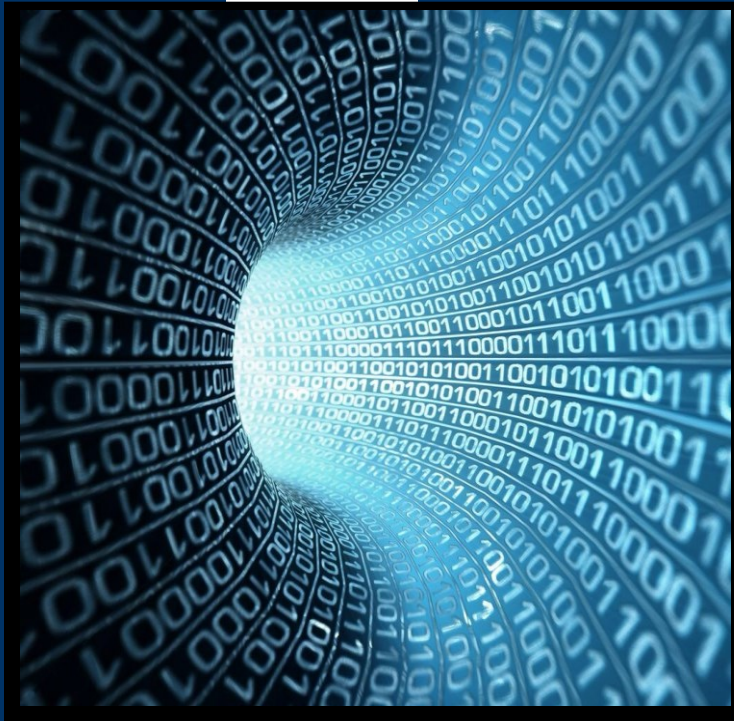


Jakobi Ákos



# Az információs kor területi kutatásának új forrásai: a „big data”

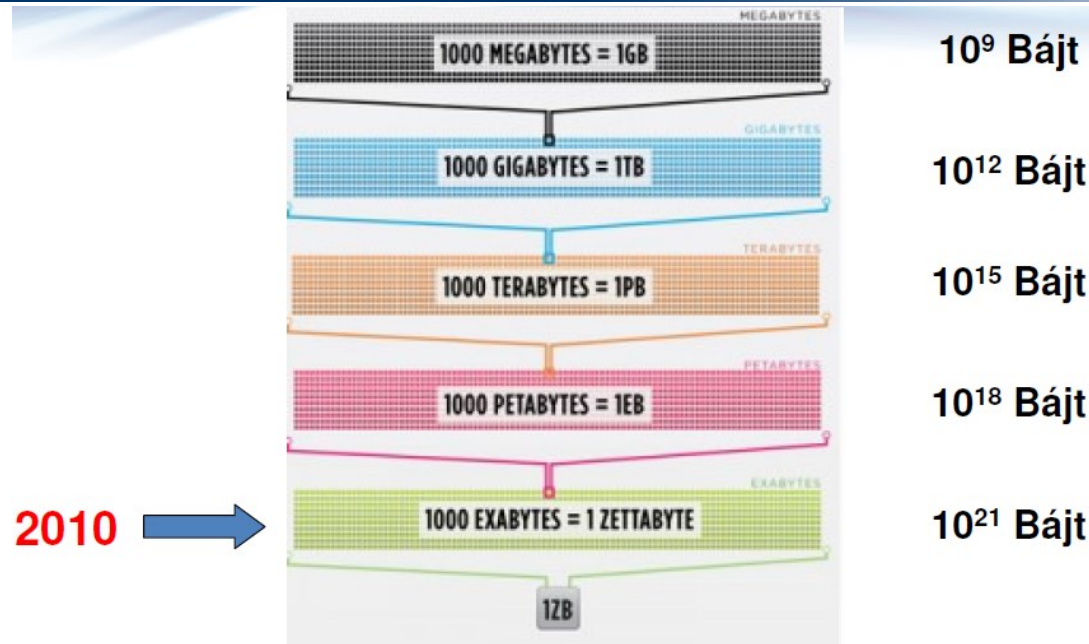
ELTE Regionális Tudományi Tanszék,  
Budapest

# Témafelvetés

- Információs fejlődés egyik következménye: a területi információk iránti igények jelentős mértékű megnövekedése
- Egyre több olyan eszköz és alkalmazás lát napvilágot, amely direkt vagy indirekt módon épp ezen területi információk hasznosítója vagy előállítója
- Milliónyi digitális nyom, mint feldolgozható új adat



# Mi az a big data?



- A globális adattömeg másfél évente megkétszereződik
- 2013-ra az Internet éves adatforgalma eléri a 667 Exabájtot (Cisco idézi Németh V. 2012).
- 2016-ra a globális mobil éves adatforgalom meghaladja a 130 Exabájtot (Cisco idézi Németh V. 2012).

# Mi az a big data?

- A „big data” (szabad fordításban „óriási adathalmaz”) arra a hatalmas adatmennyiségre utal, amely információs világunkban gyors tempóban és folyamatosan keletkezik
- Feldolgozása a hagyományos kapacitásokkal és eljárásokkal operáló módszerekkel már-már megoldhatatlan kihívást jelent
- Nagy mennyiségű, komplex adat
- Bármilyen strukturált vagy strukturálatlan adat, pl. szöveges, szenzor adatok, audio, video, streaming, log fájlok, stb.
- Nem csak mennyiségi értelemben vett adatrobbanás, de minőségi is
- A nagy adatbázisok jellemzői:
  - Nagy adatmennyiség
  - Változatosság, az adatfajták és források sokfélesége
  - Sebesség: az adatok keletkezésének gyorsaságára, a hasznosításhoz rendelkezésre álló időre vonatkozóan



**BIG DATA**

**VAST DATA**

Unstructured data  
Semi-Structured data  
Volume  
Useful Metadata  
Data analysis  
Petabytes  
Unstructured content  
People driven  
Decision making  
Zettabyte  
Framework  
Smart content database  
Text analytics  
Concept extraction  
Semantic Metadata  
Structured data

# Mi az a big data?

- A sokáig csak virtuális melléktermékként számon tartott napi információhalom épp akkor válik értékessé, amikor a különböző adatokat sikerül összekötni, köztük összefüggéseket, felismerhető mintázatokat találni, s mindebből értékelhető következtetéseket levonni.
- A társadalomtudományok számára valóságos aranybánya (vásárlási, munkabajárási, közlekedési és egyéb szokásaink adatai)
- Betekintést nyújt az emberi viselkedés egyedi és társadalmi szintjeibe
- Nagyságrendnyi előrelépést jelent az adatvolumenben, kellően nagy mennyiségű adat ahhoz, hogy a véletlen hatások szerepe mérséklődjön

# A Big Data nem csak az információs társadalom területi kutatásában hasznosítható

- Kormányzati statisztikák...
- Közlekedés...
- Üzleti világ, közvetlen fogyasztói statisztikák, kereskedelem...
- Egészségügy...
- Pénzügy, bankszektor...
- Biotechnológia...
- Tudomány...



# A digitális nyomok és a „surveillance”

- information society = transparent society
- Eltolódás az eseti megfigyelésektől a térben és időben folytonos megfigyelésig
- Majdnem minden digitális interakció és tranzakció nyomokat hagy (növekvő számban)
- Létezik egy, a fizikai test árnyékában meghúzódó „adattest” (digital body), ami nem csak követ, de esetenként meg is előz bennünket. Mielőtt megérkezünk valahova már gyakran fel vagyunk mérve és osztályozva vagyunk. (Felix Stalder, Privacy is not the antidote to surveillance, 2002)

# Néhány példa ízelítőül

- Az áramszolgáltató tudja, hogy a fogyasztó reggel mikor kel fel és este mikor fekszik le.
- Digitális TV szolgáltató: tudjuk mit nézel
- GPS, adatrögzítők az autóban: tudjuk merre jártál
- Fizetőkapuk: tudjuk mikor voltál ott
- Weboldal statisztikák: preferenciák modellezése



# Utazási kártyák

- az anonim utazás vége
- fizess többet, ha azt akarod, hogy ne kövessenek

# Digitális nyomok

- Bankkártya-  
ujjlenyomat  
(vásárlói profil)



- Password-ök,  
beléptető rendszerek,  
ujjlenyomat-leolvasók



# Az információs kor területi vizsgálatának hagyományos statisztikái

- Összetett mutatókészletek (Digital Access Index, Digital Opportunity Index, ICT Development Index, Network Readiness Index)
- Egyedi adatokat feldolgozó kísérleti modellek (Nagy és Kanalas 2003: domain név szerverek, ill. a domain regisztrációk számának területi mutatói)
- térségek vagy települések információs társadalmi pozícióit és egyenlőtlenségeit vizsgáló modellkísérletek (Jakobi 2007, Tóth 2013)
- Lényegesebb magyar adatszolgáltatók: NHH (ma NMHH), GKIeNET, KSH



# Az információs kor területi vizsgálatának hagyományos statisztikái

- személyi számítógépek száma (GKIeNET)
- mobiltelefon előfizetések száma (GKIeNET)
- telefonvonalak és ISDN vonalak száma (KSH)
- kábeltelevízió előfizetések száma (KSH)
- az e-ügyintézés szintje (GKIeNET)
- internet használók száma (GKIeNET)
- vállalati honlapok aránya (GKIeNET)
- internet előfizetések száma (GKIeNET)

# Új kihívások

- Az információs világ területi különbségeinek mérésére ezidáig lényegében csak érintőlegesen használhattunk azokat a jelzőszámokat, amelyek az információk kvalitatív részeire vonatkoznak
- Új kérdések:
  - Miként lehetne az információs társadalom területi különbségeinek mérésében szükségessé váló kvalitatív jellemzőket is kvantifikálni?
  - Honnan nyerhető adat/információ az információ-használatról, az információ-minőségről stb.?
  - Honnan szerezhetők területi információk?

# A digitális nyomok adatszerű felhasználása

- Social (human) sensors
- Térhasználati információk
- A közzétett tartalom területi statisztikái
- Az online jelenlét területi statisztikái
- A használói profilokról szerezhető területi információk

# Hol használják az emberek a teret?

## (Egyéni és csoportos térpályák vizsgálata)

Google latitude

[Dashboard](#) **Beta**

[Manage history](#)

[Settings](#)

History is Enabled

Date Range:

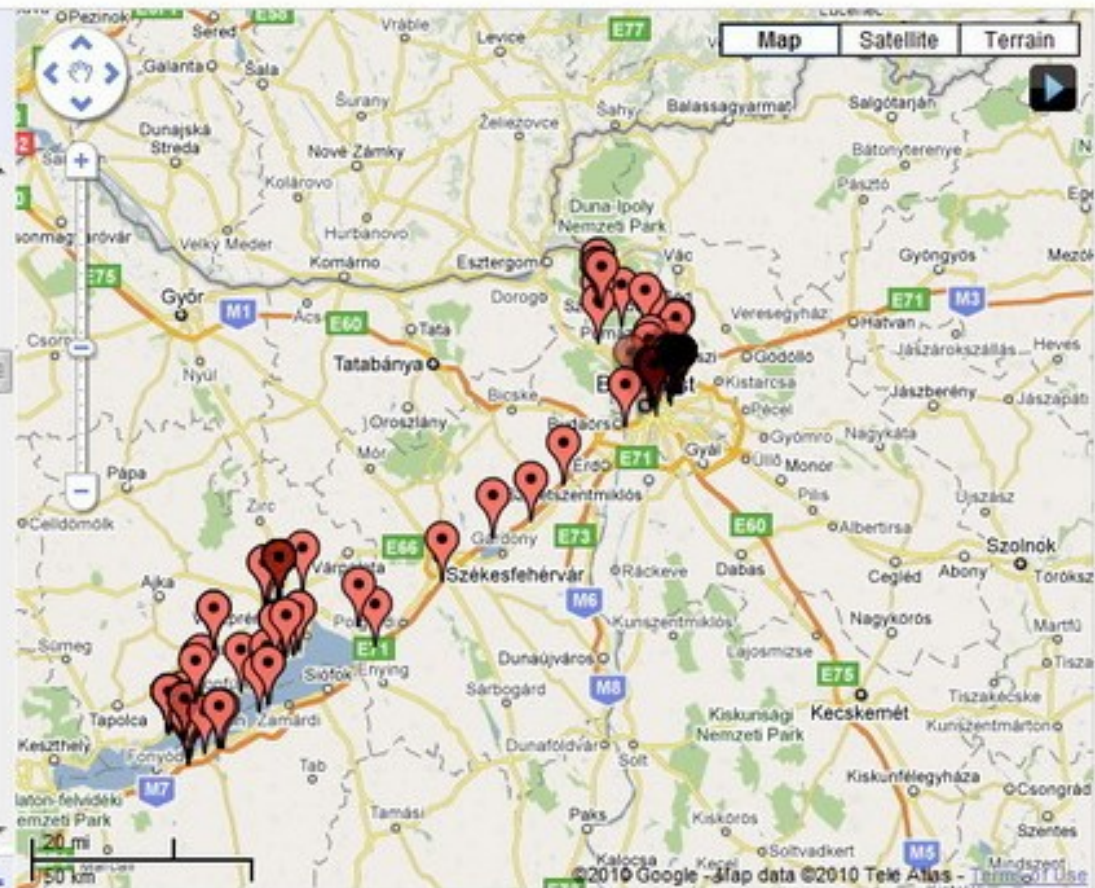
05/26/2010 - 06/06/2010

[Apply](#)

