Mapping Points of Interest through street view imagery and paid crowdsourcing

EDDY MADDALENA, University of Southampton, United Kingdom
LUIS-DANIEL IBÁÑEZ, University of Southampton, United Kingdom
ELENA SIMPERL, University of Southampton, United Kingdom

We present the Virtual City Explorer (VCE), an online crowdsourcing platform for the collection of rich geotagged information in urban environments. Compared to other Volunteered Geographic Information (VGI) approaches, which are constrained by the number and availability of mapping enthusiasts on the ground, the VCE uses *digital street imagery* to allow people to virtually explore a city from anywhere in the world, using a browser or a mobile phone. In addition, contributions in VCE are designed as *paid microtasks*, small jobs which can be carried out without any specific knowledge of the local area or previous mapping expertise in exchange for a fee. We tested the VCE in two cities to map Points of Interest (PoIs) in transport and mobility, using FigureEight to recruit participants. We were able to show that our platform enables crowdworkers to submit PoI location seamlessly, cover almost all of the tested areas, and discover several PoIs not reported by other approaches. This allows the VCE to complement existing approaches that leverage experts or grassroot communities.

Additional Key Words and Phrases: crowdsourcing, geographical information, geospatial information, micro-tasks, urban auditing, mapping

1 INTRODUCTION

We present the Virtual City Explorer (VCE), an online crowdsourcing platform for urban auditing. The VCE helps public administrations and mapping agencies recruit volunteers online and ask them to locate *Points of Interest (PoIs)* in a city, such as shops, bus stops, or disabled parking spaces. Participants explore the relevant area remotely using digital street imagery, spot the PoIs and submit their location using a built-in custom screen capture capability. Contributions are designed as *paid microtasks*, small data-collection jobs which can be carried out without having any local knowledge or mapping expertise. Each contribution is rewarded with a fee in the range of $0.05.

The Virtual City Explorer shares the goals and objectives of *Volunteered Geographic Information (VGI)* systems. VGI broadly refers to geospatial data that is voluntarily provided by individuals, and to frameworks, methods and tools that support this process [Hardy et al. 2012]. It has proven successful in a range of applications, from disaster response to urban auditing to cartography, leveraging advances in Internet connectivity and mobile sensing [Goodchild et al. 2007; Sui et al. 2012]. Several studies have compared the performance of VGI volunteers with expert mappers [Dorn et al. 2015; Fan et al. 2014; Haklay and Weber 2008], concluding that they can achieve high completeness and correctness (greater than 80%), as well as a fair positional accuracy (4 to 8 meters). In addition, VGI is sometimes used as an engagement tool by local government, encouraging residents to supply data for evidence-based policy making and improving their sense of community [Jiang et al. 2016; Ruiz-Correa et al. 2017].

The large majority of VGI projects require presence on the ground (or knowledge of the geographic area being investigated). For example, in OpenStreetMap[1], people are invited to add

---

[1]https://www.openstreetmap.org

Authors' addresses: Eddy Maddalena, University of Southampton, P.O. Box 1212, Southampton, United Kingdom, SO17 1BJ, e.maddalena@southampton.ac.uk; Luis-Daniel Ibáñez, University of Southampton, P.O. Box 1212, Southampton, United Kingdom, SO17 1BJ, l.d.ibanez@southampton.ac.uk; Elena Simperl, University of Southampton, P.O. Box 1212, Southampton, United Kingdom, SO17 1BJ, e.simperl@southampton.ac.uk.

information about shops, restaurants and other PoIs in their neighbourhood to the shared map, and are taught more advanced techniques such as GPS tracking or photo and audio mapping to take on more substantial mapping projects in under-explored areas. As such, participation is constrained by the number of people who are able to visit the relevant locations, or have the resources and expertise to contribute effectively. In addition, most projects do not use incentives to drive participation, relying on people who are intrinsically motivated to contribute, such as hobby mappers, activists, civic groups and other local initiatives [Quattrone et al. 2017]. Coverage is hence patchy [Hardy et al. 2012] and levels of engagement are difficult to manage [Mooney and Corcoran 2014].

The Virtual City Explorer follows a different approach to tackle these challenges. By using digital street-view imagery, for example the popular Google Street View service [Anguelov et al. 2010], we allow people from all over the world to explore an urban space using their personal computer or mobile phone, reducing the need to visit a location in real life. By designing contributions as paid microtasks, we provide detailed instructions and tools that help with data collection and reward participation through small payments. The use of paid microtask crowdsourcing means that we have access to potentially millions of crowdworkers on-demand, using platforms such as FigureEight or Mechanical Turk for recruitment. While financial incentives might not appeal to everyone [Mao et al. 2013], they offer city planners and mapping agencies a means to control crowd behaviour and boost contributions, especially for areas that would otherwise remain uncovered by volunteers on the ground.

This paper describes the design, implementation and evaluation of the VCE. A *requester*, the person or organisation interested in collecting the PoI data, defines an area of interest on the map and the relevant types of PoIs. They specify the number of participants to be recruited and provide instructions for the tasks they need to carry out. The VCE generates a web interface based on these parameters, which enables the crowd to explore the area of interest using digital street-view imagery and report the coordinates of the PoIs by taking a series of screenshots of the PoIs' locations. To validate the coordinates submitted by the crowd and identify which submissions refer to the same PoI, our system uses the DBSCAN clustering algorithm. Inspired by the *taboo* words technique of the ESP game for image labelling [von Ahn and Dabbish 2004], validated PoIs cannot be submitted by workers joining the task after validation took place, avoiding them to converge on the easier to locate PoIs. Participation is rewarded based on the number of PoIs submitted. To be fair with workers that arrive after all PoIs have been validated and discovered, we give the same reward to workers that have spent a pre-defined amount of time *and* have walked a pre-defined distance without submitting any PoI, *i.e.*, that trigger a *escape condition*.

VCE's exploratory approach is different from related urban auditing projects, which define itineraries [Saha et al. 2017], restricts tasks to particular locations on the digital map [Hara et al. 2015], or partition the area to explore in segments that are prioritized [Qiu et al. 2019]. These strategies require knowledge of the area to be investigated, as well as heuristics to manage the tensions between resource allocation and area coverage. By contrast, our approach can be used in any urban context with the only pre-requisite of available street-view imagery. The study was carried out on two areas of interest: the *Limited Traffic Zone* of Trento, Italy; and an area in Nantes, France constructed to roughly match the size of the Trento one, using participants recruited via the FigureEight platform, aimed at answering the following research questions:

RQ1: Feasibility Can the crowd identify PoIs effectively using the VCE?

RQ2: Completeness Does the crowd locate all PoIs in an area? How does it compare with maps generated with alternative approaches?

RQ3: Area coverage Does the crowd cover the entire area of interest in their exploration?

RQ4: Cost and time How does VCE's cost and time to generate a map compare with approaches requiring physical presence? How many workers trigger the fairness strategy, thus, being paid without ROI for the requester?

RQ5: Crowd experience To uncover issues with our task design and escape conditions, we asked the following questions: Do workers struggle with the interface? How much time crowdworkers spend in the task? How much distance they cover? How many of them trigger the escape condition? Is there any difference between workers that complete the task and workers that trigger the escape condition?

We analysed their submissions, the paths they followed to spot the PoIs, as well as the time it took to collect all PoIs and the associated costs. We benchmarked it against related expert- and VGI mapping approaches: open government datasets released by the city authorities collected by one expert on the field, as well as by volunteers coordinated by local associations. We also considered OpenStreetMap data. We found the crowd was able to effectively locate PoIs, covered most of the studied area, and discovered some PoIs that were not detected by other approaches. In turn, some PoIs discovered by other approaches were missed by the VCE.

The remainder of this paper is organised as follows. Section 2 gives an overview of related work. Section 3 describes the Virtual City Explorer, including its main components and user experience elements for requesters and crowd. Section 4 present the design and findings of the two studies, while Section 5.5 analyses the behaviour of workers when working under the basic and the optimised strategy. Section 6 discusses strategies for setting the VCE parameters and scale up on the light of our results,how to complement current approaches with the VCE, how the VCE could generalize to other geographical contexts, and how contributors' biases reported for VGI system affect it. Finally, section 7 summarises our findings and outlines areas of future work.

This work has been approved by the Ethics and Research Governance (ERGO II) committee of the Faculty of Engineering and Physical Sciences (FEPS), University of Southampton, on 16 April 2018, with submission ID 41038.

2 RELATED WORK

In this section, we give an overview of related participatory approaches to collect geospatial data. A comprehensive survey, including terminology and ongoing tools and projects is available in [See et al. 2016] or [Sui et al. 2012]. As our aim is mainly to position ourselves against other systems that use digital imagery, we will focus here on three research themes only: Volunteered Geographical Information (VGI), spatial crowdsourcing, and urban auditing. They, alongside all other themes discussed in the literature, share many commonalities in terms of principles and goals, but have emerged in different scientific communities, using different methods and focusing on specific challenges.

2.1 Volunteered Geographical Information (VGI)

Volunteered Geographical Information (VGI) enlists the help of volunteers to generate maps. It can be understood as the crowdsourcing equivalent of more traditional mapping methods that are followed by mapping agencies and other government authorities, which involve experts and established tools and methodologies [Sui et al. 2012]. The best example of VGI is probably OpenStreetMap, an open, collaborative initiative to produce a highly detailed world map, editable by anyone [Haklay and Weber 2008]. Several studies [Dorn et al. 2015; Fan et al. 2014; Haklay and Weber 2008] have observed that VGI leads to high completeness and correctness (> 80%), with a fair position accuracy (4-8 meters of difference) compared to maps produced by experts.

The main limitation of VGI is that the accuracy of the data is proportional to the number of participants [Haklay et al. 2010]. [Hecht and Gergle 2010] introduced the notion of *Spatial Content Production models* to characterise the 'localness' of VGI data, distinguishing between (i) *you have to be there* and (ii) *flat earth* models. In (i) participants are required to be physically present in a certain area of interest, while in (ii) the distance between participants and area of interest is not relevant, which means that people can contribute remotely. Our work is in the second category - the Virtual City Explorer enables volunteers to collect geocoordinates of PoIs using digital street-view imagery. This helps with recruitment, as we do not require resources on the ground.

The V in VGI is sometimes linked to an intrinsic desire to participate [Quattrone et al. 2017]. [See et al. 2016] discusses two main reasons why people get involved in projects such as OpenStreetMap: (i) contributing to a common good and (ii) the added value they may obtain in return, for example through improved public services. While this works well if there is a critical mass of participants, it leads to patchy data coverage [Hardy et al. 2012] and organic engagement which is difficult to manage or predict [Mooney and Corcoran 2014].

2.2 Spatial crowdsourcing

The crowdsourcing community is concerned with everything related to the design of crowdsourcing workflows, the allocation of people to tasks, the validation of the data they submit, and the best ways to drive participation [Estellés-Arolas and González-Ladrón-de Guevara 2012]. Spatial crowdsourcing looks at these elements for tasks that require the crowd to collect or confirm geospatial data [Zhao and Han 2016].

