EUT EDIZIONI UNIVERSITÀ DI TRIESTE

Bollettino della ASSOCIAZIONE ITALIANA di CARTOGRAFIA 2020 (168), 94-106

ISSN 2282-572X (online) ISSN 0044-9733 (print)
http://www.openstarts.units.it/dspace/handle/10077/9933

DOI: 10.13137/2282-572X/32115

Gli studenti universitari sono capaci di produrre informazione geografica di qualità? Primi risultati di uno studio empirico

Can University Students Provide High-Quality Geoinformation? First Results of an Empirical Study

Gian Pietro Zaccomer, Giorgia Bressan

University of Udine, Udine (ITALY); gianpietro.zaccomer@uniud.it, giorgia.bressan@uniud.it

Riassunto

Al giorno d'oggi è molto comune utilizzare servizi basati su mappe online per pianificare percorsi di trasporto, monitorare attività sportive o cercare l'ubicazione dei negozi, tanto per fare alcuni esempi. Ma questo crescente utilizzo della cartografia interattiva per acquisire informazioni geografiche corrisponde anche a una maggiore capacità delle persone di produrre dati spaziali? Il presente contributo mira ad indagare la familiarità di un gruppo eterogeneo di persone con le mappe online. A tal fine, è stato condotto un esperimento che richiedeva ai partecipanti di identificare la posizione di specifici luoghi su mappe incorporate in un questionario online. L'obiettivo era quello di valutare se la loro conoscenza dei GIS influisse sulla qualità della loro attività di mappatura. L'esperimento ha coinvolto circa duecento studenti dell'Università di Udine ed è stato condotto in parte di persona e in parte a distanza. Gli studenti, iscritti a vari corsi universitari, si differenziano per avere diversi gradi di esposizione alla geografia e ai GIS nel loro percorso formativo. Nell'articolo, dopo una breve presentazione dell'esperimento, descriveremo innanzitutto le caratteristiche degli intervistati, come quelle demografiche, istruzione e competenze informatiche, quindi si analizzerà l'accuratezza della posizione dei dati spaziali che hanno fornito. Infine, si valuterà la partecipazione e la qualità dei dati in relazione alle caratteristiche dell'esperimento. I risultati rivelano che il metodo di supervisione e la formazione GIS sono due fattori importanti per la fornitura di informazioni geografiche di alta qualità.

Parole chiave

Geoquestionari, Capacità GIS, Studenti universitari, Qualità dei dati, Friuli Venezia Giulia

Abstract

It is extremely common nowadays to employ online map services to plan transportation routes, track sports activities, search the location of shops, to give some examples. Does this growing use of interactive cartography to acquire geographic information also correspond to a greater ability of people to produce spatial data? The following contribution aims to investigate a heterogeneous group of people's familiarity with online maps. To this end, an experiment which requested participants to identify the position of specific geographic features on maps embedded on an online survey was conducted. The aim was to assess whether their knowledge of GIS affects the quality of their mapping activity. The experiment involved about two hundred students from the University of Udine (Italy) and took place partially in person and partially remotely. The students, enrolled in various university courses, differ in having varying degrees of exposure to geography and GIS in their educational path. In the article, after a brief presentation of the experiment, we will first describe respondents' characteristics, such as demographics, education and computer skills, and then analyze the positional accuracy of the spatial data they provided. Finally, we will assess participation and data quality in relation to the experiment's characteristics. The results reveal that the method of supervision and GIS training are two important factors for the provision of high-quality geographic information.

Keywords

Map-based survey, GIS skills, University students, Data quality, Friuli Venezia Giulia

The authors shared entirely the content of this article, however, paragraphs 2.1.1, 2.3, 3.2, 3.3 and 4 are to be attributed to Giorgia Bressan, the rest to Gian Pietro Zaccomer.

