

## TRYS SOCIOEKONOMINIŲ DUOMENŲ TAIKYMO ASPEKTAI

Vytautas Valatka, Andrius Balčiūnas

Vilniaus universitetas, Gamtos mokslų fakultetas, M. K. Čiurlionio 21/27, LT-03101, Vilnius

El. paštas: vytautas.valatka@gf.vu.lt, andrius.balciunas@gf.vu.lt

---

Valatka V., Balčiūnas A. THREE ASPECTS OF SOCIO-ECONOMIC DATA APPLICATION. *Geografijos metraštis* 45 (1), 2012.

**Abstract.** In the countries of Central and Eastern Europe, the development of quantitative and applied geography is an important issue for geographers and society as a whole. Usage of databases and positioning tools in a daily life creates opportunities for science, business and public administration. This paper reviews the best practices of high resolution socio-economic data usage. The historical review is grounded on 3 aspects of data: spatiality, diversity and temporality. A framework for the development of applied geography, based on triple helix concept and open source tools, was defined.

**Keywords:** small area data, geodemography, time geography, interactive geovisualization, open source.

Received:

---

### Įvadas

Taikomųjų geografinių tyrimų vykdymo svarbiausios prielaidos yra mokslininko gebėjimas naujai permąstyti tikrovę ir noras prisidėti prie jos perkūrimo. Siekiant taikyti tyrimų metu gautas žinias praktikoje yra svarbu, kad jos būtų detalios, išsamios ir aktualios. Tai yra įmanoma analizuojant ir kuriant atitinkamus duomenų rinkinius bei pasirenkant tinkamas analizės priemones. Programavimas ir geografinės informacijos mokslo žinios leidžia ruošti analizei, tirti ir interpretuoti anksčiau neįsivaizduotus duomenų kiekius. Pažymėtina tai, kad plėtojimosi tvarka – pirmiausia technologijos, po to metodologija ir galiausiai teorija – yra svarbi, nes geografinės informacijos mokslas yra indukcinė tyrimų sritis. Be to, geografiškai turi ribotą įtaką pagrindinių technologijų vystymui, netgi toms technologijoms, kurios savo prigimtimi yra iš esmės geografinės (Haklay, 2012). Aprašytos tvarkos paaiskinimas skatina geografus glaudžiai bendradarbiauti su informatikais. Dėl to geografiškai gali ieškoti naujų taikymų ir interpretuoti tyrimų rezultatus, o informatikai – taikyti naujus duomenų apdorojimo metodus. Metodologijos kūrimas tampa sąlyčio ir diskusijų erdve, o bendradarbiavimas leidžia kurti naudingus ir inovatyvius sprendimus. Išskirtinė programinė platforma bendram darbui – *Linux*, kurios stipriosios pusės yra nemokamos programavimo priemonės, mokslinė programinė įranga, duomenų bazės ir tinklo paslaugos.

Tarptautiniai ir neakademiniai ryšiai skatina mokslininkus dalyvauti sprendžiant visuomenės problemas ir įkvepia gyvybės mokslo ir tyrimų institucijoms. Europoje socioekonominių duomenų tvarkymo srityje toliausiai pažengę yra Jungtinės Karalystės geografiškai. Jie susitelkę Bristolio, Lydso, *University College of London* ir Šefildo universitetuose. Geografija čia yra itin stipri, ne mažiau kaip 70 % publikacijų yra vertinamos kaip aukščiausio tarptautinio lygio (RAE, 2008), o mokslininkai yra atviri bendradarbiavimui. Duomenų vizualizavimo srityje svarbūs yra 2010 m. Fraunhoferio instituto Vokietijoje vadovaujamos jungtinės 23 mokslininkų grupės pristatytos tyrimų gairės (Andrienko et al., 2010) ir Tarptautinės kartografų asociacijos geovizualizacijos komisijos darbas. Neakademiniai ryšiai šiuolaikiškiems universitetams yra būtini. Inovacijos vis labiau tampa priklausomos nuo trilypės spiralės (angl. *triple helix*): universitetų-verslo-valstybės institucijų sąveikos. Verslus universitetas užima aktyvią poziciją pritaikant žinias praktikoje ir

imasi vis naujų veiklos sričių (Etzkowitz, 2003). Mokslininkų aktyvumas yra būtinas, kad iš socialinių institucijų renkamų duomenų būtų gauta naujų taikomųjų žinių.

Per artimiausią dešimtmetį galime tikėtis reikšmingų poslinkių. Vystantis technologijoms ir keičiantis požiūriui į universitetų vaidmenį visuomenėje, aplinka taps palankesnė taikomiesiems geografiniams tyrimams. Mokslo institucijos turi išskirtines sąlygas bendrinti valstybės renkamus duomenis, sieti juos su kita informacija. Savo ruožtu toks įdirbis leidžia kurti naują bendravimo su verslo ir valstybės institucijomis kultūrą, kalbėti argumentuotai, per detalią erdvės ir laiko prizmę. Šiame straipsnyje aiškinamasi, ko reikia norint stiprinti akademinės geografijos sąsajas su verslu ir valstybės institucijomis. Pagal tris aspektus – detalumo, įvairovės ir laikiškumo – glaustai apžvelgiami socioekonominių duomenų tvarkymo geros praktikos pavyzdžiai. Aprašoma kiekvieno aspekto reikšmė ir pristatomos naujų taikymų Lietuvoje prielaidos. Straipsniu siekiama skatinti diskusijas apie taikomąją geografiją, jos specialistų poreikio darbo rinkoje klausimus.

