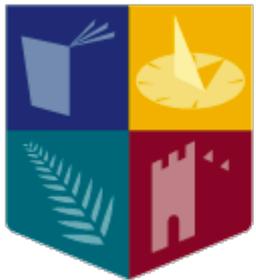


Academic Research Developments in Crowdsourcing and National Mapping

Peter Mooney

Department of Computer Science,
Maynooth University,
Ireland

Email: peter.mooney@mu.ie



**Maynooth
University**
National University
of Ireland Maynooth



Focus of this presentation

- Crowdsourcing in National Mapping: **What's happening in related academic research?**
- The focus of research: What are the burning questions? **What are the hot research areas?**
- What is the direction of research? **Is the research moving towards a common set of issues?**

A long time ago, in the universe of national mapping and cadastral agencies

**CROWDSOURCING? THE
FUTURE?**

It is the time before the geoweb, the smart phone, the always-on internet. NMCA's are working to continue to produce national mapping products to higher quality on smaller budgets/resources. This situation became even worse during the financial crisis of mid 2000s

A vision of Crowdsourcing in National Mapping in 1979?

Survey Review XXV, 195, January 1980

THE ROLE OF A NATIONAL MAPPING ORGANISATION*

A. G. Bomford

Director of National Mapping, Canberra, Australia

- “A fundamental question to be asked in deciding the area for each scale is, how much mapping will the country be able to afford to keep updated?”
- “Revision should take priority over new mapping”
- **Other people and organisations should be seen as resources waiting to be mobilised in the national cause:** (including) universities and colleges, map librarians and school teachers, map users (orienteering clubs, 4WD clubs, cycling clubs, bush walkers and ramblers)

1990s United States

Special Reports

Spatial Data Needs: The Future
of the National Mapping Program

John D. Bossler, Thomas C. Finnie, Barbara B. Petchenik,
and Thomas M. Usselman

- "It is conceivable that commercial needs for national spatially referenced digital data in transportation and marketing (census tracts, postal ZIP code data, street labeling, block addressing) will become so significant that if the government does not move to create and maintain a comprehensive, continually updated national dataset of such features a consortium of firms will have to create it."

Take home messages from 'Crowdsourcing in National Mapping' 2012

- **More focus on the social aspects of crowdsourcing of spatial data** (correct types of communication for the crowd, best channels to engage the crowd)
- **Crowdsourcing of spatial data should not be focused on maps along** - there are opportunities for using photographs, video, text, social media to assess the quality of crowdsourced spatial data.
- How does the **integration of crowdsourced spatial data** conflict with the NMCA's legal mandate and workflows?



Take home messages from 'Crowdsourcing and National Mapping' 2017

- Ensuring that contributors to VGI projects are interested and motivated to fix errors, update data and metadata, etc into the future.
- Should NMCAs become the gatherer of all crowdsourced spatial data or to take responsibility in quality assessment
- Should NMACs strive to become "less traditional" in future?
- Understanding the licensing issues and the quality requirements will be critical in how VGI/Crowdsourcing and NMCAs work together.
- VGI/Crowdsourcing and NMCAs could collaborate on mutually beneficial or interesting use-cases



The research focus of research on VGI has been moving towards...

- **The structure of the 'crowd'** and volunteers (geo-)demographic biases,
- **The impact of such biases** in different VGI projects,
- Research on how to **promote diversity of contributor communities**,
- Addressing the issues of transparency and trust while **protecting the privacy of the contributors**,
- Protecting the **intellectual property of crowdsourced data and projects**.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
2019, VOL. 33, NO. 8, 1588–1593
<https://doi.org/10.1080/13658816.2019.1593422>



Taylor & Francis
Taylor & Francis Group

EDITORIAL

Check for updates

Crowdsourced geospatial data quality: challenges and future directions

The crowdsourced approach...