1. Introduction

Map services are increasingly utilized by Internet users to carry out daily activities such as planning transportation routes, tracking sports activities, searching for the location of shops, to give some examples. Despite the fact that interactive cartography has entered people's lives, it is not clear whether this propensity for consulting online maps is consistent with people's ability to produce geographic data. Bearing this context in mind, in this paper we want to investigate what conditions favor the collection of high-quality geographic information. Although this research question is rather specific, it should be noted that this contribution is part of a wider project, where spatial data produced by ordinary citizens are at the core of the analysis. We are referring specifically to the PaRiDe project which investigates the relationship between citizens and their surrounding landscapes in the Italian region of Friuli Venezia Giulia (hereafter, FVG), which can be considered a logical continuation of the scientific work carried out for the drafting of the Friuli Venezia Giulia Regional Landscape Plan (FVG-RLP) (Guaran, Pascolini, 2019; Zaccomer, 2018, 2019). In fact, the project employed spatial data collected through both a participatory mapping approach (Maiulini, Cadez, 2019) and a survey approach, in order to obtain a statistical and cartographic description of the vulnerable and degraded landscapes present in FVG region (Bressan, Pascolini, 2019). Specifically, the project relates, amongst other things, to Volunteered Geographic Information (VGI; Goodchild, 2007), which involves the use of IT tools "to create, assemble and disseminate geographic data provided voluntarily by individuals" (Borruso, 2010, p. 245). In the research, data-produced contents concern the landscapes perceived as of exceptional beauty and those sites that are considered abandoned, degraded or deprived in FVG.

During the PaRiDe project, a map-based survey was used (Bressan, Amaduzzi, 2020) to elicit such information. However, the analysis of the spatial data obtained through this survey has prompted the formulation of a new research question: to what extent does the spatial information provided depend on the respondent's level of familiarity with these tools? In other words, how the quality of the geographic information collected depends on the ability of citizen sensors to use these new IT tools. Given that the concerns on VGI data quality are not yet comprehensively tackled (Yan et al., 2020), this paper aims to offer some reflections that might contribute to reducing such a research gap.

Taking inspiration from the work done by Poplin (2015) on the usability of interactive maps, we performed a georeferencing experiment at the University of Udine aimed at collecting spatial data and other information concerning students' demographics, education and computer skills. The study aims to verify how, and with what accuracy, students produce georeferenced data.

2. The methodological aspects

This section concerns the methodological aspects of the research. Specifically, it will present the structure of the survey, the participants in the study, and the procedure used to analyze data.

2.1 The characteristics of the survey

The geo-questionnaire is made up of three different sections. The first concerns students' personal data (sex, age, status of full or part-time student) as well as their domicile and the identification of which municipalities in the region they are more familiar with.

The second section deals with student's training and computer skills. On the first aspect, it should be noted that a question on the high school attended is included. Participants are not asked which degree course they attend, as this information is known from the design of the sample. As regards the second aspect, the survey contains a question on the average number of hours per day spent in front of a computer or smartphone and another on how many hours are spent connected to the Internet. Another question relates to their presence on the most popular social networks. A second group of questions in this section concerns the possession of basic computer knowledge (attendance of a IT course at university or of an IT course outside the university context), more specific ones on GIS fluency (we asked whether they were users of certain GIS applications)

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

and other IT skills (such as having developed an Internet web site, running a blog or actively contributing to online collaborative projects). Finally, this second section contains questions about participants' behavior regarding the possibility of georeferencing their own photographs, and their opinions on the usefulness of sharing personal geographic information.

The last section of the survey includes questions with a geographic dimension, which, being the core of the experiment, require particular attention.

2.1.1 The georeferencing experiment

As previously mentioned, this experiment stems from the willingness to explore some aspects which emerged during the PaRiDe project in greater depth. The main problems encountered in the previous survey were the limited participation of the respondents in the spatial section and, in the case of compilation, the preference of many respondents for using the textual replies instead of the map to answer the geographical questions. These facts suggest that there are difficulties in using interactive cartography. This experiment aims to explore these issues and to evaluate the accuracy of map-based contributions. Specifically, this new work is made up of three mapping exercises. The decision was to ask participants to identify the position of three well-known features in the FVG region on a map. These features are the Miramare Castle, the Tagliamento River and the Marano-Grado Lagoon.