## 1. Duomenų erdvinis detalumas

Rutiniškai renkamų, bet retai ir pagal stambias zonas arba rajonus pateikiamų duomenų naudojimas turėjo poveikį analizės būdams, taip pat lėmė kuriamų vaizdinių akiratį ir silpnynes. Žema erdvinė raiška ir reti atnaujinimai neleidžia efektyviai vaizduoti ir analizuoti procesų, slepia vidinį socialinių sąlygų heterogeniškumą rajonuose ir vaizdiniais būdingus neapibrėžtumus. Tokie suvaržymai yra nepriimtini didelio masto ir spartos pokyčių epochoje (Longley, 2003).

Erdviškai detalių duomenų naudojimas taikomiesiems tyrimams yra svarbus, nes leidžia vertinti lokalius teritorinius skirtumus. Greta abstrakčių esybių, tokių kaip apskritis, savivaldybė, aktualių valstybės valdyme galime turėti ir kasdienybėje, viešajame administravime svarbių, tokių kaip rinkimų apylinkė, kaimynystė (angl. *neighbourhood*), švietimo, prekybos, gydymo ar viešosios tvarkos palaikymo institucijų aptarnavimo arealai. Su erdviškai detalių duomenų kartoschemų pavyzdžiais galite susipažinti naudodami interaktyvias tinklapių *Maptube*, *LondonProfiler* ir *CensusProfiler* programas.

Lietuvoje geografiškai iki šiol nedirbo su visą šalies teritoriją apimančiais didelės erdvinės raiškos socioekonominiais duomenimis. Yra atlikti pavieniai seniūnijų lygmens tyrimai. Šiuose darbuose analizuojami bendriausi rodikliai – duomenys apie gyventojų, įmonių skaičių. V. Petrulis interpretuodamas Lietuvos politinį lauką (Petrulis, 2009) pasitelkė įvairesnius rodiklius (Valatka, 2010). Vis dėlto seniūnijų duomenys nėra pakankamai detalūs, neleidžia atspindėti daugelio visuomenės socialinės teritorinės struktūros aspektų ir tuo labiau lokalių procesų, todėl panašių tyrimų rezultatų taikomumas, ypač miestuose, yra ribotas. Taip pat trūksta ir sistemiško darbo. Tai įrodo faktas, kad po 2001 m. visuotinio gyventojų ir būsto surašymo (toliau – surašymas) trūko geografų iniciatyvos ir nėra parengta seniūnijų rodiklių duomenų bazė, reikalinga geografijos studijose ir ruošiant dar detalesnius duomenis.

Teritoriškai detalios analizės dalyje senųjų kapitalistinių šalių yra vykdomos keletą dešimtmečių – neabejotinai pirmauja anglosaksų šalys. Šių tyrimų priešistorė yra XIX a. pabaigos Londono tyrimai, vykdyti filantropo Ch. Booth (LSE, 2002). Jis siekdamas įrodyti, kad skurdas yra mažiau paplitęs, nei teigiama, 1886–1903 m. organizavo pastatų klasifikavimą ir kartografavimą pagal gyventojų sudėtį, tačiau įsitikino, jog situacija yra blogesnė, nei manyta. Remiantis šiuo darbu, yra vertinti ilgalaikiai pokyčiai ir nustatyta, kad Londono gyventojų geografija yra stebėtinai stabili (Orford et. al., 2002). Po Pirmojo pasaulinio karo, remiantis surašymo duomenimis, Čikagos universitete imti vykdyti surašymo apygardų (dydis < 10 tūkst. gyventojų, angl. *census tract*) tyrimai. 1960 m. tokie duomenys jau buvo parengti daugumai JAV miestų, o 1966 m. Jungtinėje Karalystėje Nacionalinė statistikos tarnyba pradėjo rengti mažų teritorijų statistikas (angl. *small area statistics*) (Vickers, 2006).

Erdviškai detalių duomenų analizės plačiai išpopuliarėjo dėl kiekybinių metodų taikymo lūžio geografijoje. Šio geografijos raidos etapo pagrindinės idėjos apibendrintos D. Harvey knygoje „Explanation in geography“ (1969). Pirmoji erdviškai detalių surašymo duomenų analizės metodika buvo sukurta Liverpulio miestui Anglijoje, 1976 m. Vėliau sekė Londono tyrimai. Galiausiai Nacionalinė statistikos tarnyba finansavo daugybę detalių visos šalies analizių 1977–1980 m.: pagal rinkimų teritorijas, surašymo apylinkes. 1978 m. erdviškai detalių duomenų klasifikacijomis susidomėjo Britų rinkos tyrimų biuras. Pradžioje naują informaciją taikė tam, kad įvertintų, kurių kaimynysčių gyventojus „atsitiktinės“ imtys reprezentuoja nepakankamai (Farr, Webber, 2001).