The relationship between VGI and spatial crowdsourcing is subtle and has been discussed in the literature [Harvey 2013]. Some authors argue that in VGI the participants are aware that they are part of a grassroots mapping project and provide and share data from an intrinsic motivation. This is one of the reasons why VGI is sometimes used as a participatory tool by public authorities - citizens collect data about their surroundings and help co-design policies, becoming more engaged with local decision makers and communities [Jiang et al. 2016; Ruiz-Correa et al. 2017]. Crowdsourcing on the other hand includes many other participatory experiences, including the use of social media feeds [Xu et al. 2018], human computation [Quinn and Bederson 2011], or games with a purpose [Ahn 2006]. The crowd helps producing maps, but their participation comes in many forms and may serve different purposes. For example, Urbanopoly is a gamified mobile platform that can be used to extend and enrich OpenStreetMap. Participants engage with it mainly from a desire to play the game and win challenges [Celino et al. 2012], and not because they are interested in open maps or helping the OpenStreetMap community. According to the game narrative, which is similar to the well-known game Monopoly, each player is a landlord that manages a given budget to buy real estate. They provide rich descriptions of PoIs and upload pictures. Cumulatively, their contributions help improve the quality of OpenStreetMap in a particular area, but the system needs to follow a different approach to drive participation compared to platforms such as OpenStreetMap or Ushahidi[2], which rely on intrinsic motivations.

Spatial crowdsourcing faces a specific set of challenges. Approaches that require presence on the ground aim to find optimal task allocation strategies that minimise the effort required from each participant, and hence the overall budget that is needed to reward participation [Wang et al. 2018]. There are also privacy concerns, as submissions can reveal a participant's location. [Kandappu et al. 2018; Tong et al. 2017] In our case, we used paid crowdsourcing to mobilise a large pool of remote participants, who collect geocoordinates of PoIs using the VCE tool. Virtual exploration means we

---

[2]https://www.ushahidi.com/

are not using any location data. However, our approach is only applicable when the data can be derived from digital street-view imagery and does not have a strong real-time element.

2.3 Microtask crowdsourcing and street-view imagery in urban auditing

Urban auditing refers to the collection of statistics about most aspects of quality of life in a city, including demography, housing, healthcare, jobs, education, environment, transport and tourism. Previous efforts have proposed the use of virtual spaces to support professional auditors with their tasks, as an alternative to field work [Badland et al. 2010]. For example, FASTVIEW [Griew et al. 2013] supports the exploration of nine street characteristics (e.g. pavement width and solidity) on top of Google StreetView, while Vanwolleghem et al. [2014] studies the reliability and validity of a Google StreetView-based safety assessment of cycling routes for children. Both papers suggest that experts are almost as accurate when using digital imagery than on the ground, and emphasise efficiency savings, for instance in travel. The Virtual City Explorer attempts to prove that paid crowd workers recruited from online platforms such as FigureEight can be as effective as experts and mapping enthusiasts in collecting information about PoIs.

There are several projects that use paid microtask crowdsourcing on top of digital street-view imagery in specific contexts. For example, [Hara et al. 2013] took a set of flat images from Google StreetView and developed an image labelling tool for Mechanical Turk, showing that untrained crowds can effectively identify accessibility issues on streets. In [Hara et al. 2015] the authors applied a similar approach to detect landmarks in order to create an accessibility map for low-vision or blind bus riders. Crowd participants are placed in the vicinity of a known bus location (using data from Google) and asked to audit it. Compared to them, our system is designed to offer participants an exploratory experience, in addition to more focused tasks which are bound to specific locations known in advance. This means that we do not need any specific knowledge about the area to be audited or preparatory work in form of pre-defined itineraries [Hara et al. 2015].

[Saha et al. 2017] presented preliminary results of *Project Sidewalk*, an accessibility auditing tool in the style of [Hara et al. 2013], with an exploratory element where participants are asked to inspect pre-defined segments of a street instead of being positioned at a location. Their task is slightly different to ours in the sense that they require complex labelling of street attributes (missing ramps, obstacles), while we focus on the location of a single type of PoI. As the pilot is ongoing, we were not able to find any information about how effective the approach is in terms of accuracy, completeness and coverage of the data.

The recent work by [Qiu et al. 2019] partitions the search space into segments of fixed length and create a priority queue to assign segments to incoming workers based on the estimation of the density of the target object across segments, calculated iteratively as more segments are processed. The density probability calculation is complemented by a geo-location prediction method that depends on the targeted urban object. However, if the distribution of target objects is known to be sparse, the density of many segments will be zero until an object is eventually found, making impossible to prioritise segments during that time. Furthermore, it might be difficult to come up with geo-location prediction method for PoIs that are irregularly distributed or with low occurrence in the area of interest. Our approach tackles the problem differently, letting workers freely explore the area and rewarding them for finding a target object instead of rewarding per segment covered.

3 THE VIRTUAL CITY EXPLORER (VCE)

The VCE is an online microtask crowdsourcing platform that facilitates the collection of urban data, in particular Points of Interest (PoIs), which are accessible via digital street-view imagery.

The platform has been developed in the context of Qrowd[3], a three years project co-funded by the European Commission, which deploys crowdsourcing to create augmented intelligence solutions in transport and mobility. It is targeted at local government looking for efficient ways to carry out urban auditing, including maps and official statistics. The public administrations that are part of Qrowd are focusing on cycling infrastructure and parking, but the system can be configured for any other type of PoI, for example potential hazards on the street, or accessibility information.

We organised a series of co-design workshops with local government to identify requirements for the VCE. We distinguish between two stakeholders: (1) the requester, an individual or organisation that is looking for PoI data and set up their crowdsourced data collection project in the VCE; and (2) participants from the crowd who explore a virtual space in search of PoIs. We use the term 'crowd' to generically refer to the collective of all participants.

The requirements are as follows:

(1) The tool should support data collection for multiple types of PoIs to meet the evolving needs by city planners and urban auditors.

(2) The tool should allow the requester to define an area of interest, so that it can be used to direct resources to where they are most needed. As such, the tool is by design understood as complementary to other data collection efforts, including VGI.

(3) The tool should provide capabilities for the crowd to report the location of the PoIs they found through exploration, as well as other, configurable visual attributes. For example, when collecting information about bike racks, a type of PoI, the crowd could be asked to describe each bike rack in terms of its shape or the number of bikes it can accommodate.

(4) The tool should provide means for quality assurance and data aggregation to compile a curated set of PoIs from the answers of the crowd. These are meant to be added to existing transport and mobility data that the requester holds to update it.

(5) The tool should allow the requester to configure the size of the crowd, the rewards for each participant, and the number of independent detections of a PoI (also required for quality assurance, see below).

Based on these requirements, we defined the high-level architecture diagram shown in Figure 1. The street-view service at the bottom of Figure 1 is external to the VCE; companies such as Google, Microsoft or TomTom, which is participating in the Qrowd programme, already offer such services, including APIs to embed imagery in 3rd party applications and access metadata. The components in the VCE box use the service in two ways: requesters need it to configure their project, while the crowd explores the city virtually to collect the PoI information. To meet our requirements, a street-view service needs to provide at least the coordinates that correspond to the street images, in order to allow the crowd to easily report PoI locations, and the timestamp of these images, to assess the accuracy of the collected data against a gold standard. Our implementation currently uses Google StreetView for this purpose.

The crowdsourcing service at the top of Figure is optional. It refers to external platforms which offer basic recruitment and payment capabilities. Currently, several such services exist, including FigureEight[4] and Mechanical Turk[5]. Our implementation uses the former. The VCE publishes tasks on FigureEight, which redirect the participants to the Virtual City Explorer.

The VCE itself is in the middle of Figure 1. It is made of the three components:

Configurator It allows requesters to specify the parameters of the exploration task the crowd will be asked to undertake: the area to be explored; the type of PoI; the number of participants

---

[3]http://qrowd-project.eu/
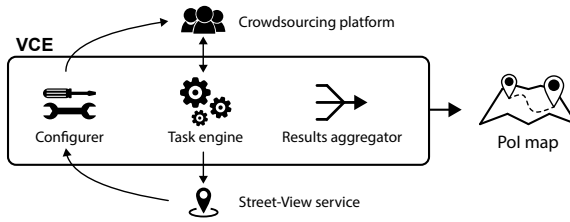[4]www.figure-eight.com
[5]www.mturk.com

Fig. 1. High-level architecture of the Virtual City Explorer

to be recruited; and their starting positions. The output of the *configurator* is passed on to the *exploration task engine*.

Exploration task engine This component creates the microtask interfaces and logs all activities by the crowd, including the information they collect about PoIs.

Aggregator This component takes all information provided by the crowd and processes it to identify which of the submitted coordinates correspond to the same PoI, validate PoIs locations, and exclude outliers. It delivers a curated PoI map, which is used to update existing geospatial datasets held by the requester.

In the following we describe each of these three components in more detail.

3.1 Configurator

The *configurator* is a web interface that allows a requester to set the parameters of the exploration task to be carried out. The requester needs to be registered with a crowdsourcing service, in our case FigureEight, to publish the tasks and pay the crowd.

Area of interest. First, the requester is presented with an embedded Google Maps view where they can draw a closed polygon that represents the area of interest that the crowd will explore. The polygon is saved as GeoJSON object and passed on to the *exploration task engine*. After selecting the area, the requester selects a type of PoI and writes instructions that will help the crowd complete their tasks effectively, as suggested in [Kazai et al. 2011].

Crowdsourcing design. Next, the requester sets the number of times they allow the task to be executed (*numExecutions*) - they have the option to disallow the same participant to repeat the task. The requester can also choose a function to assign the starting position of workers on the map (*startPoint*). In the current implementation, we assign starting points randomly; while other strategies could be implemented, e.g., based on the area covered by previous workers. In addition, when assigning the starting points, we check to see if these are located on-street rather than within buildings to make it easier for the crowd to go about their tasks. Subsequently, the requester specifies the number of PoIs that each crowd participant will be asked to locate (*numInstances*), and the reward for completing the task (*Reward*). An exploration task instance is considered complete if the participant submits *numInstances* distinct PoIs. Details on how we handle PoI duplicates are discussed below.

