

Volunteered Geographic Information (VGI)

GEOG 384

What is VGI?

Widespread engagement of large numbers of private citizens, often with little in the way of formal qualifications

Goodchild, 2007

Geographic crowdsourcing – Open Street Map (OSM)!

Geosocial media

Another source of geographic data (as was web scraping)

Also overlaps with citizen science (e.g., GalaxyZoo, Audubon Christmas Bird Count, OldWeather) and Crisis Mapping (e.g., Ushahidi, Missing Maps)

What was the pain that created OSM?

- Needed a street file from UK Ordnance Survey
- Developed out of the failures of traditional mapping authorities
- Existing data not free and open
- Data is a static snapshot
- Existing data not editable

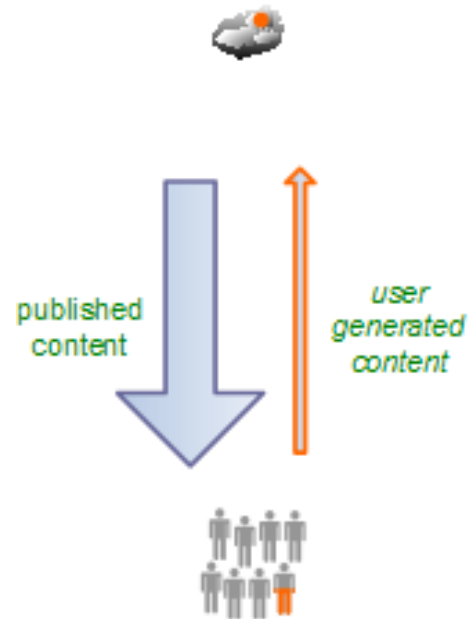


Key Features of VGI

GIS (Web 1.0)

mostly read only web

work web



45 million global users

1996

Geoweb (Web 2.0)

read/write/net



1 billion+ global users

2006

VGI as citizen sensors

large collection of intelligent, mobile sensors, equipped with abilities to interpret and integrate that range from the rudimentary in the case of young children to the highly developed skills of field scientists

Goodchild 2007, 26

Citizen sensing aided by

Tools are getting better and cheaper

Web 2.0 / Geoweb

Availability of GPS, digital cameras, sensors, smart phones

More people are getting involved

Broadband Internet access reaches even rural/remote communities

Computer literacy increases

The value of citizen sensors

Citizens can act as sensors of the environment

Closer to phenomena

More of them than scientists

Can sense change

Able to investigate

Can provide info as or more accurate than science

(Goodchild 2007)

