

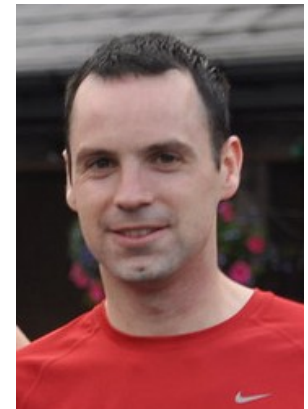
Using OSM for LBS - an analysis of changes to spatial objects

Peter Mooney

**Dept of Computer Science
National University of Ireland Maynooth**



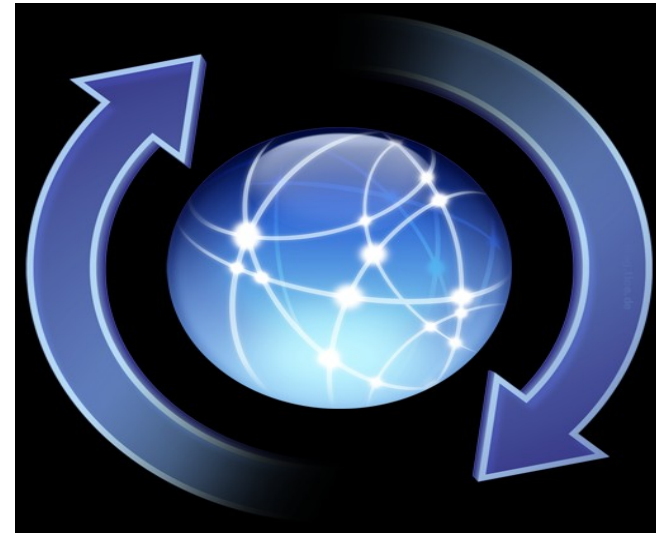
**Padraig Corcoran
School of Informatics
University College Dublin**



Spatial Data is one of the most important aspects in LBS



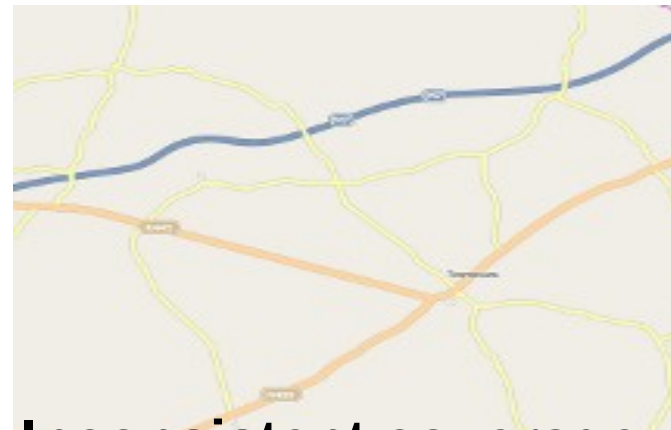
Up-to-date - current



**Frequently updated –
release cycle**

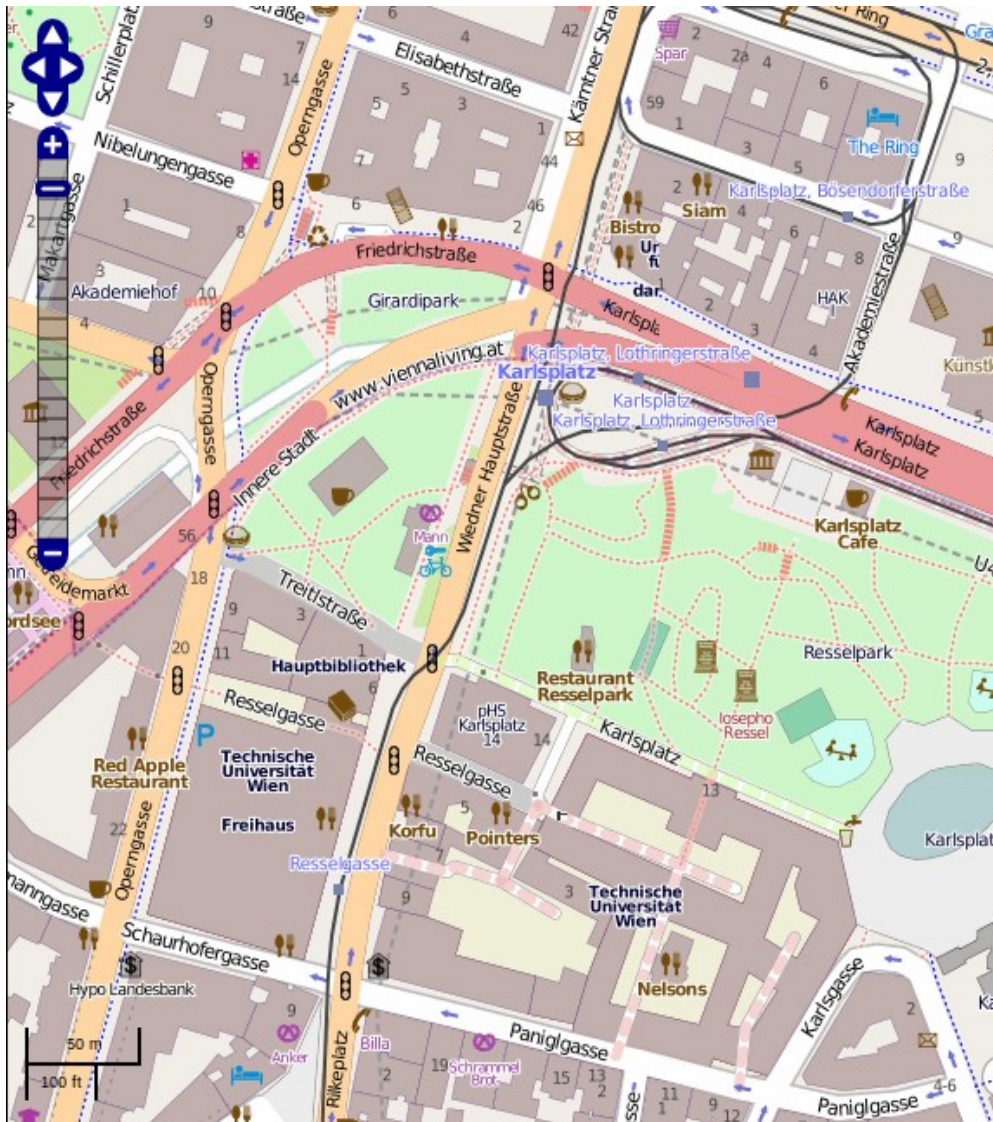


**Accurate – Geometrically and
in Attribution**



Inconsistent coverage

OpenStreetMap (OSM) is the world's leading source of VGI



- Very fast update cycle – most up-to-date data always available
- Excellent urban coverage in most large cities (Over et al, 2011)
- **How did the map evolve to it's current state?**

What causes problems for LBS applications using OSM data?

TAGGING: Incorrect (semantically or syntactically) use of tagging (key-value pairs) for features

CONTRIBUTIONS: Potentially poor quality or incorrect contributions

OSM - “Map Features” Community Agreed Ontology

wiki.openstreetmap.org/wiki/Map_Features#Shop



Shop

A shop is a place of business stocked with goods for sale.

Key	Value	Element	Comment	Rendering	Photo
shop	alcohol		Government owned or licensed shop selling alcohol. See also shop=beverages .		
shop	anime		Specials shops for anime stuff.		
shop	art		A private art gallery which sells works of art.		
shop	baby_goods		A shop that sells objects for babies (clothes, prams, cots, toys).		
shop	bakery		Selling bread		
shop	bathroom_furnishing		Selling bathroom furniture and accessories		
shop	beauty		A non-hairdresser beauty shop, spa, nail salon, etc.. See also shop=hairdresser .		
shop	bed		A shop that specialises in selling mattresses and other bedding products.		
shop	beverages		Shop focused on selling alcoholic and non-alcoholic beverages. See also shop=alcohol .		

amenity

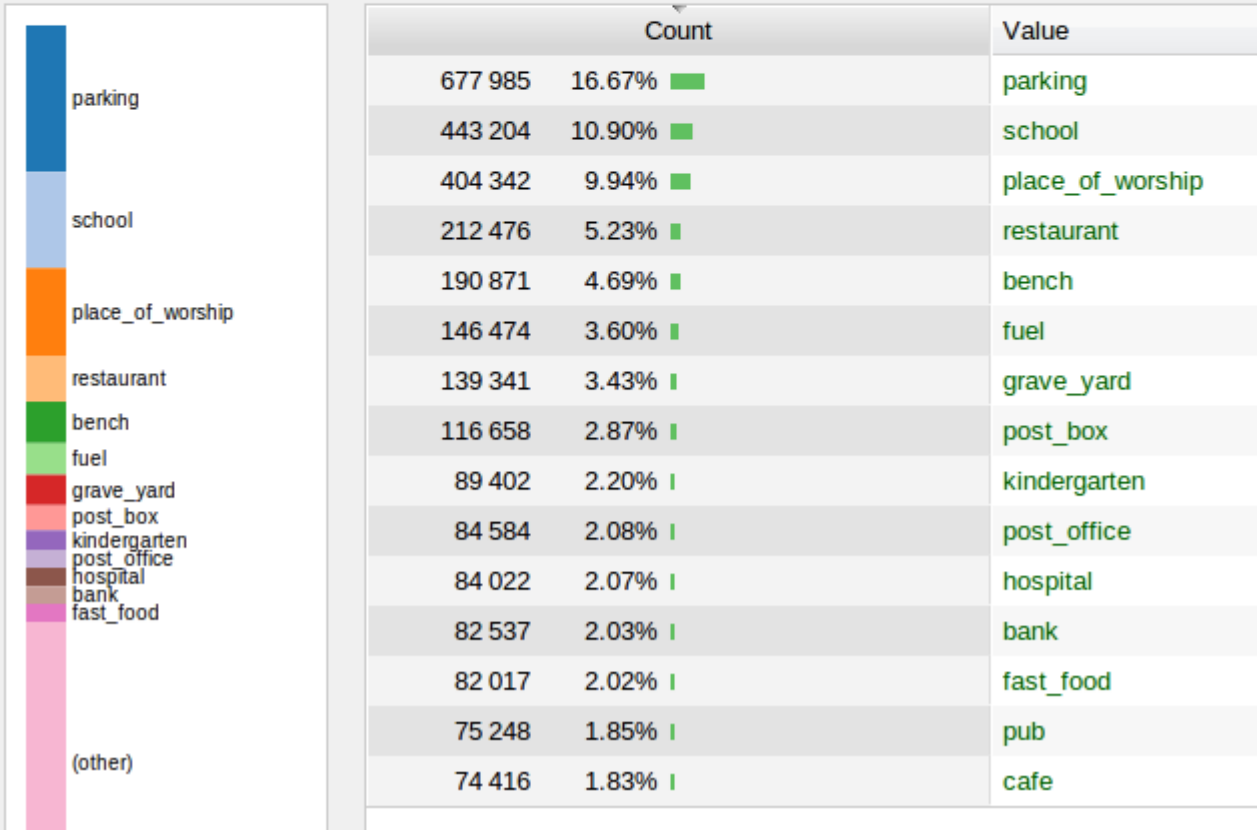
For describing useful and important facilities for visitors and residents.

