

KOMPIUTERINĖS DUOMENŲ BAZĖS PANAUDOJIMAS LIETUVOS PARLAMENTARIZMO ISTORIJOJE 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 pradėdamas rašyti parlamentarų žodyną 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 apsikeitimą tarp tyrinėtojų ir sudaryti efektyvesnes sąlygas pasiskirstyti vaidmenimis (informacijos rinkėjas, įvedimo operatorius, galutinių tekstų rašytojas) tarp bendradarbiaujančių kolegų. Į kompiuterinės DB parengimo darbus galima įtraukti tam tikrą vyresniųjų kursų studentų skaičių, tai turės nemažą edukacinį aspektą – kaip supažindinimo su reikšmingo Lietuvos istorijos aspekto problematika, taip ir kompiuterinių technologijų istorijos moksle 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.

Esminiai žodžiai: parlamentarizmas; informacinės technologijos; duomenų bazės; programinė įranga; lentelės; užklausa.

Įvadas

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 tapačios struktūros duomenų šaltiniais, kaip kad parlamentarų biografiniai duomenys, kompiuterinių technologijų panaudojimo tikslingumas tampa ypač aktualus.

Viena vertus, iš anksto nustatytus bendrus 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 jau parengtas „lentynėles“. Tokia iš anksto struktūrizuota paieška paprastai būna kur kas greitesnė ir tikslesnė už laisvą faktų medžioklę.

Kita vertus, jau pačioje DB tai leistų lengvai atrinkti rūpimus faktus iš didelio pirminės informacijos kiekio, atlikti statistinės analizės ir kitus apibendrinančius veiksmus.

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

Tikslai: sudaryti kompiuterinę duomenų bazę, apimančią informaciją apie visus XV–XXI a. Lietuvos parlamentarus. Pritaikyti ją masiniam istorija besidomintiems vartotojui, neturintiems specializuoto kompiuterinio pasirengimo. Duomenų bazę paskelbti kompaktinės plokštelės bei internetu pasiekiamos informacinio resurso forma.

Tyrimo metodai: sukurti standartizuotą klausimyną, pagal kurį tyrinėtojai rinktų 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 kryptį artimiausiems keleriems metams – keturių tomų Lietuvos parlamentarų žodyno sudarymas. Numatoma surinkti kiek galima detalesnę 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 pasitelkti kitur dirbantys kolegos. Didelė darbo apimtis, būtinybė dalytis 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 – savaime kelia mintį apie kompiuterinės duomenų bazės panaudojimo tikslingumą.

Kadangi kuriama DB – sudėtinė didesnio viso Lietuvos istorijos katedros darbo proceso dalis, 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. Todėl šio straipsnio užduotis – ne tiek atsiskaityti apie pirmąjį

parlamentarizmo istorijos DB projektavimo etapą, nes projektas dar tik įsibėgėja, bet ir išsamiau 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 valandų^[1], 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 naujovių istorijos moksle centras. Jame veikia Didžiosios Britanijos duomenų bazė^[2], skirta XIX ir XX a. statistiniams tyrimams. Apjungtos ne tik Esekso, bet ir Aberdyno, Kembridžo, Stafordšyro universitetų pastangos. Kitas padalinys – Istorinių duomenų tarnyba (History Data Service) – leidžia rekomendacijas ir metodinius patarimus, kaip projektuoti ir eksploatuoti naujas DB. Ją įsteigęs Jungtinis informacinių sistemų komitetas^[3] apibrėžė gana tikslus ir patogius vartojimo kriterijus, kuriais remiantis būtų renkami, kataloguojami, valdomi, saugomi ir palaikomi skaitmeniniai istorijos mokslo resursai. Jie gali būti pritaikyti daugeliui skirtingų duomenų apdorojimo lygmenų – nuo studentišκών tyrimų iki plataus masto tiriamųjų projektų. Pagrindinis reikalavimas – duomenis kaupti standartizuotų lentelių forma, 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 skelbia internete^[4], Danijos Duomenų archyvas (Dansk Data Arkiv^[5], specialus Danijos valstybinio archyvo padalinys) kaupia skaitmeninių duomenų banką, apimantį XVI iki XX a. Ta pačia kryptimi dirbama ir Norvegijos aukštosiose mokyklose – Oslo^[6], Bergeno universitetuose^[7]. Pradeda aktyviau dirbti šioje srityje ir Rusijos istorikai, ir nebūtinai vien didžiuosiuose mokslo centruose – paimkime 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 *lauku*, 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*, ji atitinka visą turimą DB informaciją apie atskirą 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 DB 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: pirmiausia 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 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 kalba^[9].

