kaip jie siejasi tarpusavyje, kaip siejasi tarpusavyje skirtingos lentelės, ar jose be reikalo nesidubliuos duomenys, ne kartą iš

naujo nusibraizyti duomenų bazės struktūrą, taip eliminuojant klaidas ir padarant būsimą projektą kiek įmanoma optimalesnį. Puiku, jeigu tai galima daryti kolektyviai, dalyvaujant visai projekto komandai – darbo vadovams, tyrinėtojams, ir bent daliai personalo, kuris atliks techninį, „juodąjį“ duomenų įvedimo darbą. Paprastą tekstą taisyti ir pertvarkyti galima bet kuriuo momentu, tuo tarpu duomenų bazė tokios prabangos nesuteikia – jeigu duomenys bus surinkti ir įvesti klaidingai, vėliau juos pataisyti bus labai sunku, arba visai neįmanoma.

Be to, duomenų bazės projektavimas reikalauja ne tik kompiuterinių disciplinų išmanymo, bet ir tam tikrų organizacinių gebėjimų, tvarkant panaudotus materialinius ir žmogiškuosius resursus: kompiuterinės darbo vietos, patalpos, darbuotojų skaičius ir darbo laiko paskirstymas ir t.t. Imantis didesnės apimties projekto, rekomenduotina skirstyti būsimus darbus į atskiras, pakankamai savarankiškas fazes, kurios užsibaigtų tam tikro, tegul ir tarpinio, rezultato pateikimu.

Dar prieš pradėdant konkretų parlamentarizmo DB planavimą, net nesileidžiant į konkrečias technologines specifikacijas, buvo reikalinga pirmiausia apsvarstyti šiuos klausimus:

1) Kokie duomenys reikalingi būsimam tyrimui;



1 pav. Principinė parlamentarizmo DB struktūra

Nuspręsta, kad informacijos minimumas, kurį būtina surinkti apie kiekvieną parlamentarą, yra jo vardas ir pavardė, kuriuose Seimuose ir kuriais metais jis dalyvavo, o apie naujųjų laikų parlamentarus – papildomai jų gimimo data, partinė priklausomybė. Be abejoj, tokia minimalistinė informacija tikrai būtų mažai naudinga platesnio tyrimo kontekste, todėl, esant galimybei, kiekvieno parlamentaro įrašas bus papildomas kitais anketiniais duomenimis: žiniomis apie jo socialinį statusą, politinę veiklą Seime ir už jo ribų.

2) Kokia turės būti surinktos informacijos apimtis;

Pirminiais duomenimis, reikės surinkti duomenis apie 3 000-4 000 XVI-XXI a. pradžios parlamentarus. Tačiau, net konstruojant siauros apimties, tar-kime, 10 skirtingų laukų (vardas, pavardė, partija), DB teks užpildyti virš 30 000-40 000 informacijos pozicijų.

3) Ar pakaks atitinkamos kompetencijos personalo duomenų kaupimui ir apdorojimui;

Dalis Lietuvos istorijos katedroje dirbančio personalo gali būti apmokyti pagrindinių duomenų įvedimo bei pirminės analizės veiksmų, be to, tiek duomenų rinkimui, tiek įvedimui gali būti pasitelkti vyresnių kursų IF studentai. Personalo apmokymo problema neturėtų būti labai aštri, kadangi dauguma atliekamų operacijų bus standartinės, rutininės procedūros, kurias išmokti nereikalauja daug laiko ir pastangų. Duomenų bazės įvedimo operatoriaus kvalifikacijai pilnai pakanka reikalavimų, kurie apibrėžti Lietuvoje visuotinai diegiamoje ECDL kompiuterinio raštingumo nuostatuose¹⁰⁰.

4) Kaip surinkti duomenys bus panaudoti pagrindinio Lietuvos istorijos katedros projekto rėmuose;

Pirminiu etapu surinkti duomenys bus popieriuje, jie bus padalyti galutinės parlamentarų biogramas rašysiantiems kolegoms, vėliau – svarstyti galimybė kruopščiau apdorotą parlamentarų DB išleisti kompaktinės plokštelės formatu.