Harnessing knowledge of ordinary people is scientific frontier of our time

You are a sensor

An ideology


Cheaper, more responsive

Get rid of government

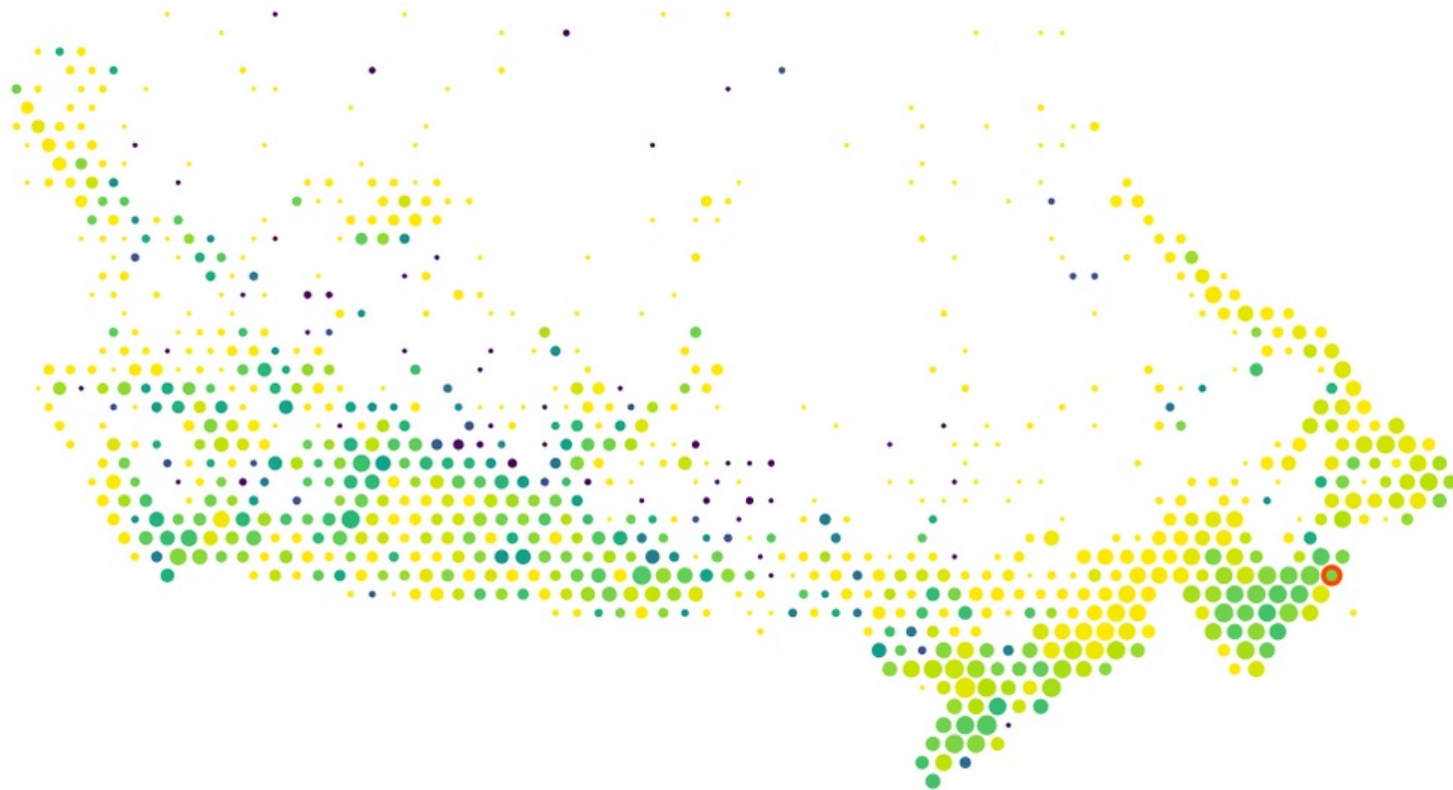
Exploitable labour force (“ghost work”)

More authentic data


Source of economic development (e.g., startups/civic tech, training data for AI)




From the sensor to the
crowd – Jeff Howe and
[http://youtu.be/F0-
UtNg3ots](http://youtu.be/F0-UtNg3ots)



Map of common nature-based names in Canada
bgrsquared.com/places/

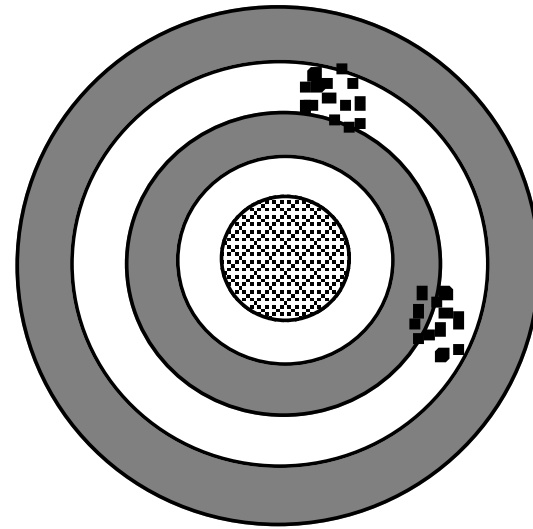
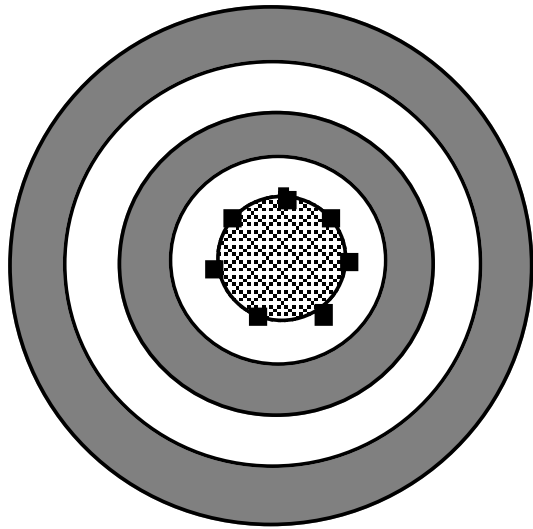


Let's rely on
crowdsourcing to make
maps. What could go
wrong?



1. How do we know the data contributed by crowd is accurate?

Accuracy and Precision




Accuracy via precision

Given enough eyeballs, all bugs are shallow

Eric Raymond

No formal coordination but large numbers of participants collaborating meaningfully, with no obvious monetary reward.