11:11 AM Zánka  
11:26 AM Balatonszemes  
11:39 AM 8638 Balatonlelle, Kossuth Lajos ...  
11:47 AM Kövágóörs  
11:51 AM Kövágóörs  
12:10 PM Kövágóörs  
12:55 PM Kövágóörs  
1:30 PM 8630 Balatonboglár, Erzsébet St 3...  
1:39 PM Kövágóörs  
2:08 PM 8630 Balatonboglár, Erzsébet St 3...  
3:04 PM Kövágóörs  
3:22 PM Balatonszemes  
3:31 PM 8630 Balatonboglár, Kórház St  
3:40 PM Balatonszemes  
3:56 PM Balatonszemes  
4:05 PM 8630 Balatonboglár, Kórház St  
4:08 PM Balatonszemes  
4:19 PM Balatonszemes  
4:28 PM Balatonszemes  
4:54 PM Kövágóörs  
5:04 PM Kövágóörs  
5:28 PM Ordacsehi  
5:38 PM Ordacsehi  
6:20 PM Kövágóörs

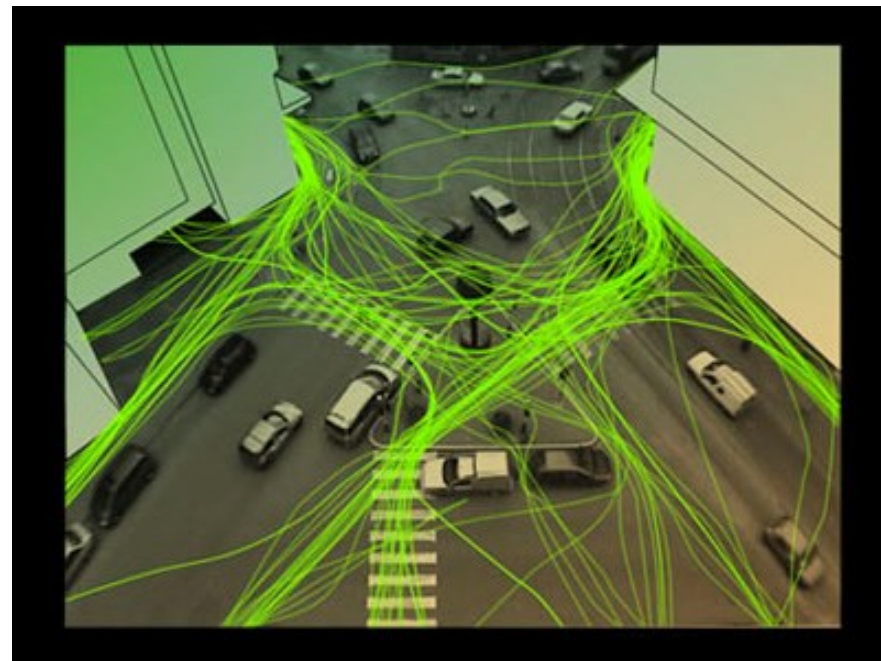
[Delete Options](#) ▼

[Export to KML](#)



©2010 Google - Map data ©2010 Tele Atlas - Terms of Use





**Városi  
térpályák**



# Csoportos térpályák „big data” adatok alapján

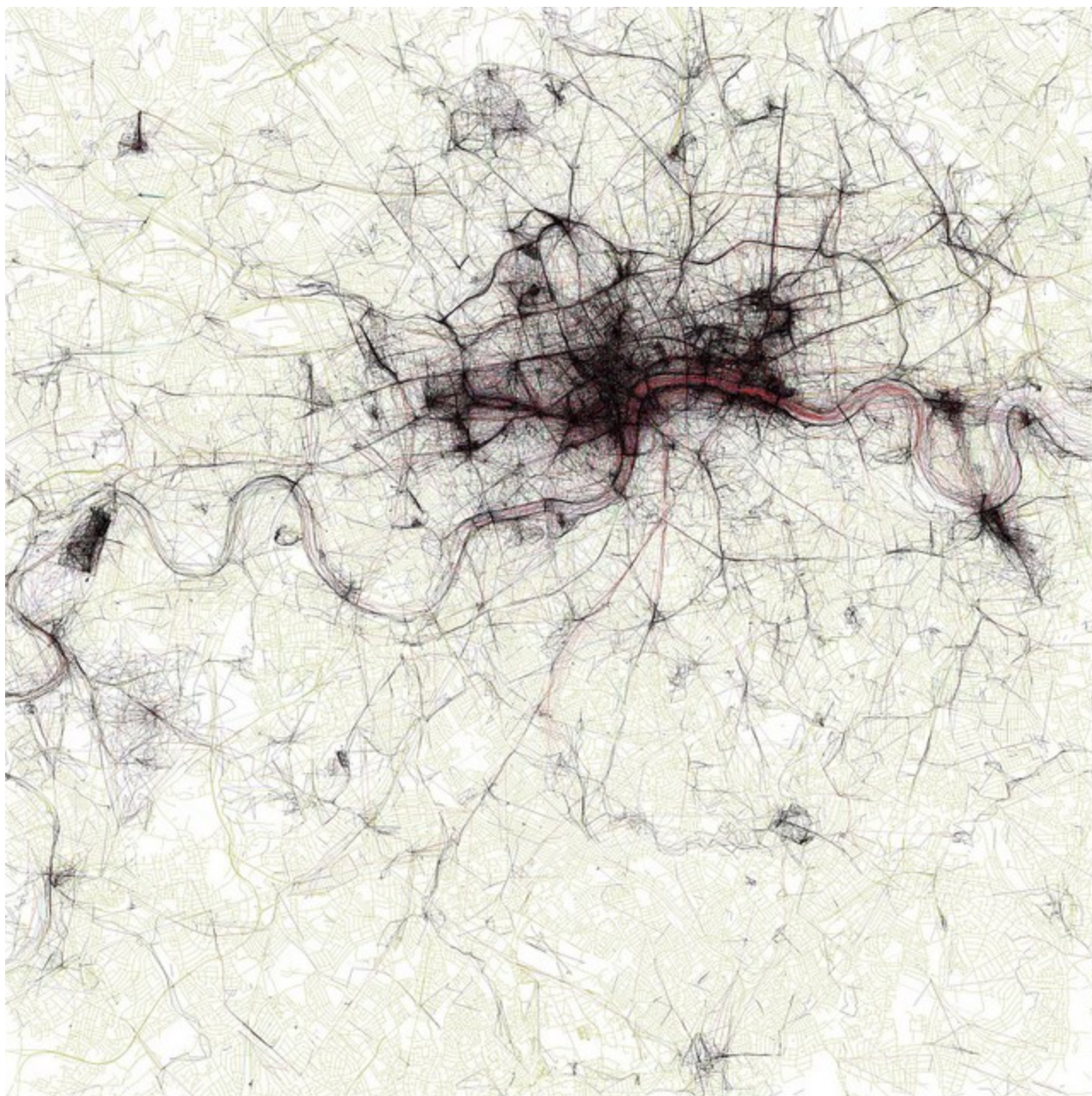


Fotósok mozgási útvonalainak területi mintázata Manhattan déli részén  
(Crandall et. al., 2009)