XAPI JOSM

Filter: No filter

Values Combinations Map Wiki JOSM

Values used with this key



TagInfo
- Visualise
the use of
key-value
pairs

Explore the
frequency of tag
usage

Example: KEY = AMENITY

amenity

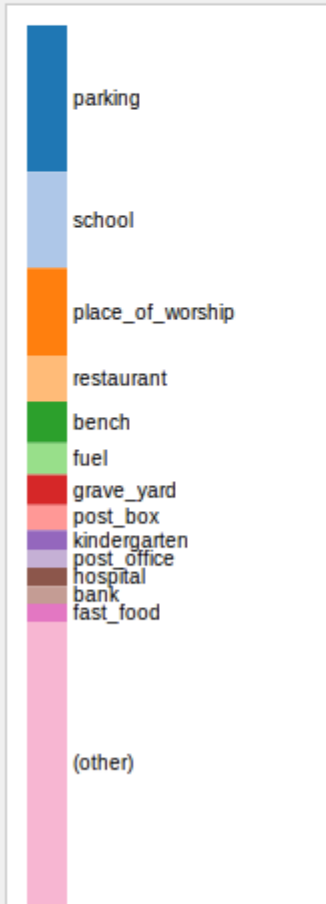
For describing useful and important facilities for visitors and residents.

XAPI JOSM

Filter: No filter

Values Combinations Map Wiki JOSM

Values used with this key



Count	Value
2 0.00%	car_recycling
2 0.00%	fontaine
2 0.00%	Modellflugplatz
2 0.00%	stamp_machine
2 0.00%	école_primaire
2 0.00%	Hair_salon
2 0.00%	HVAC
2 0.00%	trailer_rental
2 0.00%	professional_services
2 0.00%	radarska_kontrola
2 0.00%	Sportverein
2 0.00%	home_for_the_aged
2 0.00%	surf_life_saving_club
2 0.00%	vacant_block
2 0.00%	hospital:historical

OSM Map Features The "OSM Ontology"

TagInfo allows us to see 'obscure' or rarely used tag values

Working with OpenStreetMap History Database

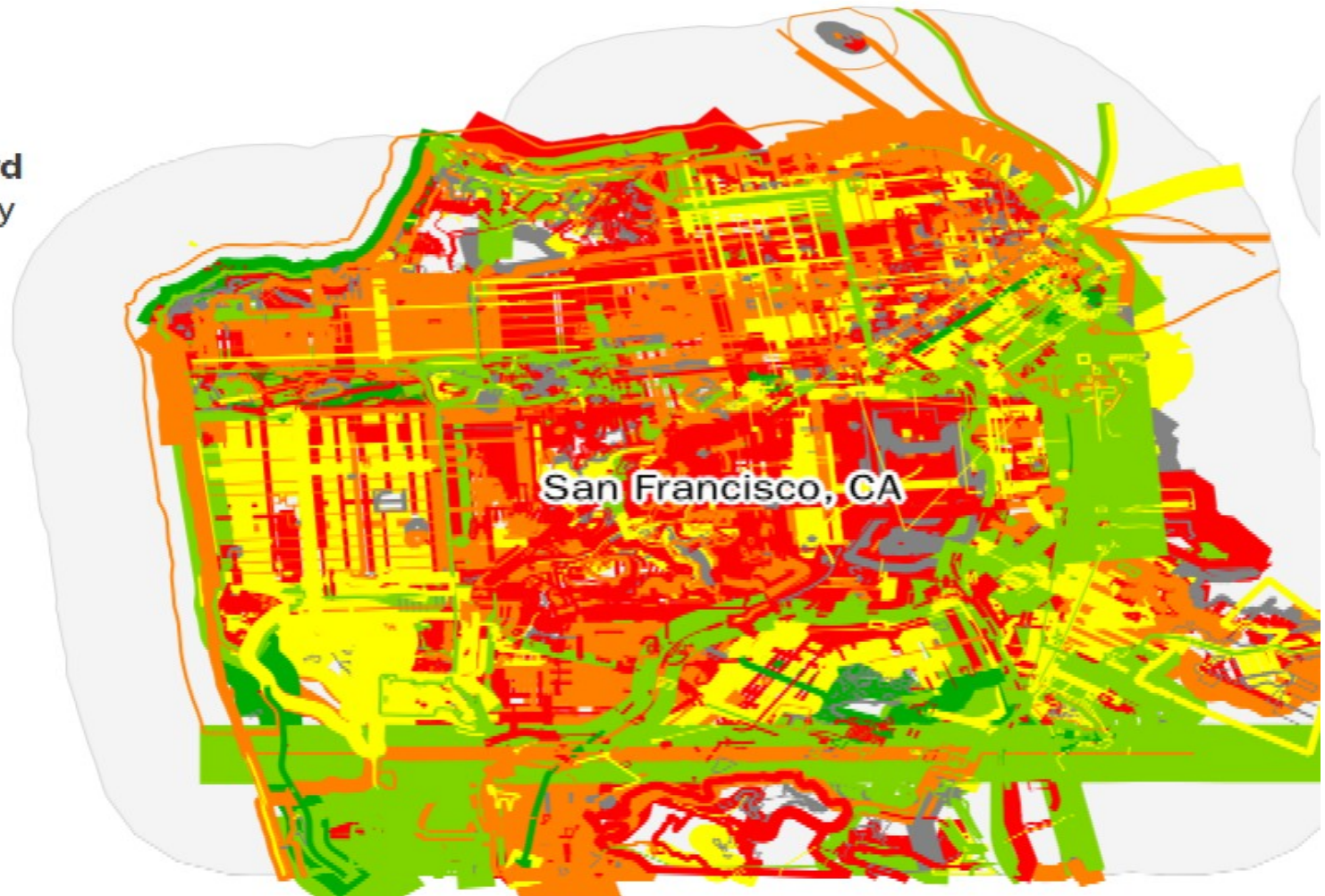
M van Exel (2011) “Taking the Temperature of OSM data”

<http://oegeo.wordpress.com/2011/09/19/taking-the-temperature-of-local-openstreetmap-communities/>

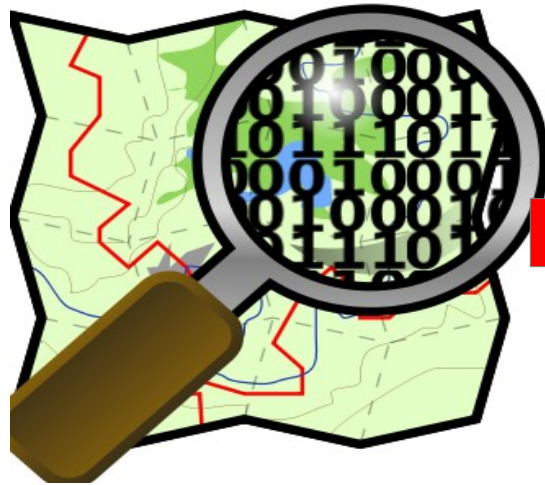
Community score card

- 95% of edits done by 12% of users
- 9% untouched TIGER
- 4.8 avg version increase of TIGER roads
- 8.3% of features touched in last 3 months, 40% in last year

77°



There is a considerable workflow in getting from OSM History to Analysis



OSM MaZderMind



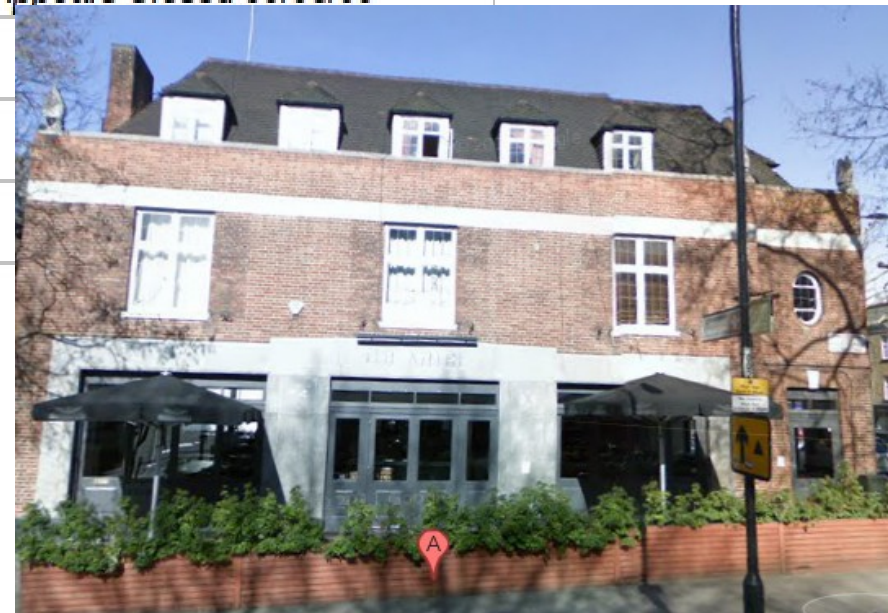
OpenStreetMap
FULL PLANET
History
Release: ~ 3 months
20GB Compressed
500GB
Uncompressed