Publication on FigureEight. Finally, when the requester launches the exploration task, the *configurator* connects to the FigureEight API to publish a task on their platform, which redirects the crowd to the Virtual City Explorer. The GeoJSON defining the area of interest, the instructions, the number of participants, and the list of starting points are passed to the *exploration task engine*.
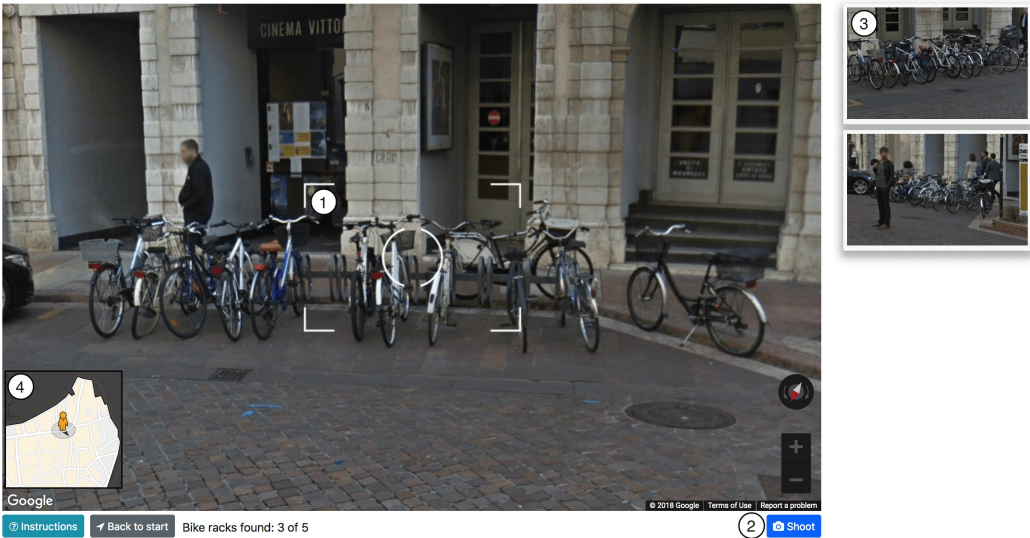
Fig. 2. The VCE interface for explorers: (1) The viewfinder combined with (2) the 'Shoot' button allows participants to take snapshots of PoIs. On the right side (3) workers can review the shots they have taken. Photos need to be taken from different angles, and when at least three are needed to enable submission. The (4) mini-map shows workers their relative position with respect to the exploration area's boundary.

3.2 Exploration task engine

This is the heart of the VCE. It implements the task interface for the crowd and stores intermediate results pre-aggregation. Participants need to be registered with FigureEight and access the VCE via a microtask published on the FigureEight platform.

Exploring the virtual city. When the participant is redirected to our system, the exploration task engine instantiates a new exploration task within the given parameters and assigns a starting point. The participant is first presented with the instructions prepared in the *configurator*. Once they acknowledge them, the component proceeds to the exploration interface, an *HTML5 JavaScript* page that connects to a *MongoDB* database. Figure 2 shows this interface as it is presented to the crowd after they have read the instructions. The page contains a main window with embedded Google StreetView content that the participants can explore using familiar Google StreetView controls located bottom right: on-screen arrow buttons to move in the arrow's direction, as well as double-click on a distant point to move to it. To restrict exploration to the area of interest, whenever a participant attempts to navigate outside its boundary, they are automatically returned to their last location within the area, and a message explaining the reason for the replacement is shown. Additionally, the interface shows participants their relative position with respect to the boundary in a mini-map located on the bottom left corner.

Submitting PoIs. To allow the crowd to report PoI coordinates, we implemented a feature similar to *photoshoots* or *screen captures*. Participants are provided with a camera-style viewfinder and a *Shoot* button (placed underneath the navigation controls in blue in Figure 2). When this button is clicked, a snapshot of the current view is taken and shown at the right side of the screen, including metadata about participant's heading and coordinates. Participants are asked to take three different pictures of each PoI, from three different angles. They can delete and update their pictures until
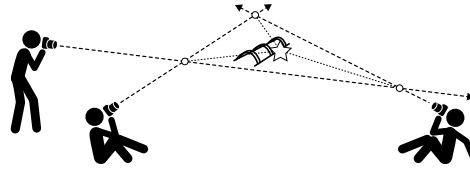
Fig. 3. Visual representation of the validation method. The centroid of the triangle formed by the intersection of the vectors of the three pictures taken by a worker (here marked as a star) is reported as the PoI location. A valid submission must have the centroid of the triangle within δ meters from all sides.

they are satisfied with them and press the *Submit* button to forward the PoI location for validation. *Submit* button is enabled when three valid pictures have been taken. In Figure 2 this is not the case - there are only two thumbnails on the right hand side.

Validating crowd inputs. When the crowd submits a potential PoI, the engine verifies that the three pictures point to the same point in space to rule out submissions with three random pictures. To do so, it computes the intersection area of the vectors that start from the registered coordinates of each participant for each picture, as shown in Figure 3. For a submission to be valid, the distance from the centroid of the triangle to each side (δ) needs to be less than a configurable distance (set to 10 meters in our studies). If the submission is valid, the coordinates of the centroid are registered as the PoI's location. Otherwise, the submission is rejected and a message is shown explaining the reason.

Completing the work and rewarding the participants. When a participant submits the required *numInstances* number of PoIs, the task finishes and they receive their reward via FigureEight. Note that the engine does not check if the pictures are of a correct PoI. Therefore, it is possible to submit any object and still get rewarded as long as the pictures pass the triangulation check. To mitigate against such malicious behaviour, we use the rating capabilities provided by FigureEight: crowdworkers are uniquely identified in the platform, workers that consistently contribute fake PoIs can be given a bad rating, affecting their reputation,and with it, their chances to get recruited by other employers. We stressed this fact in the instructions of the task. In Section 5.6, we will present the quality of the geocoordinates submitted by the crowd through this mechanism, compare them with other VGI-collected datasets, and discuss alternative methods. Once the configured number of task executions has been achieved, the task is retired from FigureEight and the collected list of coordinates is passed on to the next component for aggregation.

Taboo. During preliminary exploratory studies, we noticed that as the crowd has the freedom to choose which PoIs to submit, those that are easier to reach are submitted by a large number of participants. If a subset of the PoIs is located in an easily accessible area, participants might report those and neglect the remaining area (see also Section 5.4). This impacts on completeness and area coverage, and wastes crowd and requester resources. To mitigate this, in the spirit of other crowdsourcing scenarios using output agreement [Feyisetan and Simperl 2017; von Ahn and Dabbish 2004], we implemented a *taboo* feature which retires a PoI once it has been detected by a (configurable) sufficient number of participants to consider it *confirmed*, set in the aggregator component (Section 3.3. A retired PoI is shown in the VCE interface with a *taboo* red sign, and any subsequent attempt to submit it is blocked; an error message is displayed explaining the reasons why the PoI is no longer accessible. While this prevents the waste of resources on already confirmed PoIs, it also raises fairness concerns: participants starting the task after others have finished may find that a large share of the PoIs are already taboo and need to spend more time foraging for others.

If this takes too long, they might even abandon the task without receiving any reward [Difallah et al. 2014]. To counter this problem, we introduced an *escape condition* based on two parameters: a combination of the time spent on the task (*escapeTime*) and the distance covered (*escapeDistance*): when a participant walks a predetermined distance *and* has spent a predetermined time on the task without successfully detecting any PoI, the task finishes and the participant is rewarded.

Logging crowd activities. To improve the experience and capabilities of the Virtual City Explorer towards requesters and the crowd, we log all activities of the participants, including changes in position, pictures taken and discarded, and the time spent on each task. In the current implementation we do not store any data on tasks that the crowd abandon prematurely - this is meant to prevent malicious behaviour by some requesters, who sometimes make microtasks too difficult to complete, but take advantage of partial results without having to reward the crowd [Schmidt 2013].

3.3 Aggregator

The exploration task engine compiles a list of PoI geocoordinates submitted by each participant, using the triangulation method explained earlier. Next, individual submissions need to be consolidated. The same PoI can be detected by multiple workers through pictures shot from different angles, meaning that reported coordinates are close to each other, but not exactly the same. The aggregator clusters these points to determine which of them are likely to belong to the same PoI.

Clustering method. For this purpose, we use the DBSCAN algorithm [Schubert et al. 2017]. DBSCAN requires two parameters: the minimum distance between points in a cluster and their neighbours (*Eps*); and the minimum number of points within *Eps* for a set of points to be considered a cluster (*MinPts*). In a crowdsourcing context, *MinPts* can also be interpreted as the number of independent submissions by the crowd required to validate an answer and include it in the final result. *Eps* should be set empirically, depending on the nature of the physical characteristics of the PoI, as wider ones (e.g., buildings) enable workers to report locations that are farther apart than smaller ones (e.g., trees). In our case, we used a minimum distance of 10 meters, which we determined in an early pilot. *MinPts* is defined by the requester in the configuration step.

We now turn to the studies we carried out to evaluate the VCE and pursue the research questions listed in Section 1.

4 STUDY DESIGN

In this section, we describe the design of our study, aimed at answering research questions $RQ1$, $RQ2$, $RQ3$, $RQ4$ and $RQ5$ to assess the feasibility, coverage, completeness, generation cost and time and crowd experience of the VCE, and compare it against other VGI map generation methods.

4.1 Data

Data: points of interest. While the VCE can be used to find any type of PoI, our study focused on collecting the location of *bike racks*. This was motivated by the local government team we partnered with in the Qrowd project. Cities need to have accurate, up-to-date information about their cycling infrastructure due to the increasing interest in promoting more sustainable and green modes of transport. Many bike racks are operated by private providers and local authorities cannot afford to survey the city to catalogue them. From a crowdsourcing perspective, bike racks are challenging to detect due to their various shapes and sizes, and as they can be easily confused to other objects people use to lock their bikes on, such as poles or fences. This also means that their detection cannot be automated through machine learning without extensive training data.

Data: areas of interest. The *Zona a Traffico Limitato (ZTL)* - Limited Traffic Zone - of the city of Trento, Italy, which covers an area of 0.347 km^2, shaped as an irregular polygon. We took the polygon's definition from the open data portal of the city of Trento. The second area is a polygon from downtown Nantes, France that we created manually with approximately the same area size of the Trento ZTL.

Data: comparison datasets. We considered a combination of open government data published by the city authorities and existing crowdsourced data. In Trento, bike rack locations are collected by local government employees for the ZTL, and are published openly at the Trento Municipality website[6]. We also considered bike racks contributed by volunteers on OpenStreetMap inside the ZTL. We refer to both datasets as Trento ZTL and Trento OSM respectively. For Nantes, bike racks were collected in a *Cartopartie* (Mapping Party) organised by the Nantes municipality in collaboration with the local OSM chapter in the context of the *Nantes Digital Week* of 2015[7]. The activity lasted one day. Previous to the event, organisers partitioned the city in 78 zones that were estimated to take between 60 and 90 minutes to cover. Incoming volunteers were invited to choose a zone not covered by a previous volunteer and provided with a paper copy of its map. The task was to mark on the paper where they found a bike rack, report any segment of street that couldn't be covered and take the paper back to the command centre. At the end of the day, the organising team consolidated the results and published them both on OSM and the Nantes open data portal. Each year after the event (up to 2017) the Municipality merged any updates available on OSM into their copy of the dataset. We took the last version of the dataset published on the Nantes open data portal[8]. We refer to this dataset as *Nantes-Met*.

4.2 Methods

4.2.1 Overview. We carried out an experiment using the FigureEight platform to recruit participants. We logged crowd activities and results as explained in Section 3 and undertook a quantitative analysis based on a series of pre-defined metrics. We compared the completeness of the aggregated PoI data from the VCE against the Trento ZTL and Trento OSM datasets..

We will first describe the VCE parameters that were used to initialise the microtasks, followed by the metrics that guided the quantitative evaluation.