Párizs



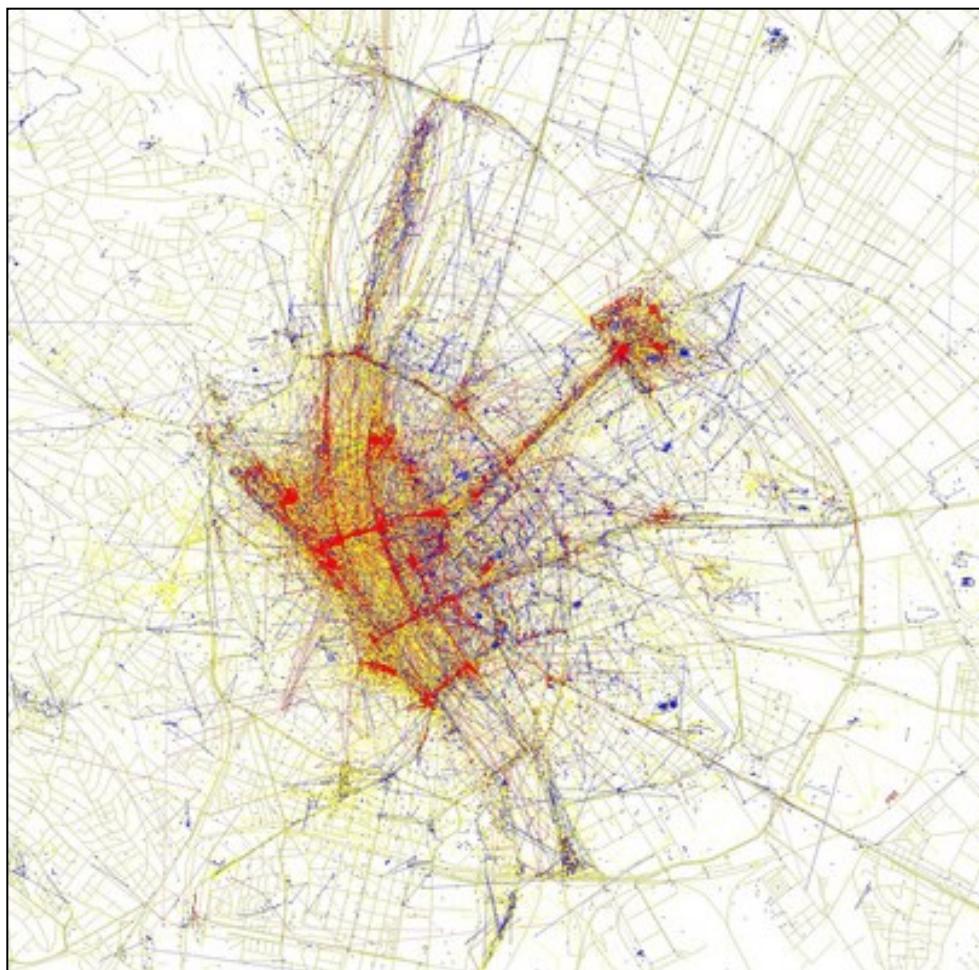


London



Budapest

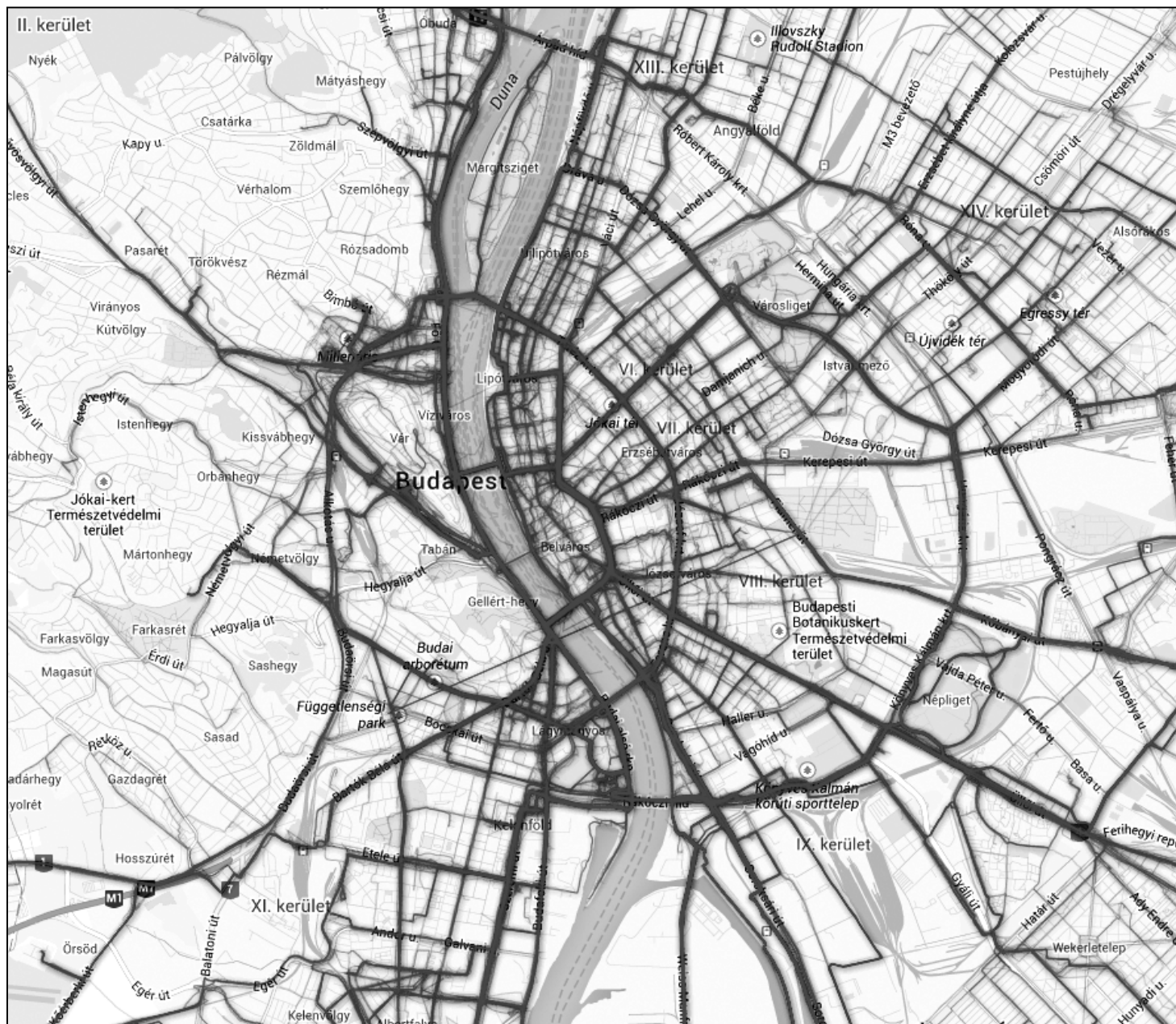




Flickr bejegyzések alapján kirajzolódó térpályák Budapesten  
(Forrás: Fisher (2011))



# Kerékpáros útvonalhasználat – kirajzolódó térpályák



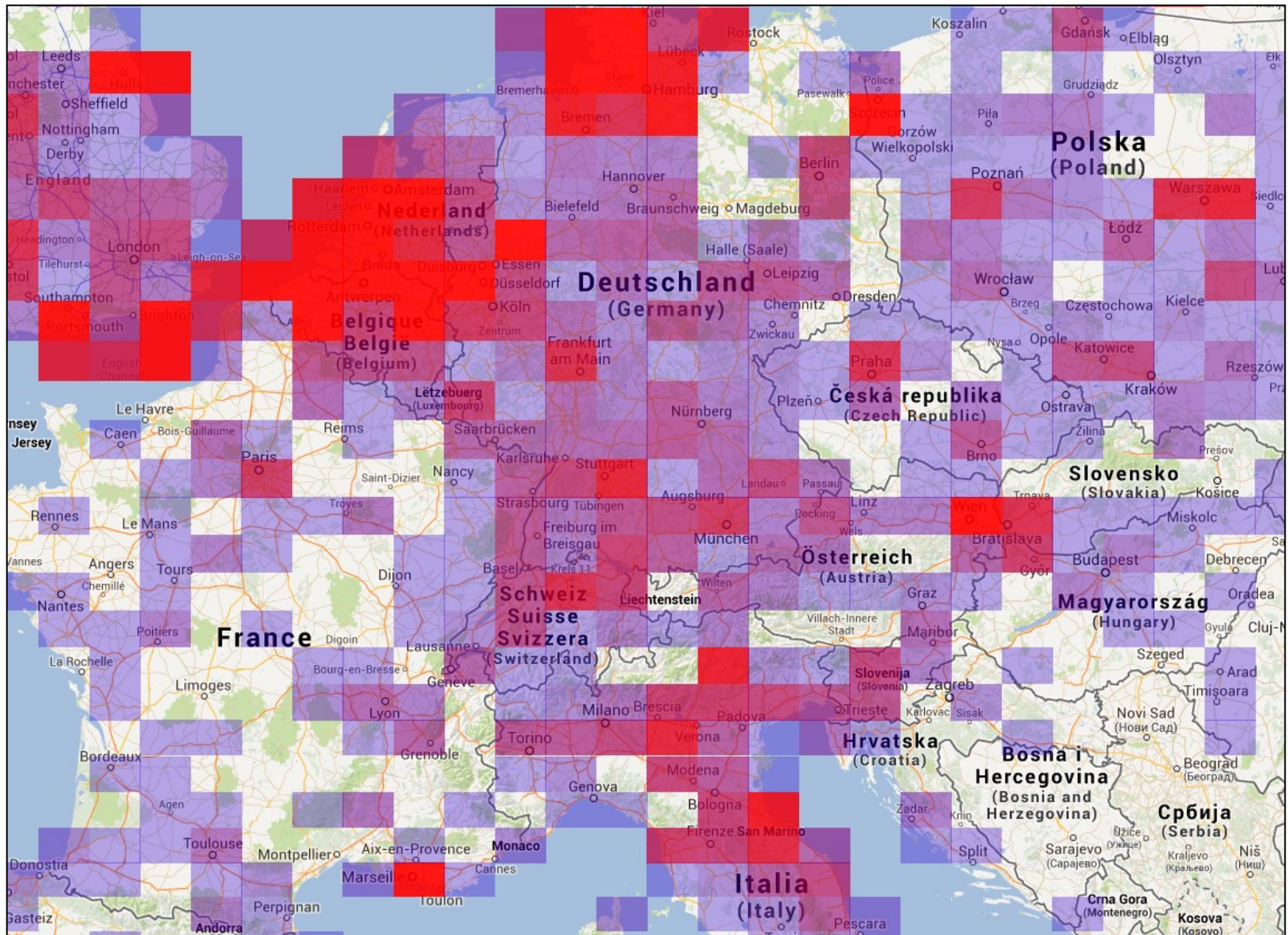
# Kerékpáros útvonalhasználat – kirajzolódó térpályák







# APRS (Automatic Package Recording System)



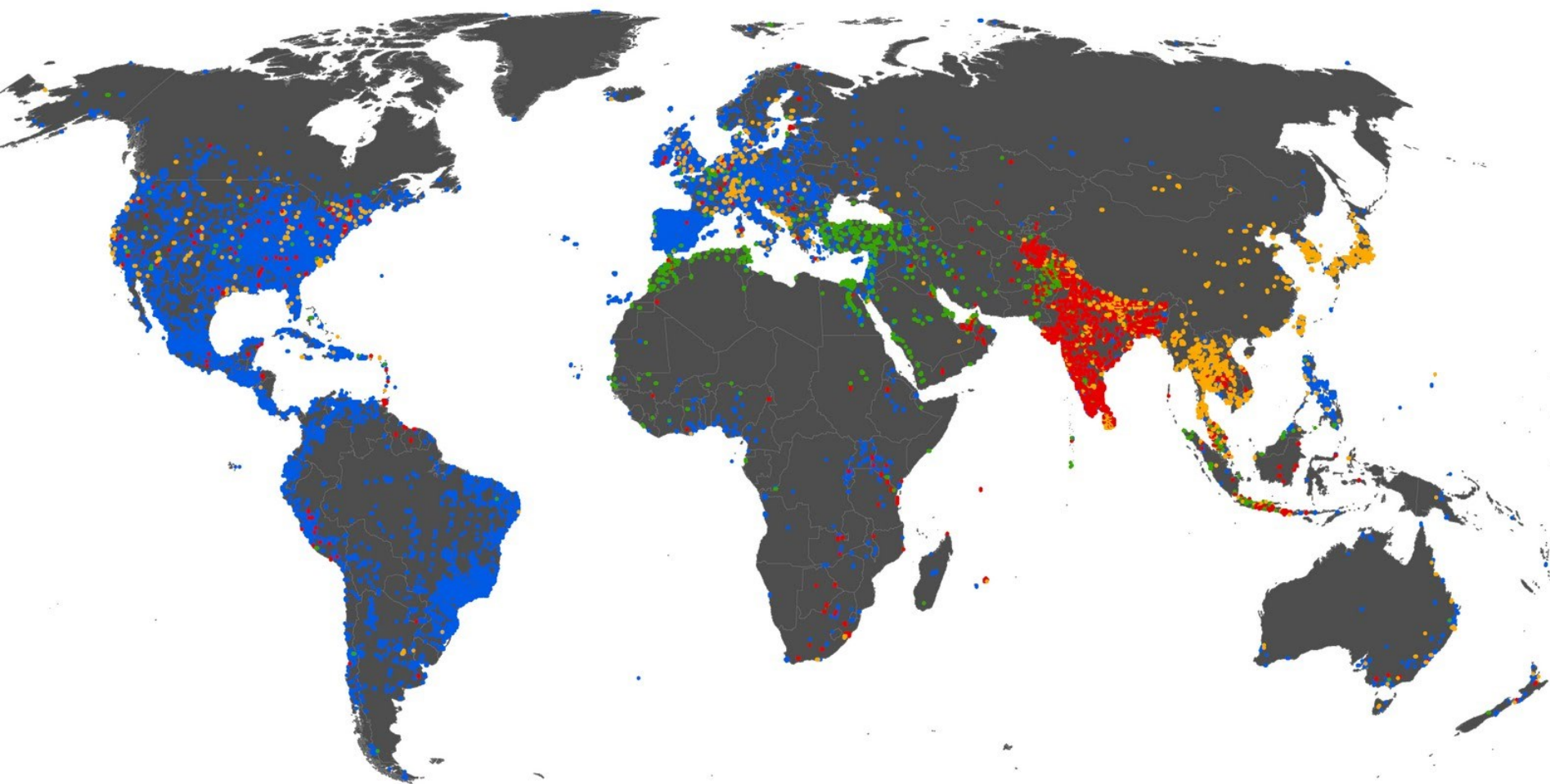
Az APRS felhasználók számának sűrűsödése Európa térségeiben 2013.11.20.



