

DR. JAKOBI ÁKOS

Újszerű területi statisztikai adatgyűjtési lehetőségek az információs világ egyenlőtlenségeinek kutatásában *

A technológiai fejlődéssel párhuzamosan egy – napjainkra már igen széles körben ismertté vált – információs robbanás tanúi is vagyunk, ami egyre többször sarkallja arra a kutatókat, hogy ne csak az eszközellátottság vagy a hozzáférés, de a tartalom és a használati szokások oldaláról is megvizsgálják korunk társadalmi-gazdasági sajátosságait.

Az információs társadalom fejlődése, a digitális világ előrehaladása egyre többek számára teszi hozzáférhetővé az új információs csatornákat, miközben persze igen sokan továbbra is kimaradnak az IKT-eszközök (infokommunikációs technológiákat) használók köreiből. Ebben a vonatkozásban új világunk digitális „bent lévőkre” és „kint lévőkre” (Mészáros 2003), avagy használókra és kimaradókra osztható, amit a szakma digitális szakadék kifejezéssel illet (Norris 2001, Nagy 2007). Manapság azonban nem elégséges mindössze „bent lenni”, egyre fontosabb az, hogy ki mennyire is vált részesévé ennek a világnak, avagy ki milyen minőségben képes e világ előnyeit kihasználni. Korunkban egyre fontosabbá válnak azok az új differenciák, amelyeket az IKT-használat tartalmában vagy minőségében érhetünk tetten. Hasonlóképpen az információs társadalomfejlődés globális és hazai tapasztalatai is azt sugallják, hogy mára számos vonatkozásban túlléptünk a kezdeti és a nekirugaszkodási fázison (Selwyn–Facer 2007, Vartanova–Smirnova 2012), s a hozzáférési differenciák helyett egyre jelentősebbé kezdenek válni a használók közötti és főleg a használat minőségéből fakadó differenciák.

Habár az információs világ területi különbségeinek mérésére ez idáig már több megoldás is született, a legtöbb közelítés lényegében csak érintőlegesen vizsgálta azokat a motívumokat, amelyek az információk kvalitatív részeire vonatkoztak. Másként fogalmazva: az információkról szóló információk, avagy metainformációk alapján kirajzolódó virtuális vagy valós térbeli területi különbségek értékelésére eddig még viszonylag kevés kísérlet vállalkozott annak ellenére, hogy információs világunk napjainkban már exponenciálisan növekvő számban ontja magából az adatbázisszerűen is vizsgálható információstatisztikai adatokat.

Sőt nem csak a mérendő jelenség, a vizsgálandó témakör frissítése, megváltoztatása is szükségessé vált, illetve nem csak a hozzáférhető adatok és információk halmaza változott az utóbbi időszakban, de az elemzési eszköztár fejlődése, a számítógépes programok, alkalmazások, de főleg a térinformatika előrehaladása is lényeges volt, ami inspiratívan hatott arra, hogy az információs világ egyenlőtlenségeinek területi kutatásában újszerű eredményekre lehessen jutni.

* A tanulmány az MTA Bolyai János Kutatási Ösztöndíj támogatásával készült.

Jelen tanulmány ezek alapján arra tesz javaslatokat, hogy miként lehetne az információs társadalom területi kutatásában olyan kvantifikálható adatokat is felhasználni, amelyek az információs kor új változásait is visszatükrözik, illetve a kialakuló területi egyenlőtlenségek magyarázatában hasznos segítségként alkalmazhatók.

Az információs társadalom területi egyenlőtlenségeinek eddigi mérési módszertanáról

Az információs kor területi folyamatainak vizsgálata – annak ellenére, hogy viszonylag új jelenségről van szó – évtizedes tapasztalattal rendelkezik. Az információs és kommunikációs technológiák fejlődése okozta változások viszonylag korán felkeltették a területi kutatói szakma érdeklődését, s ezzel egyidejűleg a társadalmi-gazdasági előrehaladás mérésének, valamint a kialakuló egyenlőtlenségek számszerűsítésének igénye is hamar megfogalmazódott.

Az információs egyenlőtlenségeket kvantifikáló módszerek jellemzően mutatóhalmozatokkal operálva, esetleg komplex indexeket formálva tettek kísérletet az új kor területi egyenlőtlenségeinek feltárására. Az integrált mutatórendszerek az egyre szélesedő lehetőségek révén mind több területi indikátort építettek modelljeikbe. Az ezredforduló környékén megszülető nemzetközi (Johoka 2000, Ramachandran 1999) és hazai (Nagy 2002, Jakobi 2002, Döry–Ponác 2003, Rechnitzer et al. 2003) tanulmányok összetett mutatóhalmozataikban legtöbbször az informatikai és hálózati infrastruktúra, az információs gazdaság (IKT-szektor) és a társadalmi adaptáció különböző indikátorait szerepeltették. Hasonlóképpen a témával foglalkozó nemzetközi szervezetek is a technológia (infrastruktúra), a gazdaság (e-kereskedelem, IKT-szektor), illetve a társadalom (képzettség, adaptáció) különböző faktorait integrálták komplex modelljeikbe (lásd: *Digital Access Index*, *Digital Opportunity Index*, *ICT Development Index*, *Network Readiness Index*).

A hazai területi adatszolgáltatás is ezen igények mentén formálódott, miközben a területi elemzők nyilvánvalóan többnyire rá voltak utalva ezen források használatára. A magyar adatszolgáltatók közül az NHH (ma NMHH), a GKIeNET, illetve a KSH tett közzé az infokommunikációs eszközhasználatra és szolgáltatásokra vonatkozó területi adatokat, ám ezek is csak részben elégítették ki az elemzői igényeket. Annak ellenére, hogy a publikált adatok (például a mobiltelefon-előfizetések, a személyi számítógépek vagy az internethasználók száma vonatkozásában) már idősoros formában is rendelkezésre álltak, az információs társadalom fejlődésének előrehaladásával párhuzamban csak itt-ott jelentek meg új statisztikák, s csak egy-egy kísérletben tértek el a hagyományos penetráció jellegű adatközléstől (például e-ügyintézés szintje). Pedig a kutatói, elemzői igény már relatíve korán megfogalmazódott, ami egyedi adatforrások feltárását, primer adatgyűjtéseket vagy adatbeszerzéseket tett szükségessé. Nagy Gábor és Kanalas Imre (2003) munkájában már viszonylag korán megjelentek az olyan alternatív, vagy a témához kiegészítő jelleggel kapcsolódó indikátorok, mint a domainnévszerverek (DNS), vagy a domainregisztrációk számának területi mutatói. A későbbi kutatások is egyre több kiegészítő indikátorral próbálták a térségek vagy települések információs társadalmi pozícióit és egyenlőtlenségeit legjobban megragadni (például Tóth 2011), ami egyre erősítette azt az igényt, hogy a mérési módszertanokban a felhasználói szokások, az információhasználat minőségével kapcsolatos területi mutatószámok is helyet kapjanak. Ráadásul maga a technikai fejlődés is mód-

szertani problémákat generált a longitudinális vizsgálatok kapcsán: a gyorsan változó technológia révén lerövidülő termék-életciklus görbék miatt az egymást váltó technológiák csak részben összehasonlíthatók, hiszen minőségi paramétereik jelentősen átalakulnak (lásd: széles sáv, mobilkommunikáció, okostelefonok megjelenése stb.), amely hozzájuk kapcsolódóan rendszeresen új adatok publikálását teszi szükségessé. Mindez állandó problematikát jelent a területi elemzések számára is. A folyamatosan fejlődő és előrehaladó információs társadalom megértésében ezért (is) célszerű az új területi adatforrások és adatgyűjtési módszerek feltárása.