4.2.2 Crowdsourcing parameters. The parameters are listed in Table 1. As discussed earlier, each participant started their exploration at a random point on the map. This is captured by the *StartPoint* function. We generated one exploratory task for the ZTL area of interest described in Section 4.1. Each participant was asked to locate 5 bike racks for each task execution, being rewarded with $0.25 after successful completion. We chose the reward amount based on the market price on the FigureEight platform, and inline with ethical research guidelines of the University of Southampton. Participants could complete more than one task. We defer to future work the study of varying *numInstances* values and the exploration of different reward schemes and starting points functions.

Both δ and *Eps* were set to 10 meters, based on preliminary evidence collected in an early pilot. Bike racks are sometimes single-slot small ones, or fairly wide, accommodating tens of bikes. They can also be relatively close to each other (e.g. on opposite sides of the same street). Setting these parameters require knowledge of the expected dimensions of the PoI. For *minPts* we chose a value of 3. Remember this value corresponds to the number of independent detections of a point to consider it a PoI. The *escapeDistance* was set as 20% of the explorable distance of Trento, while

---

Table 1. Crowdsourcing parameters set by the requester in the VCE and the values used in the first study

Parameter	Description	Value
StartPoint	Function to assign starting points to workers	*Random*
numInstances	No. of PoIs to submit per task	5
Reward	Reward after task completion	$0.25
δ	Valid distance from triangle centroid	$10m$
Eps	Distance between points considered the same	$10m$
Minpts	Cluster size for DBSCAN	3
numExecutions	No. of task executions	150
escapeDistance	Minimum distance walked without detecting a PoI to escape	$1800m$
escapeTime	Time without detecting a PoI to escape	3 minutes

the *escapeTime* was informed by the average time that workers took to cover that distance, as measured in a preliminary study.

4.2.3   Metrics and quantitative analysis. We now introduce the metrics we used to evaluate research questions.

RQ1 - Feasibility. We assessed feasibility as a function of the *precision* of the map generated by the VCE. We manually checked the pictures in PoI submissions to partition the map in true positives (TPs) and false positives (FPs) and computed the standard precision formula $prec(map) = \frac{|TP \cap VCEMap|}{|VCEMap|}$. A low precision would mean the crowd is not effective telling apart bike racks from things that look like one, or that many workers found out how to cheat the system. We also evaluated the impact on precision of adding more workers to the task, as we hypothesize that as more PoIs become taboo, workers might resort to submit "close enough" things to get the reward.

RQ2 - Area coverage. The coverage of street-view services varies from one part of the world to another. To be able to make meaningful comparisons between tasks working on different geospatial areas, we introduce the notions of *explorable path* and *explorable distance*. For a given area of interest defined by the requester (see Section 3), the explorable path is the polyline comprised by the centres of the panoramic images on StreetView. The explorable distance is the length of the explorable path. Table 1 summarises the information on the Street View version of the two areas of interest. The explorable path of the Trento ZTL has 924 points for a explorable distance of $9127m$, while the explorable path of Nantes has 1177 nodes for $12104m$ of explorable distance. To compute the Street View coverage, we manually measured the length of the segments of street not covered by Street View using the Google measurement tool.

Figure 4 (left) shows the explorable path of Trento ZTL with points as green circles. Red lines represent segments of street not covered by Street View. Almost all Trento images date from 2017 onwards. On the contrary, most Nantes' images date of 2014.

For each worker W_i, we logged the nodes visited in the explorable path $Vis(W_i)$. To measure the coverage of an area $Cover(A)$ we divided the number of nodes visited by at least one worker by the number of total nodes, i.e., $Cover(A) = \frac{|\bigcup_{W_i \in Workers} Vis(W_i)|}{N_A} \times 100\%$. We also plotted a heatmap of visited nodes to help spot any behavioural patterns.

RQ3 - Completeness. We assessed completeness according to four information retrieval metrics and inspected all errors manually to classify them. The metrics assume that A is the map generated by one of the benchmarks and B by our approach:

Table 2. Areas of interest used in the experiment

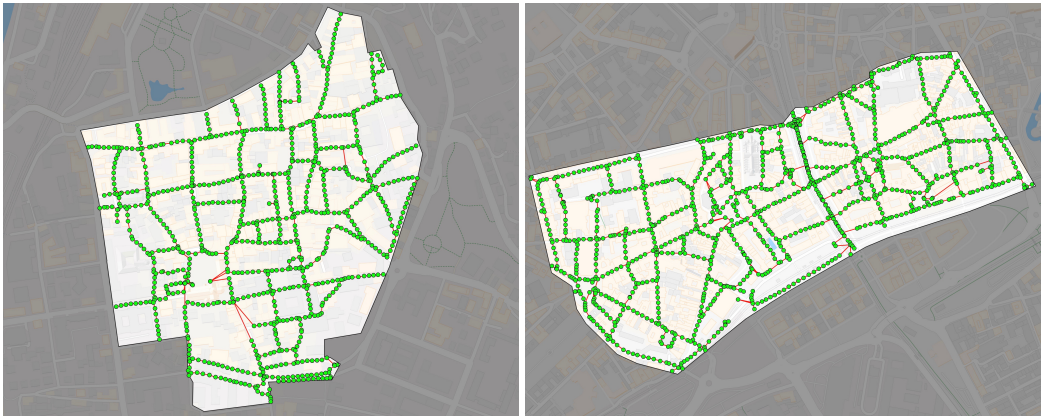|	**Trento - ZTL**	**Nantes**
Area (A)	$0.347 km^2$	$0.336 km^2$
# nodes in explorable path (N_A)	906	1177
Explorable distance ($Dist_A$)	$9127m$	$12104m$
Google StreetView coverage	93%	92%

Fig. 4. Areas of interest used in both of our studies: Trento's ZTL (left) and Nantes (right). Green dots represent the explorable path, red lines, segments not covered by Google StreetView.

The cardinality of the intersection ($|A \cap B|$) shows the number of PoIs in the benchmark that are also detected by the VCE.

The cardinality of the set difference ($|A - B|$) shows the number of PoIs in the benchmark the VCE was not able to discover. We manually classify them into: (1) *Absent*, the bike rack was not at sight in the Street View images corresponding to the coordinates in the benchmark ; (2) *Unreachable*, when the bike rack was located inside a courtyard or building not accessible by Google StreetView; (3) *Missed (Clustering)* when the crowd identified a bike rack that appears on Street View, but the coordinates computed by the VCE placed it more than 10 meters away from the coordinates in the benchmark; (4) *Triangulation impossible*, when the bike rack was in a spot where there were fewer than three images available on Google StreetView, making it impossible for workers to submit it. and (5) *Missed*, when G a bike rack at the location specified by the benchmark appears on Google StreetView, but the crowd did not detect it.

A large number of errors of type 1 and 2 would suggest a limitation of Street View for mapping bike racks. A large number of errors of type 3 would suggest that the *Eps* parameter is not well tuned. A large number of type 4 errors would suggest that there are too many objects located in places where is impossible for workers to shoot three pictures from different angles, uncovering a limitation in our strategy for calculating coordinates. Finally, a large number of type 5 errors would suggest that either not enough workers were assigned to the task or, if in combination with many crowd errors, that the crowd is not effective to find PoIs with our system.

The cardinality of the set difference ($|B − A|$) shows the number of possible false positives reported by our approach. we manually classified them as:

(1) *Crowd errors*, when they did not contain a bike rack, as in the feasibility method; and

(2) *Potential new bike racks*, when they did show a bike rack, meaning that the discrepancy is due either to the Google StreetView image being outdated, *i.e.*, the bike rack detected by the VCE crowd does not exist any more, or due to the benchmark dataset missing it.

To provide further insight, we compared the Google StreetView timestamp of the image were they appeared with the timestamp of the latest contribution in the gold standard. A later timestamp in our approach would suggest that the bike rack was missed by the other method, while the converse would suggest that the bike rack was removed at a later date than the street-view image timestamp.

The Jaccard similarity ($J(A, B)$) measures set similarity and diversity. We use it to quantify the similarity between bike rack maps. Jaccard is more informative than mere intersection, as it also takes into account the size differences between sets

RQ4 - Generation cost and time. We computed the cost of the VCE as the amount of money paid to crowdworkers that complete the task or trigger the fairness condition (aka *escape*), this is ($numExecutions \times Reward$) + $CrowdPlatformFee$.

The generation time was measured as the time elapsed between publication of the task on FigureEight until all *numExecutions* were completed. We compared cost and generation time of the VCE with those of the benchmark datasets. For Trento Open Street Map, we assumed zero cost. The generation time was computed as the time elapsed between the oldest and the newest contributions to the map. For the Trento municipality map we contacted the municipality directly. The map was generated in one working day by one employee of the municipality. The cost of the map was estimated to be the daily wage of said employee.

For Nantes, the map was generated in one day. We do not have specific cost information, as the event was organised in the context of the larger *Nantes Digital Week*. However, the official account of the event[9] reveals that the municipality contributed with the premises that served as *command center* to give instructions to volunteers and aggregate results, the printing of any support materials required by volunteers. The municipality also invested time of its geomatics department to produce the maps of the areas to be explored by volunteers.

RQ5 - Crowd experience. . As explained in section 3, the worker interface warns workers when they commit any of the following errors (a) Step out of the area of interest (b) Submit a PoI previously submitted by themselves (c) Submit a PoI marked as taboo (d) Submit three pictures that fail the triangulation check. We counted and classified the number of errors each worked committed during a task execution in both areas of interest and computed a histogram of error occurrences. A large number of errors would suggest that workers struggle with an aspect of the task.

For each of distance and time we measured the median and average of all workers, comparing workers that completed the task against those that escaped. For each escaping worker, each time the escape condition was triggered, we measured the time spent, the distance walked and the number of PoIs submitted up to that point. If the majority of workers spend more time than the set condition before walking the distance needed to trigger escape, or conversely, if they walk a much longer distance than the set condition before reaching the time limit, that would suggest these parameters need to be adjusted.

4.2.4 Summary of study design. The study explored the following research questions:

---

[9]https://wiki.openstreetmap.org/wiki/Nantes/Evenements/Cartopartie_V%C3%A9lo_19_Septembre_2015 (in French)

Table 3. Summary of methods, data and metrics used for the study

RQ	Theme	Analysis Method	Data	Metric
RQ1	Feasibility	Quantitative	VCE map	Precision as task progresses
RQ2	Completeness	Quantitative	VCE and benchmark maps	Overlap with benchmark datasets. Manual classification of reasons for missing bike racks.
RQ3	Coverage	Quantitative	Workers' log	% visited nodes of explorable path.
RQ4	Cost, time	Quantitative	Cost and time.	Cost and time to generate.
RQ5	Crowd user experience	Quantitative	Workers' log	# interface errors triggered, median time spent, median distance walked.

RQ1: Feasibility Can the crowd identify PoIs effectively using the VCE? We quantify this as how precise the VCE map is

RQ2: Completeness Can the crowd locate all PoIs in an area? We compare the VCE with maps generated with other approaches

RQ3: Area coverage Does the crowd cover the entire area of interest in their free exploration?