# Preis et al. (2013) vizsgálata

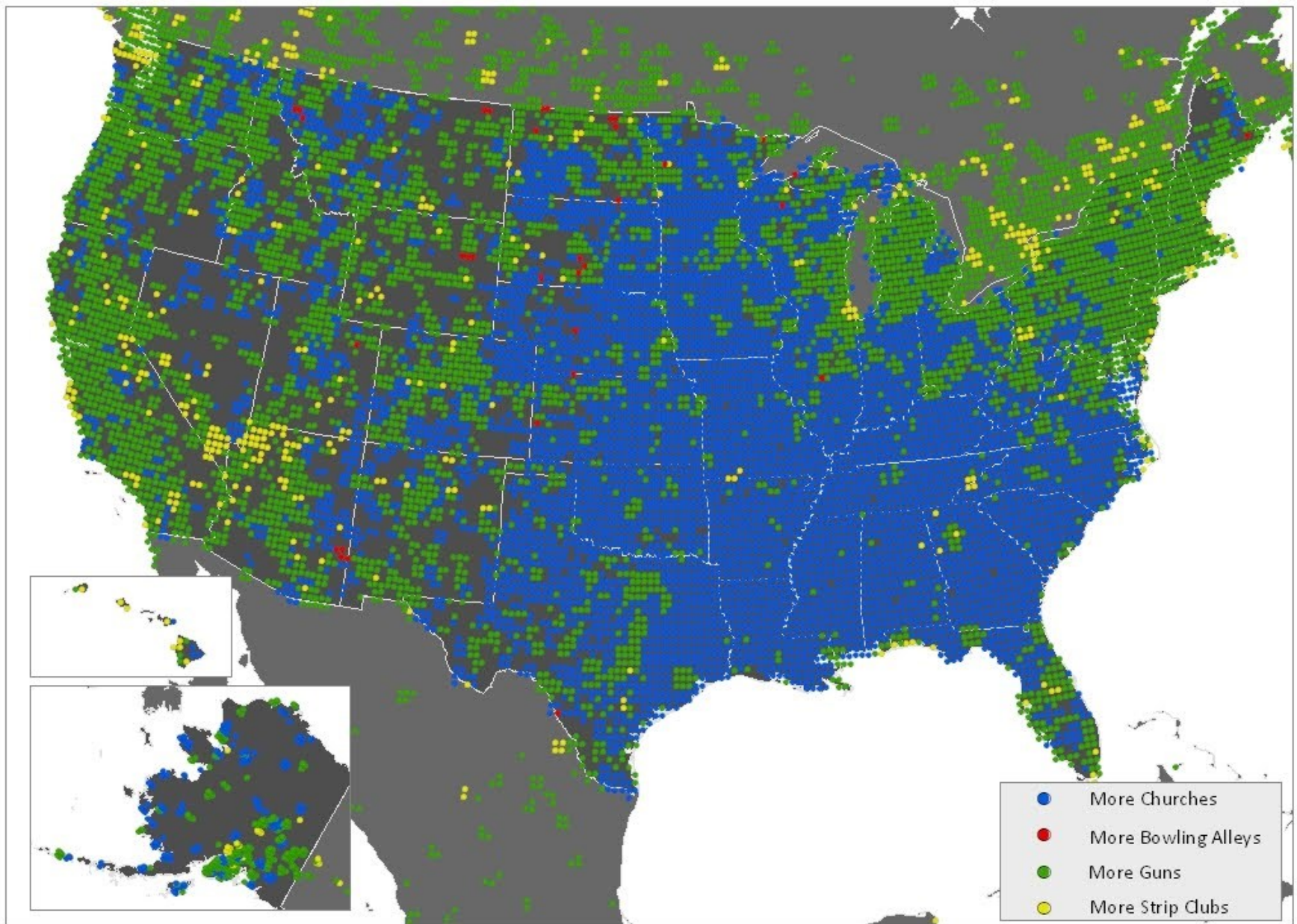
- Vizsgálatuk bebizonyította, hogy big data környezetben az internetes keresési szokásokból is stabil összefüggésekre juthatunk
- Google keresési trendek adatai alapján
  - magasabb egy főre jutó GDP-vel rendelkező országokban élő internethasználók inkább a jövőre, mintsem a múltra vonatkozó információkat keresnek
  - a jobb gazdasági teljesítményű országokban az állampolgárok információ-keresési aktivitása vélhetően nagyobb
- Az online viselkedés és a valódi világ gazdasági mutatószámai között összefüggések lehetnek

# Online keresési trendek



More Allah   More Buddha   More Hindu   More Jesus

Shading represents more references to one search term than to all others, e.g. a green circle indicates more references to Allah than to Buddha, Hindu or Jesus.  
Data collected in June 2009. [Floatingsheep.org](http://floatingsheep.org)

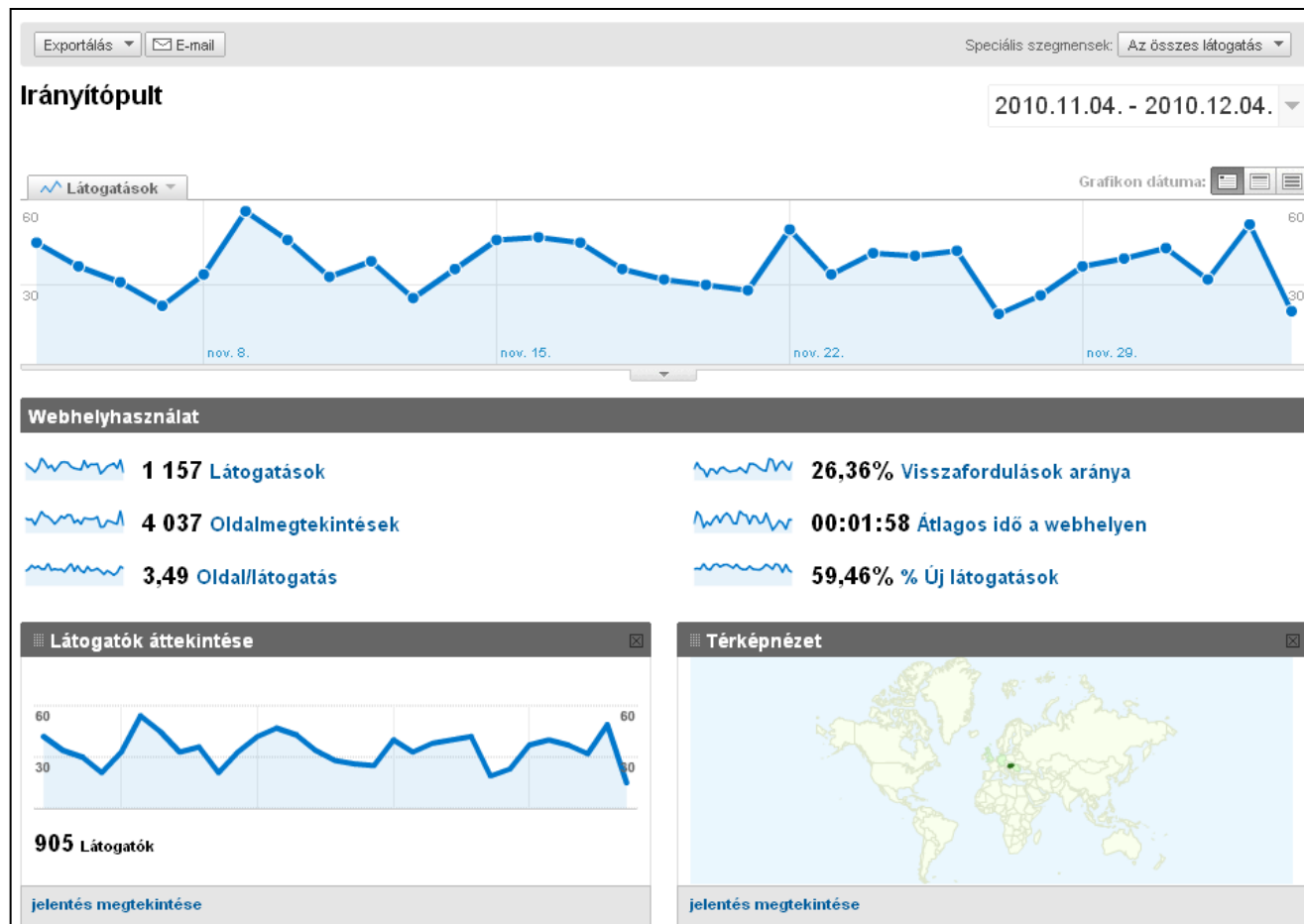


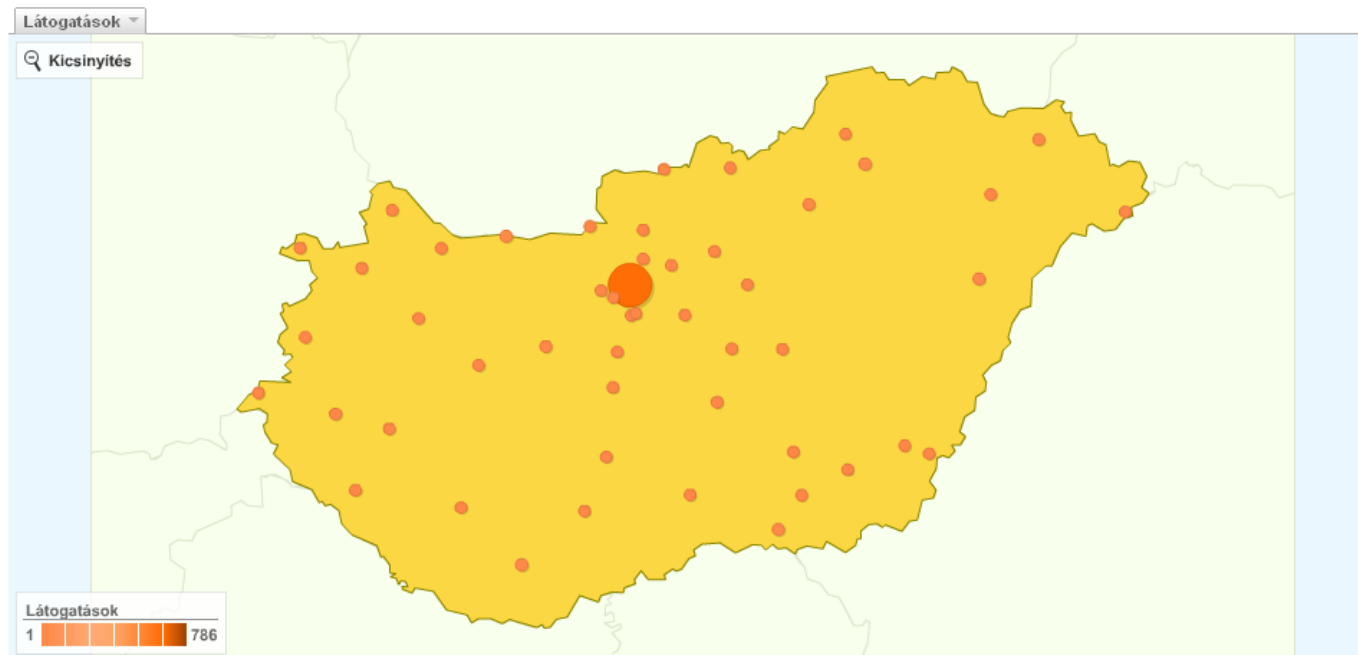
CHURCHES, BOWLING, GUNS AND STRIP CLUBS: Shading represents more references to one search term in the Google Maps directory at a particular location, e.g., a blue circle indicates more references to churches than bowling alleys, guns or strip clubs. Data collected in August 2008. Copyleft Floatingsheep.org, 2009



# Honlap-statisztikák

- Lehet tudni, hogy ki(?) és honnan, milyen paraméterekkel használja az adott oldalt





## Ez az ország/tartomány 1 042 látogatást eredményezett a következőn keresztül: 50 városok

Részletességi szint: Város Dimenzió: [Hírcs](#)

Webhelyhasználat1. célkészlet

Nézetek:

Látogatások1 042

% a webhely összesítéséből: 90,06%

Oldal/látogatás3,50

Webhelyátl: 3,49 (0,42%)

Átlagos idő a webhelyen00:01:57

Webhelyátl: 00:01:58 (-0,81%)

% Új látogatások58,45%

Webhelyátl: 59,46% (-1,71%)

Visszafordulások aránya26,10%

Webhelyátl: 26,36% (-0,98%)