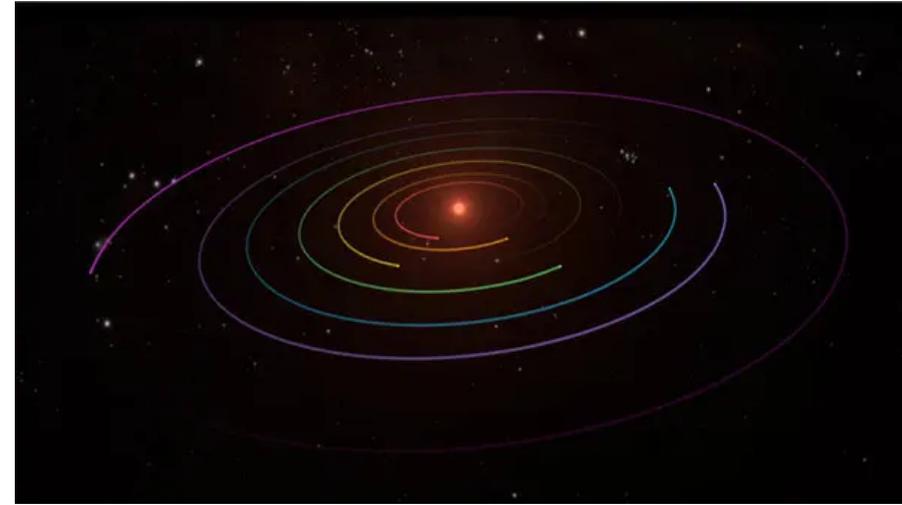
- “Those who attend are enthusiastic and generally fulfilled by their efforts. Those that attend can clearly see the benefits of their work. **Their actions offer a positive service to society** that is not often seen in professional map-making. **This is something to be worked on by critical cartographers rather than rallied against**”



Duggan, M (2019) Cultures of Enthusiasm: An Ethnographic Study of Amateur Map-Maker Communities
DOI: 10.3138/cart.54.3.2018-0002

Personal view – evolution of crowdsourcing geographic information (CGI)

- **2008 – 2010** CGI is fundamentally different to NMCA data
- **2010-2012** – CGI is looking like NMCA (fusion, integration)?
- **2014-2016** – CGI is surpassing NMCA (update cycles, etc)
- **2016 – 2018** – CGI and NMCA can work together .
- **2019** - ... divergence? – CGI on a different path due to privacy, ethics, AI, ML etc. NMCAs accessing high tech systems (3D, satellite, drone, airborne, etc)



Joining 'the crowd' and the authority

How to combine expert and crowd data to leverage the potential of crowdsourcing while maintaining the quality standard?

- **Exclusion:** differences between the crowdsourced and expert data are too large and the former must be discarded.
- **Partial conflation:** the agreement between the two approaches is discontinuous. Conflation can be applied where correlation is expected to be high. Therefore, conflation requires an understanding of where the errors are likely to occur.
- **Replacement:** the crowd and the experts strongly agree so that observations from the two groups are interchangeable,

Recommendations for hybrid data collection from Waldner et al (2019) for thematic maps.

- **An initialization set interpreted by experts to rank the volunteers** and reward them not only based on the quantity but also on the quality of their contributions.
- **The crowdsourcing campaign is run using a user-friendly data collection tool** designed to reduce task and interpretation complexity
- **Strategically allocate experts to those sampling units where the crowd lacks consensus** (targeted verification of uncertain contributions).
- **Strategic allocation of experts allows to collect multiple expert observations per sampling unit contribution**, thereby building up confidence in their interpretations (help consensus)

Quality of Information in Mobile Crowdsensing or Crowdsourcing

- Mobile crowdsensing empowers ordinary citizens (or users) with the capability to actively monitor various phenomena pertaining to themselves (e.g., health, social connections) or their community (e.g., natural environment).
- Three complex research questions:
 - How can we trust that the smartphone-human sensors will send **useful information**?
 - How can we **enforce the submission** of useful information?
 - How can we **estimate the usefulness** of the submitted information?