Közvetlen és közvetett változások a területi információk gyűjtésében: a „big data”

Folyamatosan változó világunkban az információtechnológiai eszközök és alkalmazások piacán sorra jelennek meg az újabb és újabb termékek és ezzel egyidejűleg folyamatosan módosulnak a felhasználói igények és szokások is. Ezen változások egyik nem elhanyagolható következménye, hogy a területi információk iránti igények is jelentős mértékben növekednek, és egyre több olyan eszköz és alkalmazás lát napvilágot, amely direkt vagy indirekt módon épp ezen területi információk hasznosítója vagy előállítója. Mindez legfőképp (de nem csak) a helymeghatározó eszközökkel egybeépített okostelefonok rohamos terjedésének köszönhető, amelyek révén a geoadatok használatát akár már a mindennapok részének is lehet tekinteni. Az okostelefonok és a hasonló eszközök, illetve alkalmazások révén azonban nemcsak a térbeli adatokhoz való hozzáférés lehetősége változott meg. Az alkalmazói kör akarva vagy akaratlanul milliányi digitális nyomot hagy maga után nap mint nap, amelyek jelentős része épp földrajzi tartalmú, és amelyek iránt a területi kutatói szakma érdeklődése is megélénkült.

A szakmai körökben „big data” (szabad fordításban „óriási adathalmaz”) néven ismert kifejezés arra a hatalmas adatmennyiségre utal, amely információs világunkban rapid módon és folyamatosan keletkezik, s amelynek feldolgozása a hagyományos kapacitásokkal és eljárásokkal operáló módszerekkel már-már megoldhatatlan kihívást jelent. A big data ezzel ellentétben mégis nagy lehetőségeket kínál. A sokáig csak virtuális melléktermékként számon tartott napi információhalom ugyanis épp akkor válik értékessé, amikor a különböző adatokat sikerül összekötni, köztük összefüggéseket, felismerhető mintázatokat találni, s mindebből értékelhető következtetéseket levonni. A kormányzati szerveknél, internetes és telekommunikációs cégeknél összegyűlt hatalmas adatmennyiség a társadalomtudományok számára valóságos aranybánya. Vásárlási, munkába járási, közlekedési és egyéb szokásainkról szinte korlátlanul gyűjtenek adatokat a különféle szervezetek, ezek összessége pedig betekintést nyújt az emberi viselkedés egyedi és társadalmi szintjeibe is. Az egyre szélesebb körben terjedő, térbeli információkat is használó alkalmazások révén pedig a big data a társadalom térbeli működésének megértéséhez is megszámlálhatatlan mennyiségben kínál új adatokat.

A big data komoly előnye, hogy nagyságrendnyi előrelépést jelent az adatvolumenben, hiszen így már kellően nagy mennyiségű adat áll rendelkezésünkre ahhoz, hogy a véletlen hatások szerepe mérséklődjön. Ha például egy még nem elegendően nagy csoport internetes keresési szokásait vizsgálánk, jelentős lenne az esélye, hogy téves következtetésre jutunk. Preis és munkatársainak (2012) vizsgálata viszont bebizonyította, hogy big data környezetben az internetes keresési szokásokból is stabil összefüggésekre juthatunk.

A Google keresési trendek adataira építve bemutatták, hogy a magasabb egy főre jutó GDP-vel rendelkező országokban élő internethasználók inkább a jövőre, mintsem a múltra vonatkozó információkat keresnek, továbbá azt is, hogy a jobb gazdasági teljesítményű országokban az állampolgárok információ-keresési aktivitása is vélhetően nagyobb. Megállapításaik azt valószínűsítik, hogy az online viselkedés és a valódi világ gazdasági mutatószámai között tényleges összefüggések lehetnek.

Hasonló példák bőséggel említhetők. A különböző online felületeken hagyott digitális nyomok lehetővé teszik például az elektronikus kereskedelemmel foglalkozó cégek számára az ügyfélkör részletesebb tulajdonságainak megismerését, s ezáltal az üzleti (és területi) marketingstratégiák hatékonyabb megszervezését. De említhető például az online tartalmak szövegbányászati módszerekkel történő feldolgozása is, amely lehetővé teszi, hogy számszerűsített módon is meghatározható lehessen egyes helyek és terek online reprezentációja, avagy adott helyhez kötődő tudattartalmak minősége vagy nagysága.

Ezen új adatforrások, illetve adatgyűjtési formák közvetett módon végső soron az információs kor területi egyenlőtlenségeiről adnak tájékoztatást, vagy legalábbis abból a célból is igénybe vehetők. Már az is, hogy hol, milyen országokban, térségekben keletkezik ez a hatalmas adatmennyiség, informálhat minket az egyes területek virtuális világbeli jelenlétének abszolút súlyáról. A big data felhőjében kirajzolódó digitális lábnyomok térségenkénti nagysága, minőségi eltérései pedig az információhasználó lakosság területi jellemzőiről szolgálhatnak új adalékokkal.

Térinformatikával támogatott újszerű területi statisztikai adatgyűjtés

A 21. század világát értékelő területi elemzői munkák lassan elengedhetetlen háttereként említhetők a különféle számítógépes adatbázisok, az azokat elemző szoftverek és programcsomagok, továbbá a területi szempontokat integráltan is kezelni képes térinformatikai technológiák az adatgyűjtéstől az adattároláson és -feldolgozáson keresztül egészen az elemzésig és az eredmények megjelenítéséig. Az információs társadalom területi sajátosságainak vizsgálatakor a térinformatika a területi adatelemzési folyamat minden fázisában alapvető fontossággal bír, és ez kvázi magától értetődő, hiszen az információs társadalom területi működésére vonatkozó adatok túlnyomórészt már eleve digitális formában jelennek meg. A digitális területi adatfeldolgozás legelterjedtebb eszköze manapság pedig már egyre dominánsabban a térinformatika.

A területi statisztikai adatelemzés szempontjából komoly gondot jelent, hogy az információs világ primer statisztikai alapvetően nem a területi elemzésekben megszokott hagyományos entitásokhoz kötődően látnak napvilágot, azaz alaphelyzetben gyakorta nem településekre, térségekre vonatkoznak. Mivel az új információs források az adatokat lényegében individuális szinten, legtöbbször pont-, s ritkábban esetleg vonalszerű entitásokhoz kötődően, például az utcák, a vezetékek vagy más hálózati elemek szintjén mérik, s kevésbé poligon- vagy felületszerű adattárolási formában teszik elérhetővé, a területi elemzésekhez leggyakrabban szükséges térségi adatok kialakításához területi adataggregációra van szükség. A térinformatika ebből kifolyólag lényeges szerephez jut az adatok területi elemzéséhez való előkészítésének munkafázisában is.

A területi adataggregáció térinformatikai megoldása alaphelyzetben azt feltételezi, hogy a vizsgálati alapadatok geolokalizálhatók, azaz valamilyen földrajzi objektumhoz

már eleve hozzákapcsolhatók. A virtuális térbeli teljesítmény számszerűsítésénél ez már eleve gondot jelenthet (lásd később), bár a lokalizáció több esetben – közvetett módon – itt is megoldható. Az eredményül kapott lokalizált adatokat, mint azt fentebb említettük, leggyakrabban pont- vagy vonalszerű entitásokhoz tudjuk kötni, amelyek a felületszerű entításokat tartalmazó rétegekre vetíthetők a pont a poligonban vagy a vonal a poligonban algoritmusok segítségével. Ez az eljárás a térinformatikában a helyzeti alapú adatkapcsolat (*spatial join*) módszereként terjedt el, amely a lekérdezés és az átlapolás (rétegmetszés) jól ismert technikái között nagyjából félúton helyezhető el.

A helyzeti alapú adatkapcsolatok kialakítása során egy jól definiált forrásréteg és egy célréteg azonos vetületi pozícióban (földrajzi helyzetben) lévő objektumai között hozunk létre kapcsolatot. Míg a hagyományos attribútum alapú adatkapcsolatok két tábla között a közös mező segítségével alakíthatók ki, addig a *spatial join* eljárásban az objektumok közös helyzete lesz a kapcsolat alapja. Az eljárás során a forrásréteg attribútumait a célréteg attribútum-táblájához társítjuk. A *spatial join* igazi előnye nem a kapcsolat azonosítása, hanem az új attribútum-mező létrehozása az adatbázisszerű térbeli lekérdezés eredményeként, tehát ennek révén a már ismert területi statisztikai beosztásokhoz bármilyen földrajzilag azonosított (lokalizált) adatot hozzákapcsolhatunk.

