Report on the 12th International Workshop on Location and the Web (LocWeb 2022) at WWW 2022

Dirk Ahlers
NTNU – Norwegian University of Science and Technology
Trondheim, Norway
dirk.ahlers@ntnu.no

Erik Wilde
Axway
Switzerland
erik.wilde@dret.net

Abstract

LocWeb 2022 was held as an online workshop at The Web Conference WWW 2022. It was the twelfth of its series and the third one in an online setting. LocWeb 2022 explored the intersection of location-based analytics and Web architecture with a focus on on Web-scale services and location-aware information access. Location and its integration continue to provide challenging research questions. It additionally has a role in large systems of social and societal importance and is reflected in various ways in research approaches or in the use of systems.

Date: 26 April, 2022.

Website: https://dhere.de/locweb/locweb2022/.

1 Introduction

LocWeb 2022 was the twelfth event in its workshop series and took place co-located on 26th April 2022 in conjunction with The Web Conference WWW 2021.[1] While a pure online event, WWW 2022 used the added spatial label "hosted by Lyon, France", similar to previous years when it was moved online and as an acknowledgement to the hosts. LocWeb has now been run for the third time without a physical location.

LocWeb explored the intersection of location-based analytics and Web architecture, with a focus on on Web-scale services and systems facilitating location-aware information access. Workshop details, including some presentations, can be found on the workshop homepage[2] and on the workshop series homepage.[3] The previous LocWeb report for 2021 has been published in SIGIR Forum [Ahlers et al., 2021] as well as those for previous years.

[1] https://www2022.thewebconf.org/
[2] https://dhere.de/locweb/locweb2022/
[3] https://dhere.de/locweb/

2 LocWeb Workshop Theme and Topics

LocWeb continues its main theme of Web-scale Location-Aware Information Access and spatial social computing. Subtopics include (i) geospatial semantics, systems, and standards; (ii) large-scale geospatial and geo-social ecosystems; (iii) mobility; (iv) location in the Web of Things; and (v) mining and searching geospatial data on the Web. The workshop encourages work describing Web-mediated or Web-scale approaches that build on reliable foundations, and that thoroughly understand and embrace the geospatial dimension through interdisciplinary perspectives.

The workshop's topics of interest were slightly evolved with a stronger focus on social good, pandemic impacts especially in the form of hybrid online/physical spaces:

- Location-Aware Information Access
- Location-Aware Web-Scale Systems and Services
- Physical and virtual/online spaces for collaboration and information access
- Experience of virtual and physical place
- Geospatial Data Science for Social Good
- Urban Planning and Citizen Engagement
- Geospatial aspects of Smart Cities
- Spatial Social Behavior and spatial social media
- Location in the Internet/Web of Things
- Open Geospatial Web Data
- Geo-social computing on the Web
- Evaluation of frameworks, metrics and algorithms
- Large-scale Geospatial Ecosystems
- Location-Based Social Networks
- Mobile Search and Location-Based Recommendation
- Geospatial Web Search and Mining
- Standards for Location and Mobility Data
- Modeling Location and Location Interaction

3 LocWeb Workshop Contributions

The workshop featured one keynote, two full and one short paper, forming an interdisciplinary combination of topics. Our PC members provided 3–4 reviews per paper on average, with a 75% acceptance rate. The proceedings [Ahlers and Wilde, 2022] are available in the ACM Digital Library[4] as part of the overall WWW 2022 Companion.

The papers are also linked from the conference proceedings page,[5] and should be available when linked from there under open access by the main conference. Slides are linked – where available – from the workshop homepage.

[4] https://dl.acm.org/doi/proceedings/10.1145/3487553#heading14
[5] https://www2022.thewebconf.org/companion-proceedings/

3.1 Keynote

The keynote was given by Stefano de Sabbata on *Everyday digital geographies*, with the abstract as follows:

> As a clear-cut distinction between online and physical spaces rapidly degrades, online representations of places have become an integral part of the places they represent. The content produced through online platforms has become part and parcel of our experience of place and events, as we capture our experiences and take part in other people's narratives through the medium of the internet. Despite the great interest and the large number of studies devoted to this topic in geography and related fields in the past decade, most of the research in this area has been conducted either using purely quantitative approaches, in the field of GIScience, or a purely qualitative and frequently critical perspective, in the field of human geography. In this talk, I will explore the opportunities offered by mixed methods approaches in digital geographies. Based on my work as part of the Mapping Multiculture project (in collaboration with Dr Katy Bennett and Dr Zoe Gardner), I will discuss the advantages and limitations of combining advanced artificial intelligence approaches with situated qualitative content analysis in exploring the everyday digital geographies of Leicester as a multicultural city.