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 vartotojų 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, iš esmės patenkina šiam darbui keliamus reikalavimus.

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 pagalbinių, 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 tolesnio darbo perspektyvas.

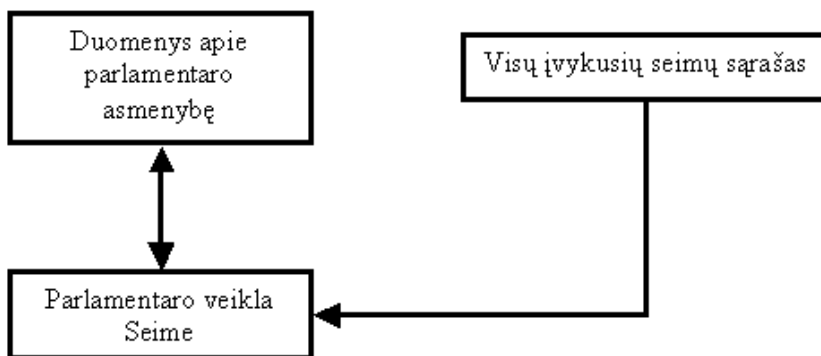
2003–2004 m. antrame semestre istorijos specialybės magistrantams buvo pradėtas dėstyti specialus kursas „Darbas su istorinių duomenų bazėmis“, sudarytojai – dr. doc. A. Ragauskas, asist. V. Stancelis. Viena iš kurso paskirčių – apmokyti studentus dirbti su atitinkama programine įranga, supažindinti su pagrindine kompiuterinio duomenų kaupimo problematika ir 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 apimtį laiko atžvilgiu, 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 antraeiliai, kaip jie siejasi tarpusavyje, kaip siejasi tarpusavyje skirtingos lentelės, ar jose be reikalo nesidubliuos duomenys, ne kartą iš naujo nusibraižyti duomenų bazės struktūrą, taip šalinant klaidas ir padarant būsimą projektą kiek įmanoma tinkamesnį. 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, o 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 kt. Imantis didesnės apimties projekto, rekomenduotina skirstyti būsimumus 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 būsimumam 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 abejo, 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 apimtys?

Pirminiais duomenimis, reikės surinkti duomenis apie 3 000–4 000 XVI–XXI a. pradžios parlamentarus. Taigi, net konstruojant siauros apimtį, tarkime, 10 skirtingų laukų (vardas, pavardė, partija), DB teks užpildyti virš 30 000–40 000 informacijos pozicijų.

3. Ar pakaks atitinkamos kompetencijos personalo duomenims kausti ir apdoroti?

Dalis Lietuvos istorijos katedroje dirbančio personalo gali būti apmokyti pagrindinių duomenų įvedimo bei pirminės analizės veiksnių, 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 išmokyti nereikalauja daug laiko ir pastangų. Duomenų bazės įvedimo operatoriaus kvalifikacijai

visiškai pakanka reikalavimų, kurie apibrėžti Lietuvoje visuotinai diegiamoje ECDL kompiuterinio raštingumo nuostatuose^[10].