Po 1980 m. surašymo erdviškai detalių tyrimų srityje ilgą laiką dominavo komercinės, rinkodarai skirtos geodemografinės klasifikacijos. Taip pat ieškota sprendimo kintamojo plotinio vieneto problemai (angl. akron. *MAUP*) (Openshaw, 1984). Su ja bent iš dalies yra susidorota nuo 2000 m. Jungtinėje Karalystėje naudojamame detaliausių geografinių duomenų teikimo karkase (Rees et. al., 2002). Informacija teikiama ne pagal surašymo apylinkes, o pagal išvesties teritorijas (angl. *data output areas*, toliau – DIT), kurios yra sudarytos taip, kad pasižymėtų kuo homogeniškesniu užstatymu, ir vidutiniškai turi po 125 namų ūkius. Galiausiai 2006 m. parengta D. Vickers disertacija, kurioje klasifikuoti DIT duomenys. Tyrimo rezultatai yra skelbiami Nacionalinės statistikos tarnybos kaip oficialios statistikos, o metodika – atvira.

Vietoje DIT surašymo duomenų analizei Šiaurės Europos šalyse yra naudojama skirtingų dydžių erdvinė gardelė. Nors iš pirmo žvilgsnio gardelių metodas turi privalumų – teritorijų plotai vienodi, ribos nekinta bėgant laikui, tačiau išsamioms socioekonominėms analizėms jį taikyti yra sudėtingiau. Visų pirma, daug ląstelių, kuriose gyventojų skaičius yra mažas, ir analizuojant didelį kintamųjų kiekį yra susiduriama su asmeninių duomenų atskleidimo problema. Statistikos tarnybos linkę taikyti paprasčiausius konfidencialumo užtikrinimo metodus – naudoti  $1 \times 1$  km gardelę ir pateikti tik pagrindinius rodiklius. Antra, gardelės ribos neturi atitikmenų tikrovėje – siekiant tikslumo reikalingas mažas ląstelių dydis, bent  $100 \times 100$  m. Pažymėtina, kad 2011 m. bene visose ES šalyse, kuriose surašymas buvo vykdomas atliekant visuotinę apklausą, teritorijos buvo planuojamos taikant GIS. Tai reiškia, kad apdorotus tyrimo rezultatus bus galima analizuoti pagal surašymo teritorijas.

Erdvinis duomenų agregavimas yra priemonė užtikrinti kiekvieno iš mūsų asmeninių duomenų konfidencialumą. Tarkime, kad statistinėje dažnių lentelėje (angl. *crossstab*) yra aprašyta dviejų požymių sąsaja tiriamoje teritorijoje. Jeigu atsitinka taip, kad tik vienam asmeniui yra būdingas dviejų požymių derinys, žinodami vieną galime rasti kitą. Iš to seka Lietuvos geografų praktikai svarbi išvada, kad jei nagrinėjami atskiri požymiai, kas paprastai ir yra daroma, tokių problemų nekyla. Kitais atvejais norėdami apsaugoti asmeninius duomenis turime didinti teritoriją, operuoti santykiniais dydžiais, atlikti mažus atsitiktinius duomenų pakeitimus ir kt.

Kiekvienam geografui, norinčiam dirbti su erdviškai detaliais socioekonominiais duomenimis, teks įdėti šiek tiek pastangų. Esant dideliems duomenų kiekiams duomenų ruošimo analizei procesas neapsieina be programavimo, tačiau kitų socialinių mokslų kryptių atstovams vykdyti erdvines analizes yra dar sudėtingiau. Daugelis problemų gali būti sprendžiamos naudojant kolegų iš kitų šalių jau sukurtą kompiuterinę techninę ir programinę įrangą. Pažymėtina, kad atviro kodo programinė įranga jau siūlo daugelį patogių sprendimų ir būtent ji paprastai yra naudojama pirmaujančiuose universitetuose; *R*, *PostGIS*, *Quantum GIS* leidžia susidoroti ir su nestandartinėmis užduotimis. Didžiausias iššūkis yra bendradarbiavimo su Statistikos departamentu gerinimas.

Pagrindinis duomenų šaltinis taikomiesiems geografiniams tyrimams yra surašymai. Norint stiprinti kiekybinės ir taikomosios geografijos tradicijas yra būtina skirti išskirtinį dėmesį 2011 m. surašymo rezultatų platinimo, analizės ir patrauklaus erdvinio pristatymo gyventojams spren-

dimams, taip pat žymiai aktyviau dalyvauti planuojant ir organizuojant 2021 m. surašymą. Tai galima atlikti vykdant keletą mokslinių projektų, į kuriuos būtų įtraukti geografs, informatikai, socialinių mokslų atstovai. Mokslininkų pareiga yra rūpintis, kad remiantis surašymo metu gautais duomenimis kuriamas tikrovės vaizdinys būtų pakankamas, įdomus ir naudingas. Kaip parodė 2011 m. patirtis, mokslininkams aktyviai neįsitraukiant į surašymo procesus, įsigali verslo ir biurokratinio valdymo principai. Tokia sistema yra nereflektivi, sunkiai gali atlaikyti kritiką žiniasklaidoje. Kyla pavojus, kad tikėjimas surašymų reikalingumu bus pakirstas. Valstybės sektorius artimiausiu metu turi skirti daugiau dėmesio didelės erdvinės raiškos socioekonominių duomenų įsisavinimui. Taip pat svarbu, kad tokie duomenys taptų prieinami verslui. Nauja informacija reikalinga siekiant didinti rinkų skaidrumą, efektyviau teritoriškai skirstyti išteklius. Neišnaudojant esminių surašymų duomenų rinkinių privalumų yra neįmanoma pagrįsti visuotinių apklausų reikalingumo.