<https://github.com/MaZderMind/osm-history-splitter>



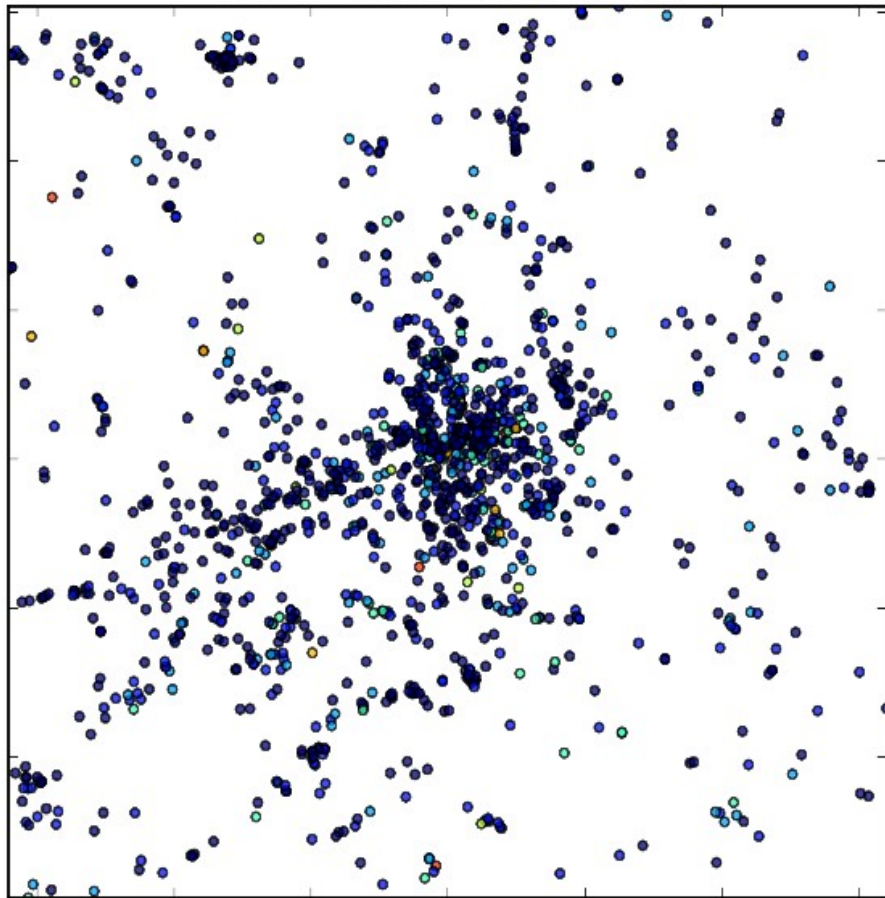
What does OSM history look like?

Version	Date	ChangeSet	USER	Amenity	Name	Note
19	2011-02-12 14:35:07	7265399	6871	pub	The Mitre	<i>reopened July 2009</i>
18	2011-01-31 10:06:25	7143502	346	pub	The Mitre	<i>reopened July 2009</i>
10	2009-07-30 08:15:18	1982204	70696	pub	The Mitre	<i>reopened July 2009</i>
9	2009-07-13 09:22:10	1814929	38244	pub	The Mitre	<i>reopened July 2009</i>
8	2009-03-31 10:18:20	871525	38244	pub	The Mitre	<i>closed pub</i>
6	2009-03-10 22:49:23	788986	38244	pub	The Mitre	<i>Former Site - Appears Closed 09/03/09</i>
4	2009-03-10 22:48:43	788986	38244	pub	The Mitre	<i>Appears Closed 09/03/09</i>
3	2009-02-16 17:51:02	486036	6871	pub	The Mitre	
2	2008-07-25 06:40:26	591192	6871	pub	The Mitre	
1	2008-06-15 12:57:44	168135	6871	pub	The Mitre	

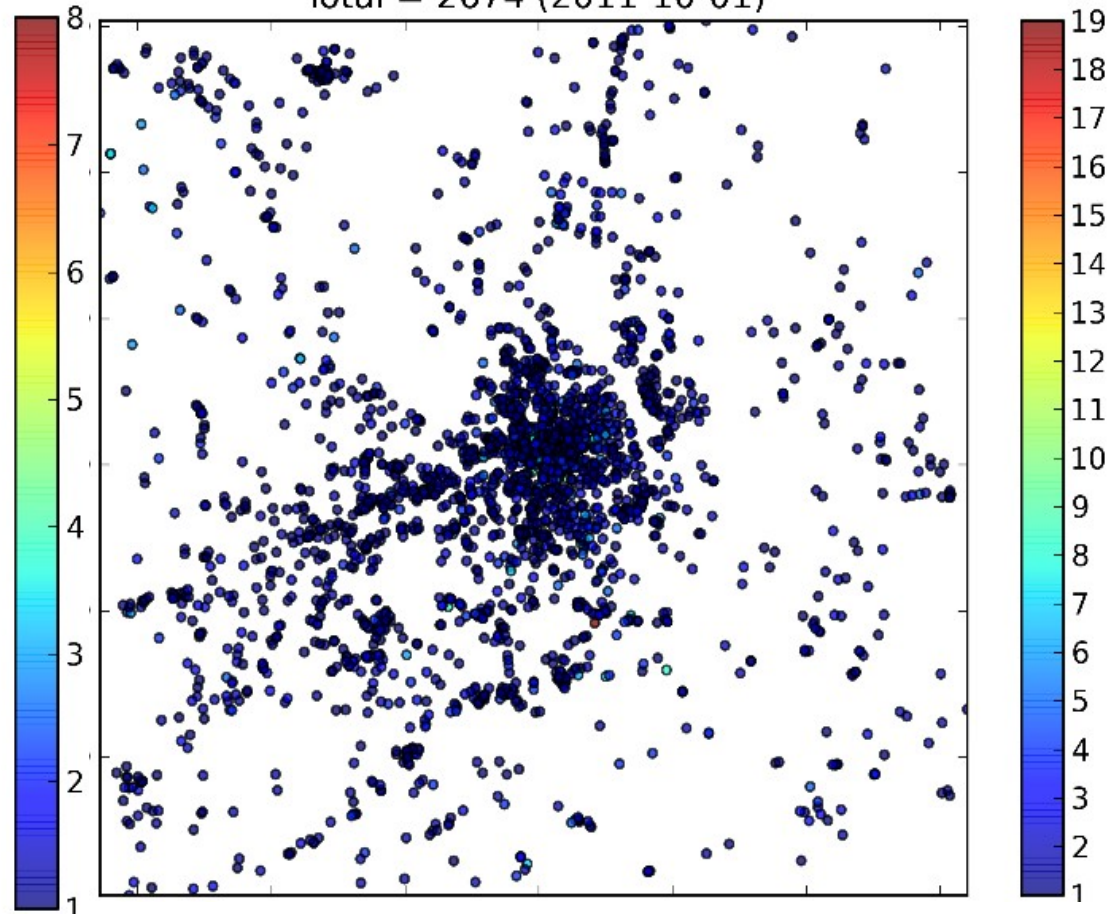


Points (amenity=pub) – London City

London City - Pubs
Total = 1560 (2010-10-01)

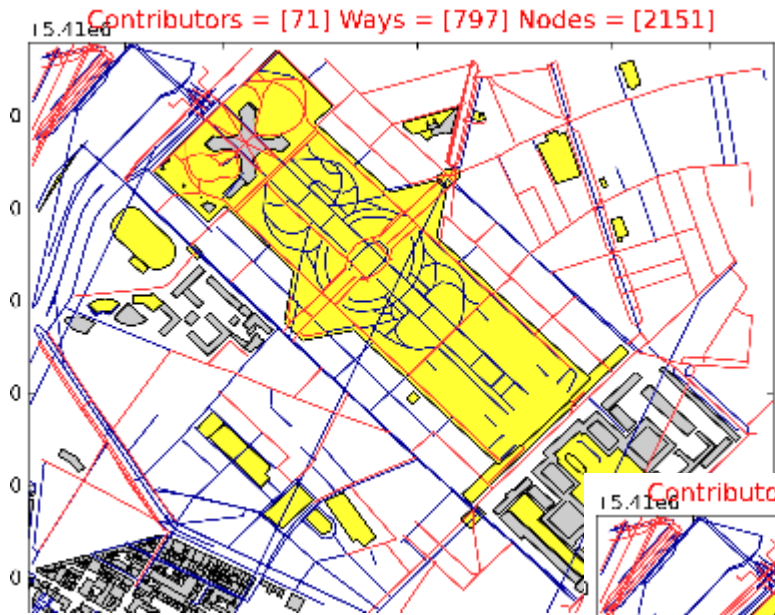


London City - Pubs
Total = 2674 (2011-10-01)

