Paper2GIS: A self-digitising, paper-based PPGIS

J.J. Huck*[1], I. Dunning[1], P. Lee[1], T. Lowe[1], E. Quek[1], S. Weerasinghe[1], D. Wintie[1]

[1] Department of Geography, The University of Manchester, M13 9PL, UK

* jonathan.huck@manchester.ac.uk

Abstract

This paper examines two established problems in public participatory GIS: affording participants the flexibility to represent geographical entities in a way that reflects how they think about the world; and the 'digital divide' in order to promote representative participation in locations that do not have good access to computing or the Internet. These two issues will be explored through the presentation of a novel prototype public participatory GIS software package, which allows the rapid automated digitisation of data collected on paper maps. This software is demonstrated in the context of a pilot study investigating perceived 'landscape value' in the village of Leh in the Ladakh region of Himalayan India, such that highly valued areas might be protected from the rapid, tourism-driven development currently taking place in that part of the Himalaya.

Keywords: PPGIS, Participatory Mapping, Digital Divide, Vagueness.

1. Introduction

This work aims to address two significant challenges to the field of Public Participatory GIS (PPGIS): the representation of geographical entities collected from the public and the 'digital divide' (Huck et al. 2014). PPGIS is the term traditionally applied to a system for the collection of geospatial data from the public. The development of participatory methods has historically been driven predominantly by decision makers' desire to increase citizen satisfaction, create realistic expectations, and produce better plans or designs by accessing local knowledge and skills (Maquil et al., 2015). However, PPGIS techniques are increasingly being applied to a wide variety of applications by researchers, as the collective knowledge of the public is increasingly seen as a valuable resource (Goodchild, 2007; Bugs et al., 2010).

1.1. Vagueness

The representation of geographical entities in PPGIS is a challenge because many of the entities perceived by the public are vague, meaning that precise and indisputable boundaries cannot necessarily be drawn for them in a justifiable way. This is particularly the case for those entities that do not have any direct physical manifestation, but instead reside within the imagination, experience and understanding of the participant in question; such as the 'rough' area of a town, or a 'beautiful' area of countryside. Such entities surely have an extent, but their boundaries either have never been fixed, or are impossible to define (Clementini and di Felice, 1996). This vagueness is pervasive and endemic in human thought and language (Williamson, 1994; Fisher, 2000; Varzi, 2001; Montello et al., 2003) and is a necessary

construct of the human experience of geography (Fisher, 2000). Accordingly, vagueness is a vital and omnipresent component of geographical information (Goodchild, 2011; Mackaness and Chaudhry, 2013), particularly that which is collected directly from the public. The creation of new interfaces that are better able to reflect the way in which the public perceive the world can therefore maximise the level of understanding that researchers are able to gain from the opinions and understanding of the public (Huck et al. 2014).

1.2. Digital Divide

Another challenge in PPGIS is that of the digital divide, a problem that is particularly important given the on-going proliferation of web-mapping platforms such as the Google Maps API (Google, 2017) and Leaflet (Leaflet, 2017). Such systems provide researchers with web-based slippy-map interfaces that are both relatively easy to set up for spatial data collection at a global scale and familiar to many participants due to their pervasive presence throughout the web and mobile applications (Huck et al. 2014). However, this means that many PPGIS applications necessarily make the assumption that users are familiar with these interfaces and have access to computers or mobile devices, high-speed internet connections and suitable language skills (normally English) (Goodchild, 2007, Riddleston and Singleton, 2014). These assumptions introduce inequitable access to many PPGIS studies and may result in a 'false authority' being placed upon them, with results in some cases appearing more representative than they actually are (Obermeyer, 1998; Green, 2010). The effect of this 'digital divide' between those with and without the access to PPGIS is important in any study, but particularly those taking place in remote areas these technologies are not widely accessible.

This paper presents a preliminary study undertaken in rural India, using a prototype PPGIS platform that seeks to address both the issues of representation and the digital divide. This platform will be presented in the context of an investigation of perceived 'landscape value' in Leh, a village in the Ladakh region of Himalayan India.

2. Method

The purpose of this research was to examine the potential of the prototype Paper2GIS software by undertaking a preliminary study in order determine what is considered as 'valuable' about the village of Leh to those people that exist within it (either as residents, seasonal workers or tourists). Leh was selected as it is currently undergoing rapid development due to the influx of tourists wishing to explore the Himalaya, and so information collected during this study could be used in order to inform the future development of the village.

The authors collected data from 94 people in Leh during the period 29th August - 2nd September 2016. Interviews (in some cases via a translator) were semi-structured, beginning with some simple demographic questions and then asking a single question: "*Where do you consider to be valuable in Leh?*". During the interview, participants were encouraged to draw the areas that they were talking about onto the map with a marker pen, in order that we could capture their spatial perceptions of places as they described them. The markup that participants added to the map could then be automatically digitised into Shapefiles using the prototype Paper2GIS software, and then analysed alongside transcribed recordings of the

interviews in order to gain deep geographical insight into their perceptions of value in their village.

2.1. Paper2GIS

The first step in the creation of a Paper2GIS survey is using the software to generate a base map. The map contains a QR code containing encoded projection and georeferencing information, which will later be used by the software to digitise participant dater. Such a map is illustrated in Figure 1.
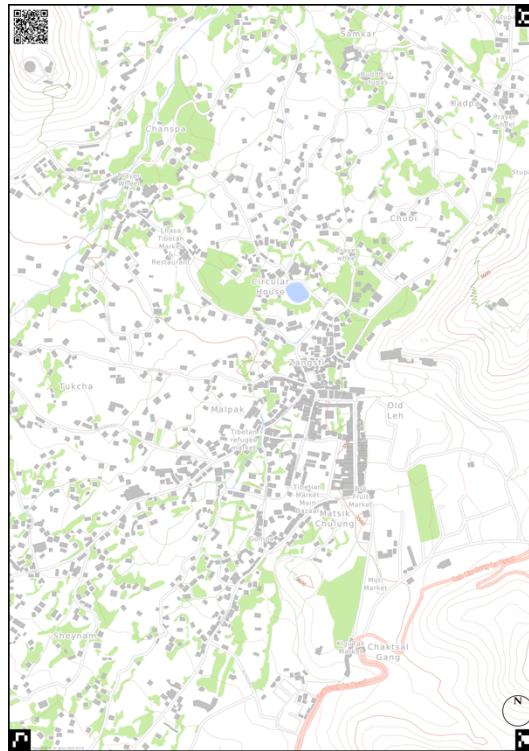

Figure 1. A map of Leh generated using the Paper2GIS software. Base map data © OpenStreetMap contributors.

The user can then print out these maps in the required quantity, and allow participants to draw information onto them during an interview using marker pens. Once the information has been collected, the user can use either a digital camera or scanner in order to convert the maps to digital images, and then pass them back to the software for automatic digitisation. Using computer vision (CV) algorithms, the software is able to find the map in the image, extract it, rectify it to its correct dimensions, isolate the participant's markup and then georeferenced and extract it to GeoTiff and Shapefile files. This automated digitisation process removes the opportunity to introduce errors or bias into the data during the otherwise time-intensive task of manual digitisation. The automatic digitisation process is summarised graphically in Figure 2.
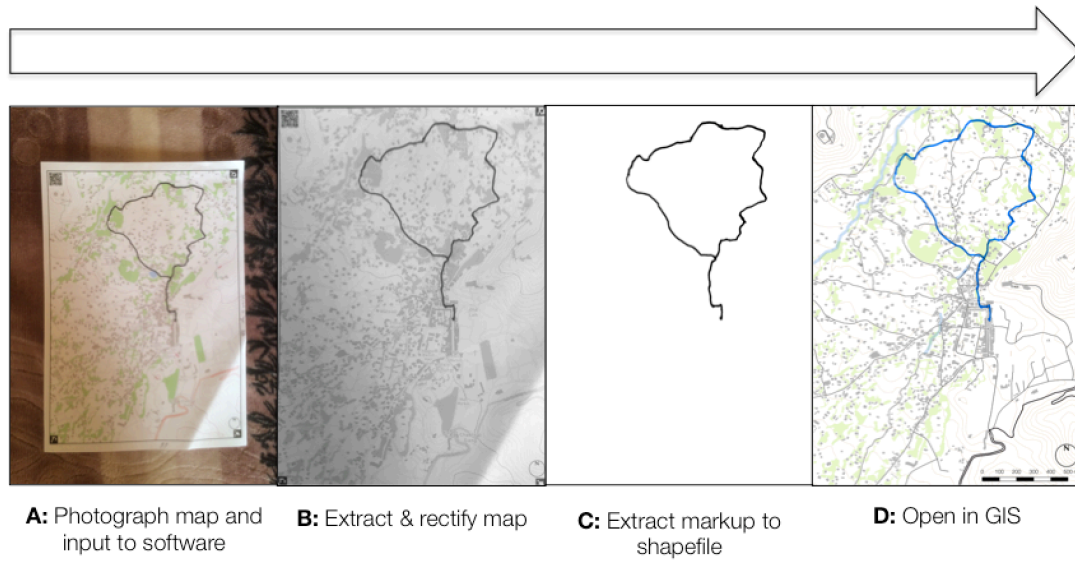
Figure 2. A graphical illustration of the process of automated digitisation in the Paper2GIS software. The data shows a favourite walking route for one of the participants. Base map data © OpenStreetMap contributors.
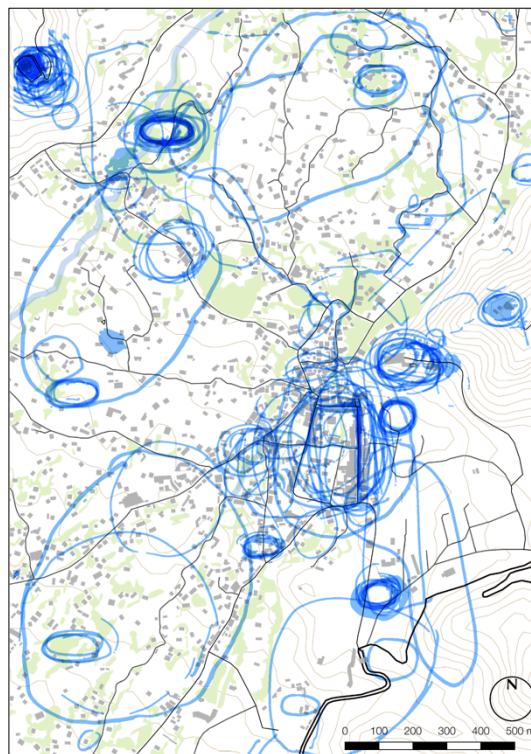
3. Results

Figure 3. The data collected from 94 participants in the village of Leh, Ladakh. Base map data © OpenStreetMap contributors.

The raw data collected from 94 participants in the village of Leh is shown in Figure 3. Though a detailed analysis of this dataset and the implications for the development of Leh is

necessarily beyond the scope of this short paper, even from the un-filtered data shown in Figure 3 it is clear which areas were considered to be the most valuable by many of the participants. These areas were typically either the main market area (towards the centre of the map), or religious buildings such as Shanti Stupa (Figure 4), which was identified by 45 of the 94 participants, and can be found in the very northwest corner of the maps in Figure 3.

Figure 4. Shanti Stupa, the 'most valuable' location in Leh according to our pilot study. Image by Ronakshah1990: https://commons.wikimedia.org/w/index.php?curid=5194701.

4. Discussion and Conclusions

The purpose of this preliminary study was to demonstrate the potential of the prototype Paper2GIS software package for the collection of data in areas where more traditional platforms are likely to be unsuitable for those who either do not have access to the computers and Internet connections, or do not have the skills or desire to use them. Whilst it is clear that further work is required in order to draw any firm conclusions about the perceptions of 'value' in the village of Leh, this paper has demonstrated a novel and useful approach by which participatory GIS data may be collected in this type of environment. Indeed, most of the participants in this study did not have regular access to computers, and any Internet connections that were identified by the researchers were insufficient for modern web-based map interfaces to be used effectively. Accordingly, it may be considered that automated digitisation technology like Paper2GIS does have the potential to mitigate the digital divide for some PPGIS applications. With respect to the representation of geographical entities, participants were free to mark-up the map in any way they liked. Whilst, as one would expect, the majority of participants' markup was comprised of points, lines and polygons, participants were at least afforded the freedom to determine what form these representations took. Further Human-Computer Interaction (HCI) analysis would be required in order to determine the extent to which this is beneficial, but the variety of approaches used (a small sample is given in Figure 4) suggests that this would be worth further investigation.
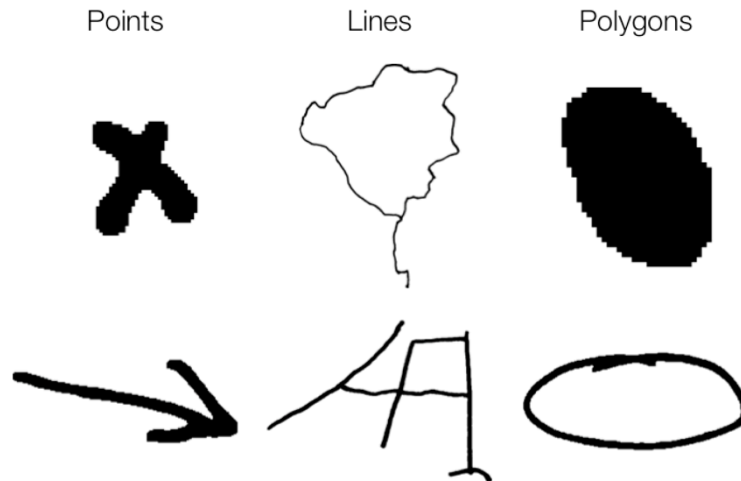
Figure 4: Individual markup 'glyphs' extracted from participants' maps illustrating several approaches to point, line and polygon representations of geographical entities.

In conclusion, the preliminary work presented here suggests that the Paper2GIS software has the potential to permit the automated collection of participatory mapping data in areas where this would not previously be possible due to the effect of the 'digital divide'. The variety of approaches used by participants in their markup of geographical entities is encouraging as it suggests that the free use of the marker pen does prevent the GIS from dictating the representations used by participants for various features in the landscape (either real or imagined). However, further research is required in order to determine the extent to which this is beneficial to both participants and researchers.

The source code for Paper2GIS is available at https://github.com/jonnyhuck/Paper2GIS.

5. Acknowledgements

6. References

Bugs, G., Granell, C., Fonts, O., Huerta, J., & Painho, M. 2010. An assessment of Public Participation GIS and Web 2.0 technologies in urban planning practice in Canela, Brazil. Cities 27, 172-181.

Clementini, E. & Di Felice, P. 1996. An algebraic model for spatial objects with boundaries. In Burrough, P., A. & Frank, A., U. 1996. Geographic Objects with Indeterminate Boundaries (pp. 155-169). Taylor and Francis, London.

Fisher, P. 2000. Sorites paradox and vague geographies. Fuzzy sets and systems, 113(1), 7-18.

Green, D. R. 2010. The role of Public Participatory Geographical Information Systems (PPGIS) in coastal decision-making processes: an example from Scotland, UK. Ocean and Coastal Management 53, 816-821.

Goodchild, M. F. 2007. Citizens as sensors: the world of volunteered geography. GeoJournal 69(4), 211-221.

Goodchild, M. F. 2011. Formalizing place in geographic information systems. In Communities, Neighborhoods, and Health (pp. 21-33). Springer New York.

Google 2017. https://developers.google.com/maps/documentation/javascript

Huck, J. J., Whyatt, J. D., & Coulton, P. 2014. Spraycan: A PPGIS for capturing imprecise notions of place. Applied Geography, 55, 229-237.

Leaflet 2017. http://leafletjs.com

Mackaness, W. A., & Chaudhry, O. 2013. Assessing the Veracity of Methods for Extracting Place Semantics from Flickr Tags. Transactions in GIS 17(4), 544-562.

Maquil, V., De Sousa, L., Leopold, U., & Tobias, E. 2015. A geospatial tangible user interface to support stakeholder participation in urban planning. In Proc. of the 1st Int. Conf. on Geographical Information Systems Theory, Applications and Management.

Montello, D. R., Goodchild, M. F., Gottsegen, J., & Fohl, P. 2003. Where's downtown?: Behavioral methods for determining referents of vague spatial queries. Spatial Cognition & Computation 3(2), 185-204.

Obermeyer, N. J. 1998. The evolution of public participation GIS. Cartography and Geographic Information Systems 25(2), 65-66.

Riddlesden, D. and Singleton, A.D., 2014. Broadband speed equity: A new digital divide?. Applied Geography, 52, pp.25-33.

Williamson, T, 1994. Vagueness. Routledge, London.

Varzi, A. C. 2001. Vagueness in geography. Philosophy & Geography, 4(1), 49-65.