Mobile crowdsourcing contains many research problems in QoI (Restuccia et al. 2017)

- **Assessing Trustworthiness of Complex Information**
 - Very difficult problem with complex Information: Multimedial (audio, video, image) and Qualitative (opinions regarding events under consideration)
- **Gamification for Quality of Information Improvement**
 - Gamification effects are greatly dependent on the context in which gamification is being implemented, as well as on the users using it.
- **Game-Theoretical Trust Frameworks**
 - development of game-theoretical trust frameworks based on behavioral game theory. Consider bounded rationality for less-than-rational behavior in participants

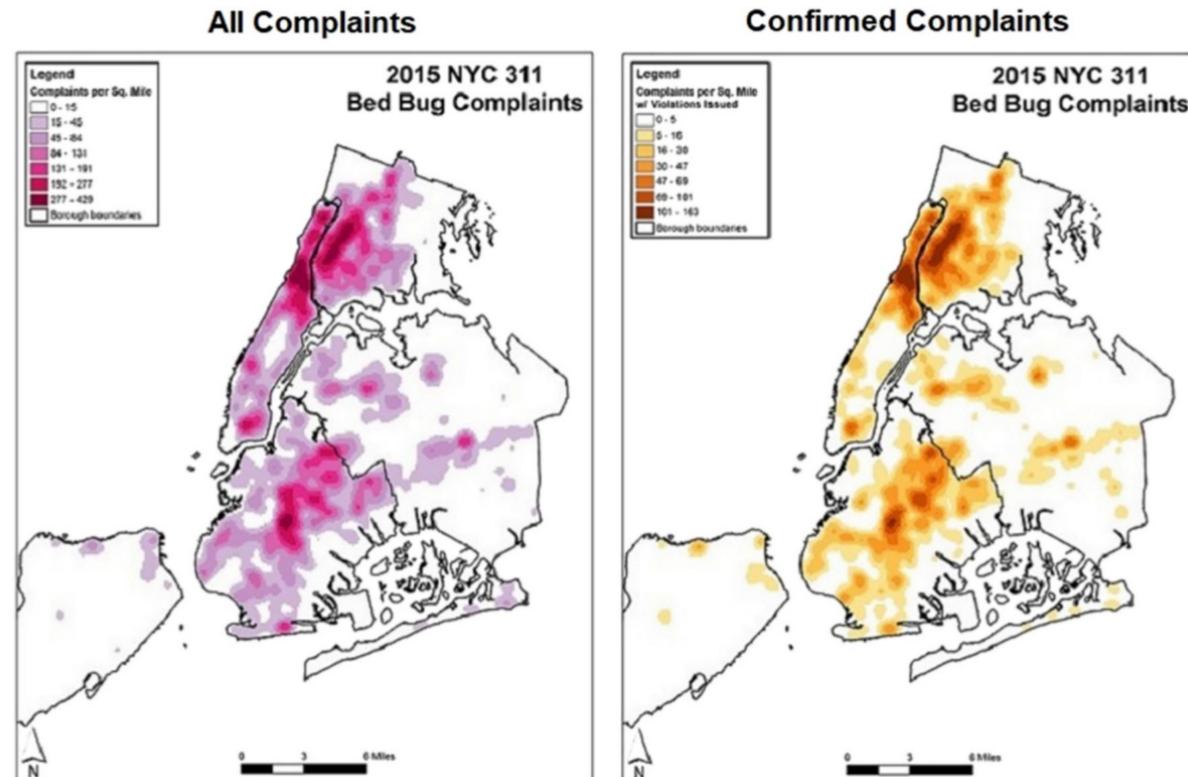
Some recent examples...



Social and behavioral processes affect the quality and quantity of health-related VGI

McLafferty et al. (2020), <https://doi.org/10.1016/j.healthplace.2019.102282>

- Strong geographic variation in credibility of bed bug reports, with negative reports generated more frequently from high-value residential buildings located in high-income neighborhoods



- VGI are embedded within, and derived from, a complex web of socioeconomic, behavioral, geographic, and technological processes. **Each health VGI data set has distinctive properties that reflect its construction.**