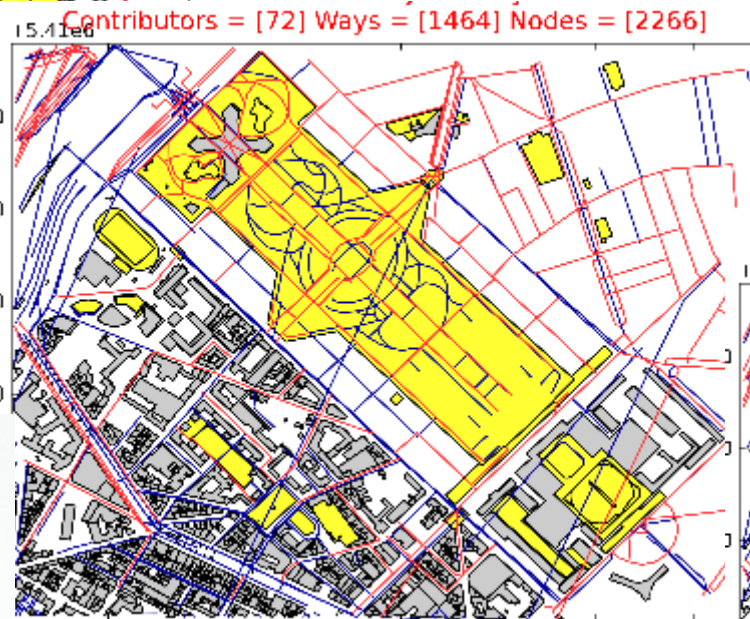


April 1st 2010

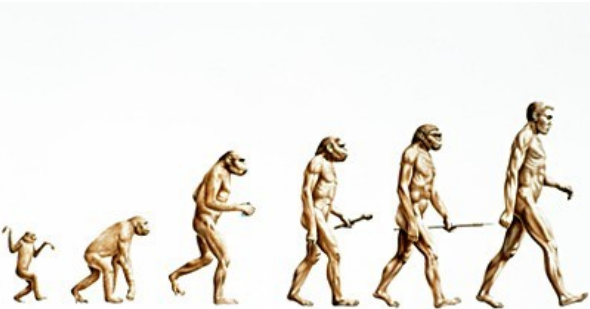
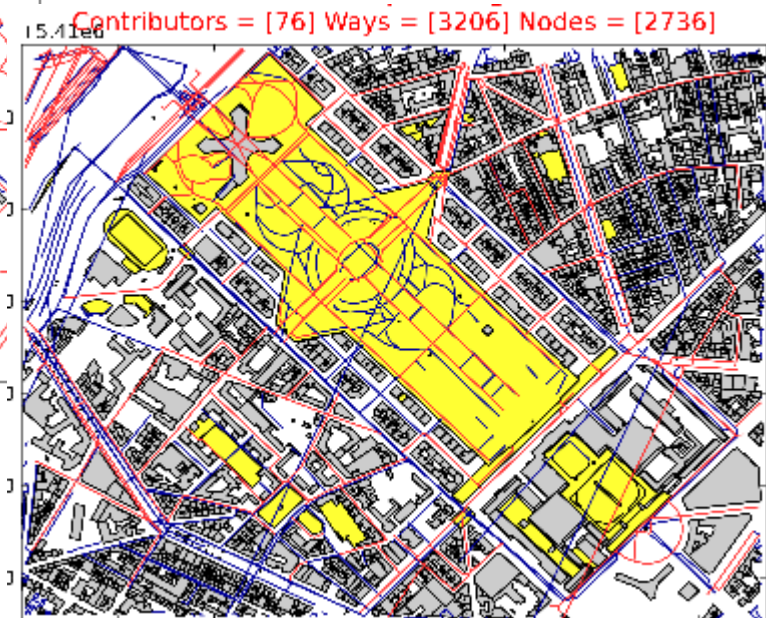
Historical Evolution of Cities – Interesting Patterns



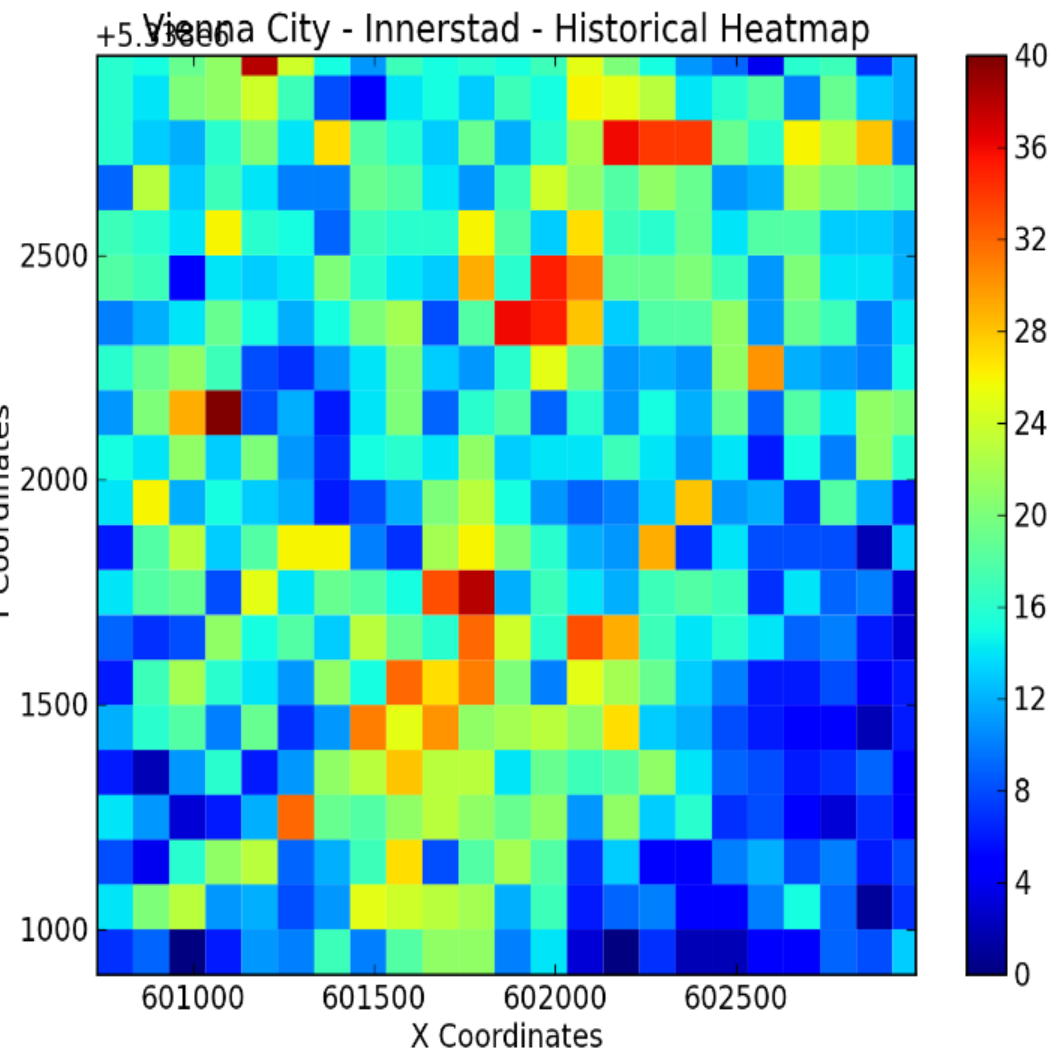
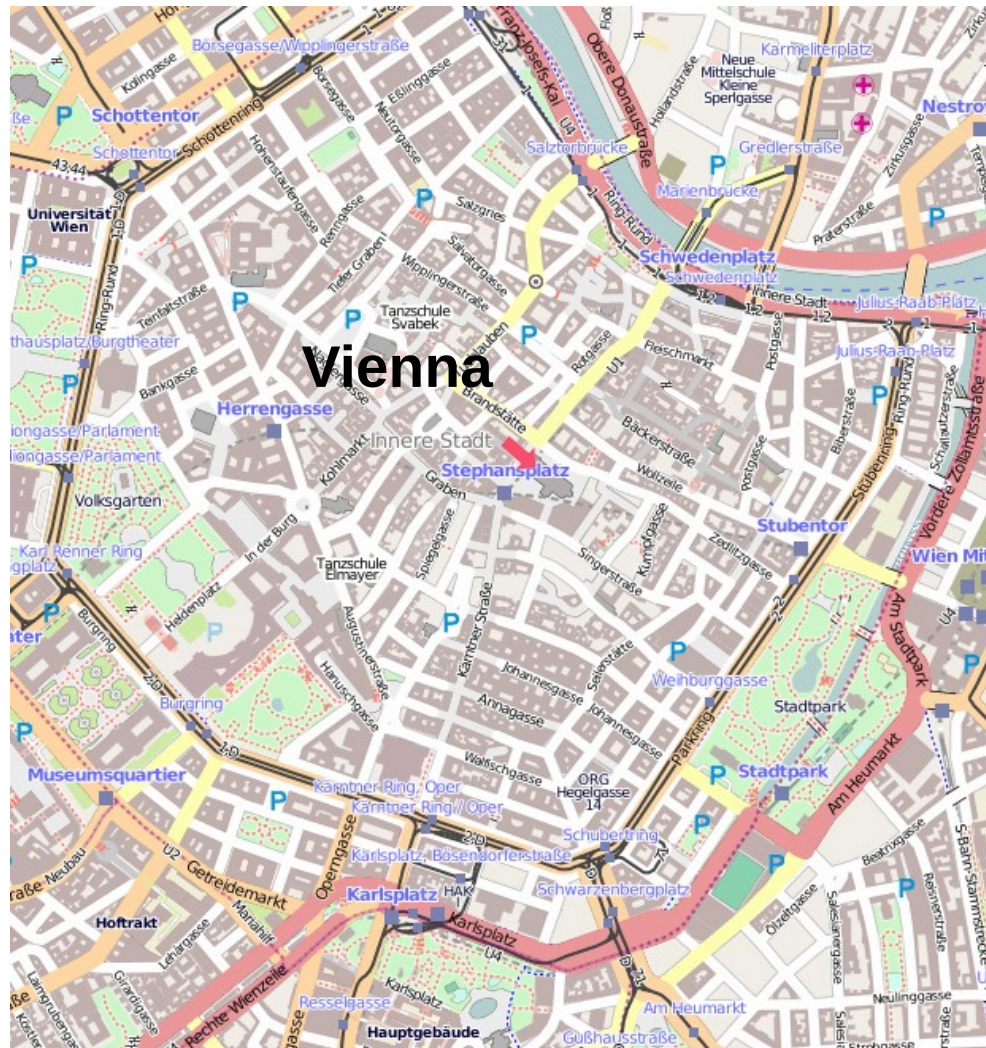
May 1st 2010