RQ4: Generation cost and time How the VCE cost and time compares to other VGI methods?

RQ5: Crowd Experience : Did workers had issues with it? How often they triggered the escape conditions? Is there any difference between workers that complete the task and workers that trigger the escape condition? Which escape condition was triggered first?

Data and methods are summarised on table 3.

5 RESULTS

5.1 *RQ1 - Feasibility*

For Trento, the VCE generated a map with 52 PoIs, 13 of them false positives. For Nantes, the VCE map returned 54 PoIs, with only 2 false positives. Figure 5 shows the cumulative precision as workers complete the task. Precision is perfect in both areas up to the 65th worker, from that point on, we observe a steady decrease in Trento, down to 0.75 after the 150th worker. In Nantes, perfect precision maintains up to the 90th worker, and the decrease is not as sharp as in Trento, being 0.95 after the 150th worker. These results suggest workers are very effective in detecting bike racks, and that our hypothesis holds: as the task progresses, the number of taboo items that workers encounter affects precision. As the bike rack density is lower in Trento than in Nantes, we observe the decrease on precision sooner.

5.2 *RQ2 - Completeness*

Table 6 shows the pairwise comparison between the map generated by the VCE and the benchmark maps for each area of interest, for each metric in section 4. For the sake of completeness, we also include the comparison between Trento's municipality and OSM maps. Overall, VCE found 52 bike racks. With respect to the Municipality map, VCE found 30 of 39 bike racks, missed 9, and added 22 potential new bike racks. With respect to the OSM map, VCE found 26 out of 54 bike racks, missed 28, and added 26 new PoIs. The low overlap (Jaccard distance) between all maps

Fig. 5. Precision of the confirmed bike racks during tasks progression

Table 4. Potential new bike racks and crowd errors wrt reference datasets

	Trento		Nantes
Type	Municipality	OSM	Nantes-Met
$|B - A|$	22	26	19
Potential new bike racks	9 (40.9%)	14 (53.8%)	17 (89.5%)
Crowd Errors	13 (59.1%)	12 (46.2%)	2 (10.5%)

Table 5. Failure reason analysis of wrong submissions and missing items wrt benchmarks

	Trento		Nantes
Failure Reason	Municipality	OSM	Nantes-Met
Absent from Street View	4 (44.4%)	13 (46.4%)	39 (61.9%)
Unreachable	2 (22.3%)	7 (25.0%)	5 (7.9%)
Missed (clustering reasons)	3 (33.2%)	4 (14.3%)	0
Triangulation impossible	0	2 (7.1%)	0
Missed	0	2 (7.1%)	19 (30.2%)
All	9	28	63

supports our assumption that approaches are complementary. For Nantes, VCE found 54 total PoIs. Compared to Nantes-Met, VCE discovered 35 out of 98, 63 were missed and 19 potential new bike racks were added.

Table 4 shows the analysis of the $B - A$ set. With respect to Trento datasets, approximately half of the set represents potential new bike racks, whilst with respect to Nantes-Met, almost 90% represent potential new bike racks. Results suggest that the VCE provides a good complement to other approaches as long as the precision has not started decreasing as we observed in Figure 5.

Table 5 shows the error types for bike racks missing in the VCE that are present in the reference datasets. With respect to both Trento datasets, the large majority of misses were caused by Street View absence and unreachability. Only two misses of the combined 66 bike racks of OSM and Municipality ($|A \cup B|$ column, third row)can be blamed exclusively on the Crowd. In Nantes, we observe more than 60% of misses due to absence, probably due to the fact that most Nantes' images are one year or more older than the collection date of Nantes-Met. There is also 30% of crowd misses in Nantes, suggesting that more workers should have been recruited.

Figure 6 shows the progression of PoI detections and confirmations as workers complete or escape the task for Trento (top) and Nantes (bottom). Each chart shows the overall number of
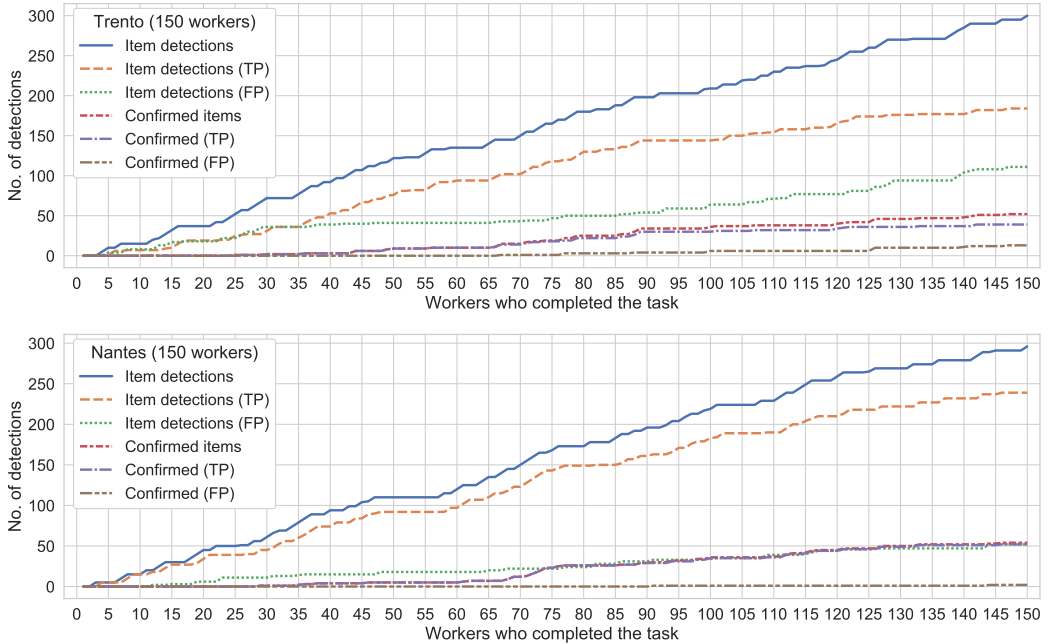
Fig. 6. PoI detections and confirmations as task progresses in Trento (top) and Nantes (bottom)

Table 6. Pair-wise comparison of the overlap and similarity between bike rack maps

Map A	Map B	$|A|$	$|B|$	$|A \cup B|$	$|A \cap B|$	$|A - B|$	$|B - A|$	$J(A, B)$
Trento								
Municipality	Crowd	39	52	61	30 (49.18%)	9 (14.75%)	22 (36.07%)	0.49
OpenStreetMap	Crowd	54	52	80	26 (32.5%)	28 (35.0%)	26 (32.5%)	0.32
OpenStreetMap	Municipality	54	39	66	27 (41%)	27 (41%)	12 (18%)	0.41
Nantes								
Nantes-Met	Crowd	98	54	117	35 (29.91%)	63 (53.85%)	19 (16.24%)	0.3

detections and confirmed items broken down by true positives (TP), *i.e.*, a bike rack, and false positives (FP). *i.e.*, not a bike rack. In both areas, we observe that detections follow a *staircase* pattern. Each step of the stair means that a number of workers (proportional to the length of the step) escaped without submitting any PoI. We observe that there is no concentration of escapes at later stages of the task, suggesting that the escape condition might have been set to a too low value. We further explore this in section 5.5.

We also observe that the rate of FPs submitted in both cities increases as the task progresses, but there are twice as more in Trento than in Nantes. In combination with the precision results in section 5.1, this suggests that precision might be an estimator for answering the question *should we recruit more workers?* In other words, when precision decreases, it might be that most or all bike racks have been found.

Fig. 7. Area coverage of Trento (left) and Nantes (right)

Table 7. Overall comparison of bike-rack maps generated with different approaches for Trento

	Municipality	OSM	VCE
Cost	€60	Free	$45
# Bike racks	39	54	52
Time to generate	Approx. 4 hours	62 months	Approx. 7 days
Generation date	Jun. 2017	2010-2015	2018
Collected by	1 expert	2 volunteers	150 crowdworkers

5.3 *RQ3* - Area Coverage

Figure 7 shows the heatmap of the crowd exploration on both areas of interest. Black circles represent points that were not visited. For Trento, 96% percent of the nodes were visited. Most of the exploration happened on the center of the area. Unexplored points are concentrated in three segments of street at the north of the area, possibly lanes were is not evident for workers that they can go in. For Nantes, 97% of the explorable path was visited at least by one worker. As in Trento, there are two full segments of street that were not visited at all. To attain full coverage, it might be needed to replace the *startPoint* function to drop incoming workers on streets that have not been visited.

5.4 RQ4 - Cost and time

5.4.1   Basic figures. Tables 7 and 8 compare the bike-rack maps available for each area against the ones generated through the VCE in terms of costs, number of bike racks, and time to collect. For Trento, the city's dataset was the fastest to generate, but for a premium price. The costs of the VCE map were computed as follows: 150 participants times $0.25 plus the 20% FigureEight fee, leading to a total of $45. The OpenStreetMap dataset was collected by two volunteers, but it required in excess of 5 years to be generated, with the last update dated October 2015. Nantes-Met was generated by 50 volunteers during one day. We do not have information about the specific operational costs assumed by the municipality, but as described in section 4.1, at least premises for a command center and advertisement costs need to be considered.

Table 8. Cost and generation time comparison between VCE and NantesCP

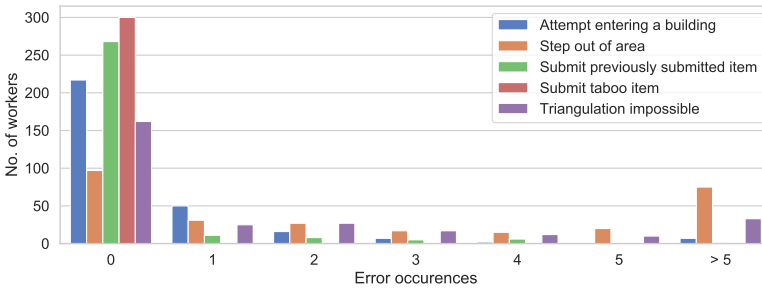|	Nantes-Met	VCE
Cost	Unknown	$45
# Bike racks	98	54
Time to generate	1 day	Approx. 7 days
Generation date	2015-2017	2018
Collected by	50 Volunteers	150 crowdworkers

Fig. 8. Number of workers that triggered 0 to >5 interface errors.

5.5 RQ5 - Crowd experience

Figure 8 shows an histogram of the number and type of errors committed by the crowd. Most workers were able to complete the task without triggering any error, or only once. The most common errors are stepping out of the area and failing the triangulation check, suggesting two areas of improvement to the interface. We also note that very few workers attempted to submit previously detected or taboo PoIs, suggesting they are honest in trying to fulfill the task. Furthermore, we looked at the difference between workers who completed the task vs. workers who escaped and found no significant differences in numbers or type or error made.

Figure 9 (top) compares for both areas of interest the distance walked by workers that escaped against those that completed the task. There is no significant difference, with both medians around the 1500m line. Figure 9 (bottom) compares the time spent by workers that escaped against those that completed the task. The median is around 3 minutes, and we do not observe significant difference between both sets of workers.