Budhathoki et al. 2010



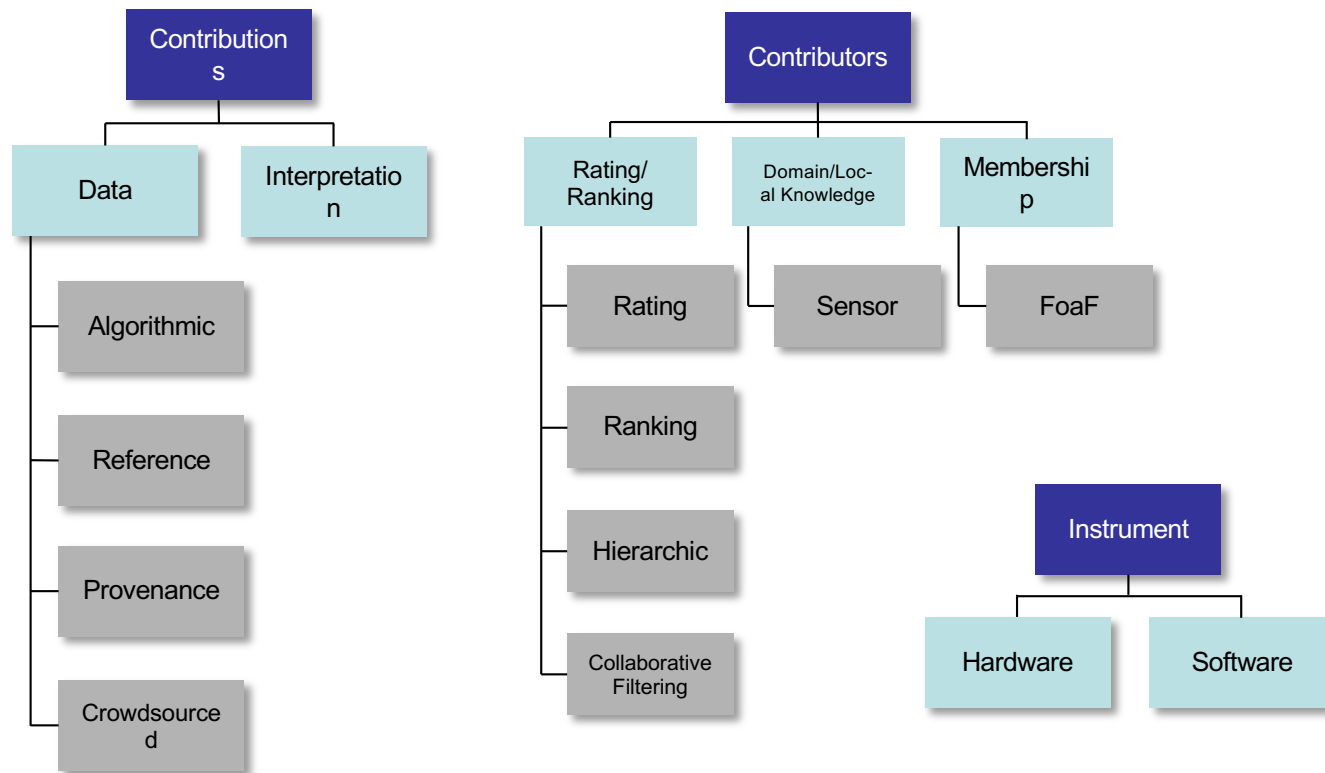
Let's take a look at OSM and accuracy. Or more broadly, data quality

- Geometry (positional accuracy) is incorrect
- Data that's never updated
- Updates from contributors that are ignored
- Geometry is correct but attributes aren't
- Can't do disambiguity
- Disconnected, complex polygons

Assessing
Spatial Data
Quality/
Accuracy in
GIS

positional accuracy
attribute accuracy
lineage
logical consistency
completeness

Assessing Spatial Data Quality/Accuracy of VGI



How do you
evaluate the
accuracy of...

Google Parking

Amateur sightings of invasive species

Yelp

OpenStreetMap

Algorithmically identifying inaccuracy of data

Buffer distances around roads, rivers

Examine precision (e.g., # points in a curve)

Investigate complexity of shapes

Klinkenberg's e-flora and e-fauna, where you spatially model to predict where contributions may be reasonably be predicted. Or not.

Assessing Contributions, cont.

