

# CASA dissertation plan submission form

Please use your UCL information (e.g. email)

Email address \*

hannah.ker.19@ucl.ac.uk

First Name \*

Hannah

Last Name \*

Ker

Student Number \*

18110511

Select your programme \*

MSc Spatial Data Science and Visualisation



Potential supervisor. Note, academics have different capacity for supervision. \*

Ana Basiri ▼

Have you contacted this supervisor \*

☒ Yes

☐ No

Do you have a 2nd choice supervisor \*

Sarah Wise ▼

Are you planning to collaborate with a partner or apply for an advertised project (listed below) \*

☐ Yes

☒ No

☐ I've found my own partner

What project are you applying for \*

I'm not, i've got my own project ▼

If you have your own partner, who are they

.....

If you have picked a partner or listed project what makes you suitable for their project (100-200 words)

.....

Have you read the mark scheme on Moodle \*



Yes

Have you read the dissertation handbook on Moodle before completing the rest of this form \*



Yes

What is the proposed title of your project (insert industry title if selected industry project) \*

Crowdsourced mapping for all? An empirical investigation into gender bias in OpenStreetMap.

What is the proposed research question - see handbook \*

How has the gendered bias in OSM contributors impacted the quality of data on OSM?

.....

What are your proposed objectives - see handbook \*

- 1) To review past and current gendered crowdsourced mapping projects, including but not limited to those conducted within the OpenStreetMap environment. The aim of this objective is to understand existing work that has been done to engage girls and women in mapping, and to illustrate the need for female-focused mapping efforts.
  - 2) To review past academic literature on bias in OpenStreetMap to identify relevant methodological approaches, past findings, and suggestions for future work.
  - 3) To develop and implement a framework for analysing points of similarity and difference in the mapping practices between male and female contributors to OSM.
  - 4) To conduct an extrinsic evaluation of the volume and comprehensiveness with which traditionally feminine points of interest (amenities) are mapped on OSM.
  - 5) To critically reflect on the results from objectives 3) and 4) with respect to my research question.
-

Provide 400–500 words of background for the research \*

It has been well-established that OpenStreetMap (OSM) contributors are predominantly male (Budhathoki and Haythornthwaite, 2013; Das et al., 2019; Gardner et al., 2019). However, there has been limited empirical research into the potential bias that this has introduced into the geospatial data in OSM. While some elements of mapping may be objective, it is acknowledged that the cartographer (or in this case, OSM contributor) plays an active role in the spatial knowledge that is represented on a map (Harley, 1989; Monmonier, 1991). As such, this research aims to continue efforts to investigate the extent to which the gender bias in OSM contributors has resulted in biased data on OSM.

Within discussions of data quality of crowdsourced geospatial data (eg. OSM), investigation into bias has also been highlighted as a key future research direction (Basiri et al., 2019). Gender-based content bias has also been explored in the case of Wikipedia (Graells-Garrido et al., 2015; Wagner et al., 2015), another crowdsourced knowledge repository with established contributor biases (Collier and Bear, 2012).

Survey data collected by Gardner et al. (2019) and Gardner and Mooney (2018) has allowed for an initial, descriptive look into the different contributing behaviours of men and women on OSM, identifying the increased likelihood of women to focus more on adding new content to the map, rather than modifying or correcting existing content. This finding points to room for more work into potentially differing geographies of editing behaviours between men and women. Additionally, Stephens (2013) has demonstrated how the granularity of amenity tags in OSM may be biased towards more traditionally masculine spaces, such as sexual entertainment venues, while feminized spaces, such as childcare facilities, may not have been given the same attention.

This initial research has laid the foundation for more work into the gendered dimensions of content production on OSM. Put simply, we know that a significant proportion of the contributors to OSM are men, but we have limited evidence of what this means for the quality of the data on the map. For example, are there certain types of amenities that are not present on the map? Are there spatial attributes, such as levels of street lighting, that are not well mapped? Given the increasing popularity of OSM as a provider of geospatial information (ref or example here), we need to continue to work to investigate any potential limitations in this vast dataset.

In conducting this project, my broad aim is to provide a stronger understanding of the data quality of OSM, hopefully leading to a better sense of the strengths and limitations of this dataset, and thus a greater understanding of its appropriate application areas. I also hope to provide an empirical justification for efforts that seek to promote greater gender inclusion in crowdsourced mapping.

.....

Provide a 300-500 word methodological outline \*

I follow Das et al. (2019) in conducting a two-part methodology, briefly outlined as follows:

1) Contributor-focused analysis

This analysis will address my research objective #3.

This analysis is enabled by access to the survey dataset from Gardner et al. (2019). I will collect data (using the tools identified in the section below) on the mapping behaviours of the 293 contributors in this dataset and investigate gendered similarities and differences in how, where, and what contributors map. Specifically, I hope to collect data that captures contributors' editing frequency, commenting patterns, engagement with other contributors, geographic locations of edits, and content of edits.

This work will build off of the initial quantitative exploration of Gardner et al. (2019) into differing mapping practices of men and women. In doing so, I hope to identify any systematic biases in the mapping practices of males and females, which could potentially point to content biases on OSM.

2) Contribution-focused analysis

This analysis will address my research objective #4.

I aim to conduct an extrinsic evaluation of the completeness of gendered points of interest (POIs) in OpenStreetMap. It is reported that female-oriented facilities, such as women's health clinics and family planning offices, are undermapped on OSM ("Tagging in Support of Women and Girls - OpenStreetMap Wiki," n.d.). I aim to compare the volume, distribution, and level of detail in attributes of various women's health facilities on OpenStreetMap with that of external data sources (see list of potential sources in the following section). Extrinsic data quality assessment of OSM follows past work, such as that from Haklay (2010). Depending on the data available, I aim to conduct this comparative analysis across a number of different geographic contexts and scales.

---

List the sources of data you are considering using \*

Following email correspondence with Dr. Zoe Gardner (via Dr. Ana Basiri), I have been given permission to use the existing survey dataset that includes demographic information for 293 OSM contributors. For each contributor, this dataset includes the following attributes: OSM username, gender, age bracket, education level, country of residence, and nationality.

OSM itself will also be a central data source for my analysis. The following tools for collecting and analysing OSM data may be relevant for my work:

- OSMCHa: <https://osmcha.mapbox.com/>
- Tag History: <https://taghistory.raifer.tech/>
- How did you contribute to OpenStreetMap? <https://hdyc.neis-one.org/>

The contribution-focused analysis will also require external POI data, which could perhaps be accessed from sources such as the following:

- <https://www.pitneybowes.com/us/data/poi-database.html>
- <https://www.ordnancesurvey.co.uk/business-government/products/points-of-interest>
- <https://developers.google.com/places/web-service/intro>
- <https://enterprise.foursquare.com/products/places>

Are you planning on conducting any field work \*

- ☐ Yes
- ☒ No
- ☐ Maybe

If yes or maybe, outline which UCL forms you will need to submit. Consult:

<https://www.ucl.ac.uk/safety-services/a-z/off-site-working>

Do you think you will need UCL ethics committee approval at this stage. Follow the link within the dissertation handbook. More information on ethics will be provided in a later lecture. \*

- ☐ Yes
- ☐ No
- ☒ Not sure yet, might do
- ☐ Not sure yet, but don't think so

What ethical considerations will consider within your research, even if you don't need formal approval (100-200 words) \*

Throughout my research, I will abide by standards of good academic practice, including proper attribution of the ideas and code of others.

I will also review with my supervisor and determine whether ethics approval is required for my use of the survey dataset from Dr. Zoe Gardner. This dataset contains personal demographic information of individuals, but does not contain any directly identifying personal information. I will be sure to properly cite and attribute Dr. Gardner's past work in collecting this data.

The other data that I plan to use from OSM and external sources is not sensitive or personally-identifying.

Are you aware of the penalties for poor academic practice or academic misconduct. Consult the UCL academic manual, Chapter 6, section 9. \*

- ☒ Yes

Have you joined the slack dissertation channel \*

- ☒ Yes



Do you have access to the CASA0004/0010/0012 Moodle page \*

☒ Yes

☐ No

Are you aware of the submission deadlines of 5pm 24th August for the digital version on Moodle and 5pm 1st September for the hard copies in the CASA office \*

☒ Yes

Bullet point (or list) an action plan for March for this project \*

- Continue to review relevant literature, with particular attention towards methodological approaches
  - Conduct exploratory analysis/review of potential data sources
  - Review ethics requirements
  - Develop more detailed outline of methodology, and iteratively adjust research questions and objectively accordingly
- .....

Do you have anything else to add

#### References:

- Basiri, A., Haklay, M., Foody, G., Mooney, P., 2019. Crowdsourced geospatial data quality: challenges and future directions. *Int. J. Geogr. Inf. Sci.* 33, 1588–1593.  
<https://doi.org/10.1080/13658816.2019.1593422>
- Budhathoki, N.R., Haythornthwaite, C., 2013. Motivation for Open Collaboration: Crowd and Community Models and the Case of OpenStreetMap. *Am. Behav. Sci.* 57, 548–575.  
<https://doi.org/10.1177/0002764212469364>
- Collier, B., Bear, J., 2012. Conflict, criticism, or confidence: an empirical examination of the gender gap in wikipedia contributions, in: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW '12*. Association for Computing Machinery, Seattle, Washington, USA, pp. 383–392. <https://doi.org/10.1145/2145204.2145265>
- Das, M., Hecht, B., Gergle, D., 2019. The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias, in: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19*. Association for Computing Machinery, Glasgow, Scotland Uk, pp. 1–14.  
<https://doi.org/10.1145/3290605.3300793>
- Gardner, Z., Mooney, P., 2018. Investigating gender differences in OpenStreetMap activities in Malawi: a small case-study 7.
- Gardner, Z., Mooney, P., De Sabbata, S., Dowthwaite, L., 2019. Quantifying gendered participation in OpenStreetMap: responding to theories of female (under) representation in crowdsourced mapping. *GeoJournal*. <https://doi.org/10.1007/s10708-019-10035-z>
- Graells-Garrido, E., Lalmas, M., Menczer, F., 2015. First Women, Second Sex: Gender Bias in Wikipedia, in: *Proceedings of the 26th ACM Conference on Hypertext & Social Media, HT '15*. Association for Computing Machinery, Guzelyurt, Northern Cyprus, pp. 165–174.  
<https://doi.org/10.1145/2700171.2791036>
- Haklay, M., 2010. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ. Plan. B Plan. Des.* 37, 682–703.  
<https://doi.org/10.1068/b35097>
- Harley, J.B., 1989. Deconstructing the Map. *Cartogr. Int. J. Geogr. Inf. Geovisualization* 26, 1–20.  
<https://doi.org/10.3138/E635-7827-1757-9T53>
- Monmonier, M., 1991. *How to Lie with Maps*. University of Chicago Press.
- Stephens, M., 2013. Gender and the GeoWeb: divisions in the production of user-generated cartographic information. *GeoJournal* 78, 981–996. <https://doi.org/10.1007/s10708-013-9492-z>
- Tagging in Support of Women and Girls - OpenStreetMap Wiki [WWW Document], n.d. URL [https://wiki.openstreetmap.org/wiki/Tagging\\_in\\_Support\\_of\\_Women\\_and\\_Girls](https://wiki.openstreetmap.org/wiki/Tagging_in_Support_of_Women_and_Girls) (accessed 2.18.20).
- Wagner, C., Garcia, D., Jadidi, M., Strohmaier, M., 2015. It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia, in: *Ninth International AAAI Conference on Web and Social Media*. Presented at the Ninth International AAAI Conference on Web and Social Media.

If you are applying for an industry project you might not be successful, what is your plan B title. If you have proposed your own project or secured your own partner (that is not from the CASA list) you do not need to complete any plan B questions

.....

What is your plan B research question

.....

What would be your plan B objectives

.....

What data would you use for the plan B project

.....

Give a brief description of the project (100 words)

.....

Do you think you will need UCL ethics committee approval at this stage for the plan B idea. Follow the link within the dissertation handbook. More information on ethics will be provided in a later lecture.

- ☐ Yes
- ☐ No
- ☐ Not sure yet, might do
- ☐ Not sure yet, probably not

Are you planning on conducting any field work for the plan B project

- ☐ Yes
- ☐ No
- ☐ Maybe

If yes or maybe, outline which UCL forms you will need to submit. Consult:

<https://www.ucl.ac.uk/safety-services/a-z/off-site-working>

What ethical considerations will consider within your plan B research, even if you don't need formal approval (50 words)

ROUND 2: What project are you applying for

Choose

ROUND 2: If you have picked a partner or listed project what makes you suitable for their project (100-200 words)

ROUND 2: What is the proposed title of your project (insert industry title if selected industry project)

ROUND 2: What is the proposed research question - see handbook

.....

ROUND 2: What are your proposed objectives - see handbook

.....

ROUND 2: Provide 400-500 words of background for the research

.....

ROUND 2: Provide a 300-500 word methodological outline

.....

ROUND 2: List the sources of data you are considering using

.....

ROUND 2: Are you planning on conducting any field work

☐ Yes

☐ No

ROUND 2: If yes or maybe, outline which UCL forms you will need to submit. Consult:  
<https://www.ucl.ac.uk/safety-services/a-z/off-site-working>

.....

ROUND 2: Do you think you will need UCL ethics committee approval at this stage. Follow the link within the dissertation handbook. More information on ethics will be provided in a later lecture. \*

- ☐ Yes
- ☐ No
- ☐ Not sure yet, might do
- ☐ Not sure yet, but don't think so

ROUND 2: What ethical considerations will consider within your research, even if you don't need formal approval (100-200 words)

.....

This content is neither created nor endorsed by Google.

Google Forms