Obtaining public domain 3D city models is challenging and limited

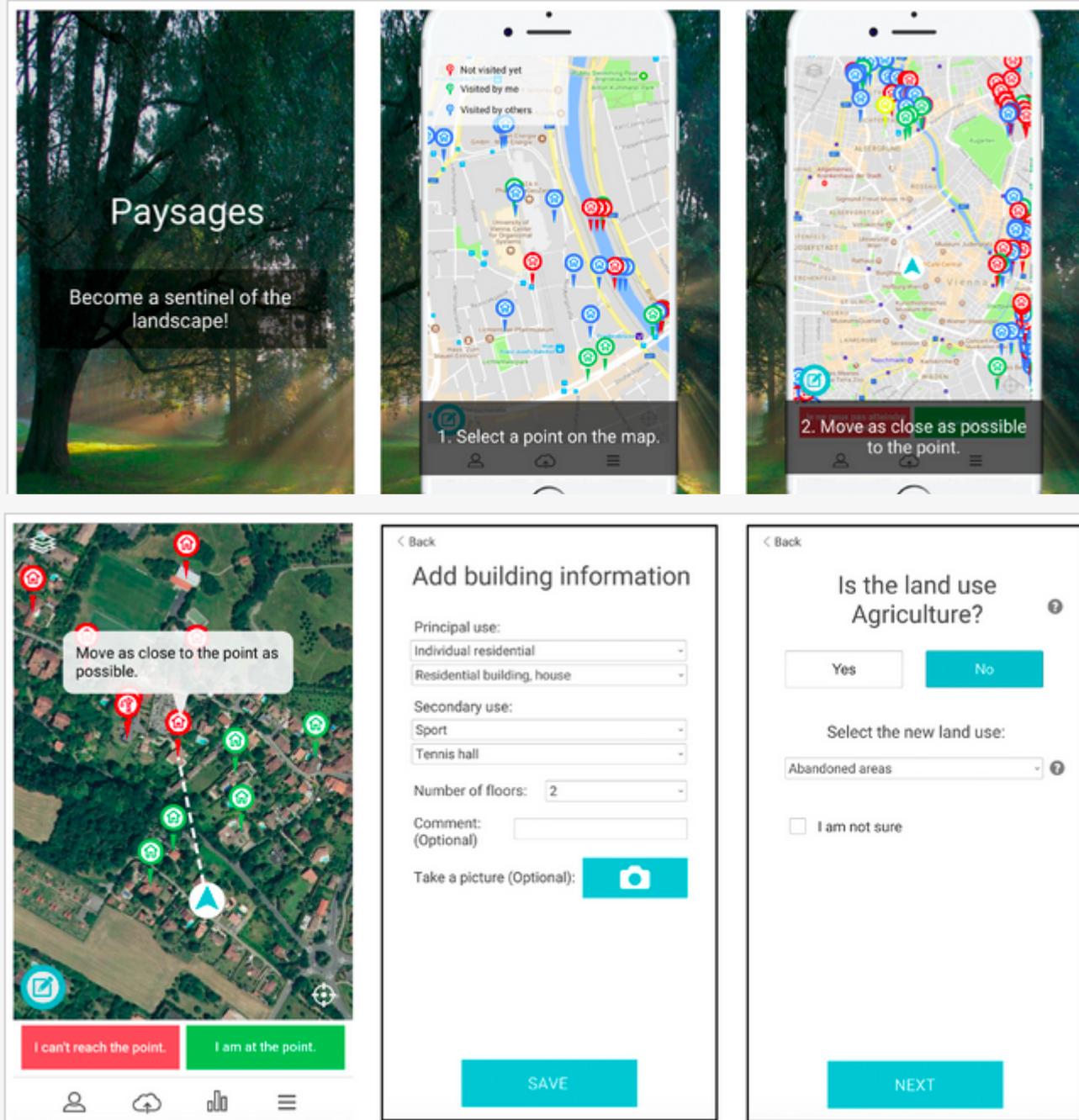
- A semi-automatic VGI-based calculation of OSM building heights from single contributed photographs
- Creation of LoD1 building models as a first step to enhancing the already established 2D OSM map infrastructure to the 3D domain.



Bshouty et al (2020), Towards the generation of 3D OpenStreetMap building models from single contributed photographs, Computers, Environment and Urban Systems, Volume 79

Integrating Community Science into Land Cover, Land Use, and Land Change Detection Processes in an NMA

- A strategy and workflow for updating and correcting LULC information on a three-year cycle and building data on a continuous update cycle for the national mapping agency, IGN France.
- One of the most challenging task is to build and maintain a community around a collaborative initiative
- Goal is to develop new methods to determine if the change is meaningful for the authoritative databases under consideration



Olteanu-Raimond et al,
Land 2018, 7(3), 103;

<https://doi.org/10.3390/land7030103>

Crowdsourcing Geographic Information

Legal, ethics, privacy.



The context for Crowdsourcing

- **Some GI scientists and legal scholars have specifically discussed the legal issues in using crowdsourcing for geospatial information** (Gundersen, 2017; Cho and Cromptoets, 2018)
- From the desks and the laboratories of researchers there are now large, **complex and rich datasets available for analysis from a variety of platforms.**
- **Personal information, collected as Crowdsourced Geographic Information, now creates data linkages between individuals, devices, organizations, the environment.** When the spatial and geographical aspects of these datasets and information streams are explored **new knowledge can be extracted.**

Mooney, P et al. 2017. Considerations of Privacy, Ethics and Legal Issues in Volunteered Geographic Information.

Table 1: Privacy, ethics and legal issues for actors involved in the collection, production and dissemination of VGI.

Who am I? I am ...	An example of a privacy issue	An example of an ethical consideration	An example of the legal issues involved
A citizen contributing my data to a VGI or Crowdsourcing Project	Ensuring that personal information is not linked to any contribution. Nor will the contribution provide the personal or private information of another individual.	A contributor will report observations honestly and truthfully, and, to the best of their ability, they will contribute accurate information.	The citizen will obtain consent or permission to survey, record or measure a specific area or geographic feature.
A National Mapping Agency, Environmental Ministry, Geological Survey, etc.	Ensuring that the organisation gives careful consideration to the scale at which it provides geographic data and to the contents of the metadata attached to those data.	The organisation will not knowingly provide false or inaccurate data or information nor report with bias on specific geographic themes.	The organisation is bound by many legal requirements to produce mapping products. The organisation is legally bound to the quality of data produced and could be liable to consequences of the use of these data.
A commercial mapping company	Privacy can mean keeping the information about the sources of the VGI hidden from public view.	Acting responsibly, ensuring the privacy of citizens is maintained and that the VGI is distributed using an appropriate licence.	The focus here is on the terms and conditions of the type of licence applied to the VGI, whether it be for the commercial usage of the data or the integration of the data with other data products.

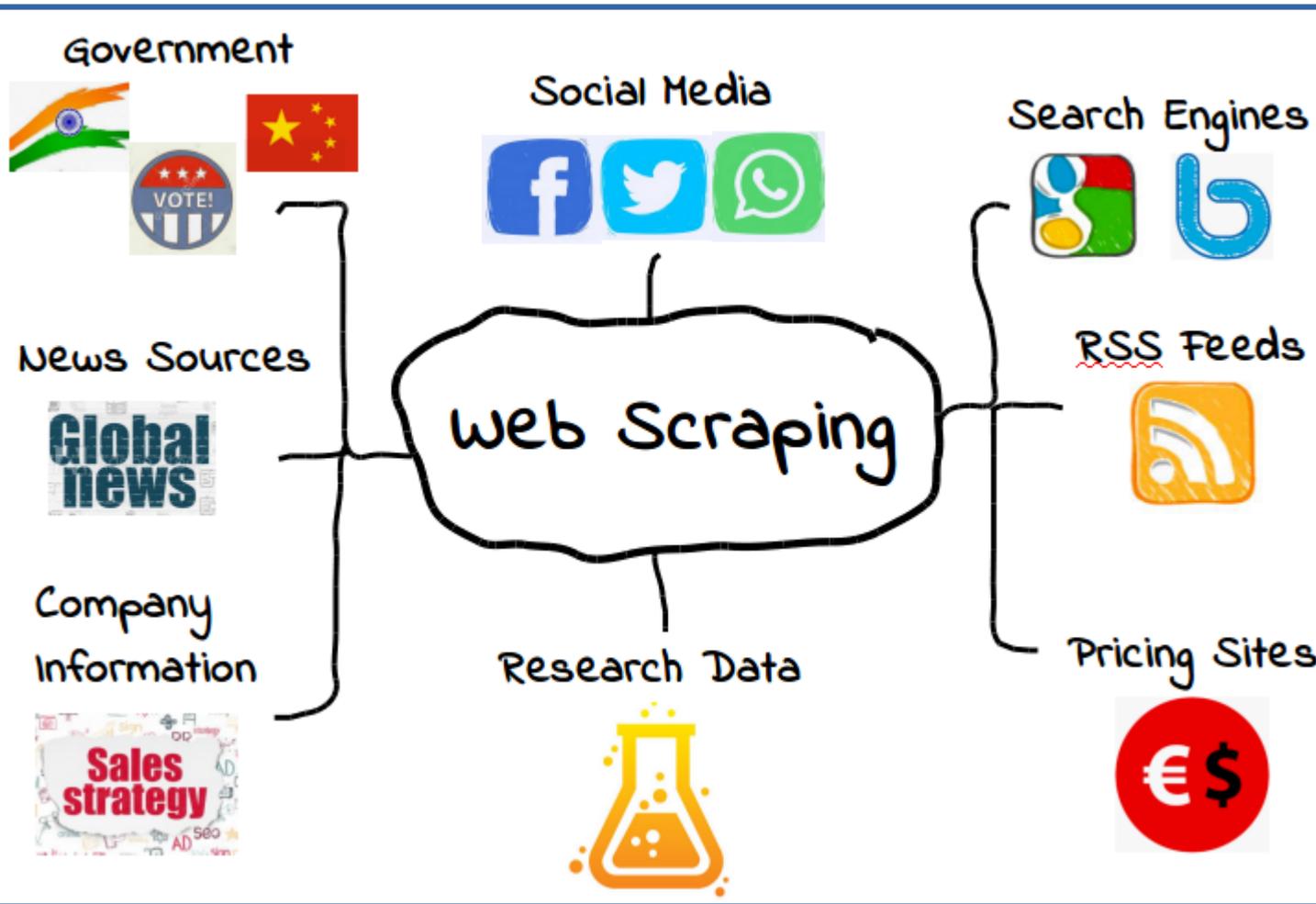
A greatly simplified view...

- **Privacy:** Ensuring an individual's right to remain private, hidden from public view, not linked to specific personal data or information
- **Ethics:** Honest and truthful reporting, acting responsibly, not knowingly acting unlawfully
- **Legal:** Obtaining consent from individuals or groups, being bound by legal requirements, understanding terms of license agreements and other international agreements.

Georgiada, de By, Kounadi (2019) – Location privacy in the wake of GDPR

- *“A privacy typology focused on the individual is both problematic and useful”.*
- **Useful:** the individual as ‘data subject’ or ‘consumer’ is subject to privacy theory and bearer of fundamental rights of dignity
- **Problematic:** An individual careless with their personal data exposes information about themselves as well as others.

Web Scraping – a legal and ethical gray area for CGI researchers



Distance and proximity: between ethics and geography

- Smith (2001) quoting Glover (1999)
- ***“Distance does not just reduce sympathy. It reduces the feeling of responsibility”***
- In particular: using data from ‘scraping’, data from crisis situations, humanitarian, etc.

The legal perspectives of the 'Prod-user' crowdsource approach will require great attention

- The terms of use of VGI data need to be accepted with **acknowledgement that sometimes (unknowingly) contributors contribute erroneous data**
- **Public sector organisations may have to use crowdsourced geospatial data** (new modes of ownership, licensing)
- **Cloud Computing – a disruptive technology**. Security, privacy, etc are issues.
- **Radical changes need to mitigate geo-liability**. The law tends to lag behind technology!

Crowdsourced geospatial data will continue to influence academic research efforts and attract NMCA attention

- **VGI/CGI research emerges from an intensive and multi-disciplinary research efforts** – multiple areas of focus. Concentration of effort needed to benefit both VGI/CGI communities and NMCAs
- **Quality assurance and assessment remains (probably) the biggest issue**
- **Legal and ethical frameworks need attention** – as crowdsourcing becomes more socio-technical



Thanks for listening....