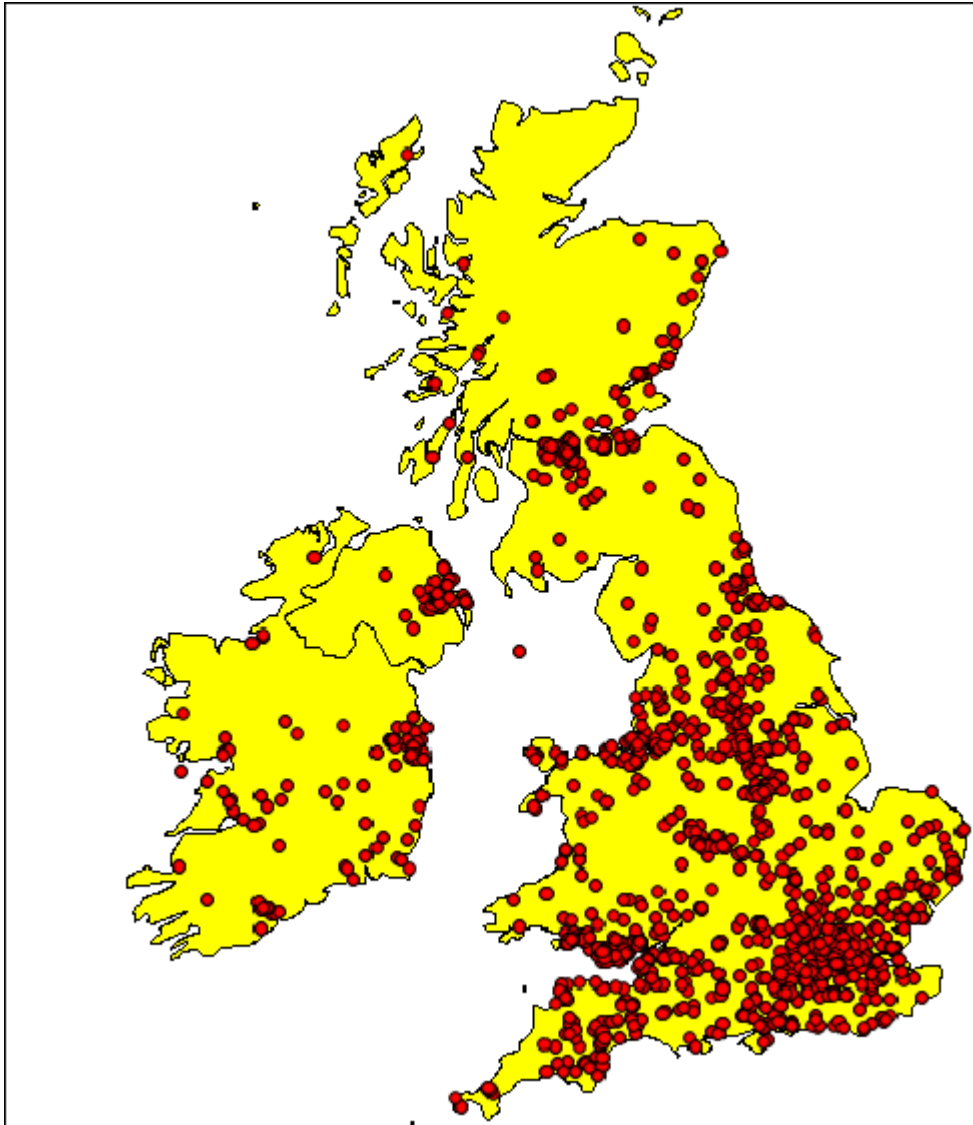
August 1st 2010



Historical Analysis – Total Contributors in 100M Grid Cells



Heavily Edited Objects in OSM



**Similar characteristics to
“Featured Articles” in
Wikipedia**

**Heavily Edited - 15 or more
edits**

**Changes to: geometry and
tagging**

**Also considered features
from Austria and Germany**

Total: 25,000 features

Mooney and Corcoran (2012 – TGIS)

Mooney and Corcoran (2012 – submitted Journal App Geog)

Table 7: Number of unique values assigned to the “highway” tag of objects in the four OSM databases. The column ‘highway’ indicates the number of unique values assigned.

Highway	UK	Ireland	Austria	Germany
1	4999 (59.4%)	298 (50.5%)	1110 (47.1%)	495 (54.8%)
2	2621 (31.2%)	222 (37.6%)	855 (36.3%)	271 (30%)
3	650 (7.7%)	60 (10.2%)	305 (12.9%)	110 (12.2%)
4	117 (1.4%)	8 (1.4%)	78 (3.3%)	22 (2.4%)
≥ 5	22 (0.3%)	2 (0.3%)	10 (0.4%)	5 (0.6%)

Number of Unique values



Database	Attribute	Compliance	Observations
UK	Landuse	39 (10)	Spelling Errors
UK	Highway	138 (101)	Spelling errors 'pedestrianised', 'tersiary' and assigning the name of the road or highway to the highway tag
Ireland	Landuse	5 (0)	All valid
Ireland	Highway	30 (0)	All valid
Germany	Landuse	105 (76)	Spelling errors 'medow', 'forrest', and invalid assignments 'fruit trees'
Germany	Highway	49 (12)	Street names assigned to highway attribute
Austria	Landuse	72 (43)	Spelling errors of core values, invalid values
Austria	Highway	118 (81)	Spelling errors, multiple value assignments, alternative values from bulk import



“NAME=” Changing Example

v2 name=**Station Road** userid = 11895 06/07/2008
v8 name=**Oswald Road** userid = 11985 06/08/2008
v11 name=**Frodingham Road** userid = 11985 07/08/2008
v13 name=**NULL** userid = 26825 10/10/2008
v23 name=**Ferry Road** userid = 11985 02/06/2009
v25 name=**Old Crosby** userid = 11985 17/06/2009
V26 – 2010-04-12 Current

“highway=” changing example: Lissinger Straße, Gerolstein, Rhineland-Palatinate



Version	Highway	Date	User_ID
1	PRIMARY	Oct 2007	16631
3	TERTIARY	Nov 2007	16631
8	TERTIARY;PRIMARY	Dec 2007	16631
9	SECONDARY	Jan 2008	16631
18	RESIDENTIAL	Apr 2008	6390
22	TERTIARY	Jan 2009	61450
29 (curr)	TERTIARY	Mar 2010	95223

But heavily edited objects ...
are often not geographically related

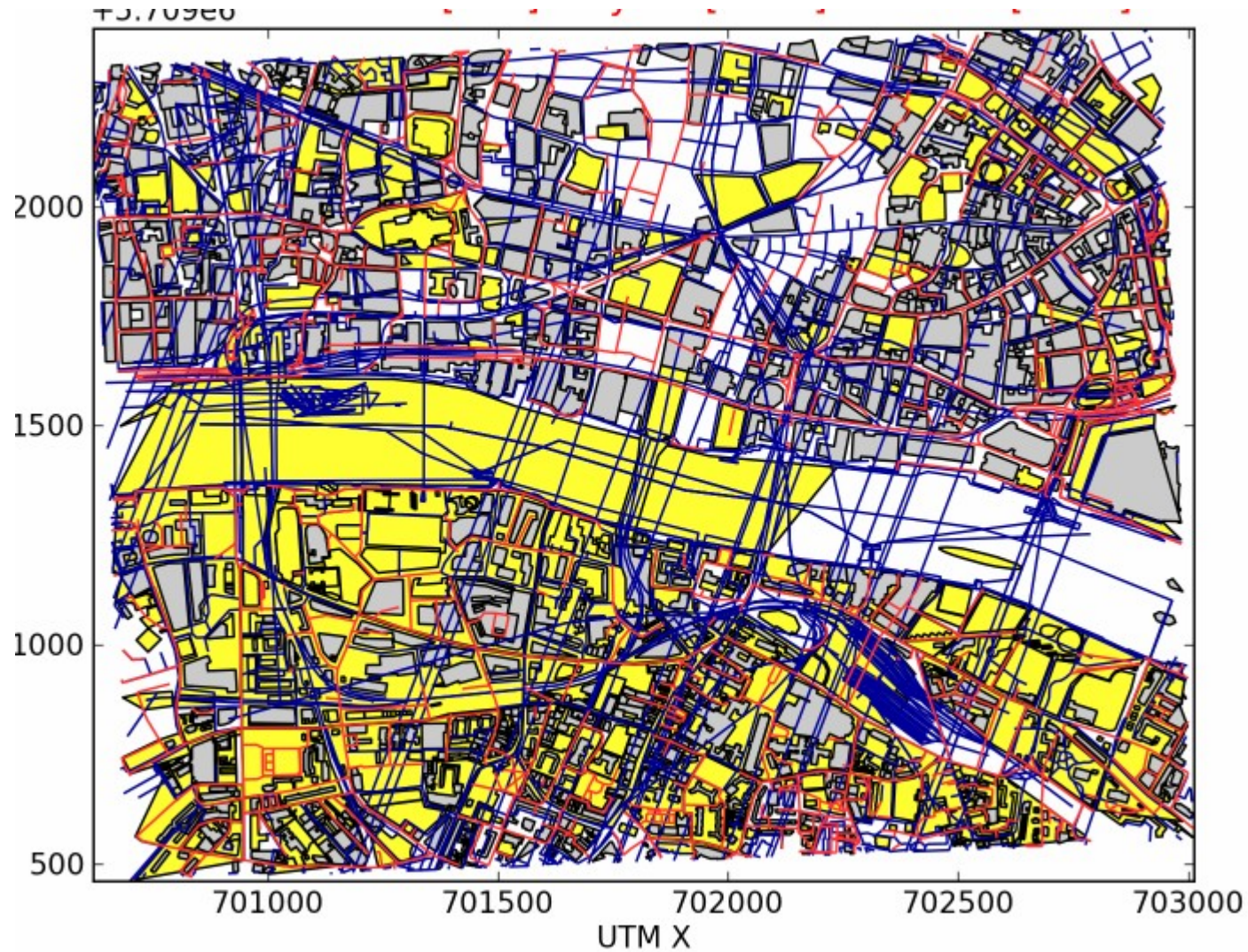
LONDON 'THE CITY' ... TOTAL: (to June 2011)

