AutoCarto 2020 Conference
Online, Nov. 18. 2020

How do people perceive the disclosure risk of maps?
- Examining the perceived disclosure risk of maps and its implications for geoprivacy protection*

Junghwan Kim, Ph.D. Candidate

Department of Geography & GIScience, University of Illinois at Urbana-Champaign

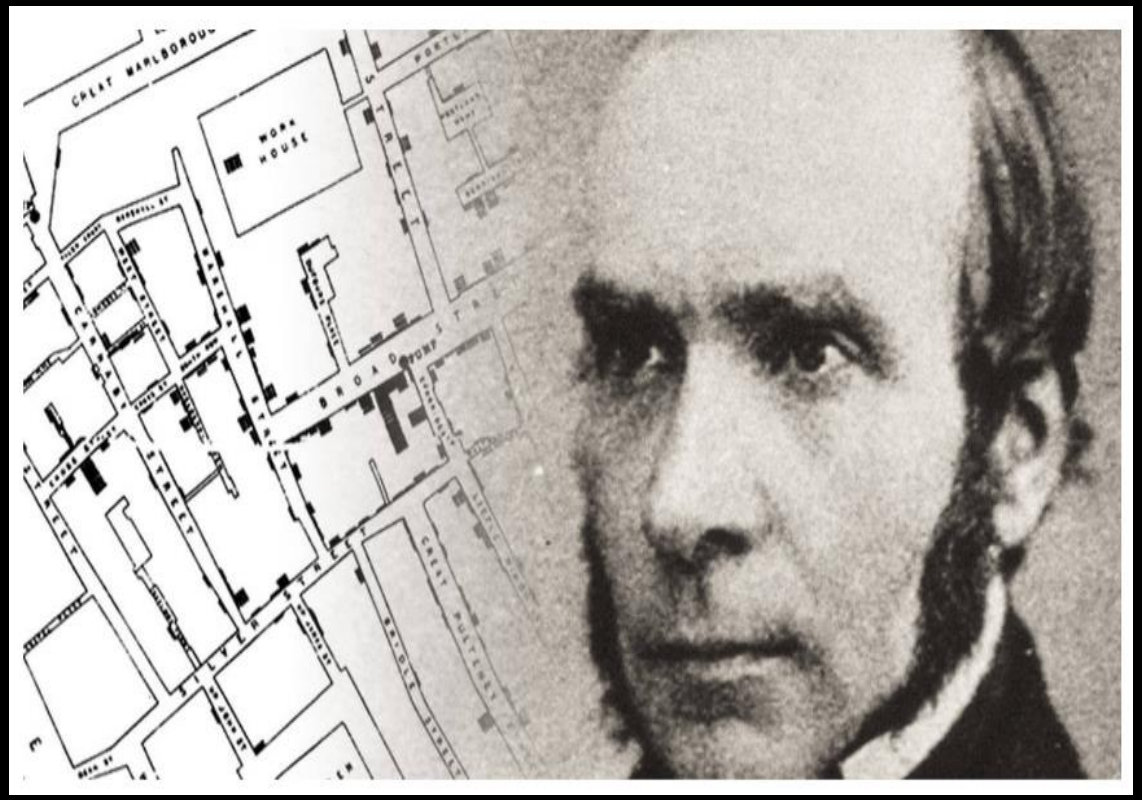
* **Co-authors**: Dr. Mei-Po Kwan, Dr. Margaret Levenstein, & Dr. Douglas Richardson

* **Corresponding Presenter**: Junghwan Kim (jk11@Illinois.edu)

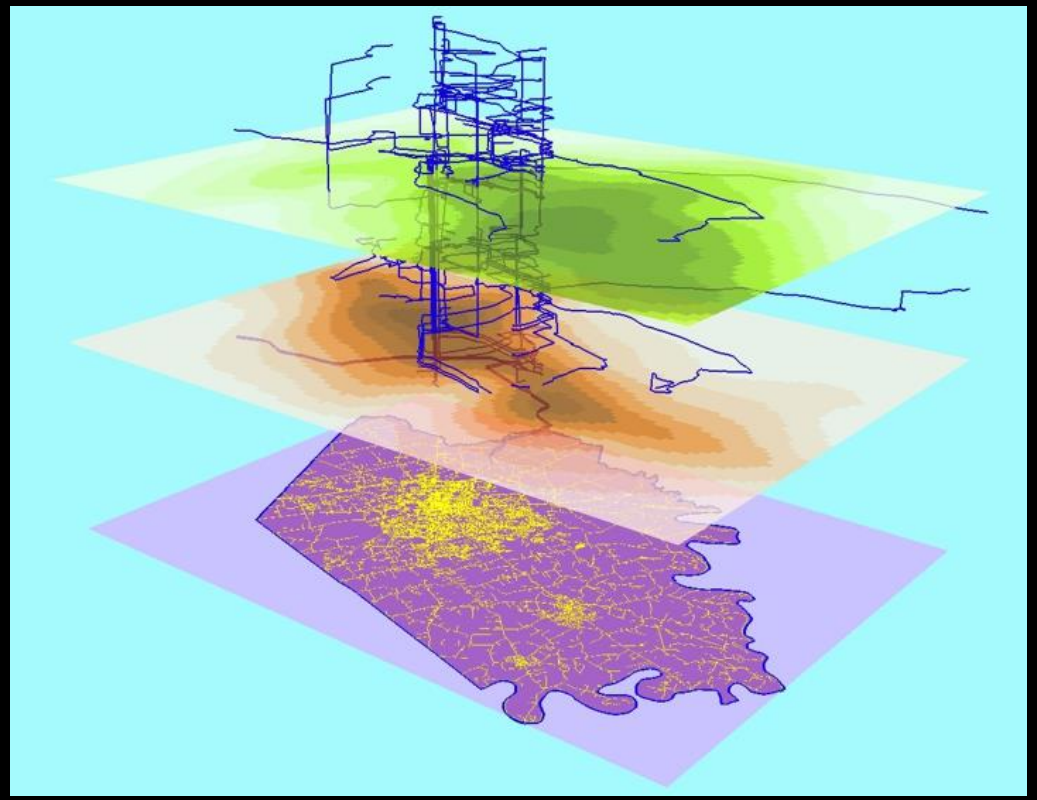
This research was supported by a grant from the **U.S. National Science Foundation (BCS-2025783).**