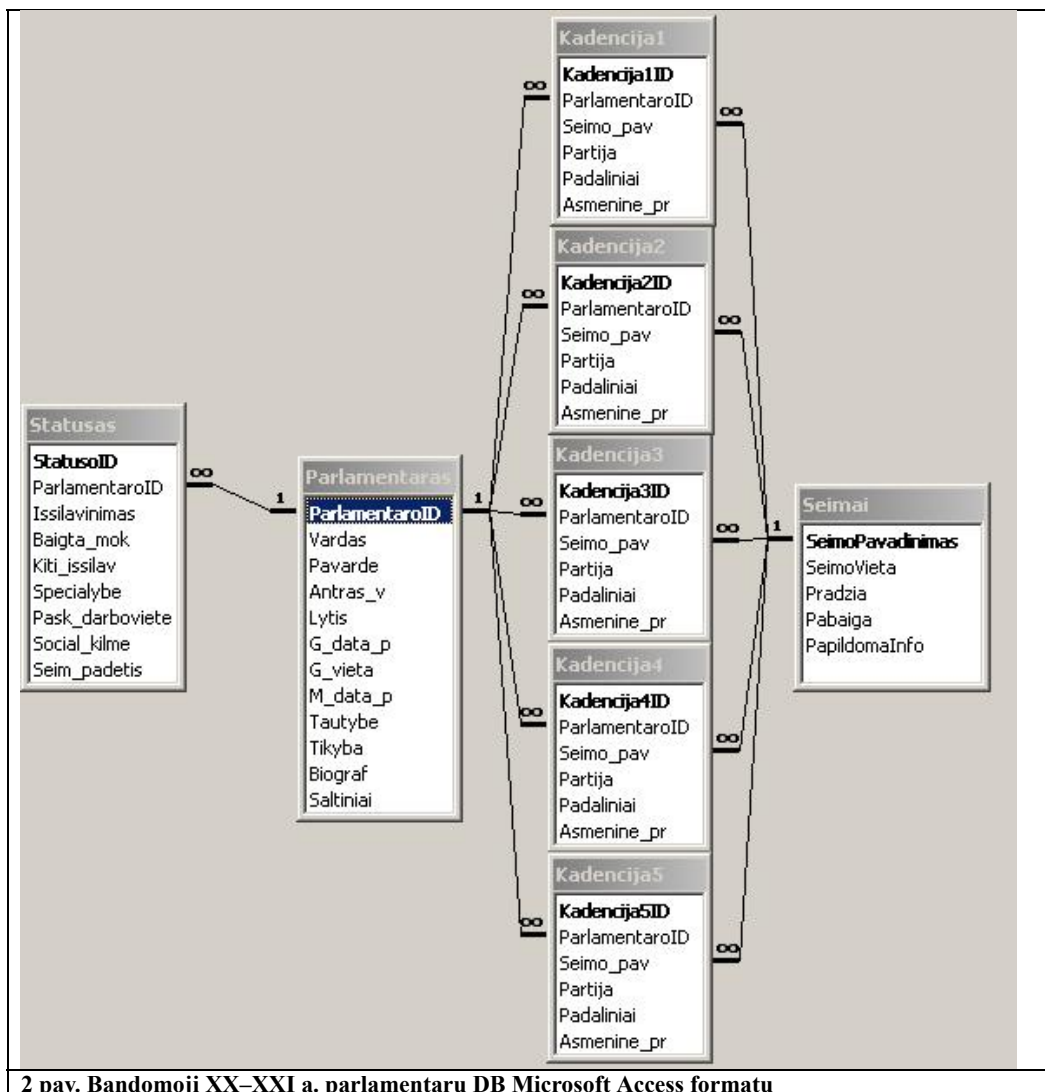
4. Kaip surinkti duomenys bus panaudoti pagrindinio Lietuvos istorijos katedros projekto rėmuose?

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

Praktinis įgyvendinimas. Aptarus minėtus 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 parlamentarų išsilavinimą, politines ir socialines pažiūras ir pan. Dėl to iš karto buvo nuspręsta, kad DB bus suskaidyta į dvi gana savarankiškas dalis: pirmą skirtą XVI–XVIII a., o antrą – XX–XXI a. Bandomajame projekte, atliktame su magistratūros studentais, siekiant patogumo buvo pasirinktas dviejų visiškai nepriklausomų ir nesusietų DB kelias, nes darbo tikslas buvo daugiau problemos analizė ir silpnųjų vietų paieška. O galutiniame variante planuojama informaciją kaupti vienoje DB, tiesiog nesuderinamą medžiagą iškeliant į skirtingas lenteles.



2 pav. Bandomoji XX–XXI a. parlamentarų DB Microsoft Access formatu

Pateiktame XX–XXI a. DB modelyje atsispindi panaudotų lentelių visuma, ryšiai tarp jų ir visi laukai, į kuriuos bus įvedama informacija. XVI–XVIII a. DB išlaikyta analogiška struktūra, tik su atitinkamam laikotarpiui 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 skelbiama lentelėje „Parlamentaras“, juos papildantys, savotiškai antraeiliai anketiniai duomenys – lentelėje „Statusas“. Visiems parlamentarams bendras įvykusių Seimų sąrašas saugomas lentelėje „Seimas“. Politinė parlamentaro veikla aprašoma lentelėse „Kadencija1“, „Kadencija2“ ir t.t.