Részletességi szint: Város

Látogatások

Oldal/látogatás

Átlagos idő a webhelyen

% Új látogatások

Visszafordulások aránya

1. Budapest

786

3,33

00:01:47

55,34%

25,45%

2. Pécs

25

3,12

00:01:50

68,00%

24,00%

3. Miskolc

23

4,30

00:03:39

56,52%

17,39%

4. Szeged

20

3,10

00:03:25

75,00%

25,00%

5. Debrecen

20

4,05

00:01:57

85,00%

35,00%

6. Komárom

19

4,32

00:03:22

63,16%

57,89%

7. Zalaegerszeg

14

3,21

00:01:02

64,29%

28,57%

8. Szekesfehelyar

10

2,20

00:00:49

60,00%

50,00%

9. Kisvada

10

4,20

00:02:59

30,00%

20,00%

10. Nyiregyhaza

9

9,11

00:07:33

66,67%

33,33%

Szűrő Város: tartalmaz

Mehet

Speciális szűrő

Ugrás ide: 1

Sorok megjelenítése: 10

1 - 10/50

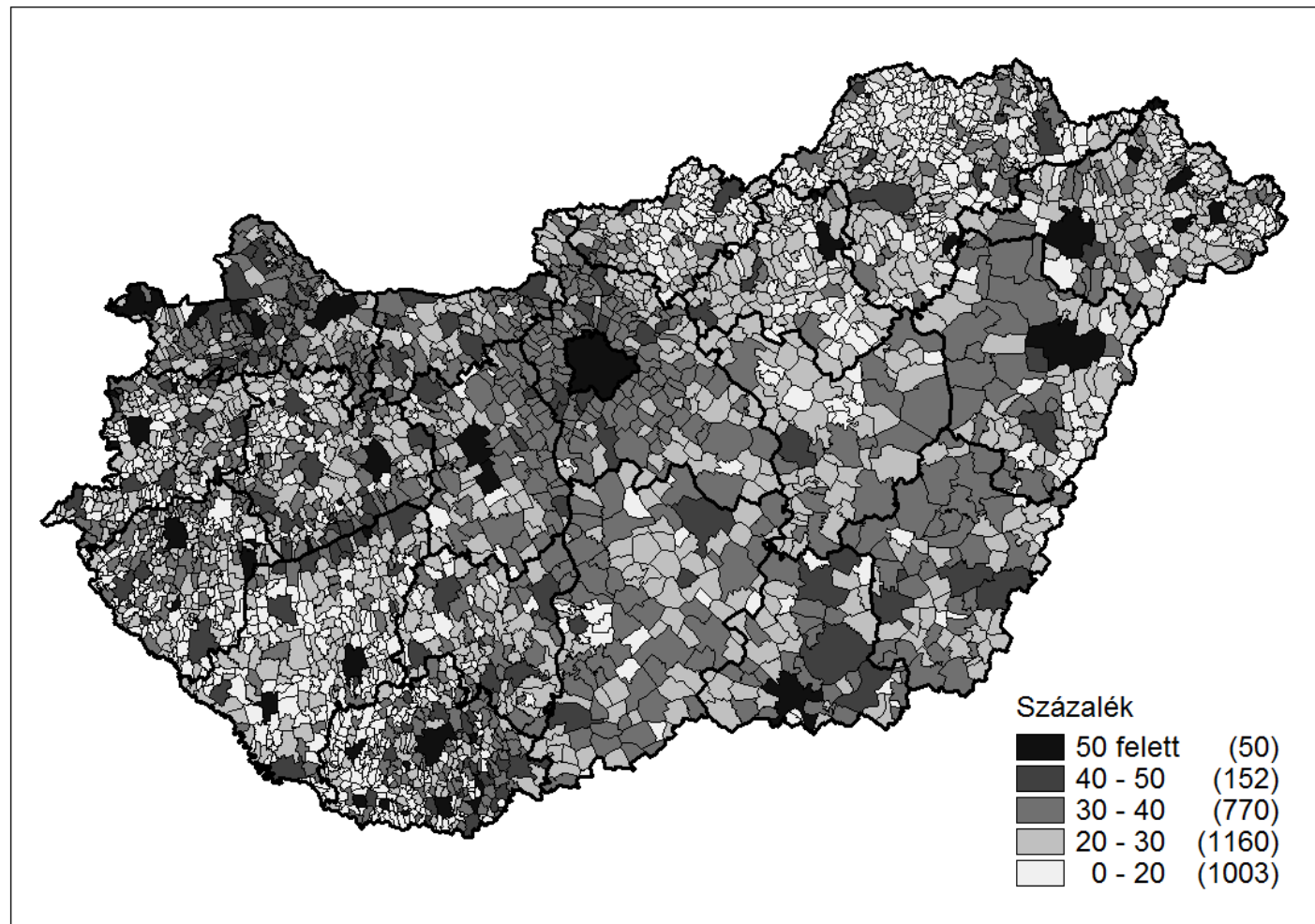
# Az online jelenlét területi különbségei: OSN statisztikák

| Ország        | Nemek |         | Korcsoportok |       |       |       |       |       |       |      | Összesen |
|---------------|-------|---------|--------------|-------|-------|-------|-------|-------|-------|------|----------|
|               | 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      |

A Facebook felhasználók százalékos megoszlása nem és korcsoportok szerint Európa néhány országában (a 13 évesnél idősebbek körében, 2013). Forrás: socialbakers.com



# Az online jelenlét területi különbségei: OSN statisztikák



Az iWiW felhasználók aránya településenként (2013. január)

Forrás: Origo Zrt. és IBS adatok alapján saját szerkesztés

# Geokódolt információk: geotag statisztikák

- A térbelivé tett térfüggetlen információk eloszlása az információk előfordulási és közvetve használati gyakoriságát jelezheti
- Geotag-elt képek globális földrajzi elterjedése
- Feltárja az egyes helyekkel kapcsolatban rögzített lokális tudás és vizuális reprezentációk sűrűségét
- Megjegyzés: a kialakított térképek közvetlen formában nem adnak információt a felhasználókról (szerzőkről), csak a közzétett tartalmak elhelyezkedéséről

# Geotag



**WIKIPÉDIA**  
A szabad enciklopédia

Kezdőlap  
Tartalom  
Kiemelték  
Friss változtatások  
Lap találmra  
Tudakozó

▼ Részvétel  
Kezdőknek  
Segítség  
Közösségi portál  
Kapcsolatfelvétel  
Adományok

► Nyomtatás/ exportálás  
► Eszközök

▼ Más nyelveken  
English  
Беларуская  
(тарашкевіца)  
Deutsch  
Esperanto  
Nederlands

szerkesztés

szócikk

vitalap

olvasás

szerkesztés

laptörténet

Keresés



Idén nyáron is (2013. július 25. – július 28. között) megrendezésre kerül a Wikitábor. Szállásigényedet 2013. június 16-ig, részvételi szándékodat június 24-ig jelezd!

A Központi Statisztikai Hivatal, a Központi Statisztikai Hivatal Könyvtár és a Wikimédia Magyarország Egyesület statisztikai szócikkíró versenyt hirdet. Régi és új szerkesztők egyaránt belevághatnak.

## Református nagytemplom (Debrecen) [bevezető szerkesztése]

A Wikipédiából, a szabad enciklopédiából

Koordináták: é. sz. 47° 31′ 55″, k. h. 21° 37′ 27″

A **debreceni református nagytemplom** Debrecen legjellegzetesebb épülete, 1500 m<sup>2</sup>-es alapterületével az ország legnagyobb református temploma. A Magyarországi Református Egyház jelképe, miatta nevezik Debrecent gyakran „kálvinista Rómának”. 1805 és 1824 között épült, *klasszicista stílusban*, a Kossuth tér és a Kálvin tér közt áll, a *belvárosban*.

**Tartalomjegyzék** [[elrejtés](#)]

- 1 Története
- 2 Az épület
  - 2.1 A templombelső
- 3 Külső hivatkozások

### Története [[szerkesztés](#)]

A nagytemplom helyén már a középkorban is templom állt, de az leégett. Helyébe *gótikus stílusú csarnoktemplom*, az András-templom épült 1297 és 1311 között, alapterülete kb. 16×43 méter volt; négy főoltárával sejteti, hogy a *Tiszántúl* legnagyobb temploma lehetett. Ez a templom 1564-ben leégett.

A templom megfeketedett falai 62 éven át álltak pusztán, újjáépítésüket *Bethlen Gábor* vezetésével indította el a már *protestáns* város 1626-ban. *Rákóczi György* támogatásával az építkezés 1628-ban befejeződött. 1628. november 26-án a debreceniek az új András-templomban gyűltek össze. A templom arculatát alaposan megváltoztatták. Téglapillérekkel alátámasztott gerendás famennyezetet kapott, tornyát széles köpárkány díszítette. A cserepes sisakkal fedett favázás tornya tetején Rákóczi György költségére készített aranyos gomb fénylett, ott mutatta az idő múlását a Rákóczi adományozta óramű is. A templomot körülvevő cinterem sarkában, 1642-ben építették fel a vakolatlan téglás Verestornyot, amelyben elhelyezték azt az 56 mázsás

### Református nagytemplom



**Hely** 4026 Debrecen, Piac utca 4-6.

#### Építési adatok

**Építés éve** 1805–1808

**Építési stílus** klasszicizmus

**Tervező** Péchy Mihály

**Építész(ek)** Thaller József

#### Hasznosítása

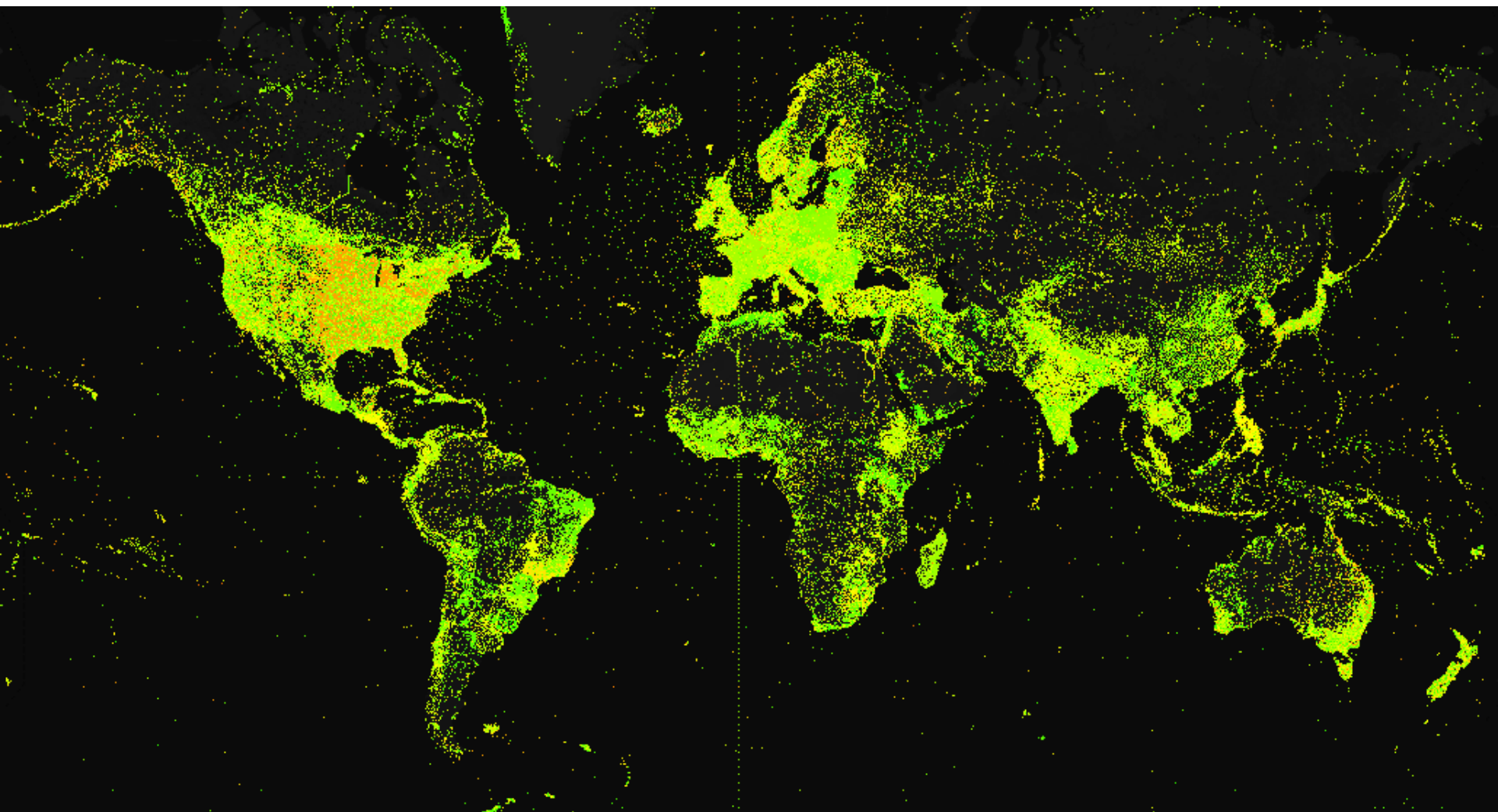
# Geotag

view-source:hu.wikipedia.org/wiki/Református\_nagytemplom\_(Debrecen)

```
124 <tr>
125 <td colspan="2" align="center">
126 <div class="poziciosTerkepDoboz" style="width: 250px;float:none;clear:none;">
127 <div class="poziciosTerkepBelso" style="position: relative; padding: 0px; width: 248px; margin-left: auto; margin-right: auto;"><a
href="/wiki/F%C3%A1jl:Debrecen_map.svg" class="image" title="Református nagytemplom (Debrecen)"></a>
128 <div class="poziciosTerkepFelirat" style="position: absolute; z-index: 2; top:42.4%; left:42.7%; height: 0px; width: 0px;margin: 0px;
padding: 0px; text-align:left; line-height:0">
129 <div style="position: relative; text-align: center; left: -4px; top: -4px; width: 8px; font-size: 8px;"></div>
130 <div style="font-size: 100%; line-height: 110%; position: relative; top: -0px;left: -105px; text-align: center; width: 210px;"><span
style="padding: 1px;"><b>Református nagytemplom</b></span></div>
131 </div>
132 </div>
133 <div style="text-align:center;"><small>Pozíció Debrecen térképén</small></div>
134 </div>
135 <span style="font-size: small;"><span id="coordinates" class="plainlinks"><a href="/wiki/F%C3%B6ldrajzi_koordin%C3%A1t-rendszer"
title="Földrajzi koordináta-rendszer">Koordináták</a>: <span class="plainlinks"><a rel="nofollow" class="external text"
href="http://toolserver.org/~geohack/hu/47_31_55_N_21_37_27_E_type:landmark?pagename=Reform%C3%A1tus_nagytemplom_(Debrecen)">é. sz.
47°&#160;31'&#160;55", k. h. 21°&#160;37'&#160;27"</a></span><span class="geo" style="display:none;"><span
cl<del>ass="latitude">47.5319444444</span>, <span class="longitude">21.6241666667</span></span></span></td>
136 </tr>
137 </table>
138 <p>A <b>debreceni református nagytemplom</b> <a href="/wiki/Debrecen" title="Debrecen">Debrecen</a> legjellegzetesebb épülete, 1500 m²-es
alapterületével az ország legnagyobb református temploma. A <a href="/wiki/Magyarorszá<del>g">C3%A1lgi_Reform%C3%A1tus_Egyh%C3%A1z"
title="Magyarországi Református Egyház">Magyarországi Református Egyház</a> jelképe, miatta nevezik Debrecent gyakran „kálvinista
Rómának”. <a href="/wiki/1805" title="1805">1805</a> és <a href="/wiki/1824" title="1824">1824</a> között épült, <a
href="/wiki/Klasszicizmus" title="Klasszicizmus">klasszicista stílusban</a>, a Kossuth tér és a Kálvin tér közt áll, a <a
href="/wiki/Belv%C3%A1ros_(Debrecen)" title="Belváros (Debrecen)">belvárosban</a>.</p>
139 <table id="toc" class="toc">
140 <tr>
141 <td>
142 <div id="toctitle">
143 <h2>Tartalomjegyzék</h2>
144 </div>
145 <ul>
146 <li class="toclevel-1 tocsection-1"><a href="#T.C3.B6rt.C3.A9nete"><span class="tocnumber">1</span> <span
class="toctext">Története</span></a></li>
147 <li class="toclevel-1 tocsection-2"><a href="#Áz_.C3.A9p.C3.BClet"><span class="tocnumber">2</span> <span class="toctext">Áz
```

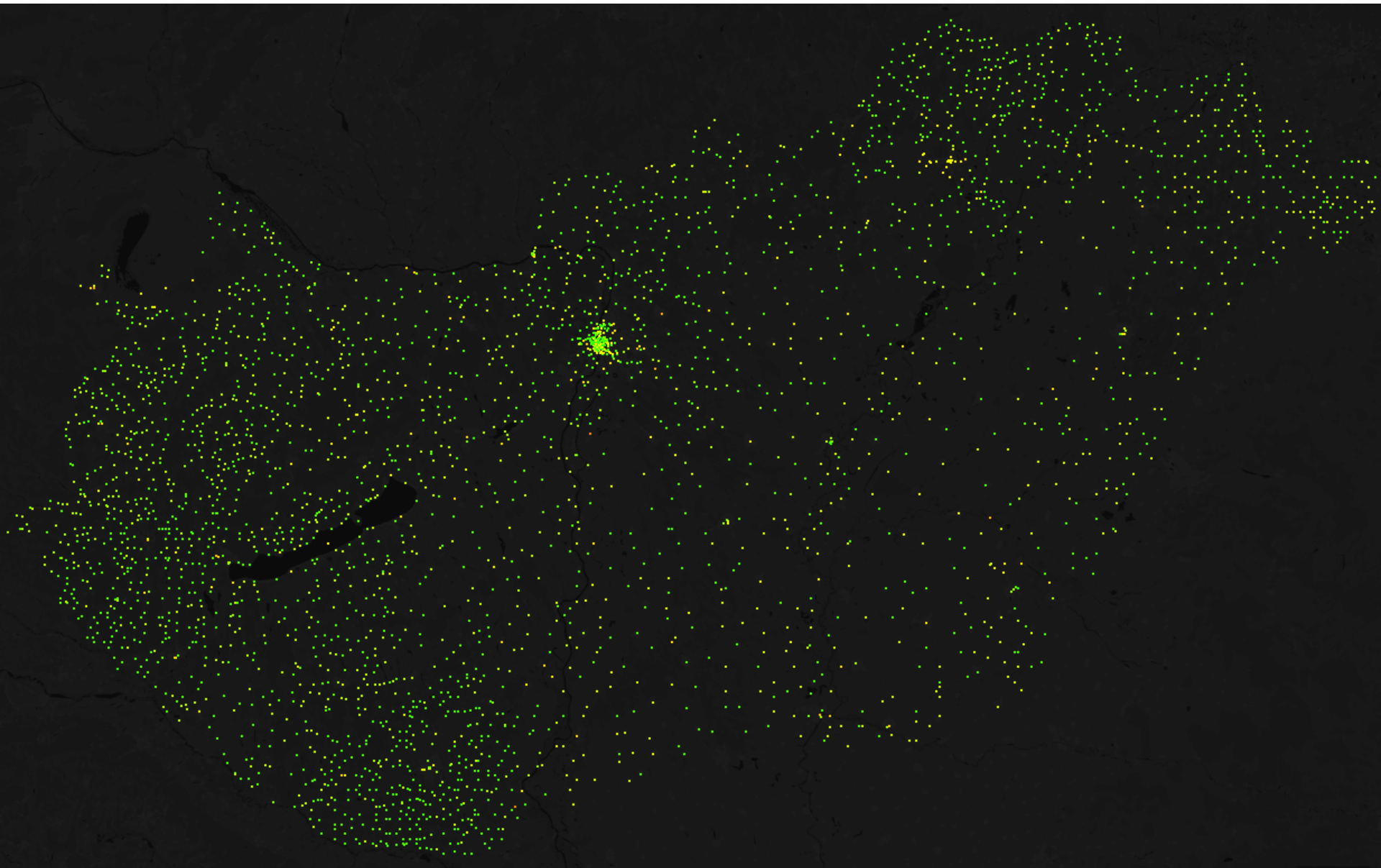


# Geotag-elt Wikipedia bejegyzések

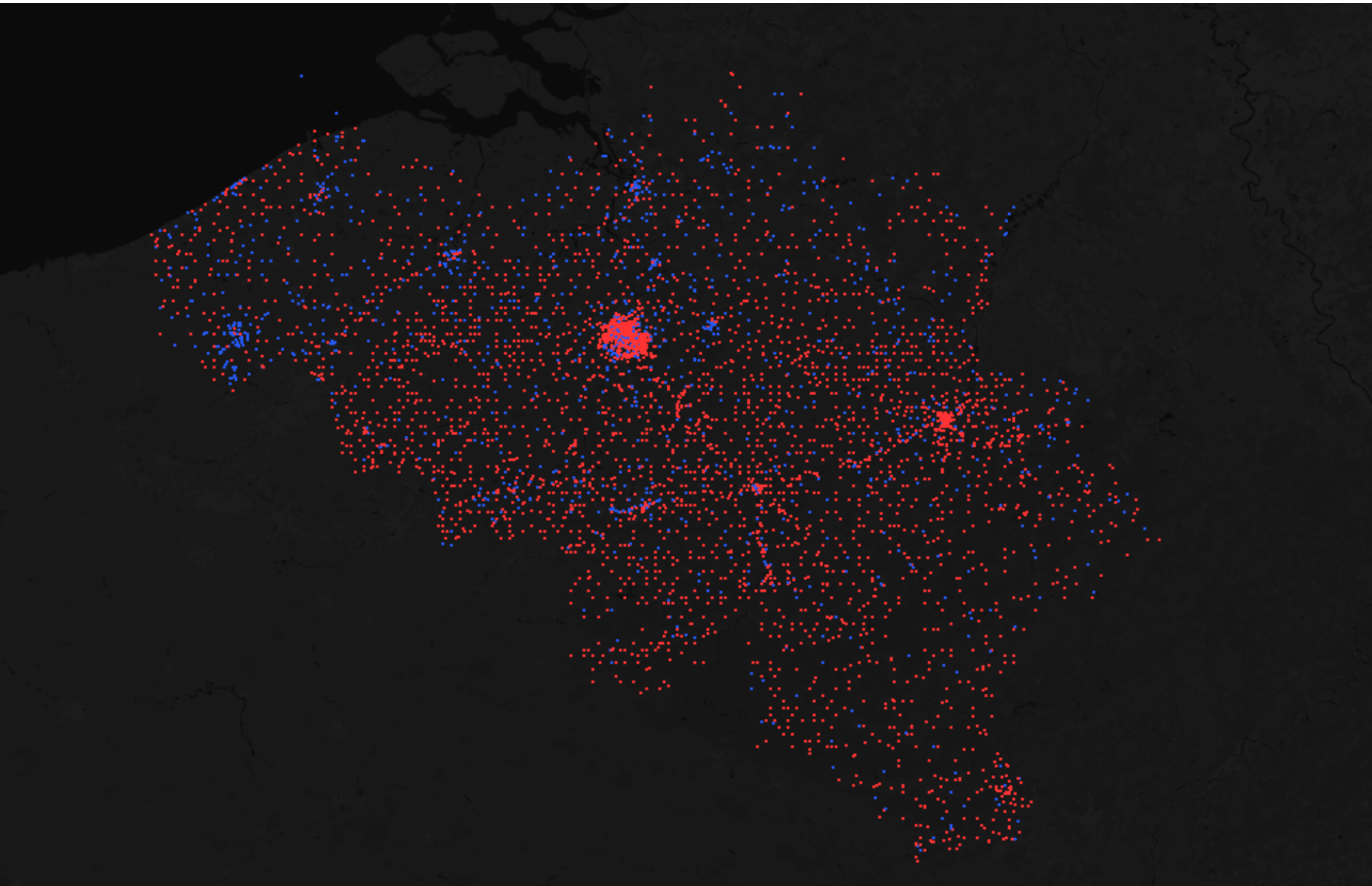


Forrás: <http://wikiproject.oii.ox.ac.uk>

# Geotag-elt Wikipedia bejegyzések (EN, FR)

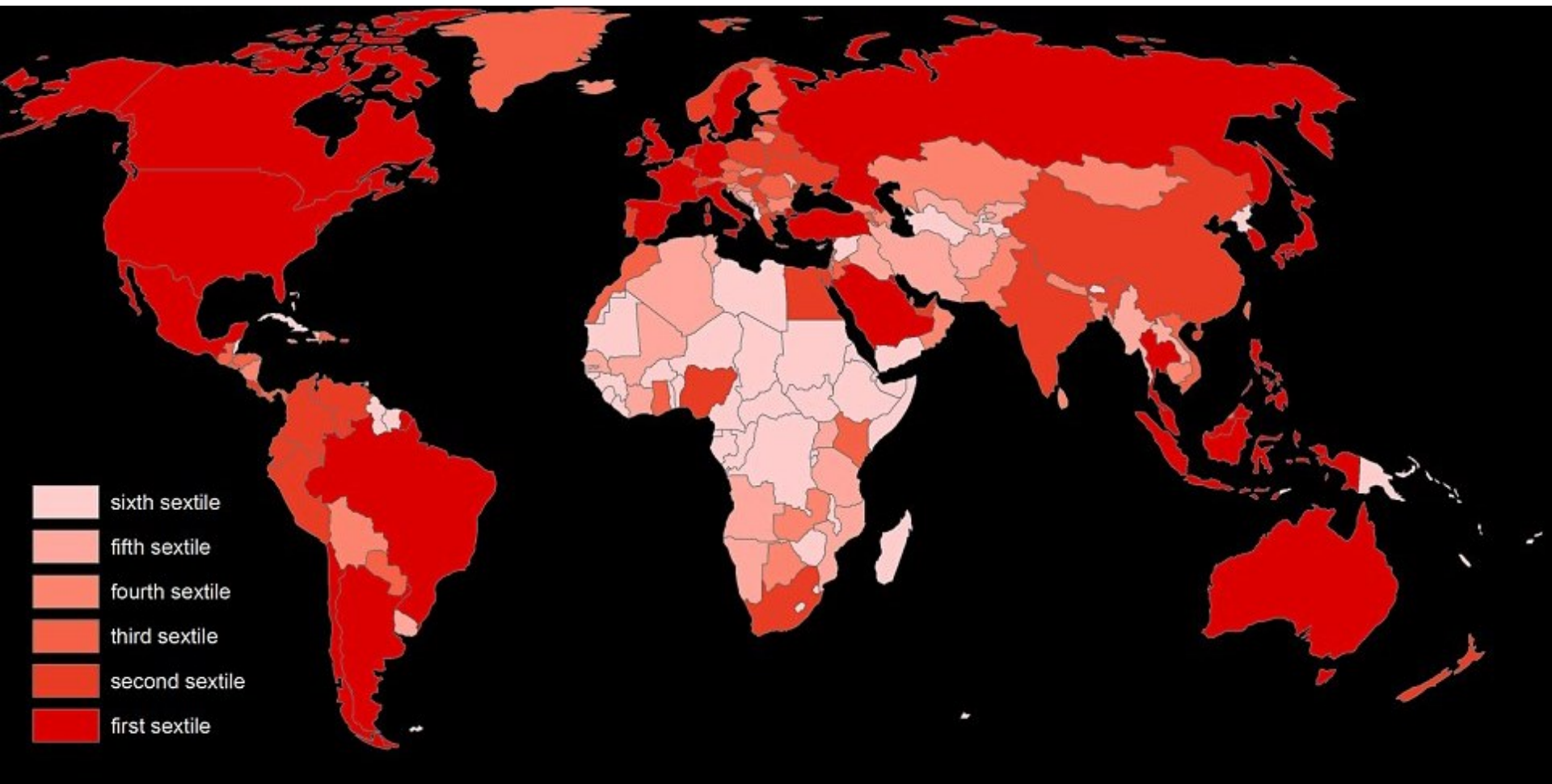


# Geotag-elt Wikipedia bejegyzések (EN, FR)



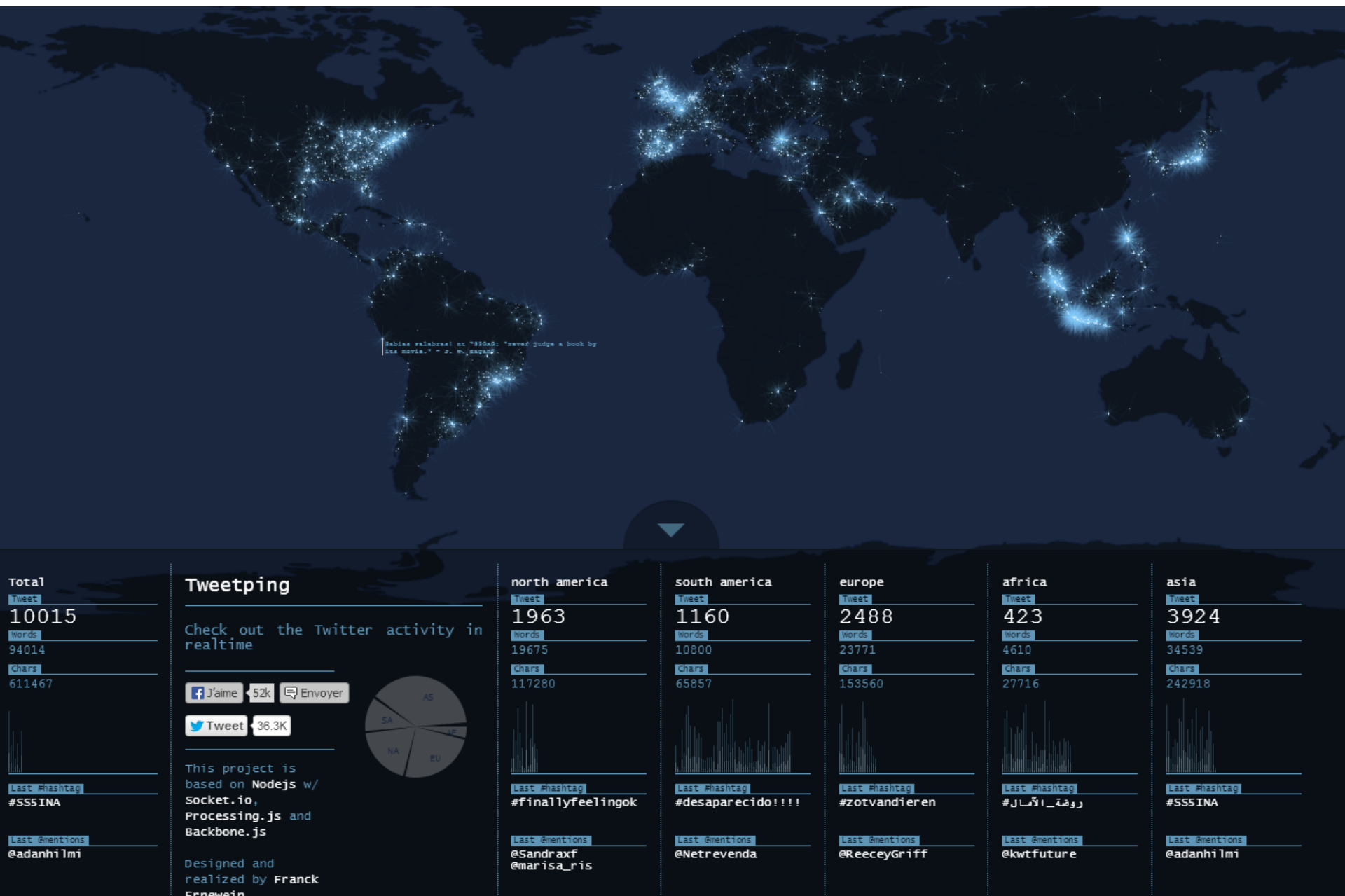