10,318 Edits

263 Contributors

6,178 (60%) made by 10 contributors

3,081 (30%) made by TWO contributors



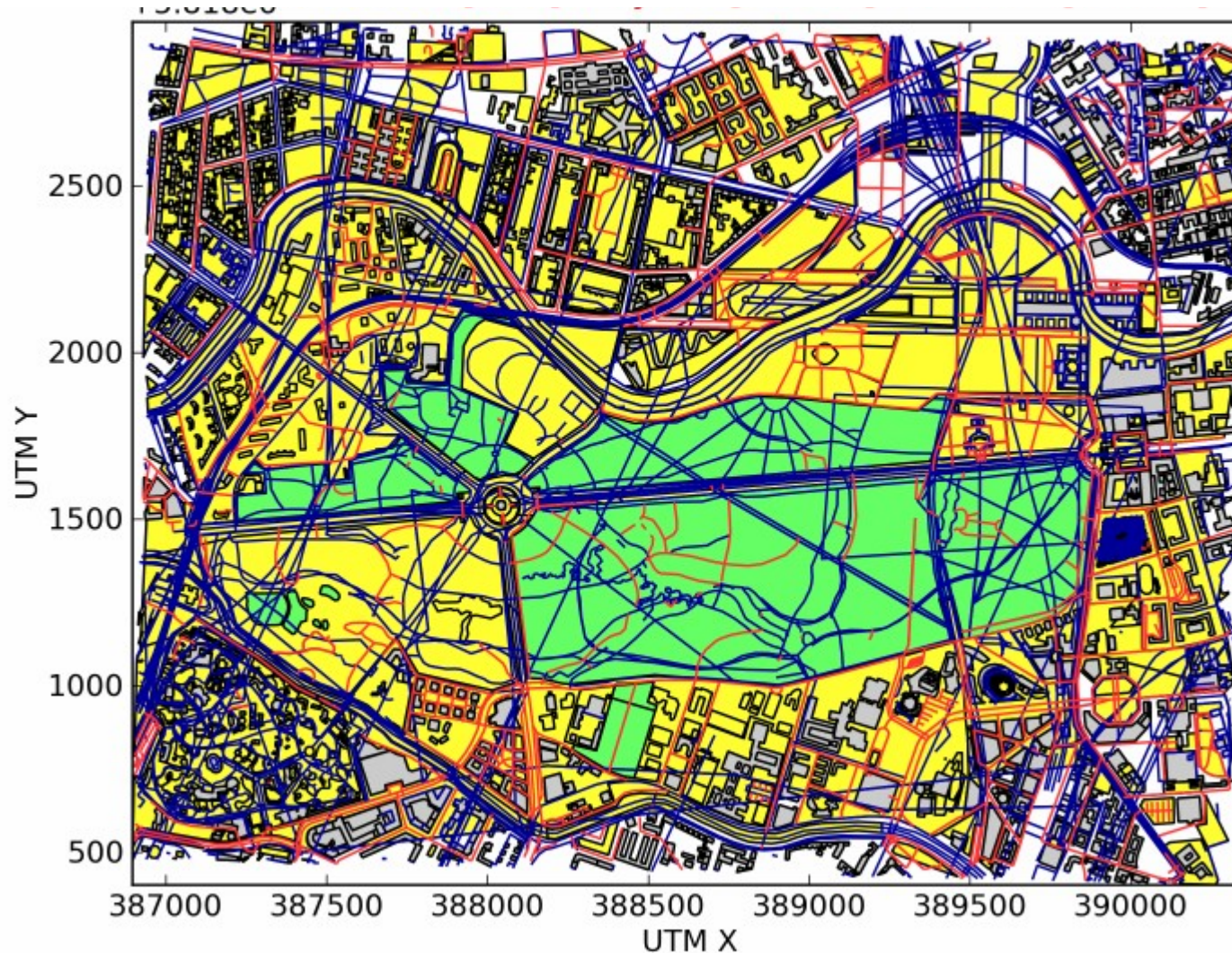
BERLIN – GERMANY - TOTAL: (to June 2011)

30,750 Edits

313 Contributors

23,484 (76%) made by 10 contributors

16,355 (53%) made by ONE contributor



How “useful” are edits over time?

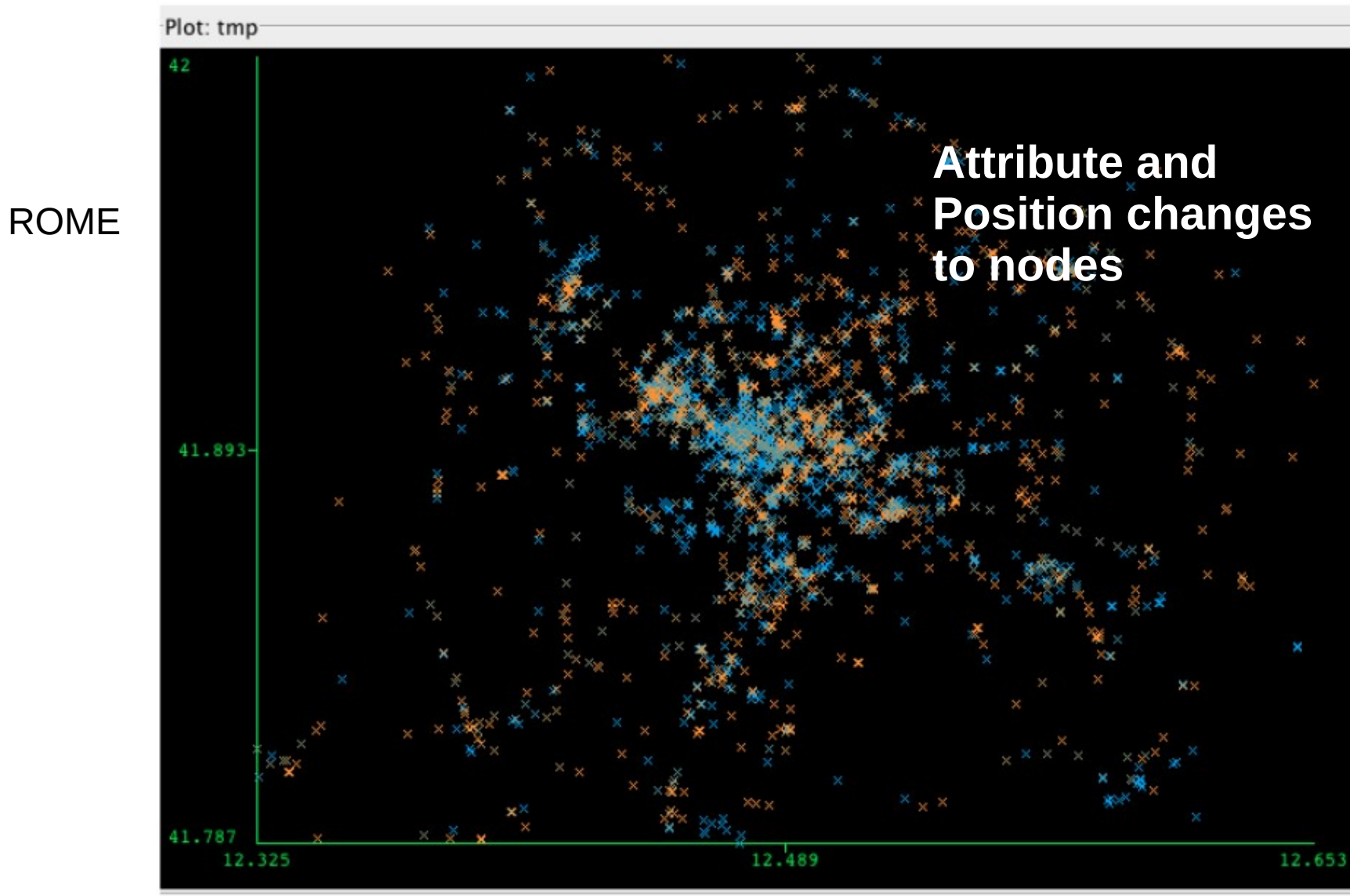


Figure 1: Geo location of POIs in Rome, where the orange colour is higher usefulness of updates on POI, and blue is least useful.

Mashhadi, Quattrone, Mooney (2012)

London – more useful edits to nodes deep in the city?