Sudarinėjant lentelių laukus teko atsižvelgti į nemažai „nemalonių“ momentų: pirmiausia, netgi lengvai prieinama medžiaga ne visada gali būti aprašyta standartizuotuose lentelėse. 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 visuomet tekdavo išsirinkti tik pačią pagrindinę šaltinio informaciją, nesvarbu, kokio amžiaus technologijomis buvo naudojama. Kita nemaloni dilema „humanitaro“ sielai – kaip tokią skirtingą ir įvairialypę medžiagą – laisvo stiliaus

prisiminimus, epistolinį palikimą, metaforišką senųjų dokumentų kalbą įsprauti į standartizuotas kompiuteriu apdorojamas lenteles.

Dažna situacija – šaltiniai iš esmės gali būti įvesti į standartinės 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 klausimų ratas padarytų ją sunkiai techniškai įgyvendinamą. 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 potencialiam vartotojui bus labai ribotos apimties, nors pati DB būtų ir neprikaištingai atlikta iš techninės pusės.

Tarkime, DB sudaryta tik iš 5 laukų, pvz.: „Vardas“, „Pavardė“, „Gimimo data“, „Partija“, „Tautybė“, taip pat turi teisę egzistuoti, tačiau akivaizdu, kad šių kriterijų pagrindu sukaupta informacija galės būti panaudota tik labai ribotos apimties studijoms. Todėl turint priėjimą prie išsamesnės informacijos, dažnai 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 kokiais nors retai pasitaikančiais 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 sunks 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 bus skirtinga informacija, o 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, apimantys 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, tačiau 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 suvokti duomenų visumą, nustatyti ir išrūšiuoti prioritetinę informaciją, o tik po turi sekti grynai technologinis įvedimo procesas.

Šis darbas reikalauja daug laiko, todėl gerokai lėtina bendrą darbo procesą. Štai bandomąją 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 (3 pav.) 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, kokios santrumpos leistinos naudoti ir pan.

3 pav. Bajoriškosios demokratijos laikotarpio DB duomenų įvedimo forma

Taip apipavidalinta 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. Turint omenyje tikslą 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 savo turimos kompiuterinės įrangos galimybių.

Išvados

1. Turint omenyje projekto komplikuotumą, sunku tikėtis iš karto tobulo rezultato, tačiau pagrindiniai užsibrėžti 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.
2. Lietuvos parlamentarizmo istorijos duomenų bazės projektas prieš pradėdant rašyti parlamentarų žodyną leidžia preliminariai išanalizuoti tyrinėtojų jau turimą medžiagą ir nustatyti kriterijus, pagal kuriuos toliau reikės ieškoti ir kaupti trūkstamą informaciją.
3. Kompiuterinės DB panaudojimas gali žymiai palengvinti informacijos apsikėtimą ir pasiskirstymą vaidmenimis (informacijos rinkėjas, įvedimo operatorius, galutinių tekstų rašytojas) tarp bendradarbiaujančių kolegų.
4. 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.
5. Duomenų prieinamumas ne tik tiesioginiams projekto dalyviams, bet ir jų publikacija didesniai tyrinėtojų ratui prieinama forma sudarytų galimybes panaudoti surinktas žinias kitų Lietuvos istorijos tyrimų kontekste.
6. 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.

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

[2] The Great Britain Historical Database – <http://hds.essex.ac.uk/gbh.asp>

[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] Генеалогия и история тверских дворян (до 1890 года) – <http://library.tversu.ru/usadyba/dvoryane/poisk.htm>
- [9] Microsoft Access 2000: žingsnis po žingsnio: priemonė mokytis pačiam. Kaunas, 2001; Minkevičius, S. Microsoft Access XP pradžiamokslis: 8 pamokos. Vilnius, 2003; Vidžiūnas, A., Marčiulynienė, 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.
- [10] ECDL programa: 3.0 versija. Lietuvos kompiuterininkų sąjunga, Informacinių technologijų institutas. Vilnius, 2001.