However, we note that almost two third of the workers triggered the escape condition in both areas. To provide further insight about which of the parameters was triggered first, we plotted the time spent and the distance walked of all workers that escaped the task. Figure 10 shows the results. The vertical dotted line represents the escape distance, while the horizontal dotted line represents the escape time. The concentration of points along the escape distance line indicates that most workers spent significantly more than 3 minutes without detecting any new bike rack before reaching the escape distance. Intuitively, we expected all points to be very close to either the escape time line or the escape distance line, however, there are some outliers that are not close to any of the lines. This is explained by the fact that we check the escape conditions only after the worker makes an action in the interface. It is then possible that a worker who has already reached the time condition, for a while, triggers the distance condition by taking a 200 meters step.
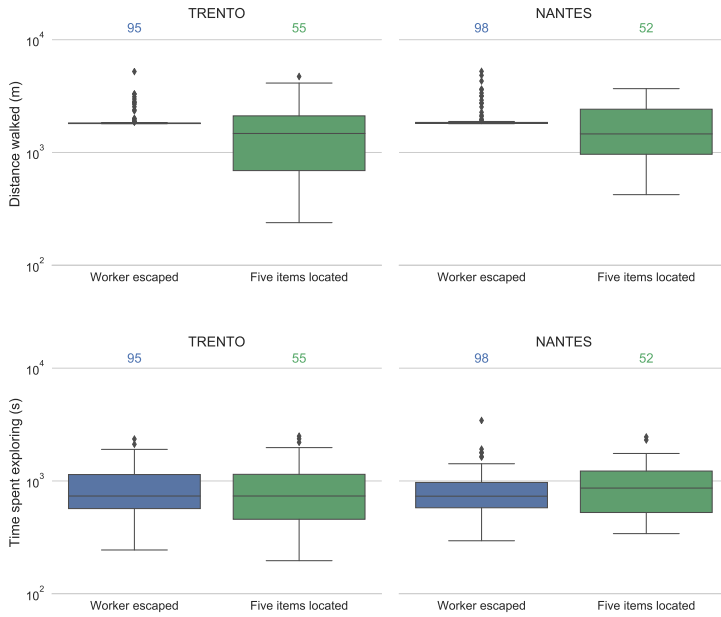
Fig. 9. Distance walked (top) and time spent (bottom) for workers that escaped and completed the task

We note that most of the workers triggered the escape condition before submitting any bike rack. This is consistent with the results of Figure 10, where data is concentrated around the *escapeDistance* value. This is a relative setback for the VCE, as we expected workers to escape because they walked more than workers that completed the task, and is probably one of the reasons behind the incompleteness of the Nantes results with respect to the reference benchmark as described in section 5.2. This suggests that to improve results, we should aim at reducing the escape rate instead of recruiting more workers.

5.6   Summary of results

Feasibility (RQ1): The crowd reached up to 100% precision up to the point where some of the bike racks became taboo. Precision decreases as workers start encountering PoIs that have already been retired through the taboo feature - suggesting they prefer to submit a false positive rather than make no submission at all. Overall, this suggests that the VCE allows workers to undertake their task well and hence that the approach is feasible. The main challenge is related to the quality of the Street View data, which needs to be considered when deciding what methods to use in an urban auditing project.

   Completeness (RQ2): VCE results overlapped with datasets produced with other approaches. In some cases, bike racks featured in other approaches could not be confirmed by the crowd. Upon inspection, this was again related to limitations of Street View - either the bike rack was not available on the map or it could not have been reached through virtual exploration. In Trento, the crowd missed few PoIs (5%), but in Nantes that figure was higher (30%), suggesting that 150 workers was enough for Trento, but insufficient for Nantes. At the same time, all PoIs accessible via Street-View imagery were found by the crowd.
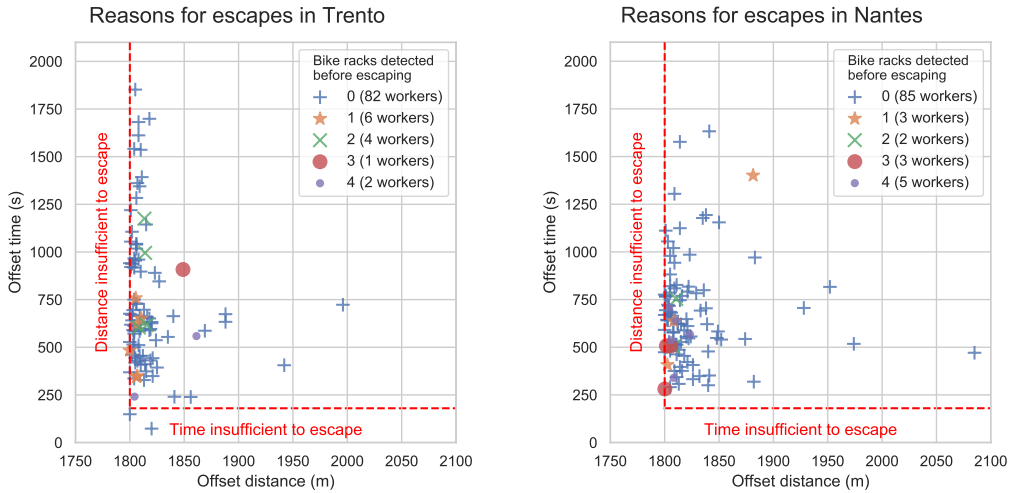
Fig. 10. Comparison of escape sub-conditions. Most escaping workers first triggered the 3 minutes time condition before the 1800m distance condition

Area coverage (RQ3): The coverage of both experimental areas was over 90%. The few uncovered segments of street could be easily handled by prioritizing them as starting positions for additional workers.

Cost and generation time (RQ4): For Trento, our approach was slightly less expensive than a surveyor on the ground, but took longer. While urban auditing does not have a time-sensitive element (hence, delivering results in a matter of days or weeks is acceptable), surveying large physical areas with traditional methods is often not possible for city administrations because they do not have the staff to do it. Compared to OpenStreeetMap, our approach clearly has its costs, but provides a greater degree of control, and is therefore faster. We did not have any information on the effort required to deliver the volunteer project in Nantes, but literature suggests it would be more expensive than ours [Handy and Mook 2011] and could not be repeated indefinitely at the same level of engagement [Hyde et al. 2016]. Their initiative was time-bound to one day, whereas we did not impose any time constraints when recruiting.

Crowd experience (RQ5): Most workers completed the task without triggering any error, suggesting that they did not struggle at all with the VCE interface. The most common error was to step out of the area of interest, which was considered in our design. The median time of exploration was around 3 minutes and the median distance explored was 1.5 km. This is clearly much faster than a surveyor on the ground and all delays incurred were related to the arrival rate of the workers on FigureEight. We noted a high rate of escapes (65%), which can be tackled by choosing a different distance parameter and/or different strategies for allocating starting points.

6 DISCUSSION

6.1 Setting the VCE parameters

Our aim was to give requesters control over the way they design their crowdsourced urban-auditing projects. Our experiments delivered additional insights into how the main parameters should be set to achieve specific results.

Minpts This parameter will have financial implications as more submissions will lead to additional costs. Our approach had a very low rate of false positives. Crowdworkers made mistakes individually, but these were successfully identified through our aggregation method. The results suggest that 3 crowdworkers are enough to identify the same PoIs based on the images they submit.

Delta and Eps These two parameters affect what PoIs will be submitted and how they will be clustered together during aggregation. As shown in Table 5 a fair number of bike racks were missed in this step. Our experiments use a type of PoI that is not without its challenges - bike racks often come in groups of different widths, therefore requesters are encouraged to try out different levels that match their local context. The lower the values of the parameters, the higher the risk to have bike racks identified more than once. Conversely, the higher the values, the more likely it will be that PoIs in close proximity will be considered as one. Setting the parameters could be done empirically, but it also depends on how the consolidated data will be used. In our case, our partner from the City of Trento, Italy was more interested in mapping the areas where bike racks were located rather than distinguishing between adjacent racks or their overall capacity. The latter could also be handled in a separate microtask, which would start with the location of the racks and then ask the crowd to count how many bikes would they accommodate. The other option we are exploring for future work is to add a bounding box question as a follow up to a valid submission to have more granular data on the position and possibly shape of the PoIs.

Escape distance and time In Section 5.5 we noted that the escape rate was quite high, with results suggesting that the time condition was set too low. As the main purpose of the escape strategy is to be fair with crowdworkers who arrive late to the task, we are considering several strategies to balance fairness with the amount of useful work that the workers submit. For example, one could set the time and distance dynamically, starting with very high values earlier in the experiment, and reducing them based on a combination of area coverage and confirmed PoIs. Alternatively, one could change the reward model, as we detail below.

Reward and numInstances In the experiments of this paper we rewarded crowdworkers with a fixed amount per batch of 5 PoIs located. Bulk payment has been reported as attracting better response rates than per task payments, coupons or material goods [Ikeda and Bernstein 2016]. However, other reward schemes may increase the number of contributions per worker, for example, paying bonuses for achieving pre-defined milestones as suggested in [Difallah et al. 2014]. In the context of the VCE, bonuses could be awarded for locating 10 or 20 PoIs. A further idea is the introduction of gamification on top of financial incentives, that may increase worker engagement and accuracy as detailed in [Feyisetan et al. 2015] for the case of image labeling. We believe that the PoI location task in the VCE is a good candidate for certain gamification elements, *e.g.*, introducing badges and leaderboards for achievements that award monetary bonuses, distance walked, venture into previously unexplored areas, discover the most PoIs in the less time, etc. However, some gamification elements, like contests, would require the adaptation of the VCE to host multiple workers exploring concurrently [Feyisetan and Simperl 2019].

By allowing workers to 'escape', and hence aiming to be fair towards them and the effort they invest, we accept that we might lose some partial results. A baseline reward model could be investigated to mitigate this, for trustworthy contributors. Alternatively, one could offer additional, related work to those crowdworkers who arrived too late to the task - previously we have spoken about follow-up questions about rack capacity or shape. *numInstances* needs to be adjusted according to the expected number of PoIs, or even dynamically, according to the intermediary coverage levels achieved, which the VCE monitors in real time.

StartPoint Our choice of random starting points showed a very good coverage of the area. However, we observed that a small number of street segments were not visited, probably because it was not evident for workers that they could traverse them. This could be solved by assigning

starting points dynamically depending on the current area coverage or to prioritise areas with PoIs that are not yet confirmed.

6.2 Scaling up

Cost and time figures needs to be understood in context. Currently, data collection by local government is patchy, due to pressure to reduce costs. Most urban auditing efforts focus hence on a small area of interest and rely on VGI projects such as OpenStreetMap to fill in the gaps. The VCE is not designed to replace volunteer initiatives, but as a tool that allows city authorities to collect data in a targeted, controlled way, which can be configured to take into account specific areas of interest, budgets, and quality requirements. In addition, when using the VCE, a requester can specify how many independent detections are required to confirm a PoI; most official surveys and VGI initiatives offer looser guarantees and cannot be carried out upon request within short periods of time. Our understanding of the total effort required to set up and manage a VGI, or any community-based initiative is still very limited. While participation of a critical mass of volunteers reduces costs, resources need to be invested in supporting tools and processes, including apps, online portals, quality controllers, marketing or prizes. The VCE cannot claim it will always deliver competitive results in terms of costs and time, but it is faster and more agile than traditional methods, and is more transparent in terms of costs than working with volunteers [Handy and Mook 2011].