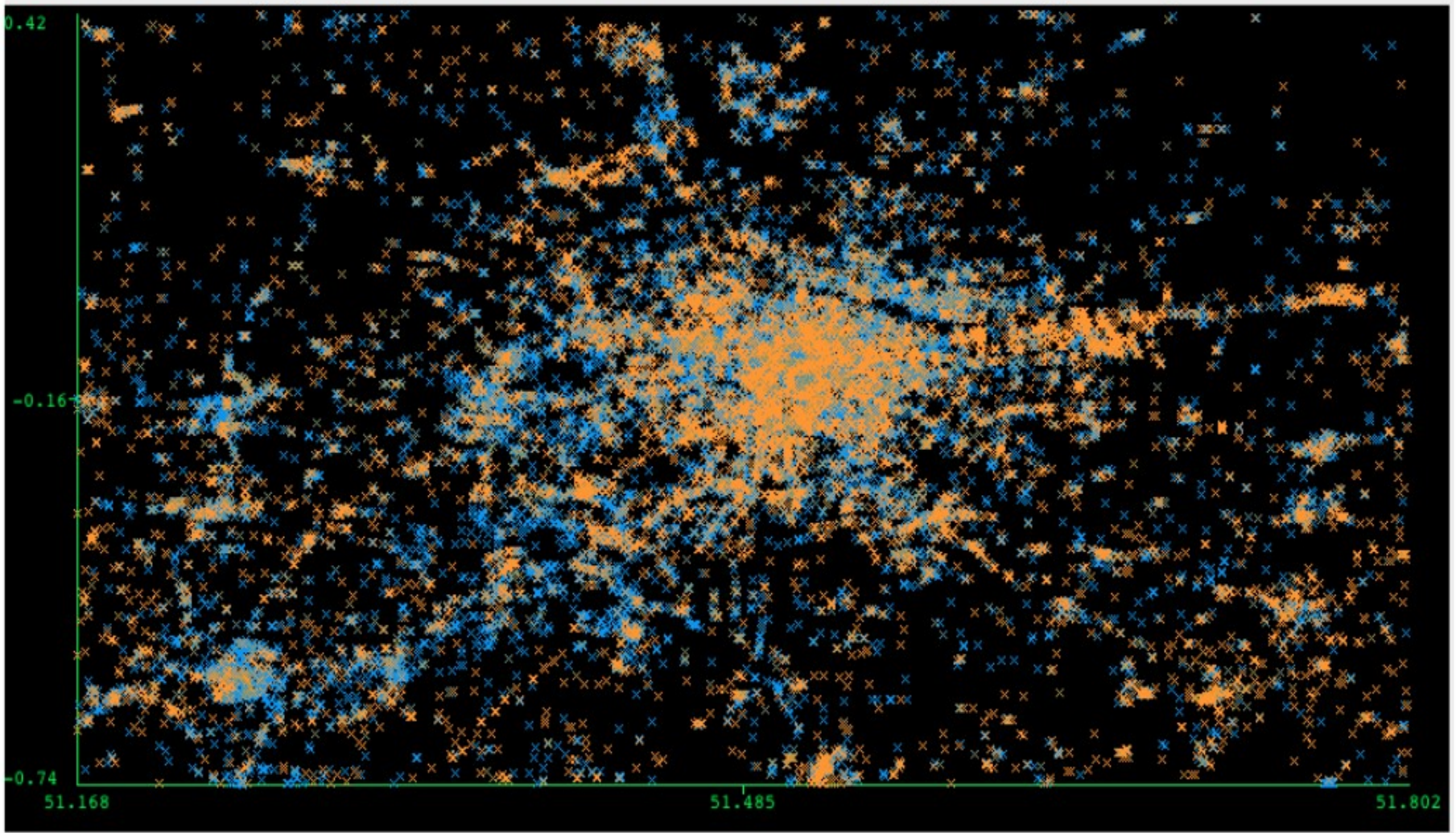
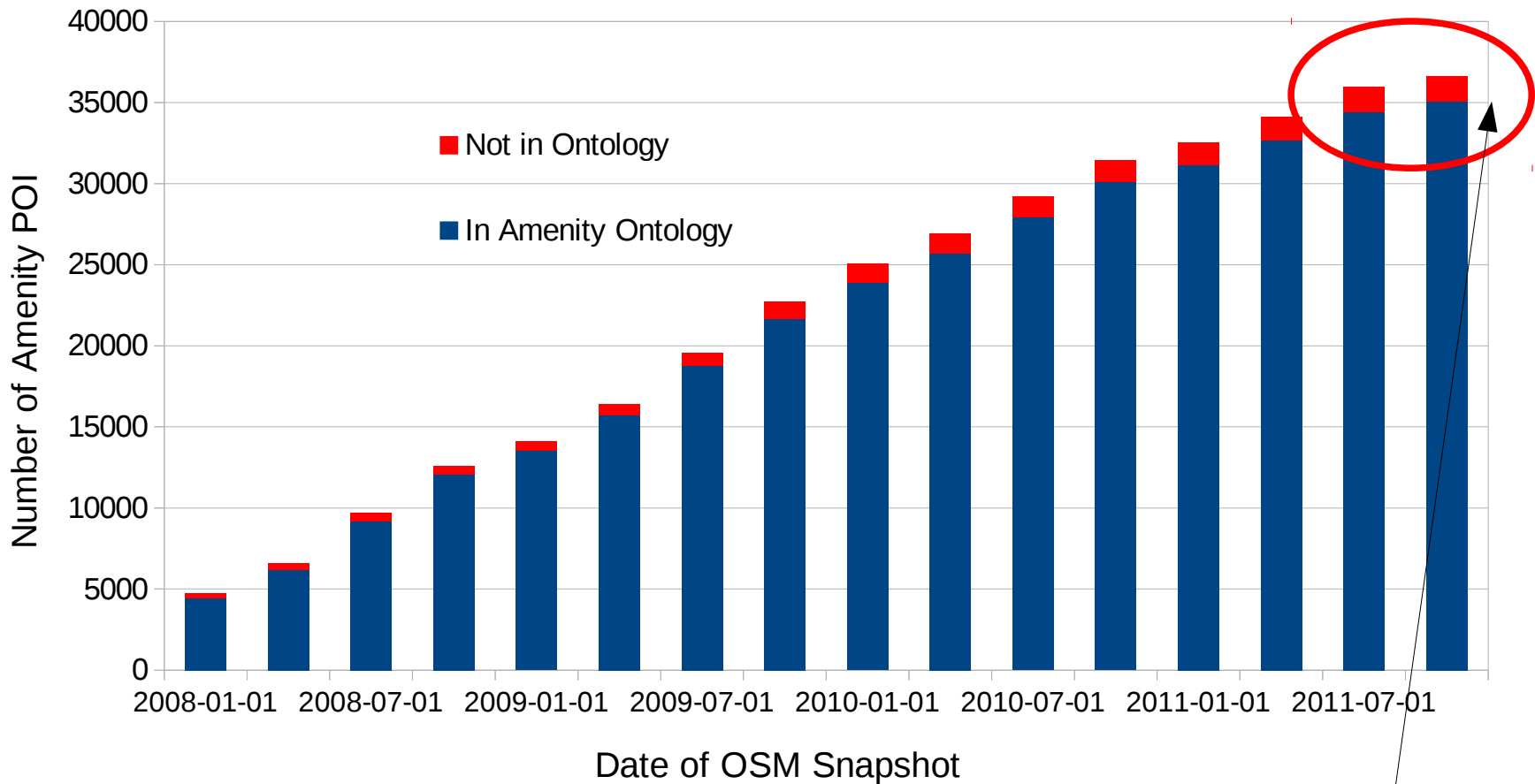


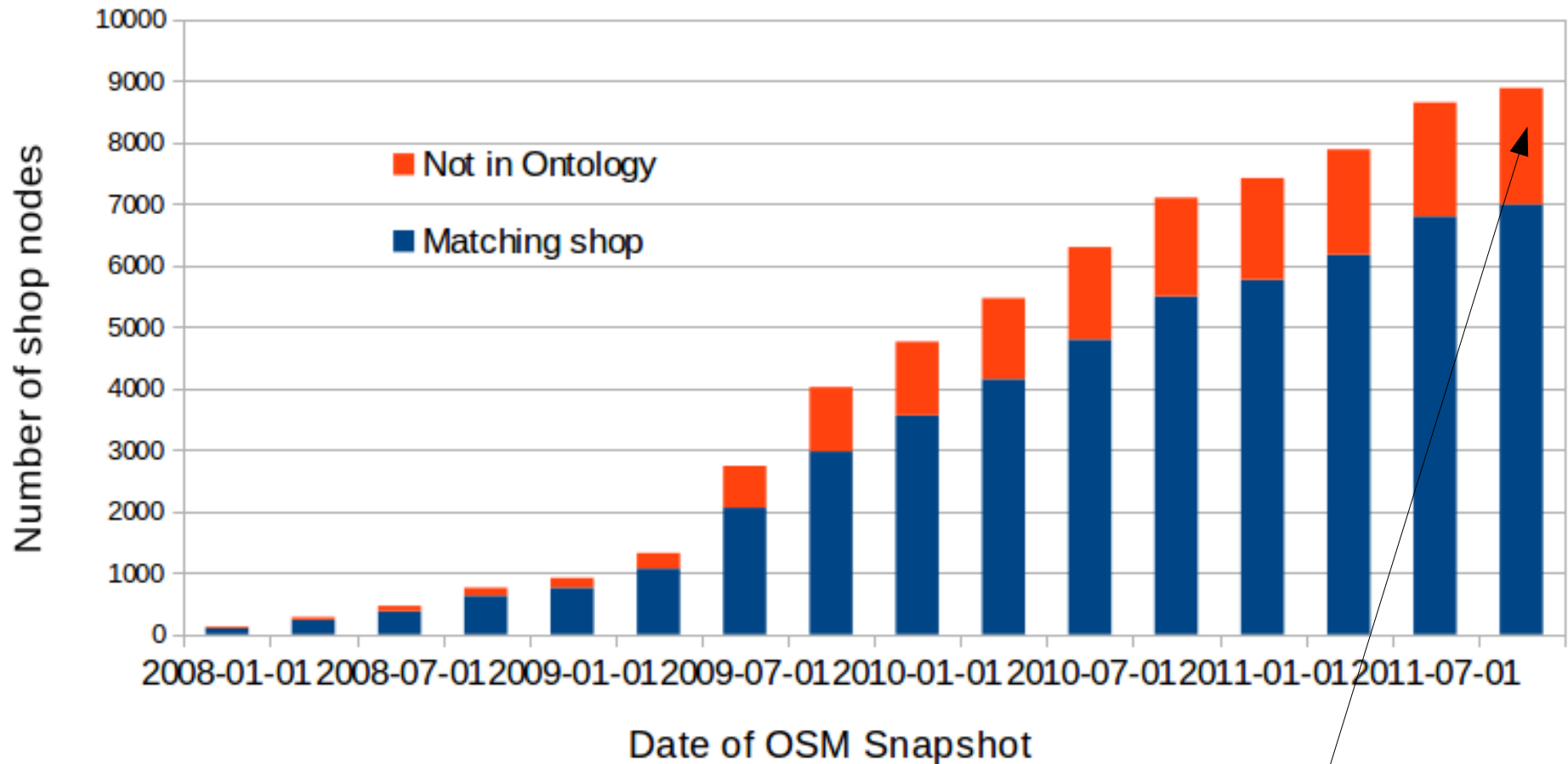
Fig. 9: Geo location of POIs in London, where the orange colour is higher usefulness of updates on POI, and blue is least useful.

London: Amenity POI – are contributors using values from the OSM ontology?



**About 7% of
AMENITY POI**

London: “shop” POI – OSM ontology provides 60 key values



27% of shop=* are not in ontology

Top 20 ranked invalid assignments to amenity tag in London

November 2008

bus_stop,91
supermarket,84
hotel,27
doctor,17
playground,16
dead_pub,14
speed_trap,13
shop,13
medical_centre,11
village_hall,10
nursery,9
health_centre,8
surgery,8
undefined,7
speed_camera,7
hide,7
garage,6
swimming_pool,5
hairdresser,5
job_centre,5

November 2010

emergency_telephone,181
bus_stop,109
supermarket,98
hotel,44
playground,35
dead_pub,28
street_light,26
shopping,22
nursery,21
marketplace,20
speed_trap,19
medical_centre,19
shop,17
motorcycle_parking,17
litter_bin,16
club,16
notice,14
health_centre,14
village_hall,13
gym,12

November 2011

emergency_telephone,175
bus_stop,109
supermarket,101
hotel,47
motorcycle_parking,44
playground,37
nursery,33
shopping,32
dead_pub,32
shop,26
street_light,26
marketplace,23
gym,21
medical_centre,21
speed_trap,19
swimming_pool,19
health_centre,18
club,18
litter_bin,16
notice,14

Application of Machine Learning Techniques

- **Identify the following:**
 - 1) Incorrect spelling of correct values
 - 2) Invalid assignments of values from other feature types
 - 3) Obscure values
 - 4) Unidentified values or rubbish values

Automatically RE-ASSIGN values to tags to improve overall compliance

Using Fuzzy String Matching to recognise spelling errors

- amenity=place_of_worship
- {*place worship, worship, place_workshop*}
- amenity=cemetery {*Cemetary, Cemmemetery*}
- Shop=optician {*optican, optition, opticians*}
- **landuse=forest** {*forrest, forestry, forist, forestt*}
-
- Levenstein Distance and Jaro-Winkler Metric

Invalid Assignment

- emergency_telephone**,175 (now emergency = phone ... former Mapnik rendering)
- bus_stop**,109 (HIGHWAY CLASS)
- supermarket**,101 (SHOP CLASS)
- hotel**,47 (TOURISM CLASS)
- motorcycle_parking**,44 (unknown)
- playground**,37 (LEISURE CLASS)
- nursery**,33 (usually kindergarten)
- shopping**,32 (unknown)
- dead_pub**,32 (It is a stop-gap pending resolution of the more generic life-cycle tagging)
- shop**,26
- street_light**,26 (highway - street_lamp)
- marketplace**,23
- Gym**,21 (too generic)
- medical_centre**,21 (unknown - too generic)
- speed_trap**,19 (still in proposal)
- swimming_pool**,19 (LEISURE)
- health_centre**,18 (too generic)
- club**,18 (too generic . . might be nightclub)

Search for mapping to other object classes in OSM Map Features

Obscure values..