Az egyszerű helyzeti adatkapcsolatban a forrásréteg egy adott objektuma a célréteg egy (vagy több) objektumával kerül kapcsolatba. Ha a forrásréteg több eleme is átfed a célréteg egy adott objektumával, már nem egyszerű, hanem összegzett helyzeti adatkapcsolati eljárást érdemes választanunk. Az összegzett helyzeti adatkapcsolat különböző függvények szerint összegzi a forrásréteg numerikus attribútumait (például átlag, összeg, minimum, maximum, szórás stb.), majd ezt az információt hozzáadja a célréteghez. Ha például az internetes hozzáférési pontokat tartalmazó réteget (forrásréteg) és a kistérségek poligonjainak rétegét (célréteg) hozzuk kapcsolatba, akkor az összegzett helyzeti adatkapcsolattal a kistérségekre eső internetes hozzáférési pontok száma lesz meghatározható. Ugyanez a módszer tökéletesen működik a vonalas vagy felületi elemeket tartalmazó forrásrétegek esetében is, azzal a különbséggel, hogy ha a beviteli objektumok egyes részei több célobjektum között is megoszlhatnak, akkor a beviteli objektumok attribútum adatait – ha annak értelme van – a célobjektummal való átfedés arányában szét kell osztanunk. Ilyen esetekben a rétegek fedésbe hozásakor már komplexebb átlapolási feladatokat is végre kell hajtaniuk.

Az adatok területi elemzésekre való előkészítése a térinformatika más, változatos technikáinak alkalmazásával is elvégezhető. Említést érdemelhet az adatok különböző rács-sűrűségű raszterizálása, illetve például az adatok gridtranszformációja, amikor nem előre meghatározott konkrét közigazgatási stb. területegységek szerint aggregáljuk az adatokat, hanem azoktól független (avagy azokhoz nem feltétlenül illeszkedő) cellákra osztott térben vizsgálódunk. A grid olyan adatmodell, amelyben a földrajzi objektumok adatait diszkrét cellák, rendszerint négyzetek szabályos (és összefüggő) hálózata formájában jelenítjük meg. Létrehozásának elvi módszerei közt haszonnal számíthatunk a pontthalmazok kvadrát analízisére vagy az átlapolási, rétegműveleti technikákra. Egy grid ugyanakkor hasonlít a kétdimenziós mátrixhoz is, amelytől csak annyiban tér el, hogy celláit a bal alsó és nem a bal felső saroktól kezdve azonosítjuk. A cellák a társadalmi térinformatikai vizsgálatokban társadalmi-gazdasági attribútumok értékeivel jellemezhetők (ellentétben a képpontok esetében használatos intenzitásértékekkel).

Új lehetőségek az információs egyenlőtlenségek területi vizsgálatában

Korunkban egyidejűleg lehetünk tanúi az adatgyűjtési módszerek fejlődésének, valamint a statisztikai vizsgálatokra alkalmas adatforrások számszerű növekedésének. A változatos kutatói fantázia, illetve a gyorsan bővülő lehetőségek révén az információs egyenlőtlenségek vonatkozásában is egyre több új vizsgálati témával találkozhatunk, s a legkülönbözőbb térségi szintek folyamatainak megértéséhez kaphatunk új ötleteket, melyek közül néhányat az alábbiakban részletesebben is bemutatunk.

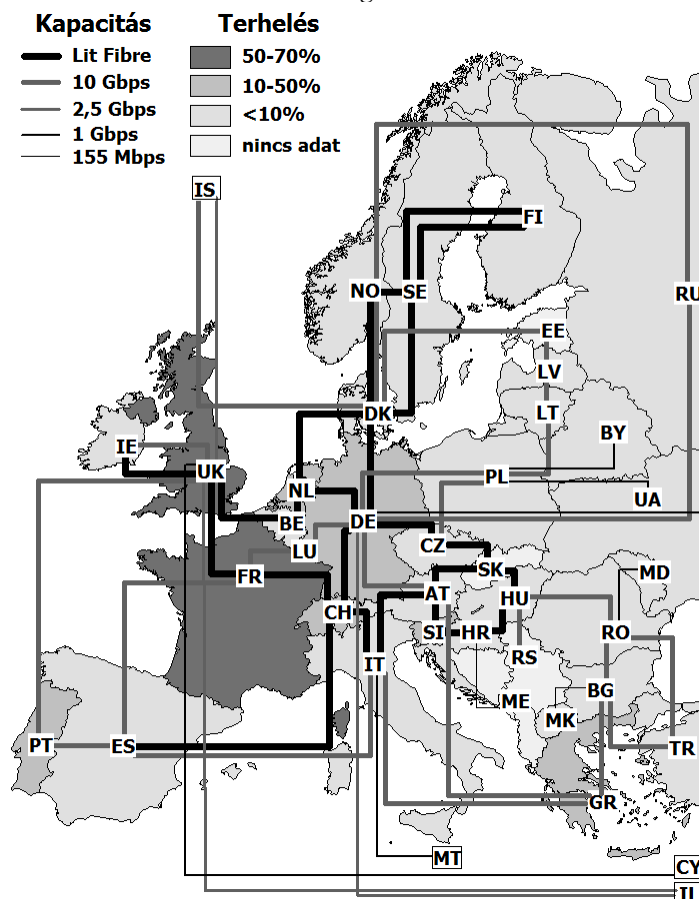
Az információs infrastruktúra újszerű statisztikái

A hagyományos statisztikai módszertanok már régóta foglalkoznak az informatikai infrastruktúra kiépültségének vizsgálatával, köszönhetően többek között annak, hogy a fizikai infrastruktúra a földrajzi térben általában jól azonosítható, könnyen lokalizálható, következőképpen a területi számbavétele is relatíve egyszerűen kivitelezhető. Nem tekinthető tehát komolyabb újdonságnak ezen elemek területi kvantifikálása, habár a technológiai fejlődés jóvoltából folyamatosan újabb és újabb infrastrukturális dimenziók felmérése válik szükségessé. Ezért tapasztalhatjuk a szakstatisztikákban az olyan új témák előretörését, mint például az optikai hálózati lefedettség területi változása (lásd például Kiss 2009, European Commission 2011) vagy a mobilszélessáv előfizetői körének regionális alakulása (ITU 2013) stb. A tradicionális adatelemzésekkel szemben az információs infrastruktúra új szemléletű vizsgálatai már nemcsak a penetráció különféle mutatóit elemzik, de például a hálózati topológiák, a gerincezeték-rendszerek területi sajátosságait (Malecki 2002), sűrűségeit (Tranos–Gillespie 2009) is vizsgálják.

A penetrációs statisztikákkal ellentétben az adatáramlási pályák és az áramlási volumenek területi aspektusú mérése viszont ritkaságszámba megy. Informatikai körökben a forgalmi statisztikák elemzése kiemelt jelentőségű, ám ezek területi vetületei csak ritkán kerülnek górcső alá, pedig az áramlási volumenek alakulása igazán sokat elárulhat az információhasználat kiemelt gyűjtőpontjairól, egyenlőtlenségeiről. Az információs infrastruktúra hálózatait nézve már önmagában az átviteli kapacitások is egyenlőtlen lehetőségeket mutatnak (nyilván persze az igényekhez igazodva alakították ki az átviteli sávszélességet), ami közvetlen vagy közvetett módon a digitális egyenlőtlenségek tükrözője lehet. A hálózati aktivitás azonban kvantitatív módon is visszatükrözheti az információs társadalmi jelenlétet (legalábbis adott hálózat viszonylatában). Az 1. ábra a GÉANT európai akadémiai gerinchálózat mint az egyik, ha nem a legfontosabb európai informatikai infrastrukturális hálózat topológiáját mutatja az átviteli kapacitások feltüntetésével, továbbá a nemzeti hálózatok felől a GÉANT hálózata felé irányuló aktuális forgalmi terhelést 2013 augusztusában. Az ábra főleg egyes nyugat-európai országokban mutat komolyabb hálózati terhelést (a kapacitások 50–70%-át elérő forgalommal), ahol így aktuálisan nagyobb relatív használói aktivitás feltételezhető.