Not all respondents access the interactive map. In fact, each exercise begins with a filter multiple-choice question asking if they know where a certain feature is and if they know its geographic position on a map[1]. Only those who claim that they can identify it on a map access the interactive map; otherwise they should move on to the next exercise relating to the successive selected feature. Respondents can choose one geometric object to map, between point, line or polygon. For an in-depth analysis of the technical characteristics of the platform used, the reader could refer to Bressan and Amaduzzi (2020). Here it suffices to say that once respondents arrive at the interactive maps, an OpenStreetMap layer is displayed by default. Clearly, the three chosen places imply a different level of complexity. The exercise on the Miramare Castle, for example, assumes that a user will zoom on the Trieste karst ridge, while for the Tagliamento River the map extent should encompass the whole study area, as this feature crosses the entire region. Once the question with the map is reached, the possible hurdles consist of reading the base map correctly and interacting with the tool. With regards to the second problem, we want to highlight, as will be illustrated later, that the compilation of the questionnaire was supervised: a researcher was available during the experiment to reply to possible technical questions.

For the sake of completeness, it should be noted that these three exercises are followed by other spatial questions where, similarly to the work done for PaRiDe, students are asked to map a place perceived as of exceptional beauty and another as degraded. Overall, students could draw on maps for a maximum of five times.

The questionnaire ends with a series of multiple-choice questions on the reasons behind the decision to use one of the geometric objects available for the mapping exercise and some other questions on the usability of the questionnaire.

2.2 Administration and reference population

The web-based survey took place during the academic year 2019-20 at the University of Udine (Italy). Six university courses were involved. The selection of these courses was based on the fact that they were attended by students having a different exposure to geography and GIS tools. Basically, the courses involved in the experiment can be classified on three different levels, each one implying a different familiarity with GIS. A first level consists of courses that do not deal with either geography or GIS aspects (Sociology, Econometrics and Economic Statistics). A second level is made up of

1 In the survey, the precise question was: "Can you localise the position of the [Miramare Castle]/[Tagliamento River]/[Marano-Grado Lagoon]? Respondents have to choose compulsorily between: "Yes, I know where it is, and I can identify it on a map"; "Yes, I know where it is, but I cannot identify it on a map"; and "I don't know where it is". It should be noted that the survey was written in Italian and the questions and answers are translated into English with the purpose of publication.

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

courses where there is only an intermediate knowledge, decidedly more theoretical than practical, of geographic information (Geomarketing and Economic Geography). Finally, the third level is constituted by a course where a formal training in geospatial data was full part of the program (Geomatics).

Given that each of the six different instructors accepted the invitation to collaborate in the experiment and no student participated twice (duplications were avoided), it is possible to consider the six groups of students involved as independent samples. Unfortunately, due to the Covid-19 pandemic, the supervision of students during the compilation of the online survey took place in the classroom for the courses held in the first semester, and remotely (using Microsoft Teams) for those of the second semester. When the supervision occurred with participants in their classrooms, the compilation was preceded by an introductory presentation illustrating the objectives of the research and how the survey was structured. During the completion of the web form students could ask questions, by simply raising their hands. In order to maintain the same conditions of the experiment, when the interaction with students took place remotely, the compilation was preceded by the same introductory presentation (even the same slides presented in the classroom were used) and during the experiment students could still ask questions in the chat box. A great effort was therefore made to avoid differences in the two types of supervision, and misunderstandings in the completion of the spatial section of the survey.

2.3 Data analysis plan

The analyses carried out in this paper are based on a dataset, which was assembled by aggregating the answers coming from the six surveys conducted for each course involved. The collected data were analyzed according to their statistical or geographic nature. As already illustrated, the first two sections of the questionnaire consist of a series of multiple-choice questions concerning demographics, education and computer skills. The replies to this "non-spatial" part of the questionnaire were imported into a statistical program to produce basic descriptive statistics and carry out statistical tests. The distributions of the daily average hours spent on the IT devices (i.e. computer and smartphone) were constructed, together with two sum indicators. These latter were obtained on the basis of the answers provided in the questionnaire, relating to GIS skills and other IT skills. As regards the statistical tests, we used the t-test for equality of means.