6.3 Complementarity to other approaches

Our analysis showed that the quasi gold standards we used to gauge completeness had their limitations. The VCE was able to find many potential new bike racks, though it also missed some of the ones that were listed in the other datasets. As mentioned earlier, an online approach depends on the quality of the Street View data. However, in areas where coverage is high, requesters have a choice of data collection options depending on available resources, budget and time constraints. The added value of a paid microtask approach using Street View or similar lies in its accessibility and the level of control it offers to a requester. Collecting data can be done at a competitive price, in time frames that fit the needs of urban auditing methodologies, at the push of a button. Ultimately, the VCE advocates an iterative, hybrid approach where coverage and completeness concerns are managed through a combination of additional microtasks and, where possible, targeted work carried out on the ground by designated surveyors or active volunteers. A requester could, for instance, first commission a pilot survey to get a sense of the data, how well it compares with Street View, then decide how to set the VCE parameters, and then undertake a large scale online project where performance can be improved on demand across several iterations. Offline, generally costlier resources could be used in a more targeted way to cover areas where online auditing fails, either because the data is likely to be outdated, or because the PoIs are not accessible through remote exploration. The analyses provided by the VCE, which are illustrated in the two experiments described in the paper, are available to requesters to support them throughout the process.

A possible direction is to evolve the VCE to support *crowd-augmented expert work*, a model where experts are simultaneously performing the same (or a superset) of tasks as the crowd, and their own work is shaped and redirected based on the crowd results. This model has been recently applied with success to the task of geolocating images using the concept of shared representations [Venkatagiri et al. 2019]. One could also envision a shared representation between a remote crowdworker and a local volunteer engaged through a spatial crowdsourcing application.

6.4 Generalizability

Generalizing to other geographical contexts: We observed a slight variability in completeness between Trento and Nantes, suggesting that the number of recruited workers needs to be tuned

according to the expected density of PoIs in the area of interest (c.f. section 5.2), and that the escape rate should be reduced (c.f. section 5.5). Future work should explore strategies for optimizing recruitment and escape conditions based on the current state of the map. Our results suggest that a decrease in workers' precision as the task progresses could indicate that most PoIs in the area have already been discovered, providing an heuristic to stop recruitment.

When using the VCE on a different geographical context, a critical limitation is the availability and freshness of street-view imagery. Today, street-view imagery is mostly provided by commercial companies who might decide on what areas to cover and refresh according to the demands of their customers, *i.e.*, more up-to-date images around touristic sites or locations often asked in search engines. In developing cities and rural areas, where access, security, or plain bias affects the availability of street-view imagery, mobile crowdsourcing approaches have been developed to engage contributors from the same communities to perform the audits [Ruiz-Correa et al. 2017]. This approach enables the creation of a dataset of images and videos focused in problems and areas relevant for the contributors. Microtask crowdsourcing has been used to collect further annotations and judgements about the state of urban infrastructure [Santani et al. 2015]. Such an approach could be adapted to engage contributors to create the required street-view imagery by themselves, feeding broader images of the community (instead of focused on specific problems) to a neutral platform like Mapillary[10]. The VCE could then be run to audit multiple dimensions, either with crowdworkers, and/or members of the community.

Dealing with contributors' bias: Contributors' bias has been extensively studied in VGI platforms. A natural theme for discussion is how some of these biases may also affect the VCE: (1) *Self-focus bias* appears when contributors focus disproportionately on information that are particularly relevant to cultural groups in the peer production community. Biases in type of contribution depending on gender have been reported for OSM [Das et al. 2019]. By design, the VCE constrains the choice of crowdworkers to the mapping of a single type of Point of Interest at a time, reducing the risk of self-focus bias. (2) *Geographic coverage bias* is when most contributors focus their contributions on specific areas. For example, analysis of OSM data suggests that contributors to USA counties undercover rural and high-poverty areas [Thebault-Spieker et al. 2018]. Strong coverage bias has also been reported when comparing power users versus casual users of OSM, with the map emerging from power users' activity being very dense around city centres, and near empty elsewhere, whereas the map constructed by casual users covers the whole country, as well as intra-urban areas, much more uniformly [Quattrone et al. 2015]. As pointed out in the previous discussion about generalizing to other geographical contexts, the main source of geographic coverage bias comes from the underlying street-imagery service. Beyond that, the area to be mapped is controlled by urban managers instead of chosen by crowdworkers. In the areas we used in our experiments we observed that coverage was near complete, but they are too small to draw definitive conclusions. A possible research direction is to characterize which parameters determine the appearance of geographic bias in the VCE, and if there is correlation with demographics and average-income as suggested by studies on OSM data. (3) *Cognitive bias* where contributors misinterpret input due to their own subjective social reality. This type of bias has been reported in the crowdsourcing literature for text annotation tasks [Eickhoff 2018]. We did not observe any evidence of cognitive bias in our study, but we believe it could appear depending on the type of PoI and the background of crowdworkers. For example, some types of bike racks are attached to existing lampposts[11], when workers that live in a city that does not use this type of

---

[10]https://www.mapillary.com/

[11]https://www.cyclehoop.com/product/cyclehoops/lamp-posts/

rack contribute in a city that does, they might miss them. To mitigate, requesters need to write instructions as clear and precise as possible.

A further generalization could be the opening of contributions to volunteers with a broad choice of areas and types of PoI to map, in a similar style to VGIs (yielding a *Virtual VGI*). We hypothesize that if the VCE is implemented in this context *as is*, coverage and self-focus biases would appear, just like on-the-field VGI. However, the online nature of the VCE offers interesting possibilities for designing and implementing interventions to mitigate biases, for example: suggest contributors a different starting point based on current map coverage; during exploration, suggest contributors to teleport to a less covered area or to look for an under-represented type of PoI; design gamified microtasks to reduce bias and promote inclusion.

7 CONCLUSIONS AND FUTURE WORK

In this paper, we have described the design, implementation, and evaluation of the Virtual City Explorer, a system to collect geo-spatial information through the generation of an exploration interface on top of digital-street imagery, through which paid crowdworkers can report the location of PoIs. The evaluation focused on the task of generating maps of PoIs within a bounded area.

We compared the VCE maps against maps generated by an expert surveyor and by volunteers (in Open Street Maps in Trento and a time-bounded campaign in Nantes, sponsored by the local authorities). The results showed that paid crowdworkers are effective at locating PoIs with the help of the VCE, and discover new bike racks that were not listed in the other datasets. The completeness of the maps generated through the VCE is nevertheless dependent on the quality of the digital-street imagery. Based on our results, we would recommend a mixed approach, where the VCE would help audit larger areas of interest, allowing public or volunteer resources to be targeted to those areas where the online data is outdated or the detection of PoIs in a virtual space proves challenging.

The experiments confirmed a tension between the exploratory nature of the task, which requires little to no upfront investment or work on the ground, and our ability to use the budget efficiently. The escape affordance we introduced to counterbalance the effects of the taboo mechanism, which ensures we do not spend resources on confirming easily accessible or otherwise popular PoIs, put treating workers fairly first. In future work, we are considering alternatives including: (1) Reward mechanisms that are adjusted during the experiment as our knowledge of the number, location and density of PoIs increases. (2) Adaptive starting positions that prioritise less explored areas to minimise the time a worker needs to invest to be able to complete a HIT, as well as interactive maps that show workers the most crowded areas. (3) Gamification strategies to be triggered when many concurrent workers are solving the task at the same time.

REFERENCES

L. von Ahn. 2006. Games with a purpose. *Computer* 39, 6 (June 2006), 92–94. https://doi.org/10.1109/MC.2006.196

Dragomir Anguelov, Carole Dulong, Daniel Filip, Christian Frueh, Stéphane Lafon, Richard Lyon, Abhijit Ogale, Luc Vincent, and Josh Weaver. 2010. Google Street View: Capturing the World at Street Level. *Computer* 43, 6 (June 2010), 32–38. https://doi.org/10.1109/MC.2010.170

Hannah M. Badland, Simon Opit, Karen Witten, Robin A. Kearns, and Suzanne Mavoa. 2010. Can Virtual Streetscape Audits Reliably Replace Physical Streetscape Audits? *Journal of Urban Health* 87, 6 (Dec. 2010), 1007–1016. https://doi.org/10.1007/s11524-010-9505-x

Irene Celino, Dario Cerizza, Simone Contessa, Marta Corubolo, Daniele DellAglio, Emanuele Della Valle, and Stefano Fumeo. 2012. Urbanopoly – A Social and Location-Based Game with a Purpose to Crowdsource Your Urban Data. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*. https://doi.org/10.1109/SocialCom-PASSAT.2012.138

Maitraye Das, Brent Hecht, and Darren Gergle. 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)*. ACM, New York, NY, USA, 563:1–563:14. https://doi.org/10.1145/3290605.3300793 event-place: Glasgow, Scotland Uk.

Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, and Philippe Cudré-Mauroux. 2014. Scaling-Up the Crowd: Micro-Task Pricing Schemes for Worker Retention and Latency Improvement. In *Second AAAI Conference on Human Computation and Crowdsourcing*. https://www.aaai.org/ocs/index.php/HCOMP/HCOMP14/paper/view/8958

Helen Dorn, Tobias Törnros, and Alexander Zipf. 2015. Quality Evaluation of VGI Using Authoritative Data—A Comparison with Land Use Data in Southern Germany. *ISPRS International Journal of Geo-Information* 4, 3 (Sept. 2015), 1657–1671. https://doi.org/10.3390/ijgi4031657

Carsten Eickhoff. 2018. Cognitive Biases in Crowdsourcing. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*. ACM, New York, NY, USA, 162–170. https://doi.org/10.1145/3159652.3159654 event-place: Marina Del Rey, CA, USA.