Forrás: <http://wikiproject.oii.ox.ac.uk>

# Geotag-elt tweet-ek száma 2012.03.05 és 2012.03.13 között

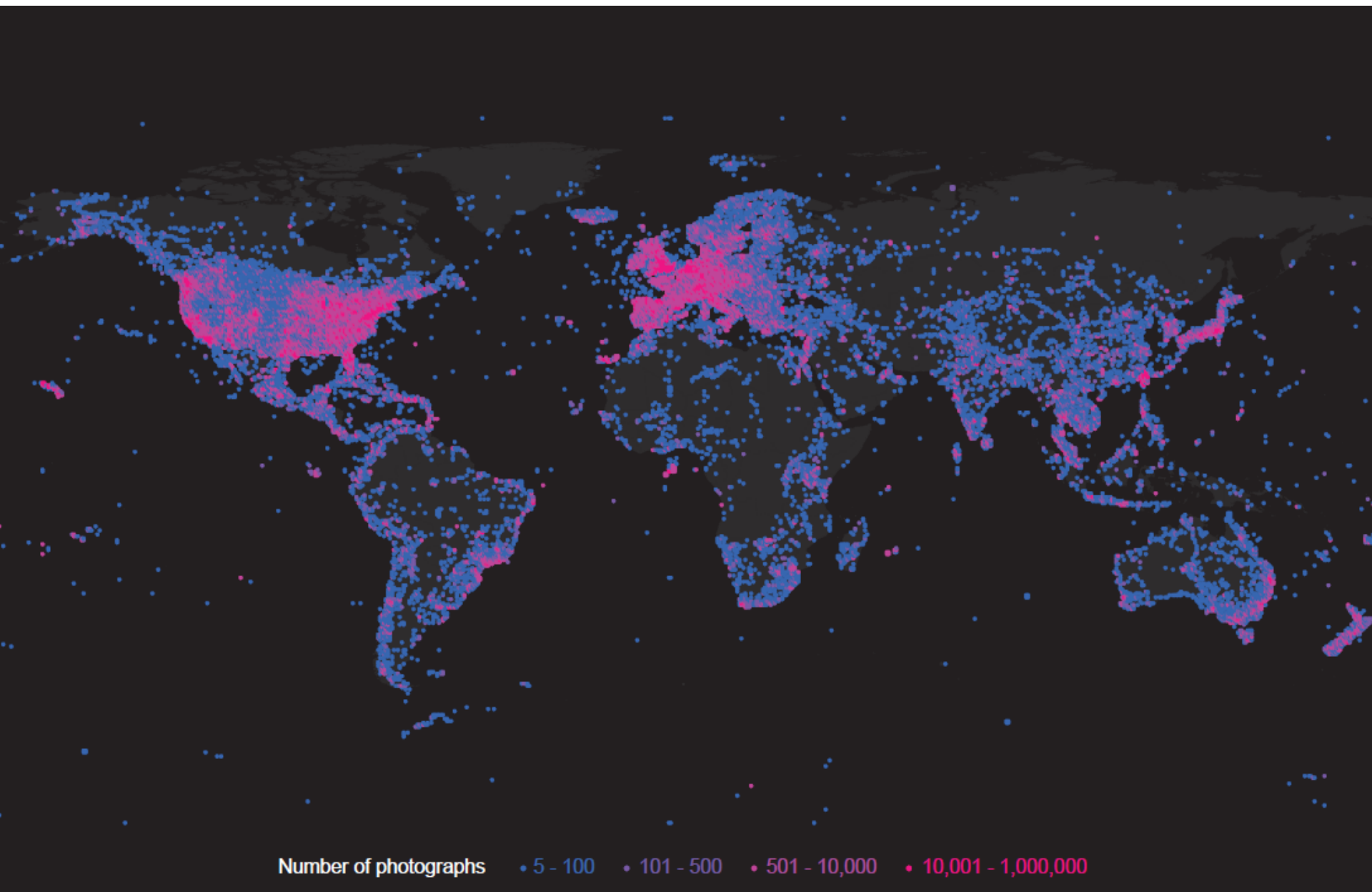




# A Twitter aktivitás globális területi mintázata (N=10015, 2013.05.07)

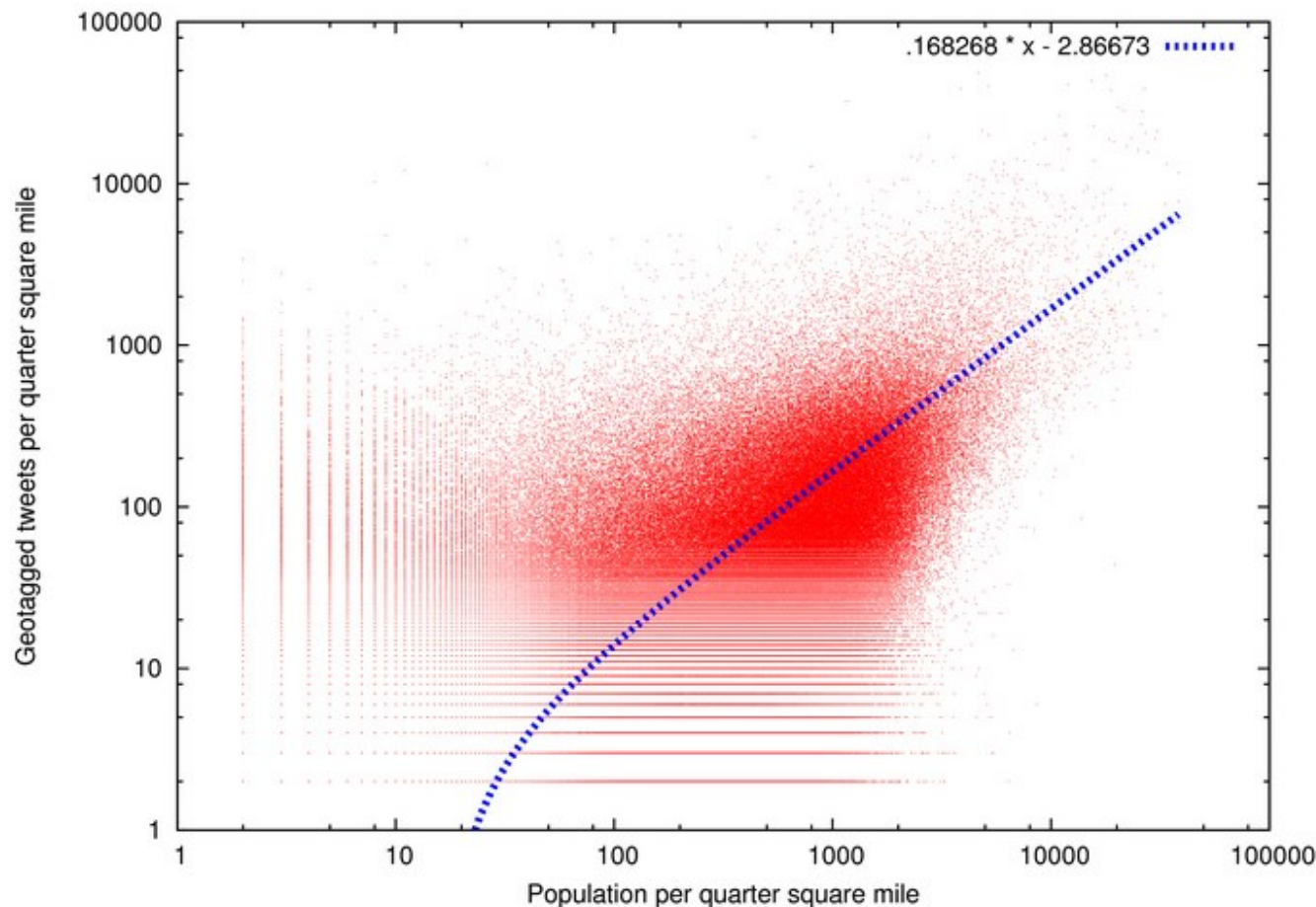


# Geotag-elt Flickr fotó bejegyzések száma



Forrás: [www.convoco.co.uk](http://www.convoco.co.uk)