The third part of the questionnaire constitutes the core of the georeferencing experiment and includes both multiple-choice questions and questions with maps. The replies to the first type of questions were analyzed using the usual descriptive tools. To evaluate the ability of respondents to map accurately, a geodatabase was created. This served us to visualize the spatial replies on GIS software and analyze the accuracy of the location of the mapped features. The request to identify three known places, having precise boundaries, makes it possible to compare them with benchmark geometries extracted from other sources and to study their positional accuracy, defined as "the accuracy of the position of features (i.e. points, lines or areas) within a spatial reference system" (Fonte *et al.*, 2017, p. 141). Operationally, to identify the quality of the data referring to the Miramare Castle, we consider the polygon extracted from OpenStreetMap dataset of the castle building (therefore squares and park were excluded) and the polygon with the entire area of the castle grounds (the source in this case is the FVG-RLP). This choice makes it possible to distinguish those who interpreted the question as a request to identify the site of the building from those who considered the castle's surroundings as a whole. For the question on the Tagliamento River, two buffers with radii of 500 and 100 meters were created around the linear feature of the Tagliamento River extracted from the dataset on Waterways from the regional infrastructure of territorial data (IRDAT). In practical terms, the decision to employ two reference polygons for the comparison has similarities with the previous exercise on the Miramare Castle. However, while in the previous case the wider benchmark polygon referred to an area that is officially labeled as Miramare, in the case of the Tagliamento the use of a large buffer means that a point can be classified as correctly positioned even if it does not lie

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

within the riverbed. As for the Marano-Grado Lagoon, the polygon of the lagoon from the IRDAT database was chosen as reference.

There is no established method in the literature for the ex post assessment of the quality of spatial data. For this article, the decision was to convert lines and polygons into centroids and to evaluate if the resulting points fall within the reference geometry[2]. A point was considered as "of quality" if it was within the reference geometry. For each exercise, information on the positional accuracy of each mapping activity exists and, therefore, it is possible to analyze the quality of user-produced contents in relation to the experiment's characteristics.

3. Results

The presentation of the results is here divided into four parts. The analysis begins with the non-spatial component of the survey and it concerns the first two sections of the geo-questionnaire (i.e. personal data, education and computer skills). The rest regards the spatial section of the questionnaire.

3.1 Personal data and computer knowledge

In total, 217 valid geo-questionnaires were collected, of which 64 belong to students of the so-called first level, 99 of the second and 54 of the third. Respondents are mostly female (64.5%) and 85.3% are between 20 and 23 years old. Furthermore, 28.1% is represented by part-time students. As for their educational background, there is equal participation from students coming from high schools and technical institutes (both 41.9%). The analysis of their geographic origin shows that mostly live within the study area, as only 21.2% state that they live outside the region (almost all come from the neighboring region of Veneto). This difference is immediately noticeable when considering the degree of knowledge of the FVG region. Students living in the study area state that they know well, on average, about 5.5 other municipalities, while others claim to be familiar to a basic extent only with the municipality hosting their main university building (Udine or Gorizia).

Passing to the second section of the questionnaire, as foreseeable from the young age of participants, more than 90% state that they have at least one profile on social media. Almost two thirds of the sample, more precisely 62.7%, spend up to five hours a day on the PC or mobile phone and up to four hours connected to the Internet. Furthermore, 70% affirm that they attend or have attended an IT course at university, and 12.9% a training course in the same field, but outside the academic environment. With regards to GIS skills, in the survey, two separate questions examine whether respondents employ ESRI software or other, unspecified, proprietary software and other two separate questions investigate whether they make use of QGIS or another, unspecified Open Source software. Our findings show that 27.2% say that they use at least one GIS application. It should be noted that 20.7% are QGIS users. This result is unsurprising because the university course of Geomatics, which we have involved, includes some laboratory hours where this open-source software is used. The users that employ other open source software are completely residual because they are only 7 students. Finally, 21.7% have developed websites, while 4.1% run their own blog. The contribution to collaborative projects concerns only 7 students and, as such, it is a marginal phenomenon. As anticipated, on the basis of these questions, two sum indicators have been constructed. The first considers the simultaneous use of one or more GIS applications and therefore varies from 0 to 4, while the second considers the three aforementioned computer activities, not related to GIS, and therefore it ranges from 0 to 3.