Mapping as a useful tool for visualizing people's detailed geospatial data

John Snow's Cholera Map (1854)



People's Space-Time Paths



https://hauntedwalk.com/news/john-snow-and-the-ghost-map/
http://meipokwan.org/Gallery/STPaths.htm

(Gutmann et al., 2008; Kwan, 2004, 2012; Richardson et al., 2013; Reich & Haran, 2018)

1-1. Background

1 However, it may violate people's geoprivacy by disclosing their private locations.

and Geographic Information Science xx (2020) 1-10

Figure 1. Study area and sample distribution in Champaign-Urbana, IL *(fictitious example)*

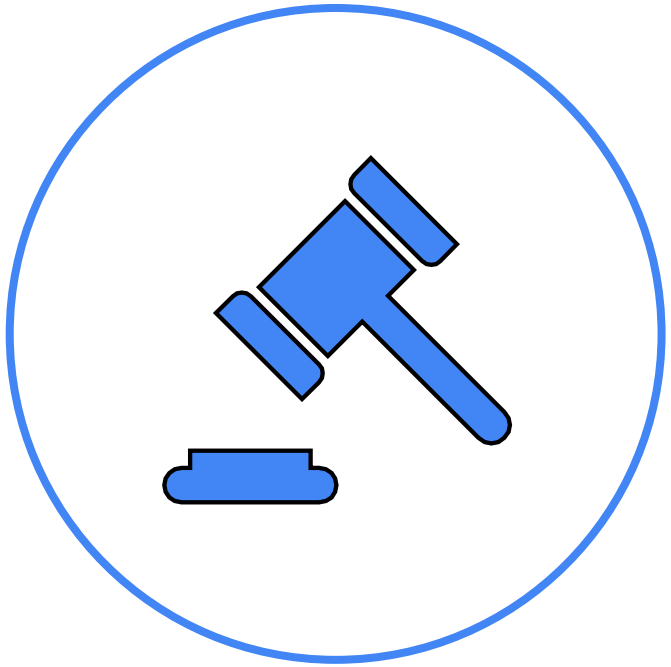
Spatial Reverse Engineering

1401 West Green St, Urbana, IL 61801
Mr. John Snow lives here.

(Brownstein et al., 2006; Curtis et al., 2006, 2011)
Image Source: Google Maps

1-1. Background

2 The disclosure of people's private locations through mapping is a serious problem.



A. Illegal
(e.g., HIPAA)

B. Unethical
(e.g., Unintended consequences)

C. Discouraging people to participate in surveys

(Gutmann et al., 2008; Kar et al., 2013; Kwan et al., 2004; Singer, 1978; Richardson et al., 2015)

1-1. Background

3 Previous studies have assessed the disclosure risk of maps **(i.e., probability of reidentification)** to help prevent privacy violations.

a Previous studies have shown that, through spatial reverse engineering, an individual's identity may be uncovered even though a point-based map does not explicitly indicate it.

- Brownstein *et al.* (2006): A street address of an individual whose home location is displayed as a point in a map may be accurately reverse engineered regardless of its resolution.
- Curtis *et al.* (2006): The accuracy of the spatial reverse engineering of a point-based map depends on the urban form (e.g., the housing pattern on a street).

b Previous studies have evaluated how geomasking methods can reduce the disclosure risk of maps and developed new measures to assess the performance of the methods.

- Kwan *et al.* (2004) observed a consistent tradeoff between analytical accuracy and confidentiality.
- Emam *et al.* (2009) concluded that there is no single threshold for aggregation because the optimal threshold is affected by the characteristics of the variables in question.

1-2. Research Questions

4 Previous studies have overlooked **how people perceive the disclosure risk of a map.**

- An individual's **perceived disclosure risk of a map** refers to how the individual subjectively perceives the disclosure risk (e.g., feels uncomfortable) of the map that displays his/her private locations (e.g., home location). Even for one map, the perceived disclosure risk can vary among individuals based on one's previous experiences and opinions.
- **Perceived disclosure risk** provides several important insights into assessing the disclosure risk of a map.
  - **1.** It can consider the **social and cultural influences on people's risk assessment**, which may not be adequately addressed by previous studies that did not consider people's perception.
  - **2.** It can be used to establish **geoprivacy protection guidelines** for mapping people's private locations.

(Benisch et al., 2011; Curtis et al., 2011; Groff et al., 2005; Keßler & McKenzie, 2018; Ketelaar & VanBalen, 2018; Kounadi et al., 2015; Slovic et al., 1980; VanWey et al., 2005)

1-2. Research Questions

4 Previous studies have overlooked **how people perceive the disclosure risk of a map.**

(Example 1) *Which map had a higher disclosure risk?*

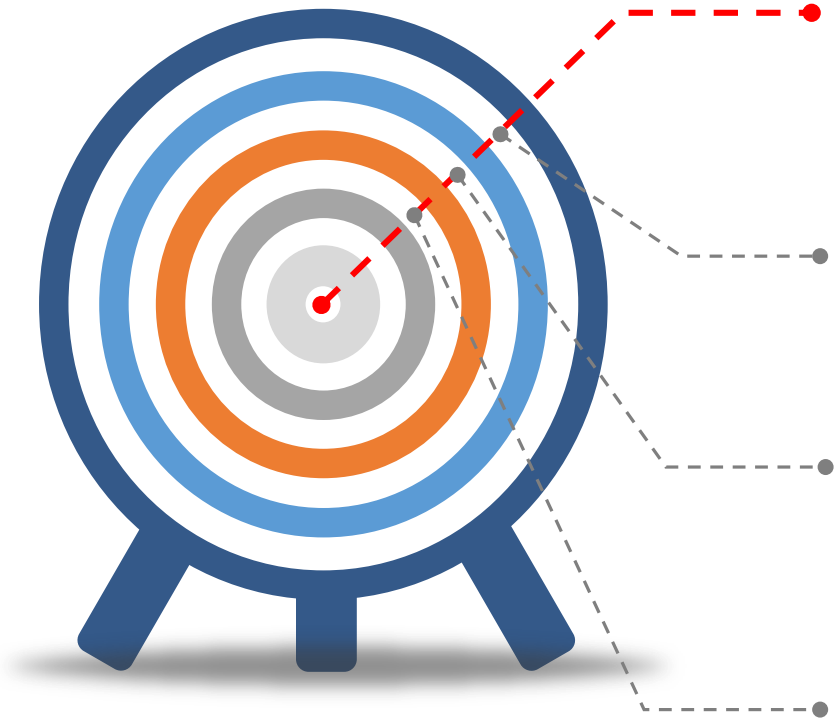
Imagine two maps where an individual can be identified at a probability of 1 in 100 (1%).
→ Map 1 displays the home locations of elementary school students.
→ Map 2 displays the home locations of lung cancer patients.

(Example 2) *Would 1:20 (5%) or 1:2,000 (0.05%) be acceptable risk level to people?*

There is no clear consensus among scientists or society at large about which risk level should be used as the safe threshold.

(Benisch et al., 2011; Curtis et al., 2011; Groff et al., 2005; Keßler & McKenzie, 2018; Ketelaar & VanBalen, 2018; Kounadi et al., 2015; Slovic et al., 1980; VanWey et al., 2005)

1-2. Research Questions



Goal. To examine how people subjectively perceive the disclosure risk of a map using original data collected in a survey.

RQ1. How do different components of a map affect an individual's perceived disclosure risk of a map?

RQ2. How do different geomasking methods affect an individual's perceived disclosure risk of a map?

RQ3. How does an individual perceive the disclosure risk of a map when the private locations of socially-vulnerable people are visualized?

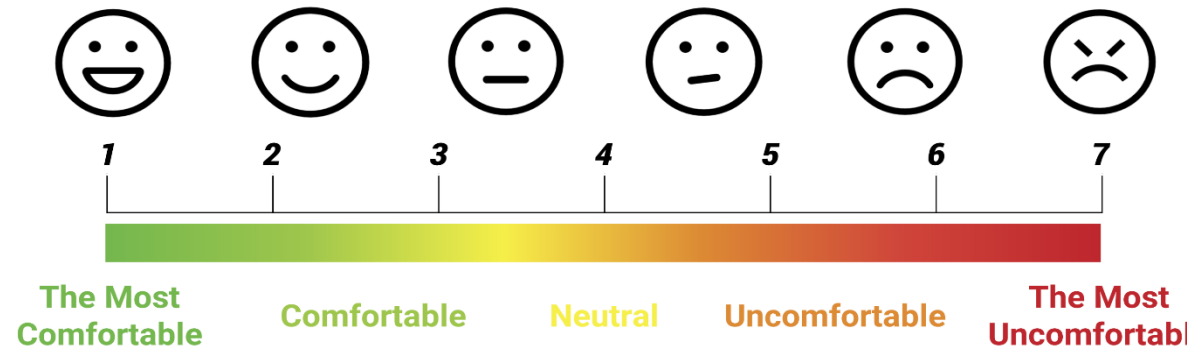
2. Data and Methods

1 Survey Questionnaire

- An online survey was implemented in Google Forms.
- 856 participants were recruited through distributing a solicitation e-mail message to 12,000 members of a university who were randomly selected in November 2019.
- The survey protocol and instrument were reviewed and approved by the IRB of UIUC.

2 The perceived disclosure risk of a map

- The perceived disclosure risk of a map represents the extent to which an individual subjectively feels comfortable with the map that displays his/her private locations.
- The feeling can be positive (i.e., comfortable) or negative (i.e., uncomfortable), and the intensity of the feeling can be strong or weak.



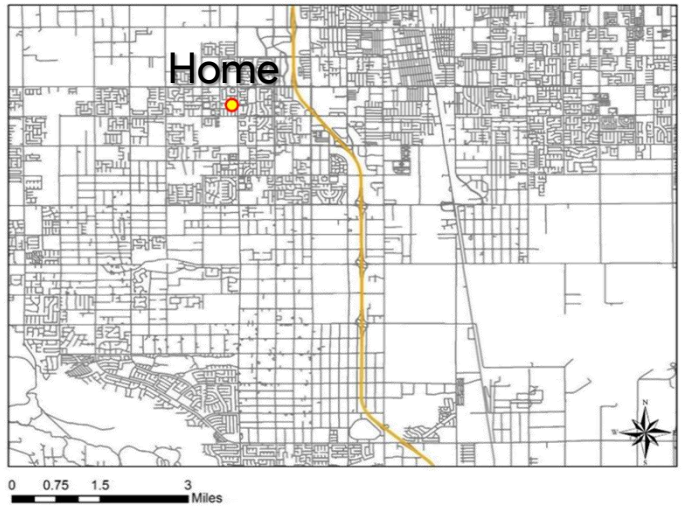
(Benisch et al., 2011; Groff et al., 2005; Keßler & McKenzie, 2018; Ketelaar & VanBalen, 2018; Kounadi et al., 2015)

3. Results

1 The effect of map attributes on the perceived disclosure risk

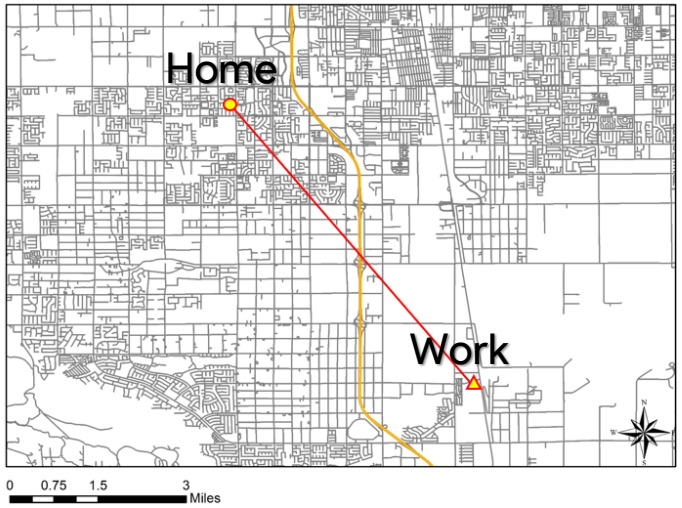
a The effects of the amount of private locational information on the perceived disclosure risk*

Map 1



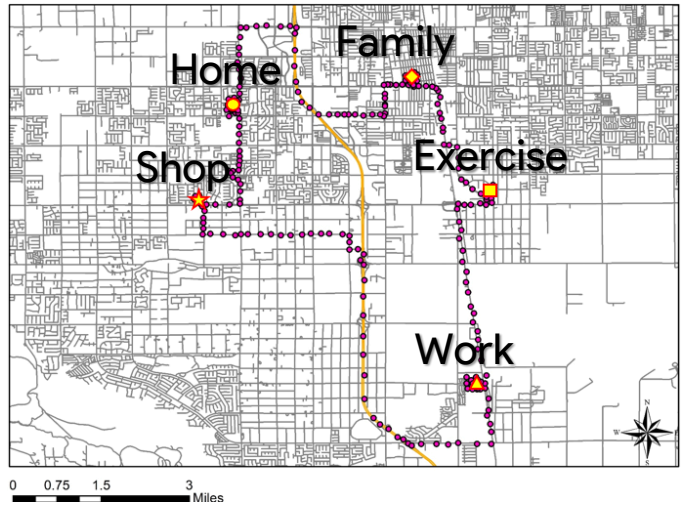
74%, Mean: 5.3

Map 2



81%, Mean: 5.7

Map 3



87%, Mean: 6.2

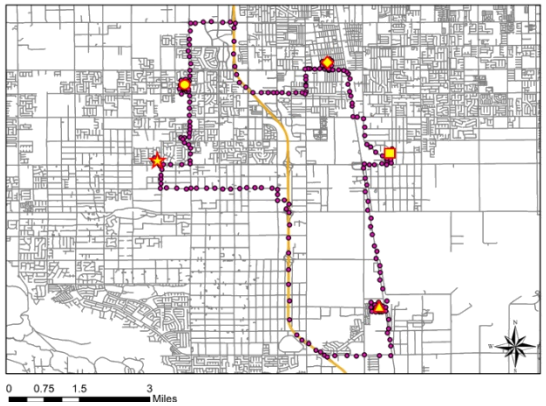
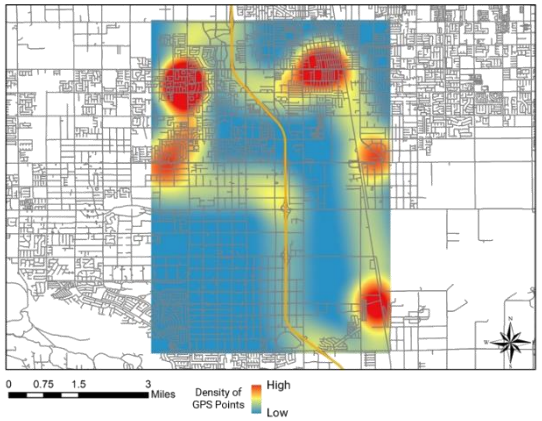
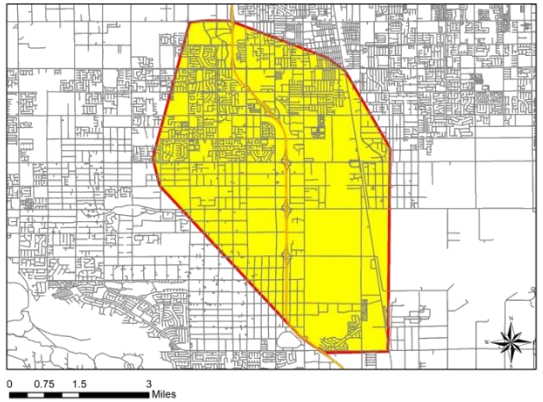
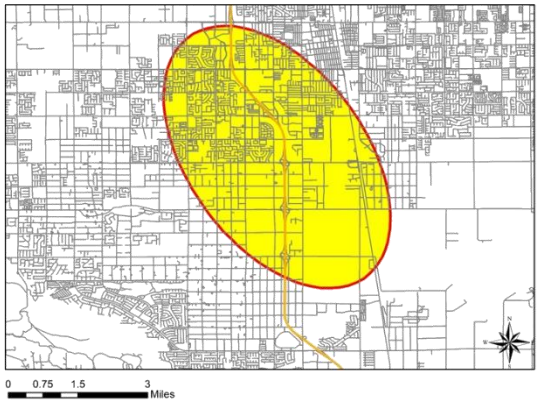
74% of people **feel uncomfortable** (i.e., responses: 5~7) with this map.

*Wilcoxon Signed-rank test with Bonferroni correction is used because of the ordinal scale of measurement (Privitera, 2011).

3. Results

2 The effect of map attributes on the perceived disclosure risk

b The effects of the type of map on the perceived disclosure risk

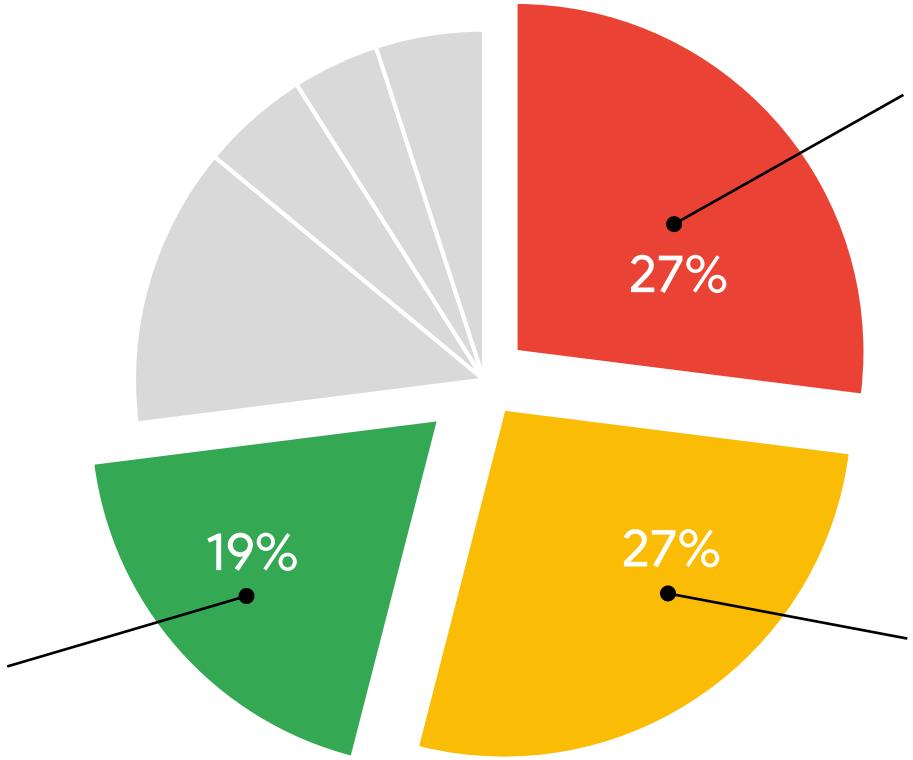
Map 3	Map 4	Map 5	Map 6
Daily GPS trajectories	Kernel Density	Minimum Convex Hull	Directional Distribution
			
87%, Mean: 6.2	56%, Mean: 4.6	43%, Mean: 4.0	39%, Mean: 3.8

87% of people **feel uncomfortable** (i.e., responses: 5~7) with this map.

3. Results

3 The effect of map attributes on the perceived disclosure risk

c Why do people feel uncomfortable with a map that discloses their home location? (n=330)
→ The results of analyzing responses to an optional open-ended question (thematic analysis).



"My identity can be easily re-identified through using web search engines or online map services."

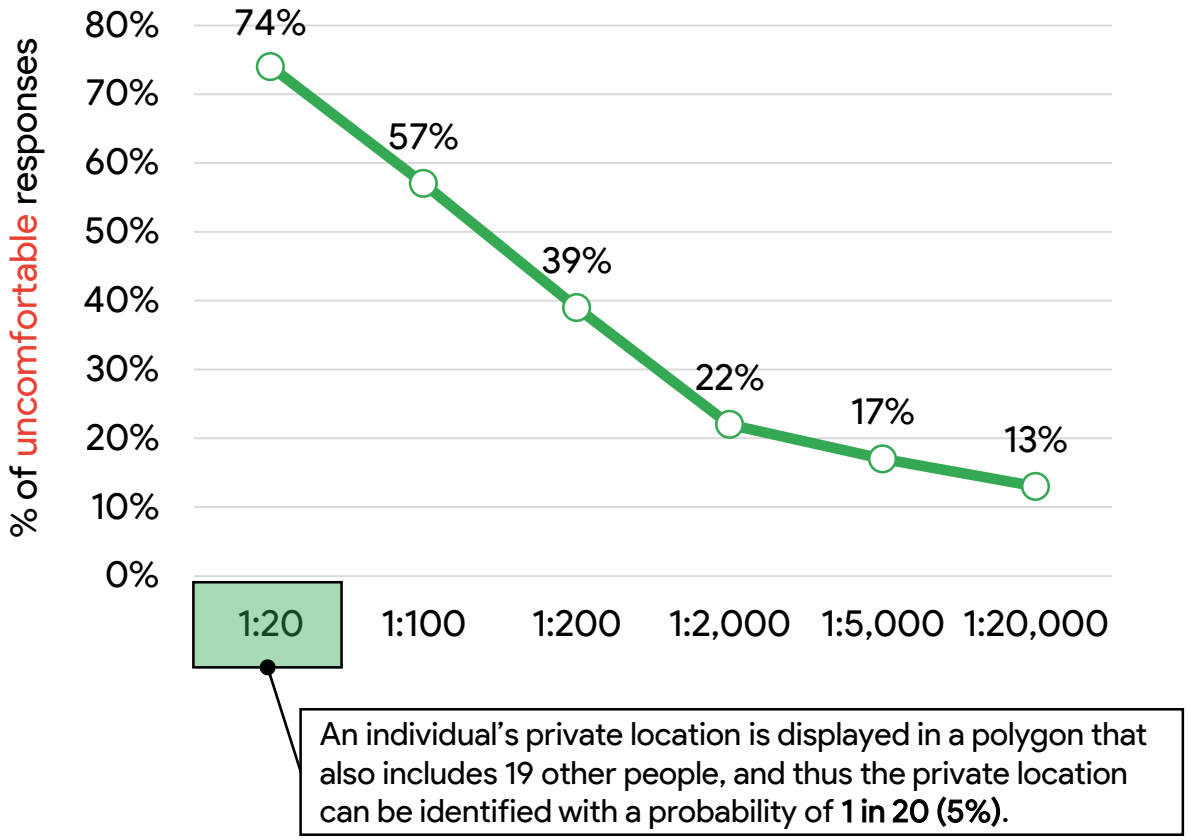
"It may lead to unintended consequences like crime and hackers."

"I dislike a map displaying my home location."

3. Results

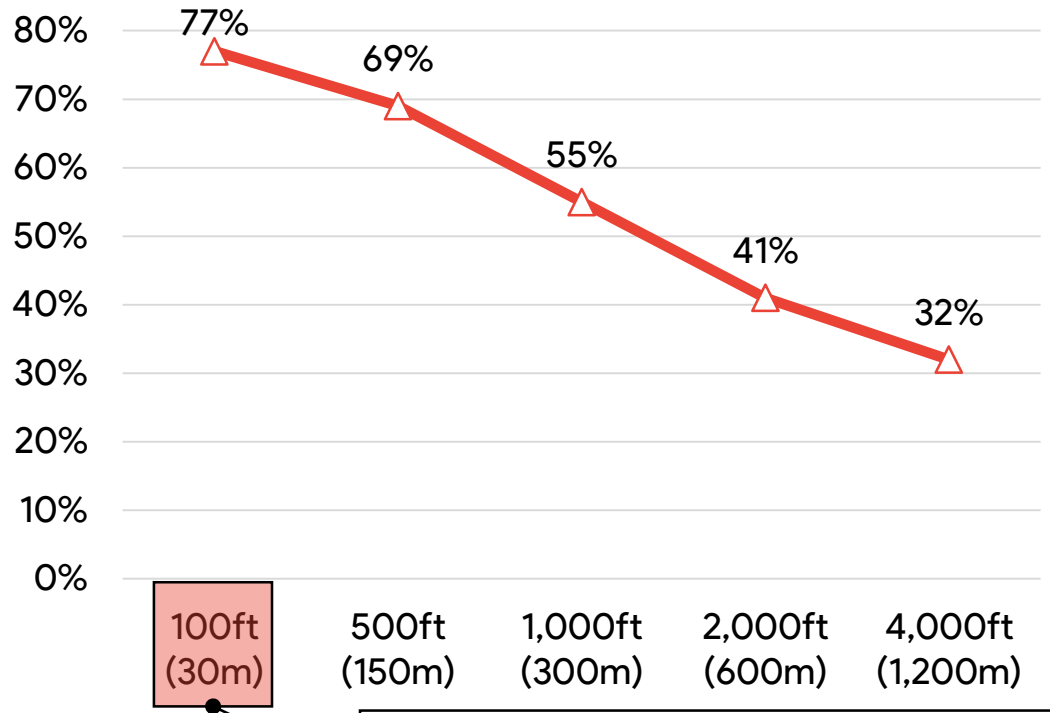
4 The effect of geomasking methods on the perceived disclosure risk

a Aggregation method (Point to Polygon)



An individual's private location is displayed in a polygon that also includes 19 other people, and thus the private location can be identified with a probability of **1 in 20 (5%)**.

b Relocation method (Point to Point)



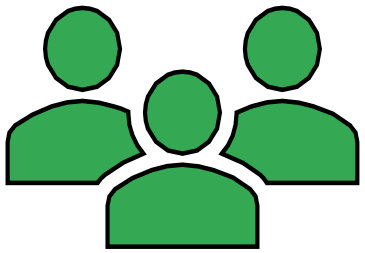
An individual's private location is displayed at a new point that is **100ft (30m)** away from the original point in a random direction.

3. Results

4 The effect of geomasking methods on the perceived disclosure risk

c Why do people feel uncomfortable with a certain aggregation level and relocation distance?
→ The results of analyzing responses to an optional open-ended question (thematic analysis).

> *"The relocation method itself may be inadequate* ***because new points can still be in the same neighborhood*** *where the original points are located. Besides, it may harm or endanger* ***the innocent people who are falsely indicated*** *by the new point."*



(Seidl et al., 2018; McLafferty, 2004)

3. Results

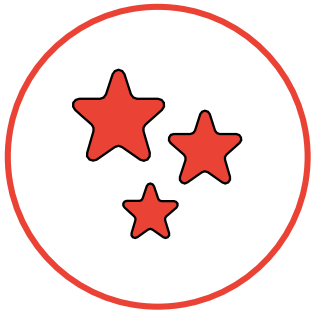
5 The perceived disclosure risk of a map when socially vulnerable people are displayed

- A scenario for which people are most uncomfortable is a map about elementary school students (S3, 79%), followed by:
  - People who are under alcohol or substance abuse treatment (S9, 78%)
  - HIV/AIDS patients (S2, 76%)
  - People who engage in sex with people of the same sex (S6, 74%)
  - Pregnant women (S4, 73%)
  - Elderly people (S5, 70%)
  - People in poverty (S7, 69%)
  - Cancer patients (S1, 68%)
  - High-income earners (S8, 61%)
  - Randomly selected people (S10, 48%)

*"I'm worry about **unintended negative outcomes**, like hate crime, discrimination, and stigmatization. […] Socially vulnerable groups **can be easily targeted** because they cannot defend themselves."*

4. Summary

1 Significances of this research



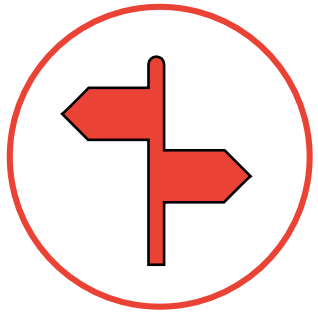
a This study significantly contributes to the literature **by systematically investigating perceived disclosure risk** and **providing solid scientific evidence** that confirms some of our common expectations.

For instance, one can easily expect that the perceived disclosure risk of a map would be reduced when the proper geomasking method is applied.

However, systematic investigations of how and to what extent people perceive disclosure risk have not been conducted to date.

4. Summary

2 Implications of the results



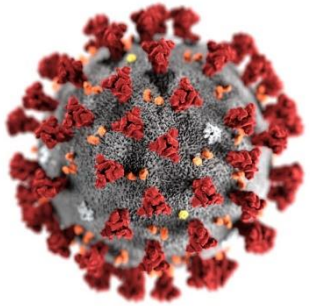
a The significant results can be used for proposing some **tentative guidelines for geoprivacy protection** that consider people's perceived disclosure risk.

(e.g.) For geomasking, the aggregation method may be preferred when compared to the relocation method for avoiding the false identification of individuals.

(e.g.) To visualize the point patterns of the location of people (e.g., survey participants), a kernel density map may be recommended rather than a point-based map that directly displays the points.

4. Summary

2 Implications of the results



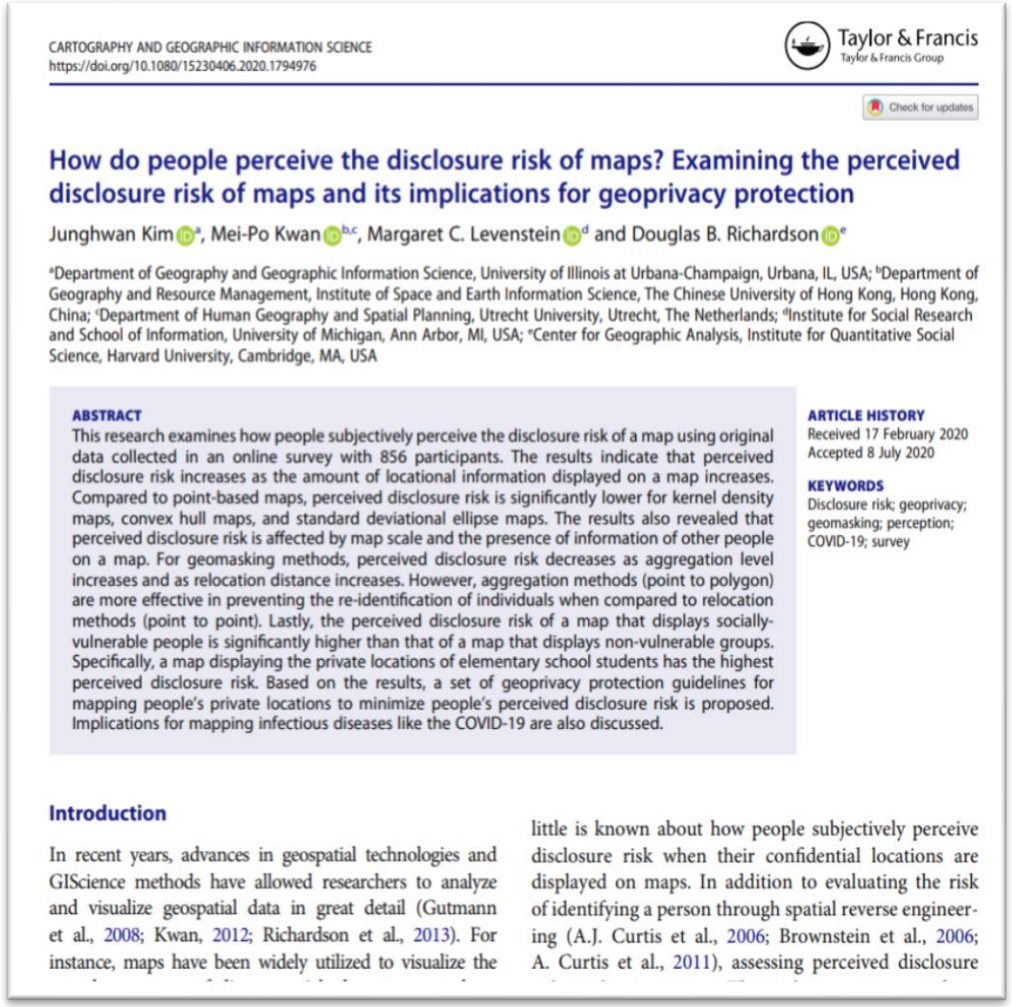
b When preparing the **COVID-19 pandemic** in 2020, conducting surveys that are similar to ours may provide important insights into **how public health agencies can wisely balance between disease control and geoprivacy protection.**

→ To effectively control a pandemic, it may be critical to balance between promoting public health (by releasing information) and protecting geoprivacy (by geomasking information) rather than exclusively choosing one over the other.

(e.g.) As suggested by the results of our survey, public health agencies may consider applying geomasking methods to maps in order to protect the geoprivacy of patients while still providing useful information to the general public.

Image Source: https://www.statnews.com/2020/02/11/disease-caused-by-the-novel-coronavirus-has-name-covid-19/

Do you want to know more about the results?



CARTOGRAPHY AND GEOGRAPHIC INFORMATION SCIENCE
https://doi.org/10.1080/15230406.2020.1794976

Taylor & Francis
Taylor & Francis Group

Check for updates

How do people perceive the disclosure risk of maps? Examining the perceived disclosure risk of maps and its implications for geoprivacy protection

Junghwan Kim[a], Mei-Po Kwan[b,c], Margaret C. Levenstein[d] and Douglas B. Richardson[e]

[a]Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA; [b]Department of Geography and Resource Management, Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China; [c]Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, The Netherlands; [d]Institute for Social Research and School of Information, University of Michigan, Ann Arbor, MI, USA; [e]Center for Geographic Analysis, Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA

ABSTRACT

This research examines how people subjectively perceive the disclosure risk of a map using original data collected in an online survey with 856 participants. The results indicate that perceived disclosure risk increases as the amount of locational information displayed on a map increases. Compared to point-based maps, perceived disclosure risk is significantly lower for kernel density maps, convex hull maps, and standard deviational ellipse maps. The results also revealed that perceived disclosure risk is affected by map scale and the presence of information of other people on a map. For geomasking methods, perceived disclosure risk decreases as aggregation level increases and as relocation distance increases. However, aggregation methods (point to polygon) are more effective in preventing the re-identification of individuals when compared to relocation methods (point to point). Lastly, the perceived disclosure risk of a map that displays socially-vulnerable people is significantly higher than that of a map that displays non-vulnerable groups. Specifically, a map displaying the private locations of elementary school students has the highest perceived disclosure risk. Based on the results, a set of geoprivacy protection guidelines for mapping people's private locations to minimize people's perceived disclosure risk is proposed. Implications for mapping infectious diseases like the COVID-19 are also discussed.

ARTICLE HISTORY
Received 17 February 2020
Accepted 8 July 2020

KEYWORDS
Disclosure risk; geoprivacy; geomasking; perception; COVID-19; survey

Introduction

In recent years, advances in geospatial technologies and GIScience methods have allowed researchers to analyze and visualize geospatial data in great detail (Gutmann et al., 2008; Kwan, 2012; Richardson et al., 2013). For instance, maps have been widely utilized to visualize the little is known about how people subjectively perceive disclosure risk when their confidential locations are displayed on maps. In addition to evaluating the risk of identifying a person through spatial reverse engineering (A.J. Curtis et al., 2006; Brownstein et al., 2006; A. Curtis et al., 2011), assessing perceived disclosure

Kim, Junghwan, Mei-Po Kwan, Margaret C. Levenstein, and Douglas B. Richardson. "How do people perceive the disclosure risk of maps? Examining the perceived disclosure risk of maps and its implications for geoprivacy protection." *Cartography and Geographic Information Science* (2020): 1-19.

https://doi.org/10.1080/15230406.2020.1794976

AutoCarto 2020 Conference
Online, Nov. 18. 2020

Thanks for your attention!

Hope you stay healthy and safe during this pandemic.

Questions? Please send an email to **Junghwan Kim** (jk11@illinois.edu)

Presenter: Junghwan Kim, Ph.D. Candidate

Department of Geography & GIScience, University of Illinois at Urbana-Champaign