1. ábra

Az európai GÉANT-hálózat topológiája és országonkénti forgalmi terhelése, 2013. augusztus



Forrás: a www.geant.net alapján saját szerkesztés.

Területi webstatisztikák

Az utóbbi évek, évtizedek egyik érdekes problematikája a virtuális térbeli jelenlét mérése. A világhálóra felkerülő tartalmak alapvetően térfüggetlenek, mégis akadnak olyan paraméterek, amelyek közvetlen vagy közvetett módon lehetővé teszik az egyes weboldalak földrajzi azonosítását.

A weboldalakhoz kapcsolódó új domaineik bejegyzése különböző internet-szolgáltatóknál (ISP) történik, amelyek a regisztráció alkalmával a bejegyzők címét (lakcímét vagy telephelyét) is rögzítik. A domainregisztrációk így – legalább a bejegyzőn keresztül – földrajzilag is azonosíthatókká válnak, s ezek az információk területi adatbázisokba is szervezhetők. Az eredendően lakcím/telephely szintű információk pontszerű elemeket tartalmazó térinformatikai rétegekben jeleníthetők meg, miközben térségekre, körzetekre vonatkozó

aggregálásuk könnyűszerrel megoldható akár térinformatikai módszerekkel, akár egyes attribútumazonosítók (például irányítószám) segítségével. A 2. ábra egy pontszerű és egy felületszerű elemekből felépülő térképen mutatja be a területi domainstatisztikák lehetséges ábrázolási módszereit. Az ábra sűrűbb, illetve intenzívebb jelölést mutató részei közvetett visszatükrözői lehetnek azoknak a területeknek, ahol az információs világbeli jelenlét dominánsabb szerepű.

2. ábra

Domain név tulajdonos eloszlása San Francisco belvárosában (1999), valamint Budapest kerületeiben (2001)



Forrás: Zook (2000), valamint Hargitai (2001) alapján saját szerkesztés.

Az esetek egy jelentős részében túlságosan körülményes, illetve nem könnyen lehetséges egy hétköznapi elemzőnek a weboldalak tulajdonosait, bejegyzőit egyszerűen felderíteni. Ilyen esetekben a közkeletűbb és hagyományos körökben is gyakrabban alkalmazott top level domain (TLD) statisztikák adhatnak legalább vázlatos képet a virtuális térbeli tartalmak földrajzi eredetéről. A generikus TLD-k (például a .net, .com, .edu végződésű domain nevek), illetve az országkód szintű TLD-k (például .hu, .uk, .de) közül ez utóbbiak csoportja földrajzi azonosítást is lehetővé tesz, de csak az országok szintjén. A TLD alapú domain statisztikákat már régóta használják (talán ez a legkönnyebben gyűjthető adat az internet világában), mégis folyamatos igény mutatkozik ezek regisztrálására és értékelésére, főleg a globális információs egyenlőtlenségek közvetett számszerűsítése és nyomon követése érdekében, ezért az alábbiakban mi is közzéteszünk egy „tradicionális” statisztikai adatbázisokból elérhető, ugyanakkor viszonylag naprakész információkat tartalmazó táblázatot.

Az 1. táblázat az elérhető területi webstatisztikák több példáját is bemutatja. A fentebb már említett domainstatisztikák mellett az internetes világ alapegységeiként is értelmezhető IP-címek statisztikája, illetve a weboldalak nyelvek szerinti megoszlása is hasznos információval szolgálhat arról, hogy korunkat milyen egyenlőtlenségek jellemzik, továbbá hogy Magyarország milyen súlyú is a virtuális térben. A weboldalak top level domainek szerinti megoszlási arányszámait alapján a hazánkhoz kötődő .hu végződésű domainek részaránya 0,3% körül alakul, miközben az ország népességszáma csak a világ 0,14%-át

ÚJSZERŰ TERÜLETI STATISZTIKAI ADATGYŰJTÉSI LEHETŐSÉGEK AZ INFORMÁCIÓS VILÁGBAN 43

teszi ki. A magyarországi IP-címek globális részaránya ezzel szemben csak 0,17%, ami a magyar internetes jelenlét ezen fizikai infrastrukturális tartalommal is bíró összetevőjében mért relatív alulreprezentáltságát vagy legalábbis alacsony szintjét jelezheti egyes országokhoz képest. Népeségarányos értékeket nézve Magyarországon száz lakosra 59 IP-cím jut, míg az e téren vezető Amerikai Egyesült Államokban 496, Kanadában 240, Dél-Koreában pedig 230, sőt sok más országban mérhetőek még 100 fölötti értékek.

1. táblázat

A magyarországi weboldalak, IP-címek, valamint a magyar nyelv súlya a virtuális térben, 2013. augusztus

Weboldalak megoszlása top level domainek szerint		IP-címek megoszlása országok szerint		Weboldalak megoszlása nyelv szerint	
domain	részarány, %	ország	részarány, %	nyelv	részarány, %
.com	53,0	USA	44,4	angol	55,4
.net	5,7	Kína	9,4	orosz	6,4
.ru	5,0	Japán	5,8	német	5,5
.org	4,2	Egyesült Királyság	3,6	japán	4,9
.de	3,0	Németország	3,5	spanyol	4,3
.jp	1,9	Dél-Korea	3,2	kínai	4,1
.uk	1,9	Franciaország	3,0	francia	3,8
.br	1,7	Kanada	2,3	portugál	2,3
.cn	1,6	Brazília	1,7	lengyel	1,8
.info	1,6	Olaszország	1,5	olasz	1,6
...
.hu	0,3	Magyarország	0,2	magyar	0,3
<i>Összesen</i>	<i>100,0</i>	<i>Világ</i>	<i>100,0</i>	<i>Összesen</i>	<i>100,0</i>

Forrás: a w3techs.com/technologies és www.whois.sc/internet-statistics adatai alapján a szerző számításai.

A lokalizált webes tartalmak területi statisztikái

A világhálón való jelenlét fontos, de nem egyetlen meghatározó eleme az, hogy hány weboldal köthető egyes térségekhez. Nemcsak arra lehetünk kíváncsiak, hogy kik és hol tesznek közzé információkat az internetes világban, de arra is, hogy miről, avagy mely helyekről közölnek tipikusan információkat a felhasználók. Természetesen az, hogy szemantikai értelemben milyen tartalmú információ kerül az internetre, területi szempontból általánosságban nehezen vizsgálható és nem is könnyen értelmezhető, ellenben az, hogy az egyes információk hol keletkeznek, illetve az egyes információ típusokra hol kíváncsiak, már elemezhető. Az információhasználat területi tartalmi minőségi különbségeiből közvetett módon például az információ-használati minőség vagy előrehaladottság szintjére lehet következtetni, ami végső soron újra az információs egyenlőtlenségek mai helyzetéről ad képet.

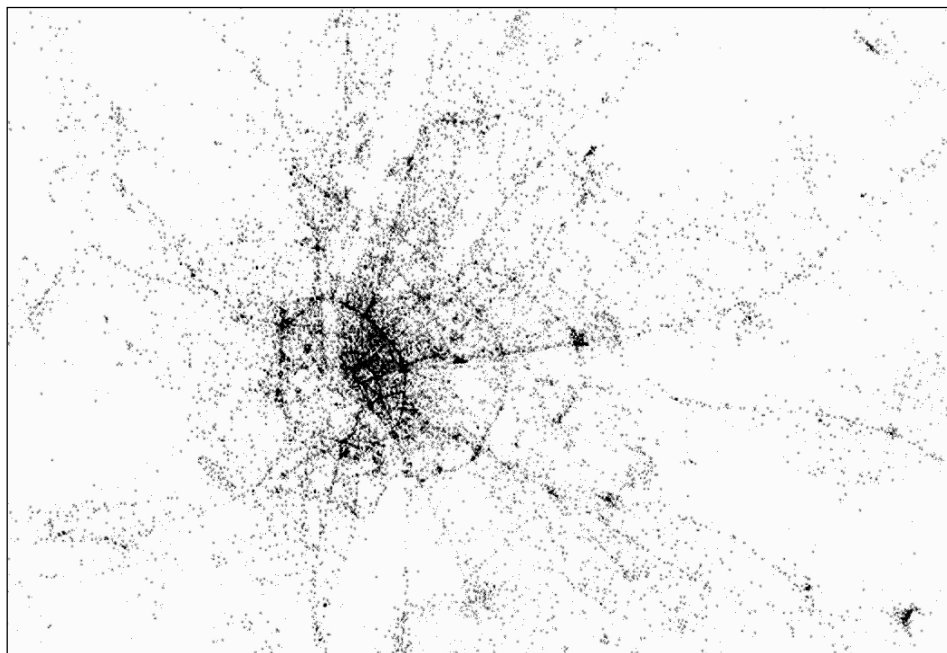
Különleges lehetőségeket kínál mindezek területi vizsgálatára az eleve térbeli azonosítási céllal, illetve földrajzi lokalizációs szándékkal létrehozott geotagek sűrűsödésének értékelése. Manapság ugyanis egyre több olyan internetes felület létezik, ahol a közzétett információk egyértelműen azonosított földrajzi lokalizációval is rendelkeznek, ami azt eredményezte, hogy ezzel az amúgy térfüggetlen információk is térbelivé váltak. A geokódolt információk a HTML-webkörnyezetben geotagek formájában helyezhetők el az

egyres weboldalakon. Ha ezen geotagek adatait összegyűjtjük és térképen ábrázoljuk, láthatóvá válnak azok a területek, amelyekről nagyobb sűrűséggel tesznek közzé a felhasználók térbeli információkat, s így közvetve az is láthatóvá válik, hogy hol használják aktívabban az információtechnológiai eszközöket.

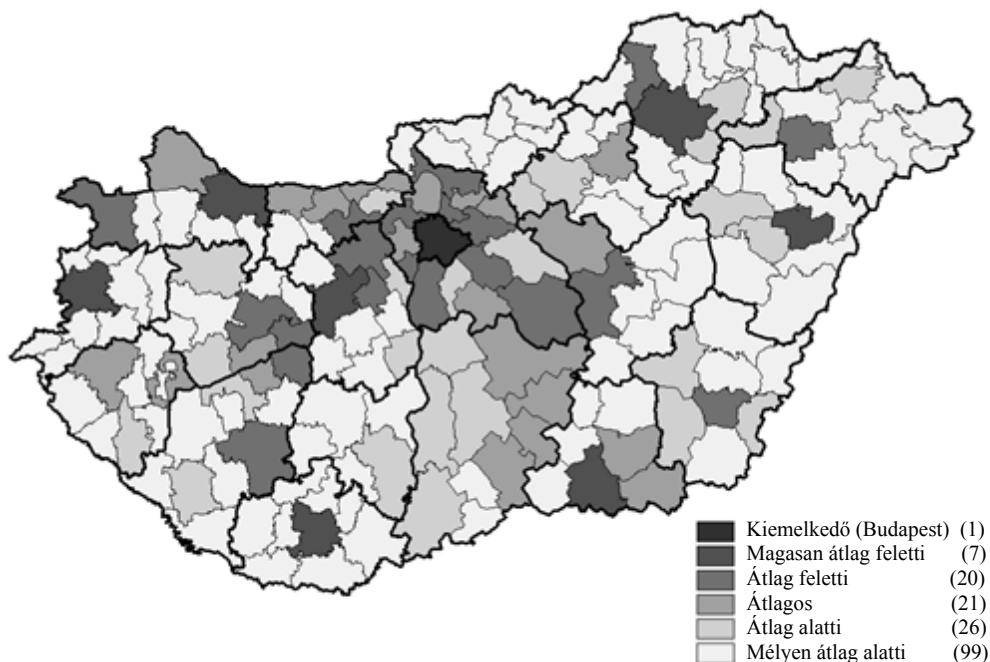
A 3. ábra például a Twitter közösségi oldalain elhelyezett geotaginformációk egyenlőtlen területi eloszlását mutatja Budapest térségében. Az ábrán dominánsabb, sötétebb színnel jelennek meg azok a helyek, ahol nagyobb gyakorisággal tettek közzé információkat a felhasználók. A térképen különösebb jelmagyarázat vagy segédfelirat nélkül is kivehető a pesti belváros, a nagykörút vagy egyes főbb utak sziluettje, de például nagyobb számban lehetett bejegyzéseket regisztrálni a Liszt Ferenc repülőtér vagy a Hungaroring környékén is. Fontos megjegyeznünk, hogy a kirajzolódó mintázatok nem feltétlenül a felhasználók számának egyenlőtlenségeit ábrázolják, inkább a térbeli tartalommal rendelkező információk relatív sűrűségeit, azaz a térségek tartalmi súlyát a virtuális térben. Persze ezek háttérben az abszolút volumenek, tehát az, hogy hol fordul elő általában több ember, meghatározók lehetnek, továbbá az is, hogy hol találhatóak a felhasználók említésre, közlésre érdemes mondanivalót, mely utóbbi leginkább az idegenforgalmi látványosságok közelében lehet felülreprezentált.

3. ábra

Geotagelt Twitter-bejegyzések sűrűsége Budapesten



Forrás: Fischer (2013) alapján saját szerkesztés.

Geotagelt Twitter-bejegyzések aggregált száma a magyar kistérségekben

Forrás: Fischer (2013) adatai alapján saját szerkesztés.

A területi statisztikai vizsgálatok többsége nem egyedi, inkább területileg aggregált adatokkal dolgozik. Erre látható kísérlet a 4. ábrán, ahol kistérségek szerinti összesítésben tüntettük fel a Magyarország területén regisztrált Twitter-bejegyzéseket. A modell kialakítása során a fentebb már említett pont a poligonban algoritmusra épülő spatial join eljárást használtuk. A számítások eredményeként létrejött térkép jelen formájában a Twitter-bejegyzések abszolút kistérségi számait tükrözi vissza (a főváros és egyes nagyvárosok magától értetődő dominanciájával), de az adatokból könnyűszerrel készíthető lehet akár népességarányos, fajlagos mutató is.

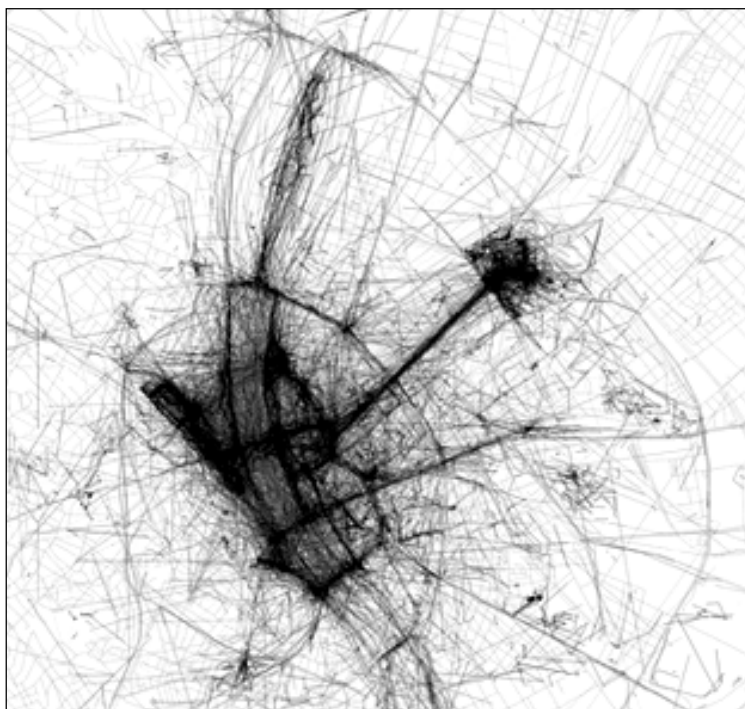
A geotaggal ellátott területi információk a felhasználók számára nyitott, azaz a sok felhasználós közösségi alkalmazások, illetve szabad szerkesztésű weboldalak esetében is visszatükrözhetik a területi információk használatának egyenlőtlenségeit. A Wikipédia honlapján bejegyzett geokódolt információk például már elegendően nagy számban állnak rendelkezésre ahhoz, hogy láthassuk, a világ mely térségei azok, amelyekről e tekintetben már több információ kering a világhálón, avagy melyek azok, amelyek nagyobb súllyal vannak jelen az internetes térben (Graham–Zook 2011). Ez esetben ismételtten arra a kettősségre kell gondolnunk, hogy az aktívabb területeken több közlendője is van a felhasználóknak, ugyanakkor általánosságban az is igaz, hogy a felhasználók nagyobb arányban tesznek közzé információkat a saját környezetükről, következésképpen általában valóban ott találjuk az információs kor jelentősebb számú és aktívabb résztvevőit, ahol ezek a térképek sűrűsödéseket mutatnak. Mindamelllett természetesen számos lokális ok is befolyásolhatja azt, hogy egyes területek aktívabbak, mások passzívabbak.

Térhasználati statisztikák

A geotagadatok alapján kialakított ábrák nemcsak a felhasználók aktivitásáról, illetve a területi szempontból érdelemes tartalmi információk közzétételének sűrűségéről tájékoztathatnak, de közvetett információkat adhatnak például az alkalmazók térhasználati szokásairól is. Erre lehet példa David Crandall és munkatársainak (2009) azon kísérlete, amelyben a felhasználók Flickrben (egy közösségi fotómegosztó oldalon) közzétett bejegyzései alapján sikerült felvázolni a fényképkészítők által bejárt legnépszerűbb útvonalakat. A Flickr weboldalán geotaggel és időbélyegzővel ellátott fotók adatainak adatbázisba szervezése a GPS nyomkövető eszközökhöz hasonló térpályák kirajzolását tette lehetővé. Az azonos felhasználók által időben nem kevesebb mint 30 perc elteltével közzétett fotók geokoordinátáit összekötve megrajzolhatóvá vált az adott személyek térpályája. A bejárt útvonalak tömeges felvázolásával végül kirajzolódtak a térhasználat legfrekvenciáltabb helyei.

Az 5. ábra Budapest esetében mutatja a Flickr-felhasználók által leggyakrabban felkeresett helyeket, a legtipikusabb térpályákat. Az ábrán általában a belvárosi zónák dominálnak, de jól kirajzolódik például a pesti és a budai Duna-part, az Andrásy út, a Városliget, továbbá népszerű helyként jelenik meg például a budai Vár vagy az összes Duna-híd is. A kirajzolódó mintázatok természetesen nem a valódi forgalomsűrűséget ábrázolják, és nem is azt, hogy mely területek aktívak az információk közzétételében, csupán a térbeli tartalommal rendelkező információk relatív sűrűségeit szemléltetik.

5. ábra

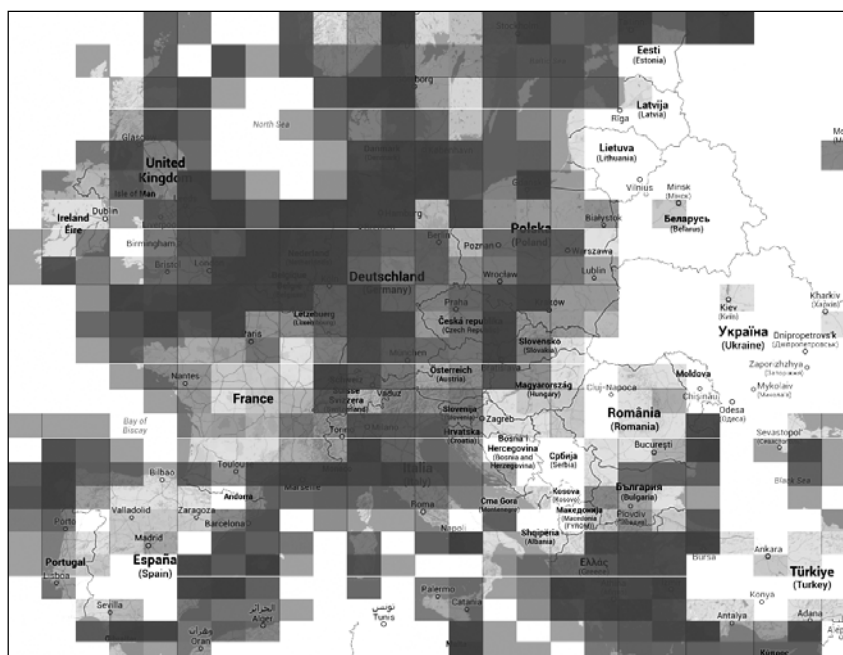
Flickr-bejegyzések alapján kirajzolódó térpályák Budapesten

Forrás: Fischer (2011).

Az új infokommunikációs eszközök egy jelentős része nemcsak információtovábbításra, de kifejezetten területi adatok gyűjtésére is alkalmas, sőt a felhasználók egyre növekvő hányada egyenesen keresi a térbeli információkat is kínáló alkalmazásokat. Bár a térbeli információhasználat lényegében eszközfüggetlen, mégis kiemelendők, illetve külön említést érdemelnek az okostelefonokhoz kötődő használati lehetőségek. Ezek az eszközök ugyanis azok számára is mindennaposá tették a térérzékeny adatok használatát, illetve annak lehetőségét, akik korábban nem sok ambíciót éreztek eziránt. Az okostelefonok az általában beépített helymeghatározó alkalmazásaikon keresztül aktív téradat-használati lehetőséget kínálnak, közelebb hozva ezzel a GPS-technológiák nyújtotta lehetőségeket a felhasználókhoz. A telekommunikációs fejlődési trendekhez az is hozzátartozik, hogy az elkülönült céleszközök (GPS-készülékek) piaca egyre szűkül, míg a mobil szoftveres piac bővül. Az Ipsos Zrt. 2010-es felmérése szerint Magyarországon a mobilkészülékek 22%-a rendelkezik már GPS-készülékkel, s a felhasználók több mint fele elképzelhetőnek tartja, hogy a mobilkészülékét használja három éven belül helymeghatározásra. A felmérés a lokalitást használó appok alkalmazásának piaci potenciálját is kimutatta, ugyanis a megkérdezett felhasználók kb. fele kisebb vagy nagyobb összeget fizetne is azért, hogy efféle szolgáltatást igénybe vehessen (Pintér 2011).

A GPS-alapú vagy egyéb helymeghatározó és nyomkövető alkalmazásokkal felszerelt készülékek a különféle alkalmazásokhoz kapcsolódóan helyadatok sokaságát generálják. Amint ezeket az adatokat nagy számban és több felhasználóra vonatkozóan területi adatbázisba is sikerül aggregálni, a puszta téradatok térhasználati vizsgálatokhoz és ezáltal statisztikai elemzésekhez is felhasználhatókká válnak. Az ilyen területi adatokat direkt módon leggyakrabban a közlekedésszervezés tudja hasznosítani, például a forgalmi terheléshez vagy az optimális útvonalválasztáshoz stb. kapcsolódó feladatokban, míg az információs társadalmi egyenlőtlenségek vonatkozásában indirekt módon csak arról kaphatunk képet, hogy a tér mely pontjait használják sűrűbben azok, akik ezeket az információtechnológiai megoldásokat igénybe veszik. Mivel ezeket a térpályákat nagyban meghatározzák az eleve kiépített vagy meglévő közlekedési csatornák (például utak), kis léptékben a frekvenciált térhasználati helyek sűrűsödése lényegében nem sokat mond az információs társadalmi aktivitásról. Nagyobb területi aggregációkban, regionális, országos vagy nemzetközi összehasonlításban azonban a különbségek összeadódása révén már szignifikáns területi differenciák kirajzolódására számíthatunk. Az „információtermelés” szempontjából aktív és passzív térségek meghatározására így az információs térhasználati statisztikák inkább közvetett segítséget nyújthatnak. A 6. ábra azt szemlélteti, miként lehet területi adatokká átalakítani az egyedi felhasználói téradatokat. A világméretű nagyszámú felhasználó által igénybevett automatikus információs csomagkövető rendszerek (*Automatic Packet Reporting System, APRS*) folyamatosan képesek rögzíteni a felhasználói pozíciókat. Az egyedi helyzeti adatokat gridtranszformációs eljárással 140x110km-es cellákba aggregálva raszterizált területi adatok képezhetők, amelyek ezután tematikus sűrűségi térképeken is megjeleníthetők. A kialakított ábrák, illetve területi statisztikai adattáblák ezáltal képesekké válnak a nagytérségi egyenlőtlenségek nyomon követésére is.

*Az APRS-felhasználók számának sűrűsödése Európa térségeiben
2013. augusztus 3-án (140x110km-es cellákban)*



Forrás: aprs.fi.

A közösségi hálók területi statisztikái

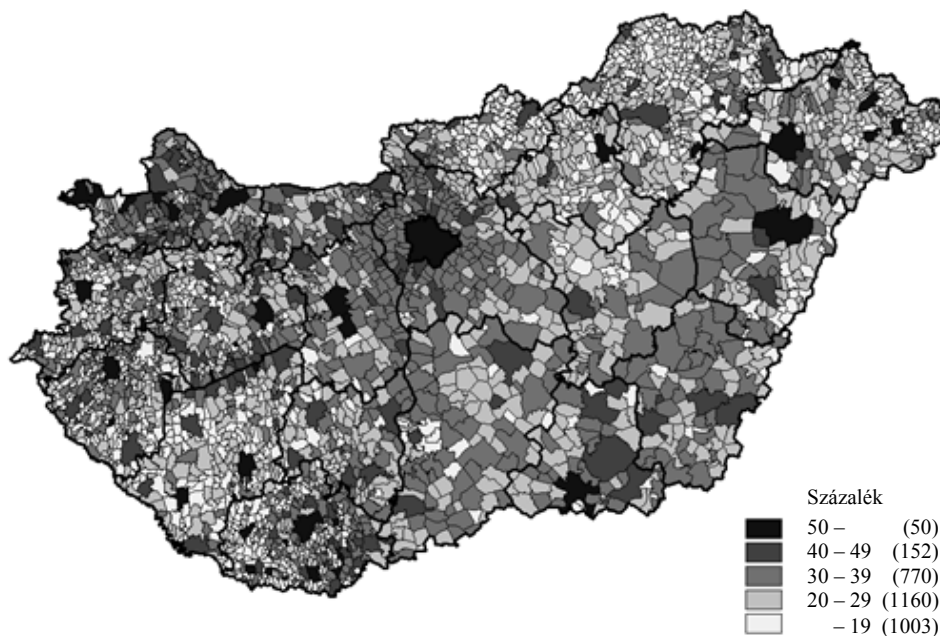
Az online közösségi hálók (OSN) népszerűsége az utóbbi időkben nagyjából töretlen. Ez a népszerűség a területi kutatói szakma számára különösen hasznos, mivel az OSN-ek az internetet használó lakosság igen nagy részére vonatkozóan képesek adatokat szolgáltatni. Az OSN kutatások a mai „big data” alapú vizsgálatok fókuszpontjában állnak. Habár a közösségi oldalakon való jelenlétnek számos különféle motivációs oka lehet, a kimondottan magas felhasználói létszámok miatt az extremitások esélye viszonylag alacsony, következésképpen az OSN-ek adatbázisai jó közelítők lehetnek a valódi információs társadalmi jelenlétnek. Természetesen a jelenlét és az aktív jelenlét között is különbség tehető, illetve nem feltétlenül igaz az sem, hogy aki nincs jelen a közösségi fórumokon, az ne lenne jelen az információs térben, mindazonáltal az online közösségi hálók adatbázisainak elemzése tekinthető az internetes világban való jelenlét számszerűsítését célzó egyik legjobb közelítő módszernek.

Manapság az OSN-szolgáltatást nyújtó weboldalak köre már igen tág, több versenytárs több különböző profillal igyekszik magához vonzani a felhasználókat. Az online közösségi hálók közül bármelyiket is vizsgáljuk, általában hasonló területi statisztikai jellegzetességekre figyelhetünk fel, ami ismételten azt jelzi, hogy ez az adatforrás az információs kor területi egyenlőtlenségeinek jó reprezentánsa. A 7. ábra például a Magyarországon a

közelmúltban nagy népszerűségnek örvendő iWiW közösségi háló felhasználóinak települési részarányait mutatja (az IBS és az Origo Zrt. adataira építve). A 2013. januári állapotokat mutató ábra összesen több mint 4 millió felhasználó adatai alapján készült. Az ábra szemléletesen mutatja a városias körzetek, a főváros környéki agglomeráció, valamint például a Balaton környékének magasabb felhasználói arány értékeit, ami az információs társadalmi jelenlét magasabb, előrehaladottabb helyi szintjére is enged következtetni (annak ellenére is, hogy a területi adatok a felhasználók saját bevallásai alapján generálódtak és bizonyos vonatkozásban torzítottak is lehetnek).

7. ábra

Az iWiW közösségi háló felhasználóinak a települések lakosságszámához viszonyított aránya, 2013. január



Forrás: az IBS és Origo Zrt. adatai alapján saját szerkesztés.

Az iWiW közösségi háló népszerűsége a versenytársaknak, köztük legfőképp a Facebooknak köszönhetően az utóbbi években lassú csökkenést mutat. Mára Magyarországon is a globális szinten uralkodó Facebook vált a legelterjedtebb közösségi hálóvá. Területi értelemben ugyanakkor ezen OSN kapcsán is a fent említett térségi trendek kirajzolódására számíthatunk.

Használói profil statisztikák

Már az online közösségi hálók felhasználói adataiból generált adatbázisok esetében is ott a lehetőség, hogy ne csak a felhasználók létszámáról, de egyéb karakterisztikáiról (kor, nem stb.) is területi információkhoz juthassunk. Egyre komolyabb az igény annak meghatározására, illetve mérésére iránt, hogy valójában kik, mely társadalmi csoportok használják az

egyres virtuális térbeli szolgáltatásokat, kikhez jutnak el az egyes konkrét információs tartalmak, továbbá milyen technológiai profilkok mellett férnek hozzá a felhasználók az egyes tartalmakhoz.

2. táblázat

A Facebook-felhasználók százalékos megoszlása nem és korcsoportok szerint Európa néhány országában a 13 évesnél idősebbek körében, 2013

Ország	Nemek		Korcsoportok								össze- sen
	nők	férfiak	13–15	16–17	18–24	25–34	35–44	45–54	55–64	65–x	
Ausztria	49	51	4	8	26	28	17	11	4	2	100
Bulgária	51	49	5	7	24	29	20	9	4	2	100
Csehország	52	48	5	7	27	28	18	8	4	3	100
Észtország	54	46	5	6	24	29	19	10	5	2	100
Horvátország	48	52	4	8	28	32	16	7	3	2	100
Lengyelország	52	48	11	9	29	29	13	5	3	1	100
Lettország	54	46	4	5	29	32	16	8	4	2	100
Litvánia	54	46	6	9	30	30	14	7	3	1	100
Magyarország	52	48	6	7	23	24	20	10	7	3	100
Németország	48	52	6	7	25	28	16	12	4	2	100
Románia	50	50	7	7	29	30	17	6	3	1	100
Szerbia	46	54	6	10	31	28	16	6	2	1	100
Szlovákia	52	48	5	8	27	29	17	8	4	2	100
Szlovénia	49	51	4	7	25	30	19	9	4	2	100
Ukrajna	53	47	4	4	26	36	17	8	3	2	100

Forrás: a socialbakers.com alapján saját szerkesztés.

8. ábra

Domináns mobilböngésző típusok Európában, 2013. április



Forrás: <http://gs.statcounter.com/>.

 ÚJSZERŰ TERÜLETI STATISZTIKAI ADATGYŰJTÉSI LEHETŐSÉGEK AZ INFORMÁCIÓS VILÁGBAN 51

A felhasználói profilt leíró minőségi paraméterek a penetrációs statisztikák helyett az infokommunikációs szolgáltatásokat igénybe vevők köréről adnak tájékoztatást, ezek pedig szerteágazó társadalmi-gazdasági vizsgálatok alapjául szolgálhatnak. A nemek vagy korcsoportok szerinti adatok területi eltérései (nők, idősek stb. részvételi aránya) a digitális egyenlőtlenségek alakulásáról tájékoztathatnak (2. táblázat). A használati eszközök minőségi technológiai paraméterei pedig többek között indirekt módon a gazdasági környezet helyi sajátosságait tükrözhetik vissza a felhasználói profilokban (lásd a Nokia dominanciáját Finnországban stb.; 8. ábra). Az egyes weboldalak üzemeltetői már a kezdetektől nagy arányban figyelik honlapjaik látogatottsági adatait. Ezekből a statisztikákból sok esetben már nem csak az derül ki, hogy hányan látogattak egy-egy oldalra, de az is, hogy kik, sőt az is, hogy honnan. Az így keletkező nagyszámú adatról a weboldalak látogatói mit sem tudnak, a „másik oldal”, az üzemeltetők és elemzők csoportja viszont eddig soha nem látott információkhoz juthat.

Összegzés

A fent ismertetett példák és gondolatok mind az irányba mutatnak, hogy az információs világ területi különbségeinek mérésére vonatkozó kísérleteinket időről időre érdemes módszertanilag átgondolnunk, finomítanunk, kiegészítenünk. A digitális megosztottság mára egyértelműen túlmutat a csak technológiai típusú elemek elterjedtségének egyenlőtlenségein, bár napjainkban is folyamatosan kerülnek elő olyan lényeges meghatározottságú eszközök, amelyek területi diffúziós és penetrációs jellemzőit még fel kell tárnunk. Komoly igény mutatkozik például az okostelefonok elterjedtségének vizsgálatára, de új mérendő tényezőként már nem csak a telefon birtoklását, de a használat minőségét, a letöltött alkalmazások számát, típusát, vagy éppen az eszközhasználati időt kellene figyelembe vennünk. Mindezek alapján már olyan új kérdésekre is választ kaphatunk, mint hogy hol található az információ termelése szempontjából aktív és passzív térségek, vagy hogy milyen fokú az egyes területek elmélyültsége, integráltsága a virtuális világba.

A területi statisztikai adatokat feldolgozó, elemző szakma számára nemcsak esély, de feladat is a régi, hagyományos közelítésű hozzáférési, használati adatok és az új adatok összekapcsolása. Bár az újfajta adatgyűjtési módszerek nem feltétlenül automatizálhatók, illetve szabványosításuk sem minden esetben oldható meg, mégis valós alternatívát adhatnak az új folyamatok kvantifikált leképezésére. A hagyományos és új szemléletmód együttes használata egyre inkább realitás lehet.

IRODALOM

- Crandall, D. – Baskstrom, L. – Huttenlocher, D. – Kleinburg, J. (2009): *Mapping the World's Photos* Proceedings of WWW 2009, April 20–24, Madrid.
- Dőry Tibor – Ponáczy György Márk (2003): Az infokommunikációs ágazatok szerepe és súlya a magyar városhálózatban *Tér és Társadalom* 17 (3): 165–181.
- European Commission (2011): *Broadband coverage in Europe in 2011. Mapping progress towards the coverage objectives of the Digital Agenda*. RESEARCH REPORT A study prepared for the European Commission DG Communications Networks, Content & Technology, Point Topic London. http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1102 (letöltve: 2013. november)
- Fischer, E. (2011): *The Geotaggers' World Atlas*. <http://www.flickr.com/photos/walkingsf/4621754005/in/set-72157623971287575> (letöltve: 2013. november)

- Fischer, E. (2013): *Locals and Tourists map*. Eric Fischer, Gnip, MapBox project. <http://mapbox.com/labs/twitter-gnip/locals> (letöltve: 2013. november)
- Graham, M. – Zook, M. (2011): Visualizing Global Cyberscapes: Mapping User Generated Placemarks *Journal of Urban Technology* 18 (1): 115–132.
- Hargitai H. (2001): Magyar honlapszolgáltatók és felhasználóik területi megoszlása. <http://www.puskas.matav.hu> (letöltve: 2002. február); másolatban elérhető: <http://jakobi.web.elte.hu/puskashirmondo.htm> (letöltve: 2014. január).
- ITU (2013): *The World in 2013. ICT facts and figures*. <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf>
- Jakobi Ákos (2002): A területi egyenlőtlenségek új elemei az információs társadalomban. *Regionális Tudományi Tanulmányok* 7., pp. 55-84. ELTE Regionális Tudományi Tanszék, Budapest.
- Johoka, H. (2000): *Computer age – Informatization White Paper*. Japan Information Processing Development Center, Tokyo.
- Kiss Gergely (2009): *A Nemzeti Digitális Közmű – mint válasz az alapproblémára*. MKE 41. Vándorgyűlés. Tihany, 2009. május 27. vandorgyules.lib.unideb.hu/sites/titan.../GKIE.NET_KisGergely_ETK.ppt (letöltve: 2013. november)
- Malecki, E. J. (2002): The Economic Geography of the Internet's Infrastructure *Economic Geography* 78 (4): 399–424.
- Mészáros Rezső (2003): *Kibertér – A földrajzi tudás új dimenziói*. Hispánia Kiadó, Szeged.
- Nagy Gábor (2002): Területi különbségek az információs korszak küszöbén (Mit mérünk, és hogyan?) *Területi Statisztika* 42 (1): 3–25.
- Nagy Gábor – Kanalas Imre (szerk.) (2003): *Régiók az információs társadalomban*. MTA RKK ATI, Kecskemét.
- Nagy Réka (2007): Új lencsék egy társadalmi jelenség vizsgálatában: a digitális egyenlőtlenségek kutatásának átfogó szemléletéről *Szociológiai Szemle* (1-2): 16–28.
- Norris, P. (2001): *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide* Cambridge University Press, Cambridge.
- Pintér Róbert (2011): *Álmodnak-e a telefonok elektronikus bárányokkal?* <http://www.slideshare.net/hirbehozo/pinter-rbert-mobil-trendek> (letöltve: 2013. november)
- Preis, T. – Moat, H. S. – Stanley, H. E. – Bishop, S. R. (2012): Quantifying the Advantage of Looking Forward *Scientific Reports* 2, Article number: 350 <http://www.nature.com/srep/2012/120405/srep00350/full/srep00350.html> (letöltve: 2013. november)
- Ramachandran, R. (1999): *The Knowledge Imperative Index* NITC, Kuala Lumpur.
- Rechnitzer János – Grosz András – Csizmadia Zoltán (2003): A magyar városálózat tagozódása az infokommunikációs infrastruktúra alapján az ezredfordulón *Tér és Társadalom* 17 (3): 145–163.
- Selwyn, N. – Facer, K. (2007): *Beyond the digital divide - Rethinking digital inclusion for the 21st century* Futurelab, Slough, Egyesült Királyság.
- Tóth Péter (2011): *Magyar települések az információs társadalomban* Doktori értekezés. Széchenyi István Egyetem, Győr.
- Tranos, E. – Gillespie, A. (2009): The Spatial Distribution of Internet Backbone Networks in Europe. A Metropolitan Knowledge Economy Perspective *European Urban and Regional Studies* 16. (4): 423–437.
- Vartanova, E. – Smirnova, O. (szerk.) (2012): *Digital Divide 2011*. Yearbook. MediaMir, Moszkva, Oroszország.
- Zook, M. A. (2000): The web of production: the economic geography of commercial Internet content production in the United States. *Environment and Planning A* 32 (3): 411–426.

Kulcsszavak: információs társadalom, térinformatika, területi egyenlőtlenségek, big data, geotag.

Resume

This paper tries to formulate suggestions on how to use quantifiable data for the spatial analysis of the information society, in order to reflect the newest changes of the information world, and to explain new spatial inequalities. Until now information society analyses applied mainly traditional regional data of statistical offices and institutes. By the emergence of the so-called big data phenomenon and by the application of modern analytical techniques of GIS a new chance is given to answer questions, which never came up before. The paper introduces new spatial statistical tools for the analysis of information infrastructure, web statistics, content statistics, spatial trajectory analysis, social network analysis and user profile statistics.