# A népsűrűség és a geotag-elt bejegyzések közötti kapcsolat alakulása

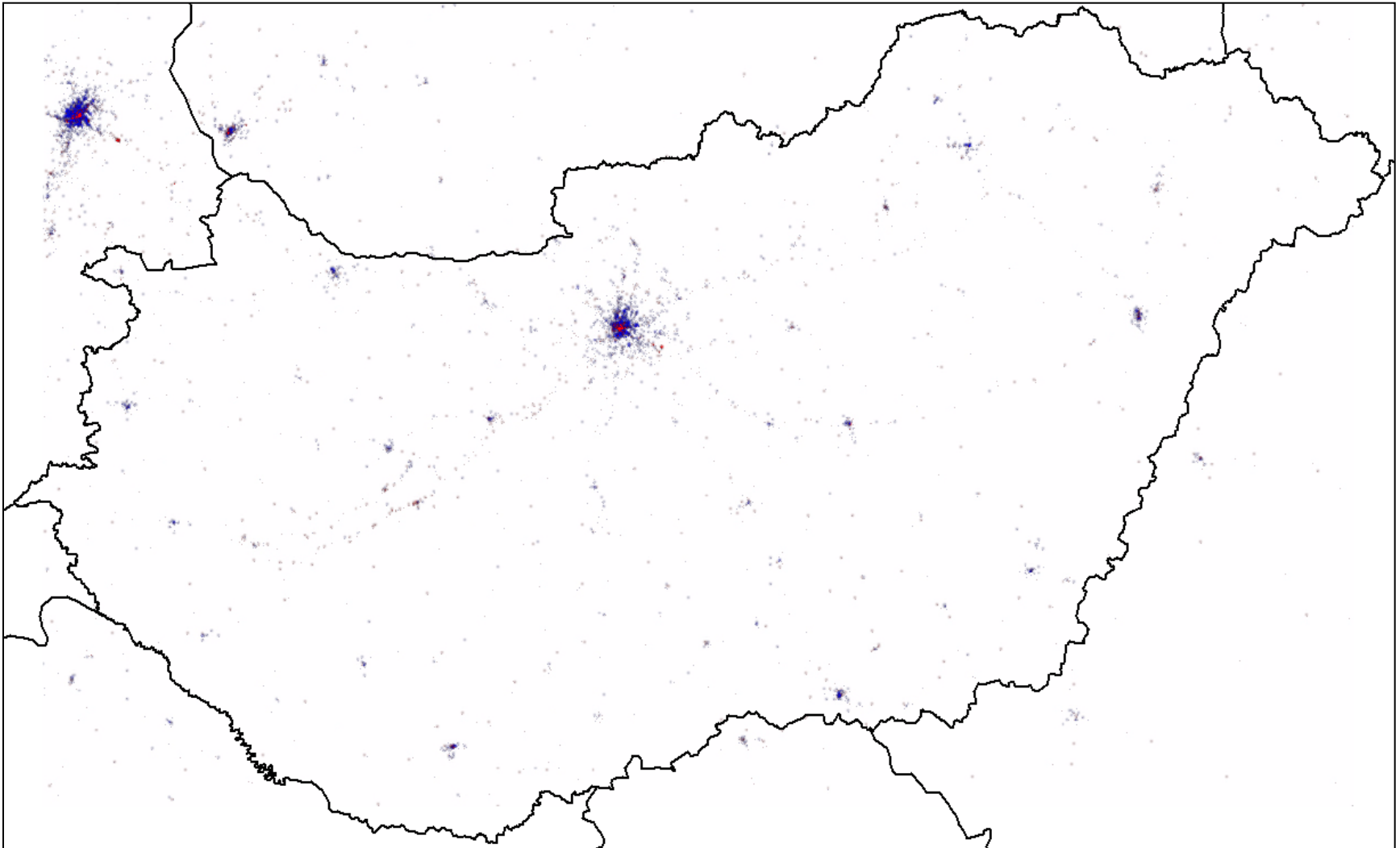


**A népességszám és a geotag-elt bejegyzések számának pontdiagramja  
(0,5 négyzetmérföldenként, az USA-ban)  
(Forrás: Fischer, E. 2012)**

Lakatlan területeken is találhatók geotag-elt bejegyzések (pl. parkok, emlékművek stb.).

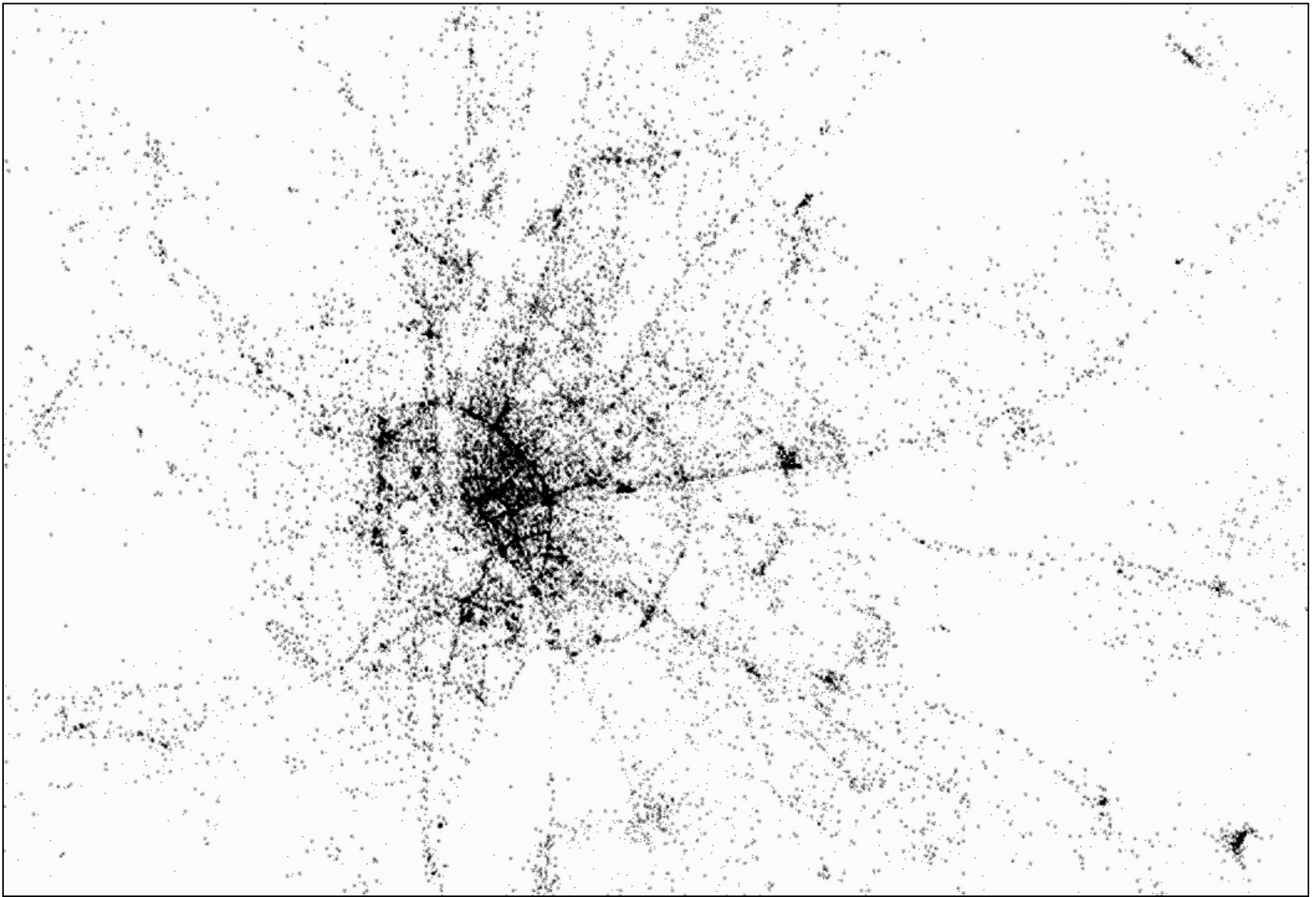
# Az online jelenlét és aktivitás közvetett területi mérése

Geotag-elt Twitter bejegyzések sűrűsödése Magyarországon

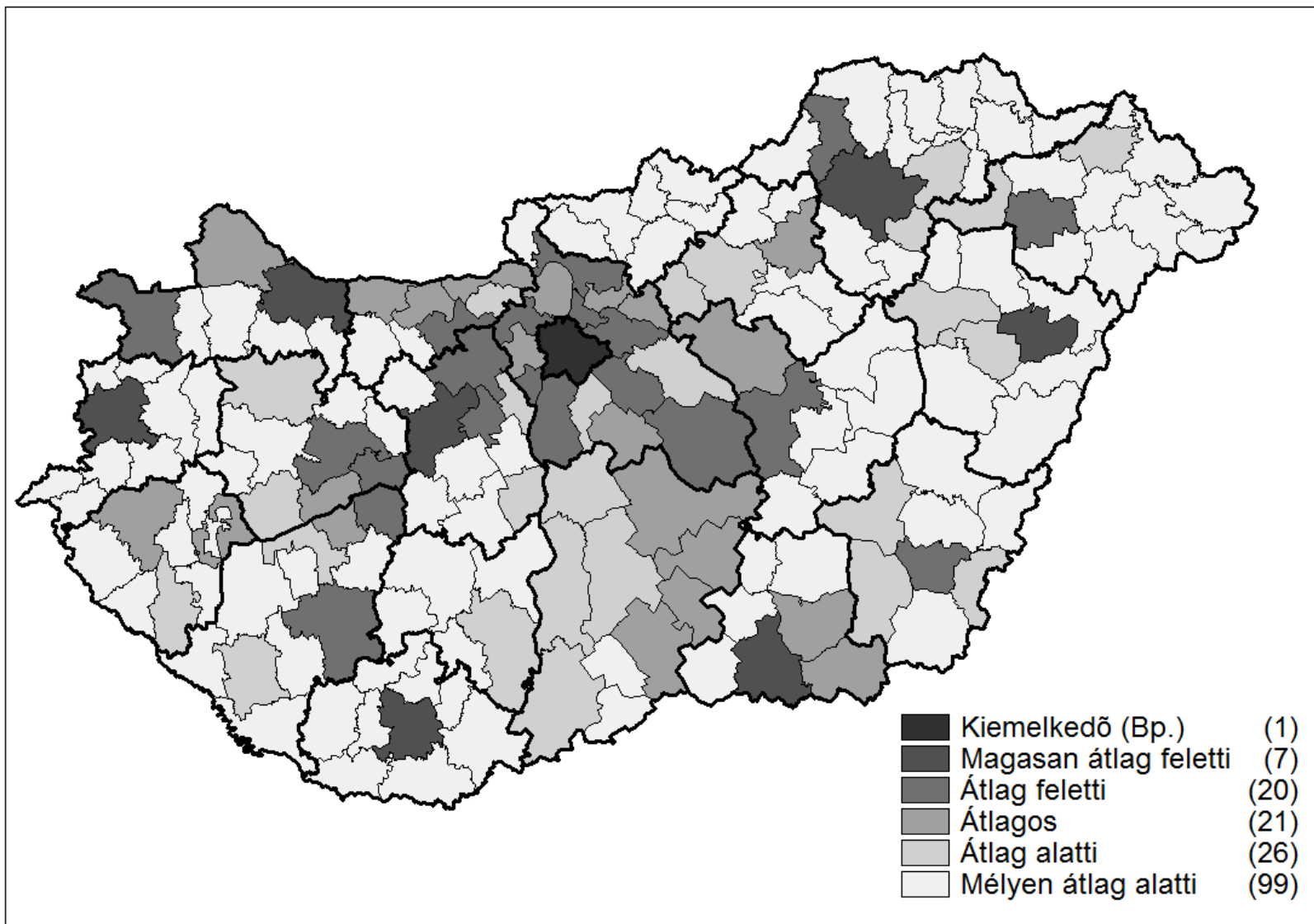


Forrás: [www.mapbox.com](http://www.mapbox.com) adatok alapján





Geotag-elt Twitter bejegyzések sűrűsége Budapesten.  
Forrás: Fisher (2013) alapján



Geotag-elt Twitter bejegyzések aggregált száma a magyar kistérségekben.  
Forrás: Fisher (2013) adatai alapján saját szerkesztés



**Köszönöm a  
figyelmet!**