3.2 Participation in the geographic section

After the presentation of the more general aspects, the attention is now drawn to participation in the geographic section of the survey. Table 1 shows the responses of

2 The "Centroids" algorithm in QGIS was used to transform lines and polygons into centroids. Subsequently, both for the points mapped by the participants and the centroids derived from lines and polygons, it was verified whether they fell within the reference geometry. The verification was always performed in the QGIS environment using the "Intersect" modality present in the "Join Attribute by Location" algorithm.

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

TABLE 1 – Replies to the questions concerning the localisation of the three geographic features

REPLIES	MIRAMARE CASTLE	TAGLIAMENTO RIVER	MARANO-GRADO LAGOON
Yes, I know where it is, and I can identify it on a map	133 (61.3%)	129 (59.4%)	116 (53.5%)
Yes, I know where it is, but I cannot identify it on a map	66 (30.4%)	70 (32.3%)	63 (29.0%)
I don't know where it is	18 (8.3%)	18 (8.3%)	38 (17.5%)
TOTAL	217 (100.0%)	217 (100.0%)	217 (100.0%)

SOURCE: Own elaboration, 2020

TABLE 2 – Replies to the questions concerning the geometric objects used to map

REPLIES	MIRAMARE CASTLE	TAGLIAMENTO RIVER	MARANO-GRADO LAGOON
Point	101 (75.9%)	26 (20.2%)	23 (19.8%)
Line	0 (0.0%)	89 (69.0%)	3 (2.6%)
Polygon	30 (22.6%)	12 (9.3%)	89 (76.7%)
No choice	2 (1.5%)	2 (1.5%)	1 (0.9%)
TOTAL	133 (100.0%)	129 (100.0%)	116 (100.0%)

SOURCE: Own elaboration, 2020

the participants to the questions on their knowledge of the selected places, while Table 2 presents which geometric objects participants used in mapping.

In each exercise, most of the respondents affirm that they know the three requested geographic features and they can identify it on a map. However, it is noteworthy that in each exercise almost a third of people answer "Yes, I know where it is, but I can't locate it on the map". This group, together with those who do not know where the features are, do not access the successive question asking them to select a geographic object to map. It should be noted, as shown in Table 2, that two respondents who claim that they know to identify the castle on the map do not choose any geometric object in the successive question. This phenomenon, that occurs also in successive exercises, probably signals some problems with the survey or unwillingness to contribute.

The preferred geometry is the point, in the case of the Miramare Castle, while with the exercise on the Tagliamento River it is the line. With regards to the Marano-Grado Lagoon, the vast majority decided to use the polygon. Overall, the point is the most used geometric object in the experiment.

3.3. Assessing positional accuracy

This section concerns the respondents' ability to correctly identify the position of the geographic feature. In fact, the filter question on the ability to locate the requested feature on the map analyzes only a theoretical ability. An assessment of the correctness of this information can only be carried out by analyzing the positional accuracy of the map-based contributions (the

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

procedure followed in this paper to assess data quality was described in 2.3).

The analysis of the positional accuracy of the mapping activity starts with the contributions that respondents provided through the point geometry. Table 3 presents the absolute number, the percentage of points mapped by the participants within the reference geometry, and those that are outside.