Fix data by comparing it to a referent system

Fix data by looking at its provenance

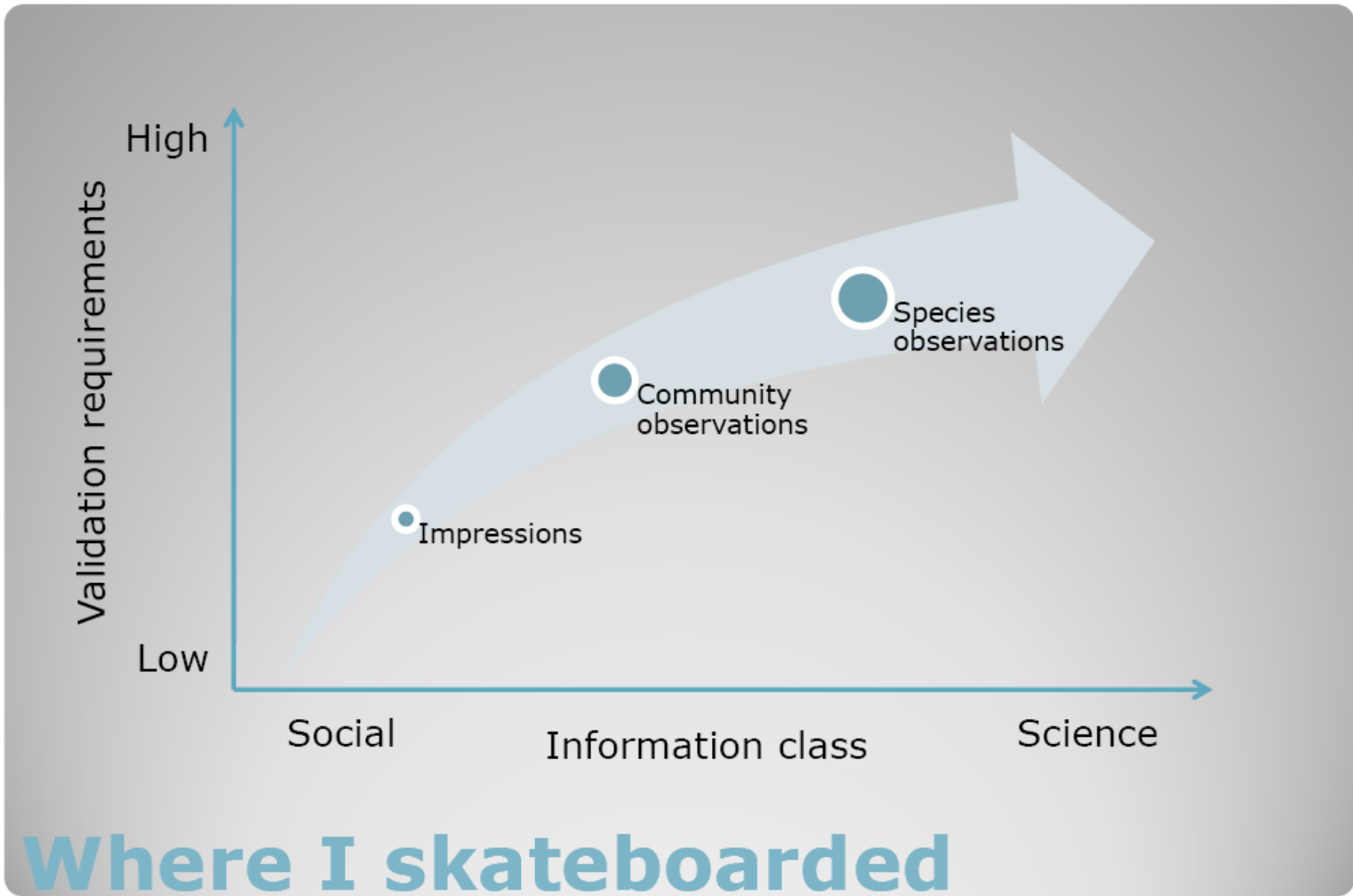
Fix data by inferring geospatial data (e.g., adding missing “roads?” by looking at crowdsourced contributions)

Is it accurate or....


Fit for a certain cause or not.

BUT every user has different requirements and expectations of quality, leading to highly differing perceptions of 'truth' (Girra & Bedard, 2009).

What do we do about assertions (after all, what's a restaurant review)?



Crowdsourced -- Klinkenberg 2010



If it's difficult to rate the contribution then what about rating the contributor?