## 2. Duomenų įvairovė

Pasiekus pakankamą erdvinę raišką tampa prasminga analizuoti sąryšius tarp įvairesnių rodiklių. Darosi įmanoma apčiuopti erdvinės socialinės struktūros detales, jos nebesislepia už didelių skaičių ir vidurkių. Taikant šiuolaikines priemones įmanoma apibendrinti dešimtis ar šimtus rodiklių iš skirtingų šaltinių. Pagal gyventojų, namų ūkių ir būstų požymius nustatoma kiekvienos mažos šalies teritorijos santykinė padėtis kitų teritorijų atžvilgiu. Išskiriamos naujos esybės, kurios apibūdina geografiškai itin artimų individų ir pastatų grupes. Pagal tokią informaciją galime ieškoti panašių ar skirtingų teritorijų, papildyti ar vertinti individualių tyrimų duomenis atsižvelgiant į artimą teritorinį kontekstą, taikyti vienur padarytus apibendrinimus taip pat suklasifikuotoms teritorijoms kitur.

Jau minėtieji Ch. Booth tyrimai nors buvo erdviškai detalūs, tačiau, galima sakyti, vykdyti analoginiu režimu. Pastatų priklausomybė kategorijai buvo nustatoma interpretuojant mokinių prižiūrėtojų, o vėliau – policijos pareigūnų užrašus. Šie asmenys gerai pažinojo savo artimąją aplinką ir kiekvienas turėjo savitą jos supratimą. Po Pirmojo pasaulinio karo ėmus remtis surašymų rezultatais tikrovė buvo abstrahuota, diskretizuota ir tapo įmanoma daugybė jos interpretacijų. Remiantis skirtingų požymių teritorine sklaida ir sklaidos diagramomis vienas kito atžvilgiu, buvo kuriamos to meto idėjas atitinkančios miesto geografijos teorijos. Čikagos universitete susikūrusi evoliucijos idėjomis besiremianti žmonių ekologijos (angl. *human ecology*) mokykla prisidėjo prie 3-jų klasikinių miesto geografijos modelių: koncentrinų zonų, sektorinio ir daugelio branduolių kūrimo. Miesto dalių raida aiškinta remiantis biologine invazijos-sukcecijos analogija (Park, Burgess, 1921).

Po Antrojo pasaulinio karo ankstesni darbai buvo tęsiami (Shevky, Williams, 1949; Shevky, Bell, 1955). Išskirti veiksniai – socialinis rangas, urbanizacija, segregacija, kurie toliau buvo tikslinami indukcinų tyrimų metu (angl. *factorial ecology*). Remiantis šių tyrimų rezultatais, kurti modeliai veikė JAV būsto rinkos vystymą, o kaimynystės gyvavimo ciklo (angl. *neighbourhood life cycle*) (Hoover, Vernon, 1959) modelis ir dabar nėra pamirštas bei kelia diskusijų priimant miesto plėtros ar atnaujinimo sprendimus. Tuo pačiu metu imtas skaičiuoti nepanašumo indeksas (angl. *dissimilarity index*), kuris šiuo metu yra taikomas Rytų ir Centrinės Europos miestų erdvių skirtumų kiekybiniuose tyrimuose (Marcinczak et. al., 2011).

Pirmosios geodemografinės, pritaikytos rinkodarai, klasifikacijos kurtos 1976–1978 m., o šiuo metu rinkoje dominuoja keletas jų kūrėjų. Tradiciškai komercinėse klasifikacijose analizuojami šimtai kintamųjų, mokslinėse – dešimtys. Komercinė geodemografijos industrija Jungtinėje Karalystėje šiuo metu yra pasiekusi brandos stadiją, stagnuoja tiek metodikos, tiek naujų įrankių

kūrimo prasme. Taikymai yra plačiai paplitę viešajame ir privačiame sektoriuose (Singleton, Longley, 2009) ir nieko nestebina.

Pabrėžtini du įvairesnių duomenų taikymo aspektai – duomenų iš skirtingų šaltinių siejimo ir daugiadimensės analizės. Geografija informacinėms technologijoms siūlo universalų duomenų siejimo būdą – geokodavimu paremtus principus. Siekiant papildyti surašymų duomenis rodikliais iš kitų šaltinių reikalingas bendradarbiavimas su valstybės ir verslo institucijomis, taip pat informatikų pagalba. Taikant pastatų pašto adresų taškinį geokodavimą, atskirus stebimus atvejus įmanoma sieti su informacija apie jų kontekstą. Problemų kelia tai, kad geokodavimui reikalingų Adresų registro duomenų nemokamo teikimo mokslo tikslais sąlygų VĮ *Registų centras* nėra numatęs. Manytina, kad tai yra viena iš priežasčių, kodėl taikomieji geografiniai tyrimai nėra paplitę. Šio straipsnio autorius 2012 m. atlikti bandymai su 2000 visą Lietuvos teritoriją dengiančių taškų parodė, kad miesto sąlygomis geokodavimo funkciją sėkmingai leidžia vykdyti *Google maps* API sąsaja. Tikėtina, kad nugalėjus kliūtis taikomieji tyrimai plėtosis panašiai kaip Jungtinėje Karalystėje: geodemografiniai GIS taikymai greičiausiai vystosi ten, kur didelės lėšos yra skirstomos lokaliu lygmeniu, tiksliau sakan, viešosios tvarkos palaikymo, švietimo ir sveikatos apsaugos srityse (Longley, 2005).