With reference to the Miramare Castle (see Figure 1), if we consider as benchmark only the geometry of the

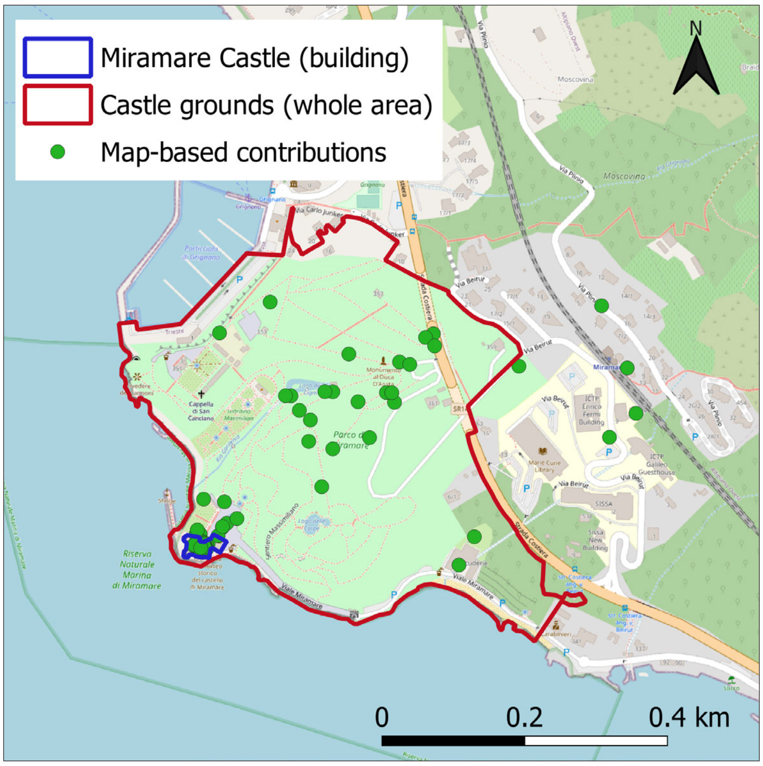
TABLE 3 – Data quality when the geometric object used by respondents is the point

Miramare Castle			
WITHIN BUILDING	WITHIN AREA	OUTSIDE AREA	TOTAL
53 (52.5%)	88 (87.1%)	13 (12.9%)	101 (100.0%)
Tagliamento River			
WITHIN BUFFER 100 M	WITHIN BUFFER 500 M	OUTSIDE BUFFER 500 M	TOTAL
14 (53.8%)	25 (96.2%)	1 (3.8%)	26 (100.0%)
Marano-Grado Lagoon			
WITHIN	–	OTHER CASES	TOTAL
21 (91.3%)	–	2 (8.7%)	23 (100.0%)

SOURCE: Own elaboration, 2020

FIGURE 1
Mapping results when participants identify the Miramare Castle with points

SOURCE: Own elaboration © OpenStreetMap contributors, 2020



ISSN 2282-572X (online)
ISSN 0044-9733 (print)

FIGURE 2
Mapping results when participants identify the Tagliamento river with points. The buffer of 500 m is considered as reference geometry

SOURCE: Own elaboration © OpenStreetMap contributors, 2020

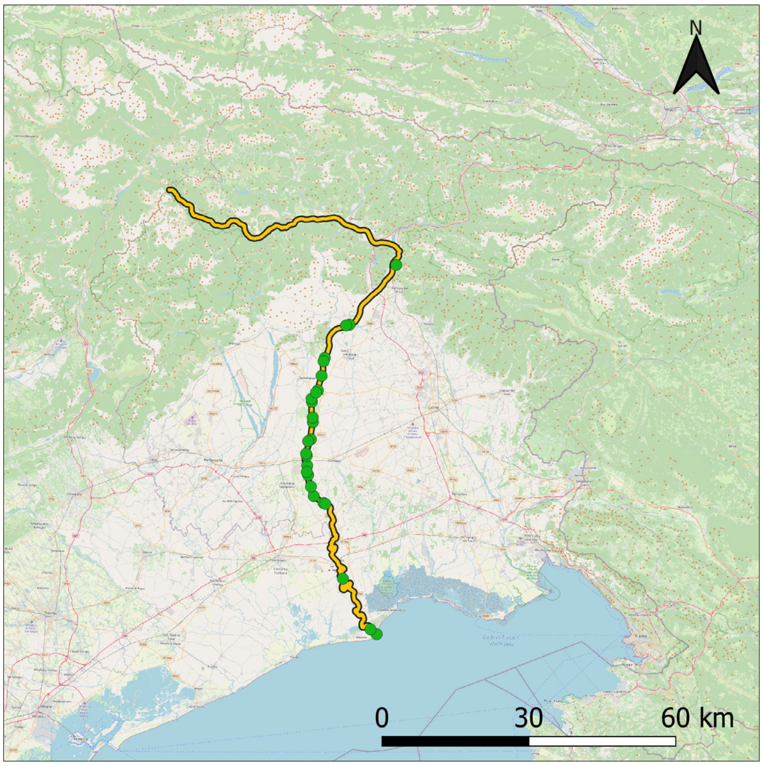
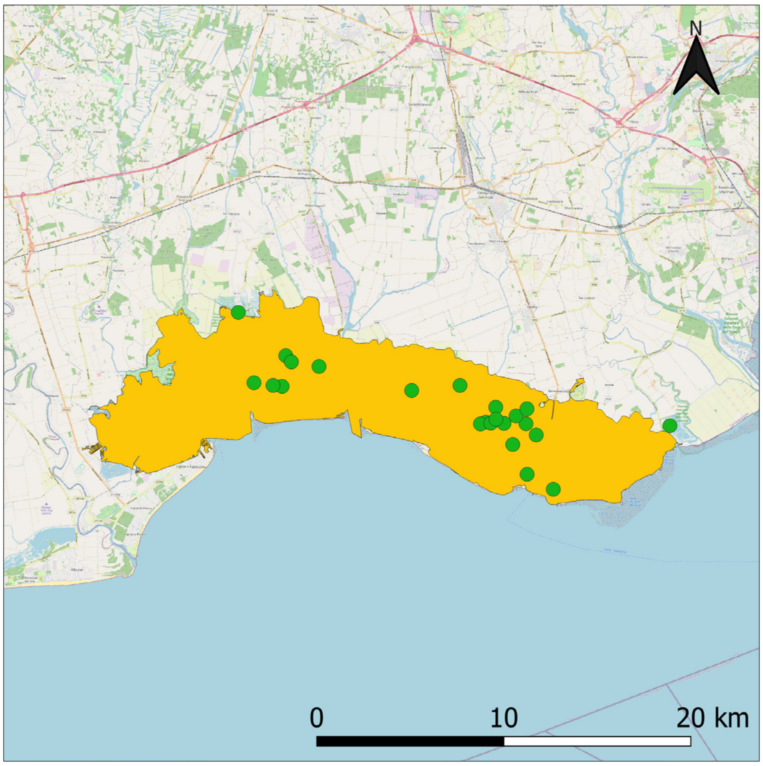


FIGURE 3
Mapping results when participant identify the Marano-Grado Lagoon with points

SOURCE: Own elaboration © OpenStreetMap contributors, 2020



ISSN 2282-572X (online)
ISSN 0044-9733 (print)

building, it can be observed that more than half of the points (52.5%) are within this polygon. When the larger polygon is taken as reference, the percentage rises greatly (87.1%). It should be noted that about one eighth of the points are completely outside the larger reference polygon (12.9%) and they are even located in the historic center of the city of Trieste and in the sea off the coast of Slovenia.

As for the Tagliamento River, if we consider the buffer with a 100 m radius, we can observe that about half of the points fall within this area (53.8%). When considering instead the buffer with a radius of 500 m, almost all the points fall within the buffer (96.2%) (see Figure 2).

As for the Marano-Grado Lagoon (see Figure 3), the overlay analysis highlights that almost all points intersect the benchmark area (91.3%).

As anticipated, for the mapping contributions made with lines and polygons the decision was to calculate their centroid and overlay such points with the benchmark polygon. Table 4 presents the absolute number and the percentage of centroids within the reference polygon, respectively when the starting geometric object is the line and the polygon. Since in the exercise concerning the Tagliamento River the greatest number of contributions are made with lines (89), and for the exercise on the Lagoon the preferred geometric object is the polygon (89), we want to focus on the quality of the data in these two cases. