

# GeoHumanities 2019 Workshop Report

Bruno Martins<sup>1</sup>, Ludovic Moncla<sup>2</sup>, Patricia Murrieta-Flores<sup>3</sup>

<sup>1</sup>Instituto Superior Técnico - INESC-ID, University of Lisbon, Portugal

<sup>2</sup>LIRIS UMR CNRS 5205, INSA Lyon, France

<sup>3</sup>Digital Humanities Hub, Lancaster University, United Kingdom  
(Workshop Co-Chairs)

## Abstract

*This article reports on the 3rd ACM SIGSPATIAL Workshop on Geospatial Humanities, held in conjunction with the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. The article outlines the objectives of the workshop, and briefly describes the technical program.*

## 1 Introduction to the ACM SIGSPATIAL Workshop on Geospatial Humanities

The 3rd ACM SIGSPATIAL Workshop on Geospatial Humanities (GeoHumanities'19) was held together with the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, on November the 5th at Chicago, Illinois, USA. The workshop addressed the use of geographic information systems and other spatial technologies in humanities research.

Scholars in the humanities have long paid attention to spatial theory and cartographic outputs. Moreover, in recent years, new technologies and methods have led to the emergence of a field that is now commonly known as the Spatial Humanities. Despite recent developments, many challenges persist in the application of state-of-the-art techniques (e.g., text geoparsing), which have been showcased in venues such as the ACM SIGSPATIAL conference. The workshop is specifically concerned with the use of geographic information systems and other spatial technologies in humanities, including history, archaeology, cultural heritage, and digital humanities research, placing a strong emphasis on new methodologies that leverage recent technical and scientific developments.

## 2 The Workshop Program

The call for papers resulted in 8 submissions describing high quality research. A program committee of 30 members, which are listed on the workshop website<sup>1</sup>, reviewed the submissions, and 6 papers were accepted for presentation. The workshop program featured two technical sessions, followed by a discussion at the end.

After a brief opening address, the workshop started with the presentation of work by Kamakura et al., concerning a digitisation project focused on the Preah Vihear Temple, i.e. a Hindu temple located on the border between Cambodia and Thailand, built in the Angkor Dynasty. The project involved specialised hardware and software for acquiring 3D point-cloud data regarding this particular structure (i.e., the authors used mobile laser range sensors developed by the University of Tokyo, suitable for measuring wide areas, as well as commercially available tripod-based sensors, together with specialised software that combines the results from multiple iterations of partial scanning). The authors also discussed the use of the 3D point-cloud data to determine the direction of the central axis of the temple, afterwards interpreting this direction in the context of Khmer legends.

The second presentation covered research by Huang et al. focused on spatial down-scaling of population data into high-resolution grids, motivated by the fact that access to this type of information is very useful in a broad

---

<sup>1</sup><http://bgmartins.github.io/sigspatial-geohumanities/program-committee.html>

range of applications. Information on building footprints, made available by Microsoft and refined through land use data from the OpenStreetMap, was used to disaggregate the census tract population data (i.e., the latest ACS 5-year estimates (2013-2017) for the conterminous United States) into a high-resolution grid of 100 meters per cell. The actual down-scaling procedure is based on dasymetric mapping, and the authors compared different weighting scenarios leveraging quantities derived from the building footprints (e.g., counts and relative sizes for the footprints within each cell) and from other remote sensing data sources (e.g., nighttime light intensity). In the discussion period that followed the presentation, Xiao Huang highlighted the fact that accurate information on building footprints does indeed provide very useful information for population down-scaling, significantly outperforming the use of other types of ancillary variables.

In the third presentation, Palumbo et al. discussed approaches for standardising category (e.g. restaurant, temple, hotel, etc.) information in point-of-interest (POI) databases, specifically in the context of data integration from multiple sources (e.g., when fusing multiple alternative data sources of POI data with the OpenStreetMap). The authors argue that having consistent category labels is an important concern when using these data for producing land use maps or population distribution estimates, discussing these two applications in detail. The process of matching the source categories, in some source, into a standardised set of categories (e.g., the OpenStreetMap tags) should involve human assessment, but systems like the SONET graph database can provide a programmatic way to manage the category mapping process, by supporting experts in the organisation, storage, and retrieval of POI categories at multiple levels of abstraction.