Assessing Quality via the Contributor

Rating/Reputation

Rating

Ranking

Hierarchical (Wikipedia model)


Collaborative Filtering

e.g., # contributions

years a member

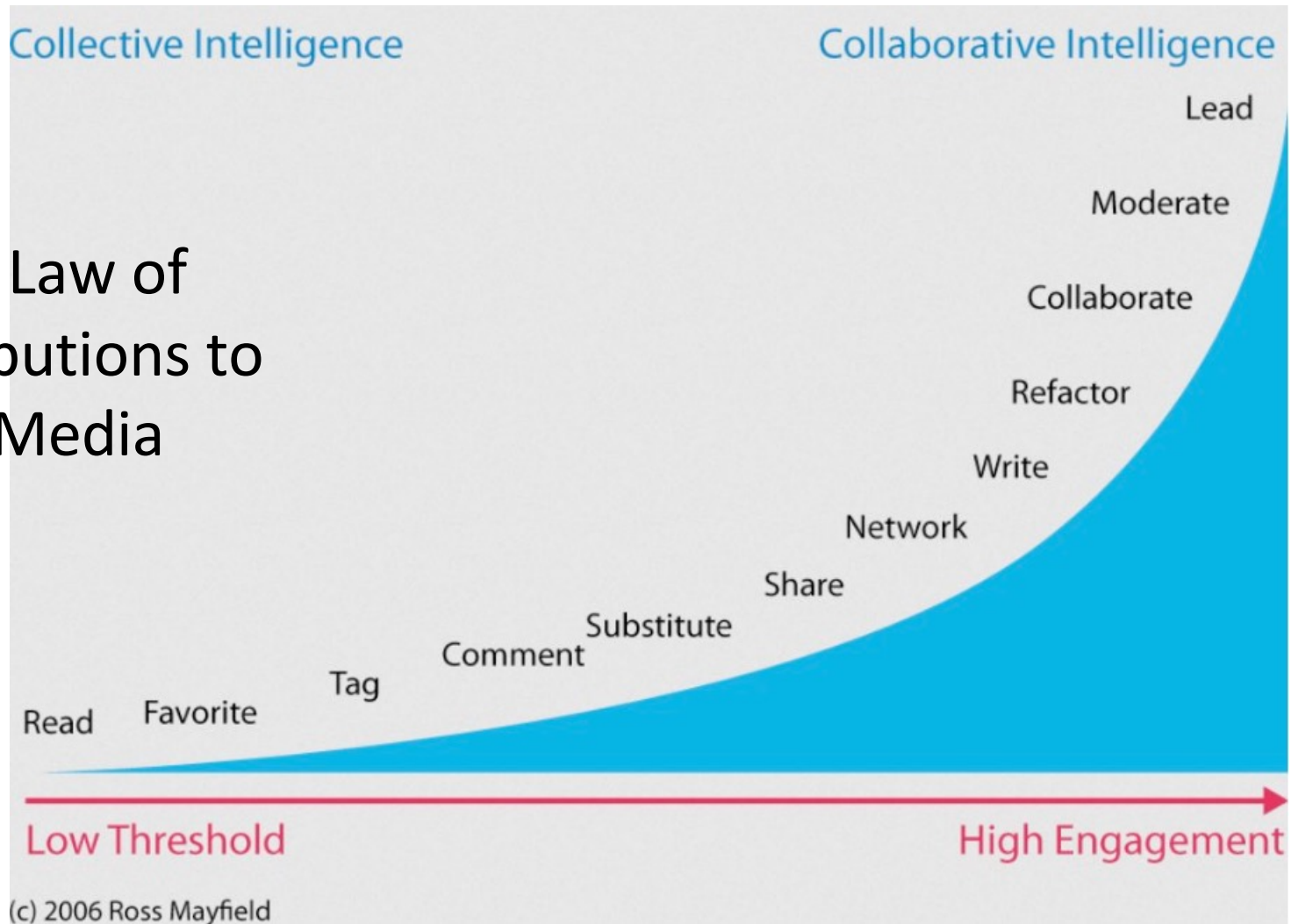
Domain knowledge--Sensor

Membership--FOAF



2. How do we motivate contributors to keep contributing?


Power Law of Contributions to Social Media



Motivations for Contributing VGI

- Strengthen social networks
- Make money
- Play
- Do it bc it' s techie/cool
- Learn
- Intellectually stimulate
- Contribute to science
- Contribute to effecting change
- Be altruistic
- Be Columbus
- **Evoke Pride of Place**
- Protect personal investment
- Enhance personal reputation (more friends and followers)
- Confirm professional interest
- Join
- Allow for self-expression
- Control agenda (a bad thing)
- Cause mischief
- Hold malice

Coleman, D.J., Y. Georgiadou, and J. Labonte. 2009. Volunteered geographic information: the nature and motivation of producers. *International Journal of Spatial Data Infrastructures Research* 4 332-358.