Enrique Estellés-Arolas and Fernando González-Ladrón-de Guevara. 2012. Towards an integrated crowdsourcing definition. *Journal of Information Science* 38, 2 (April 2012), 189–200. https://doi.org/10.1177/0165551512437638

Hongchao Fan, Alexander Zipf, Qing Fu, and Pascal Neis. 2014. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science* 28, 4 (April 2014), 700–719. https://doi.org/10.1080/13658816.2013.867495

Oluwaseyi Feyisetan and Elena Simperl. 2017. Social Incentives in Paid Collaborative Crowdsourcing. *ACM Trans. Intell. Syst. Technol.* 8, 6 (July 2017), 73:1–73:31. https://doi.org/10.1145/3078852

Oluwaseyi Feyisetan and Elena Simperl. 2019. Beyond Monetary Incentives: Experiments in Paid Microtask Contests. *ACM Transactions on Social Computing* 2, 2 (March 2019), 6. https://doi.org/10.1145/3321700

Oluwaseyi Feyisetan, Elena Simperl, Max Van Kleek, and Nigel Shadbolt. 2015. Improving Paid Microtasks Through Gamification and Adaptive Furtherance Incentives. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 333–343. https://doi.org/10.1145/2736277.2741639

Michael F. Goodchild, Pinde Fu, and Paul Rich. 2007. Sharing Geographic Information: An Assessment of the Geospatial One-Stop. *Annals of the Association of American Geographers* 97, 2 (June 2007), 250–266. https://doi.org/10.1111/j.1467-8306.2007.00534.x

Pippa Griew, Melvyn Hillsdon, Charlie Foster, Emma Coombes, Andy Jones, and Paul Wilkinson. 2013. Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *International Journal of Behavioral Nutrition and Physical Activity* 10, 1 (2013), 103. https://doi.org/10.1186/1479-5868-10-103

M. Haklay and P. Weber. 2008. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Computing* 7, 4 (Oct. 2008), 12–18. https://doi.org/10.1109/MPRV.2008.80

Mordechai (Muki) Haklay, Sofia Basiouka, Vyron Antoniou, and Aamer Ather. 2010. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *The Cartographic Journal* 47, 4 (Nov. 2010), 315–322. https://doi.org/10.1179/000870410X12911304958827

Femida Handy and Laurie Mook. 2011. Volunteering and Volunteers: Benefit-Cost Analyses. *Research on Social Work Practice* 21, 4 (July 2011), 412–420. https://doi.org/10.1177/1049731510386625

Kotaro Hara, Jon E. Froehlich, Shiri Azenkot, Megan Campbell, Cynthia L. Bennett, Vicki Le, Sean Pannella, Robert Moore, Kelly Minckler, and Rochelle H. Ng. 2015. Improving Public Transit Accessibility for Blind Riders by Crowdsourcing Bus Stop Landmark Locations with Google Street View: An Extended Analysis. *ACM Transactions on Accessible Computing* 6, 2 (March 2015), 1–23. https://doi.org/10.1145/2717513

Kotaro Hara, Vicki Le, and Jon Froehlich. 2013. Combining crowdsourcing and google street view to identify street-level accessibility problems. ACM Press, 631–640. https://doi.org/10.1145/2470654.2470744

Darren Hardy, James Frew, and Michael F. Goodchild. 2012. Volunteered geographic information production as a spatial process. *International Journal of Geographical Information Science* 26, 7 (July 2012), 1191–1212. https://doi.org/10.1080/13658816.2011.629618

Francis Harvey. 2013. To Volunteer or to Contribute Locational Information? Towards Truth in Labeling for Crowdsourced Geographic Information. In *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, Daniel Sui, Sarah Elwood, and Michael Goodchild (Eds.). Springer Netherlands, Dordrecht, 31–42. https://doi.org/10.1007/978-94-007-4587-2_3

Brent J. Hecht and Darren Gergle. 2010. On the "Localness" of User-generated Content. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. ACM, New York, NY, USA, 229–232. https://doi.org/10.1145/1718918.1718962

Melissa K. Hyde, Jeff Dunn, Caitlin Bax, and Suzanne K. Chambers. 2016. Episodic Volunteering and Retention: An Integrated Theoretical Approach. *Nonprofit and Voluntary Sector Quarterly* 45, 1 (Feb. 2016), 45–63. https://doi.org/10.1177/0899764014558934

Kazushi Ikeda and Michael S. Bernstein. 2016. Pay It Backward: Per-Task Payments on Crowdsourcing Platforms Reduce Productivity. ACM, 4111–4121. https://doi.org/10.1145/2858036.2858327

Qijun Jiang, Frank Kresin, Arnold K. Bregt, Lammert Kooistra, Emma Pareschi, Edith van Putten, Hester Volten, and Joost Wesseling. 2016. Citizen Sensing for Improved Urban Environmental Monitoring. https://doi.org/10.1155/2016/5656245

Thivya Kandappu, Archan Misra, Shih-Fen Cheng, Randy Tandriansyah, and Hoong Chuin Lau. 2018. Obfuscation At-Source: Privacy in Context-Aware Mobile Crowd-Sourcing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1 (March 2018), 16:1–16:24. https://doi.org/10.1145/3191748

Gabriella Kazai, Jaap Kamps, Marijn Koolen, and Natasa Milic-Frayling. 2011. Crowdsourcing for Book Search Evaluation: Impact of Hit Design on Comparative System Ranking. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*. ACM, New York, NY, USA, 205–214. https://doi.org/10.1145/2009916.2009947 event-place: Beijing, China.

Andrew Mao, Ece Kamar, Yiling Chen, Eric Horvitz, Megan E. Schwamb, Chris J. Lintott, and Arfon M. Smith. 2013. Volunteering Versus Work for Pay: Incentives and Tradeoffs in Crowdsourcing. In *First AAAI Conference on Human Computation and Crowdsourcing*. https://www.aaai.org/ocs/index.php/HCOMP/HCOMP13/paper/view/7497

Peter Mooney and Padraig Corcoran. 2014. Analysis of Interaction and Co-editing Patterns amongst OpenStreetMap Contributors. *Transactions in GIS* 18, 5 (Oct. 2014), 633–659. https://doi.org/10.1111/tgis.12051

Sihang Qiu, Achilleas Psyllidis, Alessandro Bozzon, and Geert-Jan Houben. 2019. Crowd-Mapping Urban Objects from Street-Level Imagery. In *The World Wide Web Conference (WWW '19)*. ACM, New York, NY, USA, 1521–1531. https://doi.org/10.1145/3308558.3313651 event-place: San Francisco, CA, USA.

Giovanni Quattrone, Licia Capra, and Pasquale De Meo. 2015. There's No Such Thing as the Perfect Map: Quantifying Bias in Spatial Crowd-Sourcing Datasets. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1021–1032. https://doi.org/10.1145/2675133.2675235

Giovanni Quattrone, Martin Dittus, and Licia Capra. 2017. Work Always in Progress: Analysing Maintenance Practices in Spatial Crowd-sourced Datasets. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1876–1889. https://doi.org/10.1145/2998181.2998267

Alexander J. Quinn and Benjamin B. Bederson. 2011. Human Computation: A Survey and Taxonomy of a Growing Field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 1403–1412. https://doi.org/10.1145/1978942.1979148 event-place: Vancouver, BC, Canada.

S. Ruiz-Correa, D. Santani, B. Ramírez-Salazar, I. Ruiz-Correa, F. A. Rendón-Huerta, C. Olmos-Carrillo, B. C. Sandoval-Mexicano, Á H. Arcos-García, R. Hasimoto-Beltrán, and D. Gatica-Perez. 2017. SenseCityVity: Mobile Crowdsourcing, Urban Awareness, and Collective Action in Mexico. *IEEE Pervasive Computing* 16, 2 (April 2017), 44–53. https://doi.org/10.1109/MPRV.2017.32

Manaswi Saha, Kotaro Hara, Soheil Behnezhad, Anthony Li, Michael Saugstad, Hanuma Maddali, Sage Chen, and Jon E. Froehlich. 2017. A Pilot Deployment of an Online Tool for Large-Scale Virtual Auditing of Urban Accessibility. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM Press, 305–306. https://doi.org/10.1145/3132525.3134775

Darshan Santani, Salvador Ruiz-Correa, and Daniel Gatica-Perez. 2015. Looking at Cities in Mexico with Crowds. In *Proceedings of the 2015 Annual Symposium on Computing for Development (DEV '15)*. ACM, New York, NY, USA, 127–135. https://doi.org/10.1145/2830629.2830638 event-place: London, United Kingdom.

F. A. Schmidt. 2013. The Good, The Bad and the Ugly: Why Crowdsourcing Needs Ethics. In *2013 International Conference on Cloud and Green Computing*. 531–535. https://doi.org/10.1109/CGC.2013.89

Erich Schubert, Jörg Sander, Martin Ester, Hans Peter Kriegel, and Xiaowei Xu. 2017. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.* 42, 3 (July 2017), 19:1–19:21. https://doi.org/10.1145/3068335

Linda See, Peter Mooney, Giles Foody, Lucy Bastin, Alexis Comber, Jacinto Estima, Steffen Fritz, Norman Kerle, Bin Jiang, Mari Laakso, Hai-Ying Liu, Grega Milčinski, Matej Nikšič, Marco Painho, Andrea Pődör, Ana-Maria Olteanu-Raimond, and Martin Rutzinger. 2016. Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. *ISPRS International Journal of Geo-Information* 5, 5 (April 2016), 55. https://doi.org/10.3390/ijgi5050055

Daniel Sui, Sarah Elwood, and Michael Goodchild. 2012. *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice.* Springer Publishing Company, Incorporated.

Jacob Thebault-Spieker, Brent Hecht, and Loren Terveen. 2018. Geographic Biases Are 'Born, Not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork (GROUP '18).* ACM, New York, NY, USA, 71–82. https://doi.org/10.1145/3148330.3148350 event-place: Sanibel Island, Florida, USA.

Yongxin Tong, Libin Wang, Zimu Zhou, Bolin Ding, Lei Chen, Jieping Ye, and Ke Xu. 2017. Flexible Online Task Assignment in Real-time Spatial Data. *Proc. VLDB Endow.* 10, 11 (Aug. 2017), 1334–1345. https://doi.org/10.14778/3137628.3137643

Griet Vanwolleghem, Delfien Van Dyck, Fabian Ducheyne, Ilse De Bourdeaudhuij, and Greet Cardon. 2014. Assessing the environmental characteristics of cycling routes to school: a study on the reliability and validity of a Google Street View-based audit. *International Journal of Health Geographics* 13, 1 (2014), 19. https://doi.org/10.1186/1476-072X-13-19

Sukrit Venkatagiri, Jacob Thebault-Spieker, Rachel Kohler, John Purviance, Rifat Sabbir Mansur, and Kurt Luther. 2019. GroundTruth: Augmenting Expert Image Geolocation with Crowdsourcing and Shared Representations. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article Article 107 (Nov. 2019), 30 pages. https://doi.org/10.1145/3359209

Luis von Ahn and Laura Dabbish. 2004. Labeling Images with a Computer Game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04).* ACM, New York, NY, USA, 319–326. https://doi.org/10.1145/985692.985733 event-place: Vienna, Austria.

L. Wang, Z. Yu, Q. Han, B. Guo, and H. Xiong. 2018. Multi-Objective Optimization Based Allocation of Heterogeneous Spatial Crowdsourcing Tasks. *IEEE Transactions on Mobile Computing* 17, 7 (July 2018), 1637–1650. https://doi.org/10.1109/TMC.2017.2771259

Z. Xu, Y. Liu, N. Yen, L. Mei, X. Luo, X. Wei, and C. Hu. 2018. Crowdsourcing based Description of Urban Emergency Events using Social Media Big Data. *IEEE Transactions on Cloud Computing* (2018), 1–1. https://doi.org/10.1109/TCC.2016.2517638

Y. Zhao and Q. Han. 2016. Spatial crowdsourcing: current state and future directions. *IEEE Communications Magazine* 54, 7 (July 2016), 102–107. https://doi.org/10.1109/MCOM.2016.7509386