To highlight main points, they discussed how geographers are interested in human geography. "The digital" is then not just what info is about, but where it is created, how it relates to humans etc. We need to make room for the "everyday" and the mundane on the Web, which will need mixed methods. "Digital Geographies" and "of the digital" are ways to see things from a different perspective.

A grounding reminder was that "most of our lives are made up of 'everyday life'" but "most of our research focuses on the exceptional".

Such work can reveal new aspects of social interactions, but it is more challenging to handle algorithmically. A main message methodologically was that mixed methods require dialogue. On the technical side, a reminder and experiences were given regarding tools that "off the shelf" != "ready to use" and lots of adaptations are needed.

As examples from geographers' interest, they presented the ideas of chain restaurants as a 'third place' in recent work [Bennett et al., 2022]. Regarding the mixed methods, here the approach built in dialogue was using deep learning to complement manually annotated tweets.

3.2 Paper session

The first paper, *Exploiting Geodata to Improve Image Recognition with Deep Learning* [Arbinger et al., 2022] by Christian Arbinger, Martin Bullin, and Andreas Henrich, used deep learning to integrate auxiliary geodata into image classification tasks by integration of geo features into a CNN fusion model. Their work includes both image and geo-coordinate chains. They use mainly addresses as textual features, but also include satellite images into the network. It discusses likely error cases; some are common to image analysis, some are due to baselines, where for example standard cases (e.g. cars, fauna and flora, etc.) are already well covered. An interesting

observation was that comparisons to related work are difficult, as this is a rather unique approach on data sets and algorithms.

In *Anonymous Hyperlocal Communities: What do they talk about?* [Reelfs et al., 2022], Jens Helge Reelfs, Oliver Hohlfeld, and Niklas Henckell discuss findings from an analysis of the Jodel location based social network and content. Their example of Saudi Arabia gives different cultural environments and independent hyperlocal communities. They use both a content classification schema, and a crowdsourcing campaign, to show how reactions to content can vary between locations.

Finally, in *Predicting Spatial Spread on Social Media* [Rimjhim and Dandapat, 2022], Rimjhim and Sourav Dandapat present their work which aims at predicting the spreading of topics on social media. In their specific case, they aim to predict the final spread of a given hashtag as its spatial distribution within the social network, and they aim to do this in the early stages, e.g. when there are yet only around 50 tweets using the hashtag. With that, they try to prediction differences in general importance and in spatial importance of local vs regional or national spread.

4 Discussion, Conclusion, and Future Directions

The discussion started with issues from the workshop contributions and then expanded to wider issues. A general theme was that datasets are still an issue, regarding quality, granularity, and bias as a strongly emerging topic of attention.

It was remarked that the granularity of datasets can still be challenging for many advanced questions, and that the combination of the spatial and temporal dimension often exposed different quality or granularity levels where some participants remarked that timestamps are often of better quality than geo-'stamps'. Based on the dataset question, we then discussed: What sort of questions can we even ask? What problems can we address / study? How do we understand better what we work on or what we avoid working on due to such issues? This evolved into a discussion on how we could set a threshold between quantitative and qualitative approaches.

Our keynote speaker Stef noted an important difference in two views of *geographies of the digital* (understand data that you have), and *geographies through the digital* (understand general phenomena through data)

This reflects the earlier questions of what we really want to understand in our work. As a practical example, this can be the reactions of people on twitter to an earthquake vs. the actual earthquake itself. In many such cases, using geospatial and temporal dimension of the content comes very naturally. Data can still be a bad proxy for real-life phenomena, and this difference is often not made explicit.

Regarding bias, there is a clear experience from participants that bias also appears in data distribution or the the size of places and the amount of data available for them, for example comparing data in (mega)cities versus smaller places, which is not just a function of number of inhabitants; an issue continuing from very early mentions in the literature, and discussions at previous workshops.

Someone brought up the xkcd joke that many heatmaps of alleged phenomena are only population maps[6], but it was also remarked that this may be less strong outside the US and in other

[6] xkcd.com, Heatmap, 2012, https://xkcd.com/1138/

places. However, that can often be the case due to unequal access to technology or connectivity, as well as other factors. It was also remarked that high data production and availability of geo data can still be very biased to certain countries. Bias can then also appear due to less (or less equalised) data production or data production in a different way, which can throw off analysis approaches. For example, people may react or act differently on social media depending on culture, country, or other background, and in general, people are less homogenous than assumed by some methods. Participants noted a gap in the literature on how to better assess biases and use them to understand the impact of developed solutions and their implicit bias better. While often authors or developers may be aware of the bias, it is often not openly discussed or acknowledged. The question was then asked of how much bias, uncertainty, etc. is in the actual input data, even before any processing? In other words, it is often not obvious if bias comes from processing, or from the underlying data, and participants feel unsure about how to understand and address it.