Interpretuojant daugiadimensius duomenis ir jų analizės rezultatus yra būtina erdvinė vizualizacija. Geografų turima kartografavimo patirtis čia yra nepamainoma. Vis dėlto atkreiptinas dėmesys į tai, kad daugiadimensių duomenų atveju yra būtina pereiti nuo statinių žemėlapių kūrimo prie interaktyvių kartografavimo programų naudojimo. Tinkamos yra tos, kurios vienu metu geba vaizduoti keletą skirtingų vaizdų su pažymėtais geografiniais objektais (Monmonier, 1989), operatyviai keisti kategorijų simbologiją (*CommonGIS*), leidžia patikrinti skirtingų geostatistinių metodų tinkamumą turimiems duomenims (*GeoDa*). Be paminėtų, dar yra keliolika universitetuose sukurtų programų. Interpretuojant rezultatus geografi gali padėti kitiems specialistams, tačiau klasifikuojant daugiadimensius duomenis yra vertinga informatikų pagalba. Kaip rodo pradinė erdviškai detalių 2001 m. surašymo duomenų klasifikavimo patirtis, standartiniai klasterizavimo metodai užima per daug laiko arba neleidžia išryškinti daugiadimensiuose duomenyse atspindėtų struktūrų. Svarbu yra parinkti tinkamą skaičiavimo algoritmą. Pažymėtina, kad pradėjus taikyti erdviškai detalių surašymo duomenų klasifikacijas, sudaromos sąlygos daugelio sričių geografizavimui. Toks „erdvės standartizavimas“ darbo rinkoje kuria naujas nišas kiekybinius metodus taikantiems geografiams.

### 3. Duomenų laikiškumas

Visa laiko geografija yra taikomoji geografija, o erdvė be laiko neturi energijos ir yra keistas darinys (Thrift, 1977). Laiko dedamoji yra itin vertinga taikomuosiuose geografiniuose tyrimuose. Taikant naujausias technologijas ir turint operatyvius bei erdviškai detalius duomenis yra įmanoma daryti apibendrinimus, kurie leidžia kaupti stebimus procesus. Iš kitos pusės, ne geografinės operatyvių duomenų analizės papildžius geografinė dimensija tampa paprasčiau interpretuoti pirminius duomenis ar skaičiavimus.

Vienu metu grafiškai vaizduoti laiko ir erdvės dimensijas nėra paprasta. Trys jau esami techniniai sprendimai yra: keletas susietų vaizdų, animacija ir erdvėlaikio kubas (Andrienko et al., 2010). Keleto susietų vaizdų metodas vaizduoja laiko dedamąją kaip vieną ar keletą papildomų grafikų. Animacija yra mažai interaktyvus, tačiau geriausiai erdvėlaikio srautą perteikiantis vaizdavimo būdas. Socioekonominių duomenų atveju bene aktualiausias yra trečiasis būdas, kurio pavyzdžiai yra pateikti internete (Kraak, 2003; Andrienko, 2012). Erdvėlaikio kubo metodas vertingas siekiant papildyti surašymų teikiamą informaciją. Nors daugelyje šalių ji yra pagrindinis duomenų apie geografinį gyventojų pasiskirstymą šaltinis, tačiau siekiant geriau suvokti procesus kyla dvi



problemos. Pirmoji, nors dauguma veiklų ir išlaidų patiriama ne namuose, asmenų padėtis apibrėžiama pagal jų gyvenamąją vietą. Antroji, susijusi su pirmąja, dauguma geografinių tyrimų apsiriboja gyventojų naktine padėtimi ir nesutelkia dėmesio į padėtis dienos metu arba kitaip – judėjimą paros metu (Janelle et. al., 1998).

Dar iki kiekybinių metodų taikymo lūžio geografijoje Janelle minėtus trūkumus stengtasi kompensuoti siejant laiką ir erdvę. Švedijoje, Lundo universitete sėkmingai tai atliko T. Hagerstrandas. Tyrimuose jis analizavo žmonių veiklas erdvėlaikyje per visą jų gyvenimo ciklą. Šio autoriaus prieigai, vadinamai laiko geografija, būdinga tai, kad laikas yra traktuojamas kaip visuomenės funkcionavimui būtinas išteklius, panašiai kaip erdvė. Siekiant maksimalaus paprastumo ir aiškumo tiriami reiškiniai aprašomi absoliutinėje erdvės ir laiko koordinacių sistemoje. Laiko geografijos modelio pagrindinės sąvokos yra projektai, suvaržymai, trajektorijos, punktai ir pluoštai. Kiekvienas žmogus tikslui pasiekti dalyvauja projektuose, kuriuos sudaro užduočių sekos. Užduotims atlikti reikalingi žmonės, erdvė, laikas ir materialiniai ištekliai. Suvaržymai, dėl kurių užduotys gali būti neįvykdomos, yra skirstomi į tris grupes: gebėjimų, susijungimo ir valdžios. Visuomenė ir individai susiduria su problema, kaip sutalpinti kuo daugiau projektų turint ribotus erdvėlaikio išteklius. Erdvėlaikio kube kiekviena esybė suvokiama kaip trajektorija, o sociumas įsivaizduojamas kaip trajektorijų, tekančių per erdvėlaikio padėtis – punktus – visuma. Pluoštuose individai sąveikauja atsižvelgiant į suvaržymus ir punkto padėtį. Tiriama iš trijų perspektyvų: individo, punkto ir erdvėlaikio išteklių (Thrift, 1977).

Laiko geografija kartais yra suvokiama kaip visuomeninės geografijos laiko-erdvės teorija, tačiau ji taip pat gali būti priimama kaip hipotezių tikrinimo ir teorijų vystymo prieiga. <...> Iš vienos pusės, yra atvaizduojami individai, įvykiai ir jų erdvėlaikiniai ryšiai. Iš kitos, tampa akivaizdi erdvėlaikinių apribojimų svarba. Tokia prieiga, pavyzdžiui, leidžia tirti socialinių ryšių erdvėlaikines pasekmes (Stjernstrom, 2004). Taip pat pažymėtina, kaip informacinių technologijų plėtra veikia laiko geografijos tyrimus: pirma, silpsta ryšys tarp veiklų ir punktų; antra, tas pats pažymėtina apie veiklas ir laiko apribojimus; trečia, kelionių sprendimus dažniau priima atskiri individai, o ne namų ūkiai. Žmonių naudojančių ir nenaudojančių informacines technologijas mobilumo pobūdis skiriasi (Couclelis, 2005?). Sukūrus laiko geografijos modelį jį taikant buvo sudėtinga fiksuoti individo buvimo vietą. Šiandien tai padaryti paprasčiau, tačiau individo buvimo vieta apie jo veiklą ne visada pasako taip daug, kaip anksčiau. Pažymėtina, kad įspūdingiausius duomenų kiekius laiko geografijos tyrimams tvarko *Google* korporacija. Kiekvienas vartotojas sutinka su šiomis sąlygomis: „kai naudojate vietą leidžiančią nustatyti paslaugą, galime rinkti ir apdoroti informaciją apie faktinę jūsų vietą, pvz., padėties nustatymo sistemos (*GPS*) signalus, kuriuos siunčia mobilusis įrenginys. Taip pat galime naudoti įvairias technologijas, kad nustatytume vietą, pvz., jutiklių duomenis iš jūsų įrenginio, kurie gali, pvz., teikti informaciją apie netoliese esančius bevielio (*Wi-Fi*) ryšio prieigos taškus ir mobiliojo ryšio bokštus“ (Google, 2012). Sėkmingai mobilumo tyrimus vysto ir Tartu universiteto geografi (Ahas, 2011).

Švedijoje svarbiausia laiko geografijos tyrimų prielaida buvo ir yra išskirtinai centralizuota valstybės administracinių registrų sistema. Šiuo metu Umea universiteto Socialinės ir ekonominės geografijos fakultete dirbama su *ASTRID* duomenų baze. Ji leidžia vykdyti atskirų šalies gyventojų laiko geografijos tyrimus. Erdvinės gardelės dydis yra 100 m., o apimamas laiko intervalas – 40 metų. Tačiau ir Lietuvoje naujausių techninių priemonių visuma leidžia atlikti tai, kas anksčiau buvo sudėtinga. Svarbiausia yra tinkamai atrinkti respondentus, sukurti pasitikėjimą ir susitarti dėl dalyvavimo tyrime. Atlikę apklausą ir gavę žinių apie asmens projektus bei suvaržymus, išmaniajame telefone įdiegę programą, teikiančią tyrimui informaciją apie trajektoriją bei interneto paieškas, o stacionarioje darbo vietoje GIS priemonėmis parengę skirtingo lygmens punktų sąrašus galime vykdyti skirtingus taikomuosius tyrimus.

Svarbu laikiškų duomenų šaltinius suvokti maksimaliai plačiai. Akivaizdu, kad administraciniais tikslais ir vykdant pirkimo-pardavimo sandorius kaupiami duomenys nuolat yra atnaujinami, taip pat jie dažniausiai turi ir geografinę dedamąją. Kaip pažymi Longley (2005), „akademikams ir GIS praktikams reikia suprasti, kad teikiant viešąsias paslaugas sukuriama daug kokybiškų duomenų, kurie galėtų suteikti vertingą foną priimant sprendimus, tačiau ligi šiol per apsileidimą buvo nenaudojami. Galbūt tie duomenys yra slepiami“. Geografai, norintys dalyvauti verslaus universiteto kūrime, privalo veikti kartu su kitais mokslininkais, kūrybingai taikyti naujas programines priemones ir aktyviai bendradarbiauti su verslo ir viešojo administravimo institucijomis. Detalūs laike ir erdvėje apibrėžti valstybės registru ir įvairiais būdais verslo renkami – mokėjimus kortele atliekant, mobilaus ryšio paslaugas teikiant fiksuojami duomenys yra XXI a. taikomųjų geografinių tyrimų dirva. Standartinė GIS programinė įranga yra gremėzdiška ir sudėtinga, netenkina poreikių dirbant su dideliais duomenų kiekiais ir susidūrus su sudėtingomis analitinėmis problemomis (Andrienko et. al., 2010). Išėitis – naudoti ir universitetuose kuriamą programinę įrangą, operatyviai rengti žemėlapius, greta jų vaizduoti laiko dedamąją, taikyti duomenų agregavimo ir apibendrinimo algoritmus.

## Išvados

Remiantis geriausiais pavyzdžiais straipsnyje aprašytos glaudesnės mokslo, verslo ir valstybės sąveikos galimybės. Analizė pagal tris aspektus leido Lietuvos geografų bendruomenei glaustai pristatyti verslesnės geografijos modelį. Pažymėtina, kad:

1. Vakaruose mokslininkai buvo priversti daug laiko skirti analizės priemonėms tobulinti. Lietuvos geografai būdami atviri ir bendradarbiaudami su informatikais gali sparčiai vyti kolegas, o kartu siūlyti naujų sprendimų viešajam sektoriui bei verslui, įsijungti į tarptautinę tyrimų erdvę.

2. Erdviškai detalių duomenų tvarkymas keitėsi nuo „analoginės“ pastatų klasifikacijos, prie darbo su surašymo apygardų (< 10 000 gyv.) ir surašymo apylinkių (< 1000 gyv.) duomenimis. Galiausiai, kuriamos sistemos: santykinai homogenišku užstatymu pasižyminčių ir panašų gyventojų skaičių turinčių duomenų išvesties teritorijų arba vienodu plotu ir stabiliomis ribomis pasižyminčių erdvinių gardelių. Socioekonominiams tyrimams parankesnis yra duomenų išvesties teritorijų naudojimas.

3. Remiantis erdviškai detaliais surašymo duomenimis ir taikant šiuolaikinius duomenų siejimo metodus yra kuriamos konkrečioms poreikiams pritaikytos daugiadimensės mažų teritorijų klasifikacijos. Rodiklių atrankai ir interpretuojant rezultatus reikalingos interaktyvios geovizualizacijos programos.

4. Visuotinai plintant ryšių prietaisams su geografinės padėties nustatymo įranga, duomenų bazėms, geografinė informacija tapo laikiška. Tokių duomenų tvarkymas, operatyvus rezultatų pateikimas atveria didžiausių galimybių ir reikalauja aukščiausios kompetencijos.

5. Taikomosios geografijos plėtrai nepamainoma yra atviro kodo programinė įranga. Ji yra tarsi greitkelis tarp universitetų, valdžios institucijų ir verslo. Sukaupia atviro kodo priemonių naudojimo patirtis leidžia išvengti finansinių sąnaudų plėtojant verslą ar diegiant naujus taikymus valstybės institucijose.

6. Skatinant taikomosios geografijos vystymąsi reikalinga kiekybinių metodų taikymą pasirinkusių studentų specializacija: automatizuoto duomenų rengimo tyrimui, apdorojimo, rezultatų pateikimo interaktyviomis kartografavimo priemonėmis srityse.

## Literatūra

- Ahas R.**, 2011. *Mobile methods. Mobile positioning*, p. 183–199. London. New York: Routledge.
- Andrienko K., Andrienko N.**, 2012. *Space-time cube in visual analytics*. <http://www.itc.nl/personal/kraak/move/andrienko.pdf> [prisijungta 2012 m. lapkričio 9 d.]
- Andrienko G., Andrienko N., Demšar U., Dransch D., Dykes J., Fabrikant S., Jern M., Kraak M., Schumann H., Tominski Ch.**, 2010. Space, time and visual analytics. *International Journal of Geographical Information Science* 24 (10), p. 1577–1600.
- Couclelis H.**, 2005?. *Activity modeling with GPS tracking data: new assumptions for the age of ICT*. [http://www.csiss.org/events/meetings/time-geography/docs/Helen\\_bio\\_and\\_position.pdf](http://www.csiss.org/events/meetings/time-geography/docs/Helen_bio_and_position.pdf) [prisijungta 2012 m. lapkričio 12 d.]
- Etzkowitz H.**, 2003. Innovation in innovation: the triple helix of university-industry-government relations. *Social Science Information* 42 (3), p. 293–337.
- Farr M., Webber R.**, 2001. MOSAIC: from an area classification system to individual classification. *Journal of Targeting, Measurement and Analysis for Marketing* 10 (1), p. 1–15.
- Google**, 2012. Privatumo politika. <http://www.google.com/intl/lt/policies/privacy/> [prisijungta 2012 m. lapkričio 9 d.]
- Haklay M.**, 2012. Geographic information science: tribe, badge and sub-discipline. *Transactions of the Institute of British Geographers* 37 (4), p. 477–481.
- Harvey D.**, 1969. *Explanation in geography*. New York: St. Martin's Press.
- Hoover E., Vernon R.**, 1959. *Anatomy of a Metropolis*. Cambridge: Harvard University Press.
- Janelle D., Klinkenberg B., Goodchild M.**, 1998. The temporal ordering of urban space and daily activity patterns for population role groups. *Geographical Systems* 5 (1–2), p. 117–138.
- Kraak M.**, 2003. *The Space-time cube revised from a geovisualization perspective*. [http://www.itc.eu/library/Papers\\_2003/art\\_proc/kraak.pdf](http://www.itc.eu/library/Papers_2003/art_proc/kraak.pdf) [prisijungta 2012 m. lapkričio 6 d.]
- Longley P.**, 2003. Geographical information systems: developments in socio-economic data infrastructures. *Progress in Human Geography* 27 (1), p. 114–121.
- Longley P.**, 2005. Geographical information systems: A renaissance of geodemographics for public service delivery. *Progress in Human Geography* 29, p. 57–63.
- LSE**, 2002. *Charles Booth and the survey into life and labour in London (1886–1903)*. <http://booth.lse.ac.uk/> [prisijungta 2012 m. spalio 7 d.]
- Marcińczak S., Musterd S., Stępnik M.**, 2011. Where the grass is greener: social segregation in three major Polish cities at the beginning of the 21<sup>st</sup> century. *European Urban and Regional Studies* 19 (4), p. 383–403.
- Monmonier M.**, 1989. Geographic brushing: enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis* 21 (1), p. 81–84.
- Openshaw S.**, 1984. *The modifiable areal unit problem*. Concepts and techniques in modern geography. Nr. 38. Norwich: Geo Books.
- Orford S., Dorling D., Mitchell R., Shaw M., Davey S.**, 2002. Life and death of the people of London: a historical GIS of Charles Booth's inquiry. *Health and Place* 8, p. 25–35.
- Park R., Burgess E.**, 1921. *Introduction to the science of sociology*. Chicago: University of Chicago Press.
- Petrulis V.**, 2009. *Lietuvos politinio lauko teritorinė struktūra (elektorinio metodo pagrindu)*. Doktoro disertacija. Vilnius: Vilniaus universitetas.
- RAE**, 2008. *Research Assessment Exercise quality profiles*. Geography and Environmental Studies. <http://www.rae.ac.uk/results/qualityProfile.aspx?id=32&type=uoa> [prisijungta 2012 m. gruodžio 26 d.]
- Rees P., Martin D., Williamson P.**, 2002. *The census data system*. London: Wiley.
- Shevky E., Bell W.**, 1955. *Social area analysis: theory, illustrative application and computational procedures*. Stanford: Stanford University Press.
- Shevky E., Williams M.** 1949. *The social areas of Los Angeles: analysis and typology*. Berkeley: University of California Press.
- Singleton A., Longley P.**, 2009. Geodemographics, visualisation, and social networks in applied geography. *Applied Geography* 29 (3), p. 289–298.
- Stjernstrom O.**, 2004. Theory and migration: towards a framework of migration and human actions.



*Cybergeo: European Journal of Geography. Espace, Société, Territoire.* Nr. 254. <http://cybergeo.revues.org/3827> [prisijungta 2012 m. lapkričio 9 d.]

**Thrift N.**, 1977. *An introduction to time-geography.* Concepts and techniques in modern geography. Nr. 13. Norwich: Geo Abstracts Ltd.

**Valatka V.**, 2010. Visuomenės socialinės teritorinės struktūros pažinimas visuotinio gyventojų surašymo kontekste. *Filosofija. Sociologija* 21 (4), p. 357–367.

**Vickers D.**, 2006. *Multi-level integrated classifications based on the 2001 census.* PhD thesis. United Kingdom: University of Leeds.

## THREE ASPECTS OF SOCIO-ECONOMIC DATA APPLICATION

**Vytautas Valatka, Andrius Balčiūnas**

*Vilnius University, Faculty of Natural Sciences*

*E-mail: vytautas.valatka@gf.vu.lt*

### *Summary*

This paper places an Eastern European case of high resolution socio-economic data usage practice in the international perspective. A review of applied and quantitative geography traditions in the Lithuanian academic geographers' community was made. The instrumental division of complex spatial-temporal data analysis aspects was done using spatiality, diversity, and temporality dimensions.

Firstly, till now an assumption in the academia was made that a high resolution spatial data is not available for the research. But the problem is rather about the data usage agreements, a right choice of software tools, digitizing of census maps and a fear of large and complex datasets. It is complicated to make an analysis in a Western applied geography tradition without an appropriate background, but still geographers have basic skills for small area data analysis.

Secondly, data diversity issue is overlooked, because data fusion techniques and multidimensional analysis are topics more common for computer scientists than geographers in Lithuania. In spite of that, geographers' experience with spatial data analysis is on demand. Interactive geo-visualization techniques make data interpretation easier.

Lastly, dealing with the data temporality is a hot topic in international geographers' and computer scientists' community. A time geography space-time aquarium is a good choice to start with socio-economic data analysis. Mobility data is available through mobile phones tracking applications. State registers and finally, business databases could be the next steps.

The key advantage of Lithuanian case is an opportunity to use the existing open source and free software made by universities. Cooperation with computer scientists and business sector is needed. More emphasis should be made on the data mining techniques and interactive visualization. The weak capabilities of public sector to use socio-economic data for decision making should be compensated for by incentives of scientists.