After the coffee break, the second session started with Brynne Godfrey presenting a paper by Andris et al., in which the authors described the use of a large call detail record (CDR) dataset, from the Trk Telekom mobile operator, to examine how refugees from the Syrian Refugee Crisis are integrating into Turkish society. Individuals were classified according to their call patterns (i.e., refugees who often call Turkish nationals, refugees who do not call Turks, Turks who often call refugees, and Turks who do not call refugees), and they were also geo-referenced with basis on the cell towers that they use. With this information, the authors analysed spatial patterns in the locations that are more commonly associated to the *bridging* members of Turkish society, and they looked for correlations between (a) the location of the different classes of individuals, and (b) different types of infrastructural and socioeconomic indicators (e.g., distance from the Syrian border, intersection with urban amenities, etc.). Bridging Turks were found to be located near particular infrastructural elements (e.g., places of worship or community centers) more often than their non-bridging counterparts, highlighting the significance of social amenities and meeting places for refugee integration. In the discussion session that followed the presentation, Brynne Godfrey and some of the participants discussed challenges involved in using CDR data for research purposes, while complying with privacy safeguards.

Afterwards, Ludovic Moncla described work developed with his collaborators towards the identification of qualitative *neighbours* for toponyms referenced in an eighteenth-century French encyclopedia (i.e., for place mentions in Diderot and d'Alembert's *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers, par une Société de Gens de lettres* (1751-1772, EDDA)). The authors proposed a graph-analysis method that draws on the co-occurrence relations between toponyms that are present in the corpus of encyclopedia articles, arguing that the co-occurrence network can act as an alternative to geospatial representations based on latitude and longitude coordinates, and as a useful proxy to geospatial proximity when no historical gazetteer exists for the source material's period. In his presentation, Moncla also discussed the use of the qualitative neighbours for toponym disambiguation applications.

Finally, Wang and Hu presented their recently proposed EUPEG benchmarking platform for evaluating text geoparsing systems (i.e., systems for recognising place mentions in textual documents, and for assigning each place name to the corresponding location coordinates). In its present version, the EUPEG platform integrates eight annotated datasets, nine baseline geoparsers, and it reports on eight different performance metrics. The authors used the platform to evaluate the different systems that participated in the SemEval 2019 shared task that focused on Toponym Resolution in Scientific Papers, against the baselines included in the platform. Although some challenges remain, the results show that some of the new geoparsers introduced in the SemEval

competition, e.g. based on state-of-the-art neural models, indeed improve the performances on multiple datasets.

A discussion period followed the last presentation of the second session, covering ideas for future developments and common aspects between the different contributions presented at the workshop (e.g., discussing the possibility of integrating other datasets and systems into the EUPEG platform from Wang and Hu, or discussing the possibility of using large POI datasets for population down-scaling).

In total, the workshop had 9 officially registered participants and, on average, 12 attendees were present at each of the session presentations. We believe GeoHumanities'19 was a successful event that, although small, allowed the participants to explore the contributions that modern GIS and geo-spatial analysis technologies can enable within and beyond the digital humanities.

### **3 Acknowledgments and Final Remarks**

The organizers would like to thank the authors for submitting and presenting their contributions, and also the program committee members for their commitment to the paper review process. We hope that the proceedings<sup>2</sup> of GeoHumanities'19 will inspire new research ideas, and that you will enjoy reading them. The organizers would also like to acknowledge the support provided by Fundação para a Ciência e a Tecnologia (FCT), through the project grant with reference UID/CEC/50021/2019, and also by the Trans-Atlantic Platform for the Social Sciences and Humanities, specifically through the 'Digging into Early Colonial Mexico: A large-scale computational analysis of 16th century historical sources' project with reference HJ-253525-ES/R003890/1.

---

<sup>2</sup><http://dl.acm.org/citation.cfm?id=3356991>