For certain distribution-dependent bias types, even if we narrow it down to a single city, we may get a different type of bias again (just at a different scale). Participants wondered if there may be a detectable step change when we move from large-scale spatial data to results of local participatory data collection, or if one may be more dominant at certain scales, but we could not find an answer to that. The question was of course linked with larger streams of bias in AI, ethics, or privacy (especially for location), and also reflected a lot on the inputs given by the keynote. As a possible way out, the question was posed on whether we can better quantify the uncertainty and thus find a different way of working with uncertainty and biased data.

The discussion touched on many important topics, which each in themselves carry important research questions, which hopefully will be inspiring to the next iteration of the workshop.

Acknowledgements

We thank the authors of submitted papers for their efforts, our keynote speaker Stefano de Sabbata, and all workshop attendees. Special acknowledgement go to the members of the LocWeb program committee for their hard work in reviewing the contributions and providing substantial feedback on a tight deadline:

Andreas Henrich (University of Bamberg), Auriol Degbelo (University of Münster), Bruno Martins (IST and INESC-ID – Instituto Superior Técnico, University of Lisbon), Carsten Keßler (Aalborg University Copenhagen), Chien-Wen Shen (National Central University Taiwan), Christopher Jones (University of Cardiff), Claudia Hauff (TU Delft), Clodoveu Davis (Universidade Federal de Minas Gerais), Francisco J. Lopez-Pellicer (University of Zaragoza), Jalal S. Alowibdi (University of Jeddah), Leonidas Anthopoulos (University of Thessaly), Luca Aiello (IT University of Copenhagen), Marc Spaniol (Université de Caen Normandie), Pieter Colpaert (Ghent University), Rainer Simon (AIT – Austrian Institute of Technology), Ross Purves (University of Zurich), Torsten Suel (New York University), Yana Volkovich (Xandr).

We further like to thank the organizers of WWW2022 for hosting the workshop and providing the online meeting facilities.

References

Dirk Ahlers and Erik Wilde. LocWeb2022 Workshop – Proceedings of the Twelfth International Workshop on Location and the Web. In *WWW '22: Companion Proceedings of the Web Conference 2022*, WWW '22, New York, NY, USA, 2022. ACM. ISBN 978-1-4503-9130-6. doi: 10.1145/3487553.3524871. URL https://doi.org/10.1145/3487553.3524871.

Dirk Ahlers, Erik Wilde, Marc Spaniol, Ricardo Baeza-Yates, and Omar Alonso. Report on the 11th International Workshop on Location and the Web (LocWeb 2021) and the 11th Temporal Web Analytics Workshop (TempWeb2021) at WWW2021. *SIGIR Forum*, 55(2), 2021. URL https://sigir.org/wp-content/uploads/2022/02/p06.pdf.

Christian Arbinger, Martin Bullin, and Andreas Henrich. Exploiting geodata to improve image recognition with deep learning. In *Companion Proceedings of the Web Conference 2022 (LocWeb2022 workshop)*, WWW '22, page 648–655, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391306. doi: 10.1145/3487553.3524645. URL https://doi.org/10.1145/3487553.3524645.

Katy Bennett, Zoe Gardner, and Stefano De Sabbata. Digital geographies of everyday multiculturalism: 'Let's go Nando's!'. *Social & Cultural Geography*, pages 1–20, April 2022. doi: 10.1080/14649365.2022.2065699. URL https://doi.org/10.1080/14649365.2022.2065699.

Jens Helge Reelfs, Oliver Hohlfeld, and Niklas Henckell. Anonymous hyperlocal communities: What do they talk about? In *Companion Proceedings of the Web Conference 2022 (LocWeb2022 workshop)*, WWW '22, page 639–647, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391306. doi: 10.1145/3487553.3524644. URL https://doi.org/10.1145/3487553.3524644.

Rimjhim and Sourav Dandapat. Predicting spatial spread on social media. In *Companion Proceedings of the Web Conference 2022 (LocWeb2022 workshop)*, WWW '22, page 656–659, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391306. doi: 10.1145/3487553.3524646. URL https://doi.org/10.1145/3487553.3524646.