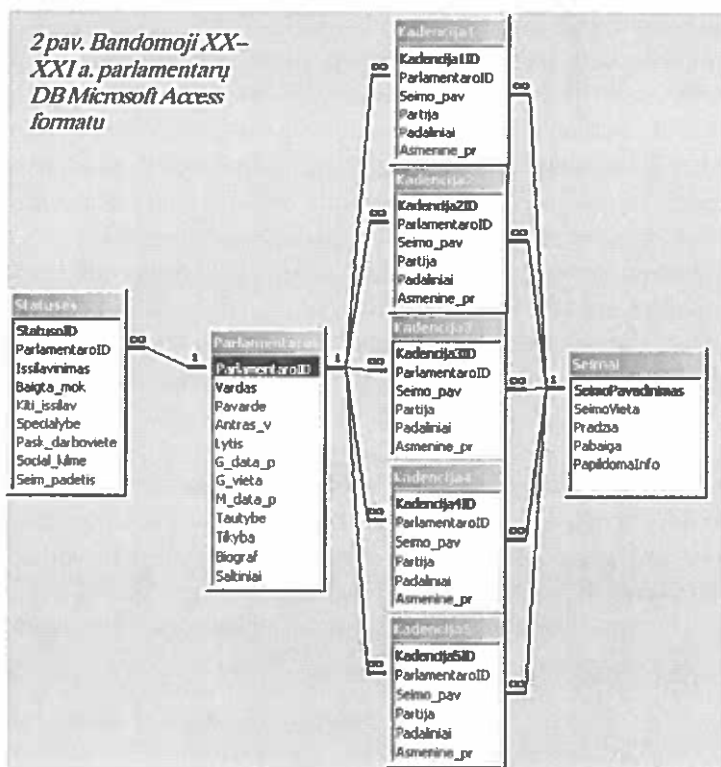
5) Kaip bus užtikrinta informacijos apsauga nuo praradimo.

Aktyvaus duomenų įvedimo laikotarpiu periodiškai bus daromos DB ko-

¹⁰⁰ ECDL programa: 3.0 versija // Lietuvos kompiuterininkų sąjunga, Informacinių technologijų institutas. Vilnius, 2001.

pijos tiek į susietus tinklo kompiuterius, tiek ir nuolatinio saugojimo laikmenas (CD-ROM).

Praktinis įgyvendinimas. Aptarus aukščiau paminėtus klausimus, galima pereiti prie nuoseklaus duomenų atrinkimo kriterijų svarstymo. Turint omenyje labai plačias parlamentarizmo projekto chronologines ribas, akivaizdu, kad praktiškai neįmanoma sutalpinti vienoje ir tose pačiose lentelėse žinias apie Abiejų Tautų Respublikos periodo ir XXI a. pradžios parlamentų išsilavinimą, politines ir socialines pažiūras ir pan. Dėl to iš karto buvo nuspręsta, kad DB bus suskaidyta į dvi gana savarankiškas dalis: pirma skirta XVI–XVIII a., o antra – XX–XXI a. Bandomajame projekte, atliktame su magistratūros studentais, siekiant patogumo buvo pasirinktas dviejų visiškai nepriklausomų ir nesusietų DB kelias, kadangi darbo tikslas buvo daugiau problemos analizė ir silpnųjų vietų paieška. Tuo tarpu galutiniame variante planuojama informaciją kaupti vienoje ir toje pačioje DB, tiesiog nesuderinamą medžiagą iškeliant į skirtingas lenteles.



Pateiktame XX–XXI a. DB modelyje atsispindi panaudotų lentelių visu-
ma, ryšiai tarp jų ir visi laukai, į kuriuos bus įvedama informacija. XVI–
XVIII a. DB išlaikyta analogiška struktūra, tik su atitinkamam laikmečiui
pritaikytais atskirų lentelių laukais: pvz., atsisakyta tokių laukų kaip lytis,
specialybė, socialinė kilmė, papildomai įvesti giminystės ryšiai ir pan. Kaip
matome, pagrindinė informacija apie parlamentarą talpinama lentelėje „Par-
lamentaras“, juos papildantys, savotiškai antračiliai anketiniai duomenys –
lentelėje „Statusas“. Visiems parlamentarams bendras įvykusių seimų sąrašas
saugomas lentelėje „Seimas“. Politinė parlamentaro veikla aprašoma len-
telėse „Kadencija1“, „Kadencija2“ ir t.t.

Sudarinėjant lentelių laukus teko atsižvelgti į nemažai „nemalonių“ mo-
mentų: pirmiausia, netgi lengvai prieinama medžiaga ne visada gali būti
aprašyta standartizuotų lentelių pavidalu. Svarbu iš karto suvokti, kad net
labai kruopščiai remiantis dokumentine medžiaga, sukurta DB jau nebus
istorinis šaltinis. Kompiuteris negali atspindėti visų originalaus dokumento
savybių ir tobula 32 bitų spalvinės kokybės skaitmeninė nuotrauka negalės
perteikti kiekvieno niuanso – popieriaus tipo, būklės, dydžio, tekstūros ir t.t.
Tačiau ši problema nėra jokia naujiena istoriko darbe – tyrinėtojams visuo-
met tekdavo išsirinkti tik pačią esmingiausią šaltinio informaciją, nesvarbu,
kokio amžiaus technologijomis buvo naudojamosi. Kita nemaloni dilema
„humanitaro“ sielai – kaip tokią skirtingą ir įvairialypę medžiagą – laisvo
stiliaus prisiminimus, epistolinį palikimą, metaforišką senųjų dokumentų
kalbą įsprausti į standartizuotas kompiuteriu apdorojamas lenteles.

Dažnai besikartojanti situacija – šaltiniai iš esmės gali būti įvedami į stan-
dartinės struktūros lenteles, tačiau esama elementų, kurie ten niekaip nedera,
ir visa problema, kad šie elementai svarbūs būsimam tyrimui. Vienas iš galimų
sprendimo kelių – vis dėlto pervesti ir šią informaciją į skaitmeninį formatą,
tačiau palikti už tipinės DB ribų. Kuriant bet kurią DB tenka sąmoningai
atmesti dalį potencialiai naudingos informacijos, nes pernelyg platus klausi-
mų ratas padarytų ją sunkiai techniškai įgyvendinama. Dėl to būtina išsirinkti
tik svarbiausius, esminius dalykus, kurie yra kritiškai svarbūs, tokius punktus,
kurie būdingi praktiškai kiekvienam DB aprašomam objektui. Kita vertus,
neapgalvotai siaura DB problematika ves prie labai ribotos potencialių varto-
tojų apimties, nors pati DB būtų ir nepriekaištingai atlikta iš techninės pusės.

Tarkime, DB sudaryta tik iš 5 laukų, pvz.: „Vardas“, „Pavardė“, „Gimimo data“, „Partija“, „Tautybė“, taip pat galima, tačiau akivaizdu, kad šių kriterijų pagrindu sukaupia informacija galės būti panaudota tik labai ribotos apimties studijoms. Todėl turint priėjimą prie platesnės informacijos, verta į DB įtraukti ir aprašyti daugiau laukų, negu yra būtina konkrečiam, šiuo metu vykdomam projektui, taip suteikiant platesnę erdvę panaudojimui vėlesniuose tyrimuose ar dalijantis medžiaga su kolegomis. Kita vertus, DB su pertekliniu laukelių skaičiumi – ypač nepatogi ją pildant. Jeigu rezervuosime laukelius kokiai nors retai pasitaikančiai informacijai, sąlyginai pavadinkime tokį laukelį, pvz. „Ar buvo apdovanotas III laipsnio Vyčio kryžiumi“. Šia informacijos laikmena bus naudojama itin retai, bet dėl to, kad ji turės būti privalomai įtraukiama kiekvieno asmens byloje, nepateisinamai sunkins tiek duomenų įvedimą, tiek analizę.

Antra, dėl patogumo verta iš anksto numatyti, kurie atsakymų variantai pasirinktuose laukeliuose bus iš esmės unikalūs, o kurie nuolat kartosis. Taip, tarkime, laukai „Pavardė“, „Gimimo vieta“, jei nekreipsime dėmesio į retas išimtis, kiekviename naujame įrašė talpins skirtingą informaciją, tuo tarpu laukai „Lytis“, „Tautybė“, „Partija“ dažnai kartos identiškus atsakymus. Microsoft Access programa sudaro patogią galimybę iš anksto užprogramavus pasikartojančius atsakymus juos vėliau įvedinėti vieninteliu „pelės klaptelėjimu“. Tačiau šis akivaizdus patogumas bus deramai išnaudotas tik tada, kai galimi atsakymų variantai bus galutinai apgalvoti, tikslūs, apimančys visą įmanomų pasirinkimų spektrą – kitaip nepakankamai kruopščiai iš anksto numatytų pasirinkimų ratas tik siaurins informacijos įvedimą, per platus – įvedimo operatorius per daug laiko užtruks kiekvieną kartą turėdamas išsirinkti „teisingiausią“ alternatyvą iš per ilgo sąrašo.

Trečia, popieriuje turima ar tiesiog dabar surasta informacija retai kada gali būti tiesiogiai įvedinėjama į DB be tarpinės, pirminės analizės. Konkretus asmuo gali būti baigęs keletą mokymo įstaigų, įgijęs skirtingas specialybes, dirbęs daugelyje skirtingų darboviečių, tuo tarpu mašinos „logika“ reikalauja griežto ir tiesmuko atsakymo: kokia mokykla baigta ir kokia specialybė įgyta. Dėl to, jeigu duomenis suvedinėja net ir juos rinkęs asmuo, o tuo labiau – kitas, su tirama medžiaga mažai arba visai nesusijęs asmuo (žinoma, dirbant su istorinėmis DB tai nerekomenduotina praktika), reikia laiko

duomenų visumos suvokimui, prioritetinės informacijos nustatymui ir išrūšiavimui, o tik po turi sekti grynai technologinis įvedimo procesas.

Šis darbas reikalauja daug laiko, todėl gerokai lėtina bendrą darbo procesą. Štai bandomąjį XX–XXI a. parlamentarų DB iš viso sudarė 55 laukai. Žinoma, tai teorinis žinių apie vieną asmenį maksimumas, kadangi ne apie kiekvieną parlamentarą turime patikimus duomenis: socialinę kilmę, baigtas lavinimo įstaigas, religines pažiūras ir pan., todėl vidutiniškai vieno asmens įrašas būdavo sudarytas iš 20-25 laukų. Praktika parodė, kad ir tokio sutrumpinto įrašo užpildymas, turint omenyje dar ir prioritetinės informacijos atranką įvedimo proceso eigoje, užima 25-30 min., tai yra turintys palyginti neblogus darbo su kompiuteriu įgūdžius magistratūros studentai per vieną užsiėmimą (2 akad. valandas) suspėdavo įvesti informaciją tik apie 3, daugiausia 4 asmenis.

Remiantis bandymo duomenimis, nesunku prognozuoti būsimo parlamentarizmo DB darbo apimtis: ~ 4 000 x 30 min. ≈ 2 000 darbo valandų. Vadinasi, gauname kasdieninį 5 žmonių grupės 8 darbo valandų trukmės darbą, turėsiantį tęstis maždaug 8 savaites, žinoma, apsidraudžiant reikalinga daryti prielaidą apie tam tikrą, tarkime ± 25 proc. paklaidą – skirtingo sudėtingumo duomenys, nevienoda darbuotojų kvalifikacija. Svarbu, kad visas šis laikas – tik techniniam, rutininiam įvedimui, o dar reikia turėti omenyje duomenų paiešką archyvuose ir literatūroje, klaidų patikrą, vėlesnę įvestos informacijos analizę ir apdorojimą iki galutinio varianto, patogaus bet kuriam šia tema susidomėjusiam vartotojui.

Siekiant mažiau įgudusį vartotoją apsaugoti nuo potencialių klaidų, Microsoft Access, kaip ir kitos šiuolaikinės DB, naudoja įvedimo ir peržiūros formas. Žemiau (p. 189) (pav. 3) pateiktas XVI–XVIII a. periodo parlamentaro asmens duomenų įvedimo skydelis. Kaip ir pridera bandomajam, eksperimentiniam variantui, jis nepasižymi išskirtiniu išbaigtumu, tuo tarpu galutiniame DB variante įvedimo laukai bus papildyti vartotoją informuojančiais užrašais, kurie patars, kaip tiksliau formuluoti įvedamą tekstą, kurios santrumpos leistinos naudoti ir pan.

Parlamentaras

Vardas: Jan Antras vardas

Pavardė: Sokolniski Gimimo data: Marties data: 1630

Gimimo vieta

Giminytės ryšiai: Sokolniskio Jontušo - Polocko pakamarininko pusbrolis

Tautybė: lenkų Tikyba: katalikų Papildoma tikybinė inf.

Biografija
1608m. pasirašė paktą tarp Abiejų Tautų Respublikos ir Maskvos. 1627 m. dalyvavo diskusijose dėl piniginių reformos.

Saltiniai: Nagleski P. Sokolniski Jan. //PSB,2000, T.XL/1, S.164, P.35 - 37.

Record: 3 of 6

3 pav. Bajoriškosios demokratijos laikotarpio DB duomenų įvedimo forma

Tokio formato ir pagaliau užpildyta DB įgys pavidalą, tinkamą vidiniam, darbiniam naudojimui, leidžiančią visą surinktą apie konkretų asmenį informaciją išspausdinti atskiros kortelės pavidalu. Tačiau norint galutinai pritaikyti DB platesniam vartotojų ratui, bus sudarytos iš anksto numatytos užklausos (Queries), atspindinčios dažniausiai tyrinėtojus dominančius klausimus – tautinė, socialinė, politinė parlamentarų priklausomybė, statistiniai duomenys apie išsilavinimą, politinį aktyvumą ir t.t. Siekiant galutinius duomenis atsieti nuo jų apdorojimo priemonės, šiuo atveju MS Access, užklausomis atrinkta informacija bus pervesta į kiek įmanoma plačiau prieinamą formatą: delimituotą tekstą, HTML formato lenteles. Taip suinteresuotas vartotojas galės naudotis atlikto tyrimo rezultatais, nepriklausomai nuo turimos kompiuterinės įrangos galimybių.

Išvados

Turint omenyje projekto komplikuoumą, sunku tikėtis iš karto tobulo rezultato, tačiau pagrindiniai tikslai – pakankamai detalaus Lietuvos parlamentarų sąrašo sudarymas, jo panaudojimas bendrame platesnio kolegų rato darbe, o vėliau – kaip atskiro, savarankiško ir patogaus apdorojimui informacijos šaltinio platinimas yra neabejotinai pasiekiami.

1. Lietuvos parlamentarizmo istorijos duomenų bazės projektas prieš pradedant parlamentarų žodyno rašymą leidžia preliminariai išanalizuoti tyrinėtojų jau turimą medžiagą ir nustatyti kriterijus, pagal kuriuos toliau reikės ieškoti ir kaupti trūkstamą informaciją.

2. Kompiuterinės DB panaudojimas gali žymiai palengvinti informacijos apsikėtimą ir pasiskirstymą vaidmenimis (informacijos rinkėjas, įvedimo operatorius, galutinių tekstų rašytojas) tarp bendradarbiaujančių kolegų.

3. Į kompiuterinės DB parengimo darbus planuojama įtraukti vyresnių kursų studentų, o tai turės nemažą edukacinį aspektą – kaip supažindinimo su reikšmingo Lietuvos istorijos aspekto problematika, taip ir kompiuterinių technologijų istorijos moksle propagavimo prasme.

4. Duomenų prieinamumas ne tik tiesioginiams projekto dalyviams, bet ir jų publikacija platesniam tyrinėtojų ratui prieinama forma sudarytų galimybes panaudoti surinktas žinias kitų Lietuvos istorijos tyrimų kontekste.

5. Naujų informacinių technologijų panaudojimas istorijos tiriamajame darbe kol kas tebėra nepakankamai perprasta sfera, todėl šis DB kūrimo bandymas gali būti naudingas ir kaip naujos metodikos kūrimo ir bandymo dirva.

USE OF COMPUTERISED DATABASE FOR INVESTIGATIONS INTO THE HISTORY OF PARLIAMENTARIANISM OF LITHUANIA

VIGINTAS STANCELIS

Summary

Having in mind complexity of the project it is difficult to expect to receive a perfect result at once, however, the main objectives set – compiling a sufficiently detailed list of the parliamentarians of Lithuania, its use in joint work of a wide circle of colleagues, and later, distributing it as a separate, independent source of information, which is convenient to process – are possible to be attained without doubt. The project of the database of the history of parliamentarianism of Lithuania, prior to starting compiling the Dictionary of Parliamentarians, enables us to preliminary analyse the material collected by scholars and to reject the criteria on the basis of which lacking information shall be sought and stored. The use of the computerised database can greatly facilitate the process of exchanging information and distributing the roles (a collector of information, an input operator, a writer of final texts) between the co-operating colleagues. It is planned to engage graduate students in the pre-

paratory work of the computerised database, which will have a significant educational aspect – both in the sense of acquainting them with the problems of a significant aspect of the history of Lithuania and from the point of view of propagating computer technologies in historical sciences. Accessibility of the data not only for direct participants in the project but also their publication in the form accessible for a wider circle of investigators would raise possibilities to use accumulated knowledge within the context of other investigations into the history of Lithuania.

The use of new information technologies in the research work of history has been an insufficiently understood sphere thus far; therefore this attempt at creating the database (hereinafter referred to as the DB) can be useful as the sphere of creating and testing new methodology. One of the main trends in the work of the Department of History of Lithuania of Vilnius Pedagogical University for the immediate five years is compiling the

Dictionary of Parliamentarians of Lithuania consisting of four volumes. A large scope of work, the necessity to share work and exchange information, the standardised requirements for the end product, when it is necessary to create thousands of biograms of the same structure – the date of birth, the party affiliation, election to the Seimas, the political activity within it – naturally raise the idea about the expediency of the use of the computerise database.

Attempts will be made to achieve that computer data should be collected and presented to other colleagues in an accessible format as early as possible so that on the basis of the available entries it would be possible to compile final texts of biograms of the parliamentarians. Historical information recorded in a digital format has a potentially great and long-term value to the community of investigators and pedagogues, however, rationally it will be used only if not only the person who directly stores the data but also other interested users will have access to it.

Foreign experience. The use of the DB in Western Europe and the United States of America in historical investigations has long-established traditions. However, the concept of the database used in these countries is rather undefined, often the same term is applied to both the simplest lists of some officials and holdings and strictly standardised electronic tables suited for the computerised analysis – what

specialists in informatics consider as databases. In this article, however, only the second “real” DB variant will be considered.

In reviewing the work done by the historians of foreign countries, the greatest attention had to be paid to the material published in the English language. Though the USA has done tremendous work in applying the DB to historical investigations, for example, over 60 000 working hours¹ were spent on creating the database of the Civil War of America managed by Historical Data Systems, the investigations carried out in Great Britain and methodologies publicised have to be treated as the most useful in the sense of the history of parliamentarianism. This is determined by the very specificity of investigations into the history of the USA where practically all attention is focused on the modern times, whereas British scholars cover a much wider chronological and problematic space.

Special mention must be made of Essex University in Great Britain, which is a peculiar centre for computer innovations in historical science. The Database² of Great Britain devoted to statistical investigations of the 19th and the 20th centuries in which efforts of not only the University of Essex but also those of the Universities of Aberdeen, Cambridge, Staffordshire are united, operates there. Another division – the History Data Service – works in issuing recommendations and methodological advice on how to design and exploit the new DB.

The Joint Committee on Information Systems³, which established it, defined the exact and convenient criteria for use on the basis of which digital resources of the historical science could be collected, catalogued, managed, stored and maintained. They can be adapted to many different levels of processing data, from student investigations to the research projects on a larger scale. The main requirement is to store the data in the form of standardised tables, which later will be used in computerised databases, calculators or statistical programs.

A lot has been done in converting historical information into an electronic format and adapting it to systematic computerised processing in the Scandinavian countries. The archive of the city of Stockholm publishes some part of its material on the Internet⁴, the Data Archive of Denmark (Dansk Data Arkiv⁵, a special division of the state archive of Denmark) stores the bank of digital data covering the period from the 16th to the 20th centuries. Higher education schools of Norway work in the same direction – at Universities of Oslo⁶ and Bergen⁷. Russian historians are also starting to work more intensively in this sphere, and not necessarily in large scientific research centres – let us take, for example, the databases prepared at the University of Tver: „Тверская усадьба. Век за веком“, „Тверское дворянство. Гербы“⁸.

Preparatory works. The creation of digital resources on the basis of his-

torical documents or research material as an independent, valuable research project is not a common thing, most often this work is undertaken as a constituent part of some larger research program rendering it an auxiliary, additional role. On the other hand, storage of resources as a separate task, which is not related to research, generalising work can become an ambitious occupation, which does not pay or justify itself. Therefore, following the principle of an integral link between storage of information and its processing, the work of creating the database of the history of parliamentarianism of Lithuania has been commenced, intending in advance to integrate it into the program of compiling the Dictionary of Parliamentarians of Lithuania carried out by the Department of History of Vilnius Pedagogical University.

Starting compiling an accomplished “perfect” DB by basing oneself in essence on the resources of the staff of the Department of History alone would not sound as a well-considered idea. Therefore a lot of preparatory work has been undertaken whereby it was sought to check technical and human possibilities, to acquaint the colleagues with the main working methods, to carry out the first tests thus elucidating the problems and perspectives of further work.

1) Preliminary planning is a critically important stage of the creation of the database.

According to the preliminary data, data about 3 000-4 000 parliamentarians of the beginning of the 16th-21st centuries will have to be collected. Hence, even in case the database of a narrow scope, for example, containing 10 different fields (a name, a surname, the party) is constructed, 30 000-40 000 information positions will have to be filled in.

Will the personnel of relevant competence be available for storing and processing the data;

2) In what way will the collected data be used within the framework of the main project of the Department of History;

The data collected during the initial stage will be on paper, they will be given to the colleagues who will write final biograms of the parliamentarians, later the possibility to release a thoroughly processed DB in the form of a compact disc will be considered.

3) In what way will information be protected against losing it.

During the period of active input of information, copies of the DB both in the related network computers and main storage (CDR) will be made.

The work takes plenty of time therefore a general process of work slows down. For example, the experimental DB of the parliamentarians of the 20th - 21st centuries consisted of 55 fields. Of course, this is only a theoretical maximum of knowledge about a single person because reliable data are not available about every parliamentarian:

his social background, institutions of general education, his religious views, etc., therefore on average the entry of a single person consisted of 20-25 fields. Filling in this shortened entry takes 25-30 minutes, which means, that students of the master's course who have comparatively good computer skills managed to enter information only about three or four persons at the most during one class (two academic hours).

On the basis of the data of the experiment, it is easy to forecast the future work volumes of creating the DB of parliamentarianism: ~ 4 000 x 30 min. = 2 000 working hours. Hence, we have daily eight-hour work of a group of 5 people, which will have to last approximately 8 weeks. Of course, one must make an assumption about a certain, for example, ± 25 per cent error - data of different complexity, different qualification of workers. The important thing is that all this time is devoted only to technical, routine input, however, one also must have in mind a search for data in archives and literature, verification of errors, a later analysis and processing of information entered to have the final variant, convenient to any user who might become interested in this topic. Conclusions:

1. The project for the database of the history of parliamentarianism in Lithuania, prior to starting compiling the Dictionary of Parliamentarians, enables the investigators to make a preliminary analysis of the material available and to establish criteria on

the basis of which missing information will be further sought and stored.

2. The use of the computerised DB can considerably facilitate exchange of information and distribution of roles (a collector of information, an input operator, a writer of final texts) between the co-operating colleagues.

3. It is planned to engage graduate students in the preparatory work of the DB, and this will have a significant educational effect – both in the sense of acquainting them with the problems of a significant aspect of the history of Lithuania and from the point of view of propagating computer technologies

in historical sciences.

4. Accessibility of the data not only for direct participants in the project but also their publication in the form accessible for a wider circle of investigators would raise possibilities to use accumulated knowledge within the context of investigations into the history of Lithuania.

The use of new information technologies in the research work of history is the sphere that has not been understood to the full thus far, therefore this attempt at creating the DB can be useful as the sphere of creating and testing new methodology.