Always constructive
motivations to contribute,
right?

When is a
volunteer not a
volunteer?

Digital and other Divides

Selection Bias (e.g., Socio-economic)

Location of volunteer, site

Passive VGI

Coerced VGI

What “users” of VGI must consider?

What is the organization's rationale for considering VGI?

Additional data, faster data, more “accurate” data from citizen sensors, engaged public

To what extent, if at all, VGI should be adopted?

How will credible VGI contributors be qualified?

How may incorrect, misleading or damaging contributions be identified and excluded?

How much control over content and quality are organizations prepared to relinquish?

Legal Issues in VGI

IP and Copyright
Example: Postal Codes
Example of Liability

Intellectual Property & Copyright

IP: Property that results from original creative thought, as patents, copyright material, and trademarks


Copyright Law: exclusive right to make copies, license, and otherwise exploit a literary, musical, or artistic work, on various media (e.g., printed, audio, video)

Intellectual Property & Copyright

Who owns the referent data? Who has the right to copy?

Geolytica's use of Canadian postal codes
OpenStreetMap and sources of data

“Most information that you find on the web is copyrighted, including "free" maps like Google Maps. You may **never use copyrighted resources** because it can cause a lot of trouble to OSM. As a rule of thumb, use no external resources except those available in the editors. If you think you found a non-copyrighted resource that isn't available in the editors, please discuss it first with local mappers using [contact channels](#).”



Street View wasn't just a neat way to get imagery to accompany the data already found in Google Maps. As it happens, it was a way to drive the same roads that were already in Google Maps, tracing them with Google's own road teams, and—through efficiency and brute force—do away with those costly map licenses. Google has mapped the US and will surely map the rest of the world soon enough.

<http://gizmodo.com/5391966/google-and-the-deadly-power-of-data>

Intellectual Property & Copyright

Who owns the VGI and what is owned?

Remember TripAdvisor's ToS? Ownership is transferred from contributors to company.

OSM: "free to copy, distribute, transmit and adapt our data, as long as you credit OpenStreetMap and its contributors. If you alter or build upon our data, you may distribute the result only under the same licence."


In VGI, what are we owning: a single point or an entire dataset? The point or the polygon?

Then there's liability for VGI

Case 1. Resident of City of Regina reports a pothole via an Open311 app. Few days later, same resident drives car over pothole, wrecks universal joint. He sues. Who wins? Can depend on # lawyers.


Case 2. Heavy rains and flooding force people from their homes and concentrate them in a storm shelter. HOTOSM volunteers collect SMSs and map them. They receive an SMS that a levy has ruptured near the shelter and the shelter will soon flood. Throughout that day, they have received similar false reports submitted as pranks. They map the shelter as being safe. It wasn't. They are sued by a family member of one of the victims. The group or the individuals themselves might be held liable. Good Samaritan laws won't protect them.

Edward Robson. 2012. Responding to Liability: Evaluating and Reducing Tort Liability for Digital Volunteers. Commons Lab, Woodrow Wilson Center



Problems could arise from aggregating data (e.g., cutting and pasting from Twitter, SMS), disseminating false information, developing software in a sloppy manner, failing to act in a matter commensurate with similarly situated professionals, or failing to properly vet and supervise volunteers. Digital volunteers may also be subject to liability if they fail to act when they have a duty to do so. Such a “duty to rescue” can arise if a digital volunteer creates a hazardous condition, begins to render assistance, or forms a special relationship with survivors.

Edward Robson. 2012. Responding to Liability: Evaluating and Reducing Tort Liability for Digital Volunteers. Commons Lab, Woodrow Wilson Center



*Before publishing the report of the ruptured levy, the digital volunteer group must weigh the **precautions** against the **probability/magnitude** of the potential harm. Although the probability that there has been a rupture in the levy is low (because of the volume of phony reports), the potential magnitude of injury is high and could be compounded by the delay associated with extra precautions. A group might be able to discharge its duty by **aggressively attempting to verify the information before posting it to the map**, reporting the information to professional responders rather than directly to the public, or posting the information with a conspicuous disclaimer indicating that the report is unverified.*