- **AMENITY**

Shortwave, **swimming_outdoor_shower_thingy**, micro_sc
ooter_parking, Aeroporto do Galeão do Rio de
Janeiro, **big_postal_sorting_office**, Greater University of
Ruana, dead_post_office, dog_bin, disused_post_box, Yell
ow-fever Centre, pub_being_built

-

- **HIGHWAY** = local_knowledge; **Surrey_Aerial**, junction,
minor, gate

-

- **LANDUSE** = **hackerspace**, ventilation shaft, observator

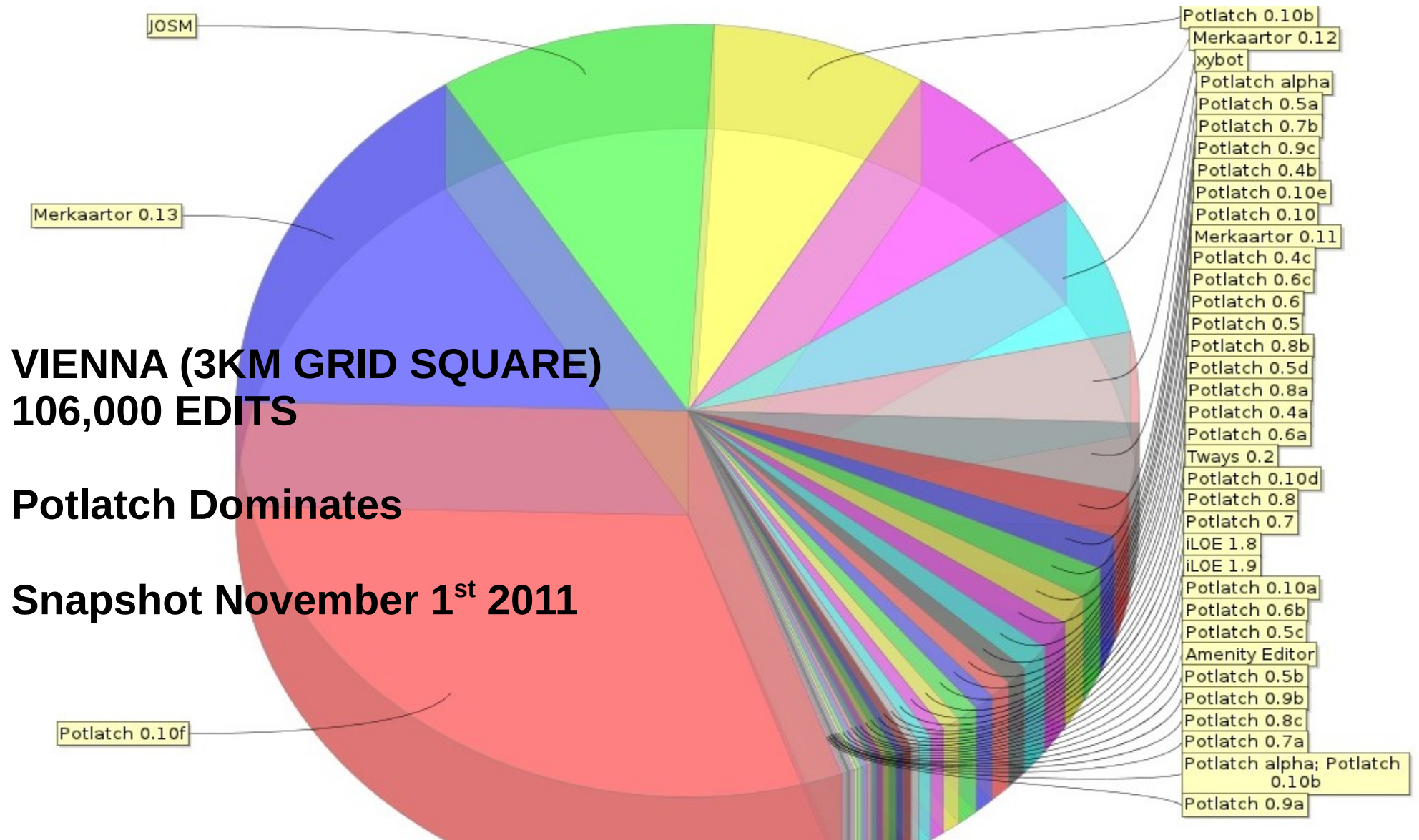
-

- **SHOP** = **not_sure**, bells, fruiterer, **generic**, bakery

Semantic Matching of Incorrect Values – much more difficult!

- Incorrect spelling – easy
- Completely different spelling – same meaning
-
- Example:
- **shop=estate_agent**
- Possible equivalents: UK and IE
(rental_agency, letting_agency, auctioneer
(IE),real_estate) + spelling mistakes

Influence of OSM Editor Software: Smith and Mooney (2012)

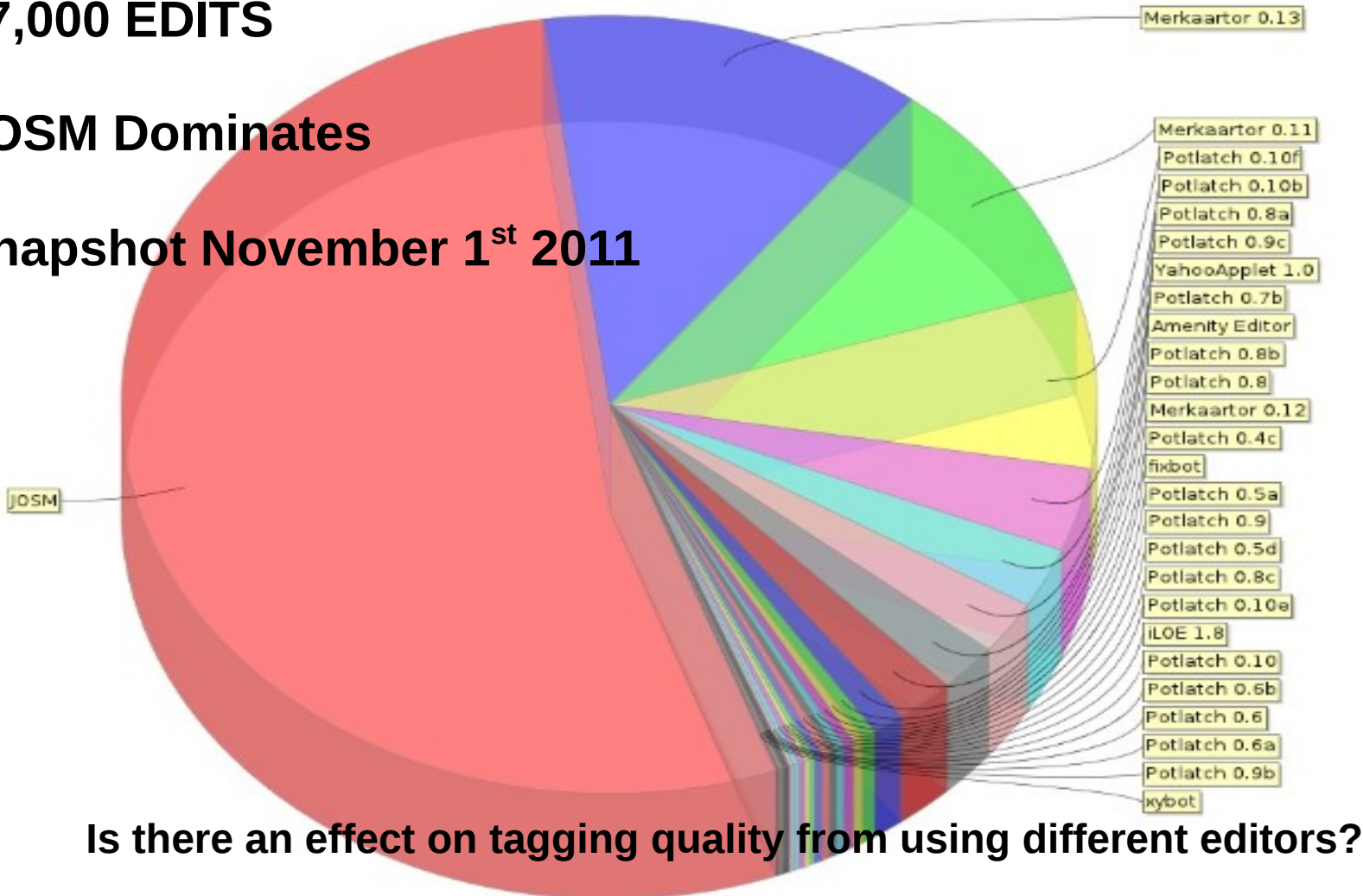


In a similar example for Berlin – JOSM becomes the dominant editor

Berlin (3KM GRID SQUARE)
87,000 EDITS

JOSM Dominates

Snapshot November 1st 2011



Is there an effect on tagging quality from using different editors?

Conclusions: OSM is a wonderful organic & evolving VGI database!

- **TAGGING VALUES:** Our approach of “fixing the fixable” tag problems could help to improve overall quality of tagging in OSM (good for LBS) – reduction of 90% (mean) invalid values
- **TAG STABILITY:** Examples of “name” attribute and “highway” attribute changes – potential to provide a *stability score* for features
- **RESPONSIBILITY:** Who is responsible – the editing software? The “super” contributors? The occasional contributors?

Thanks! Email: peter.mooney@nuim.ie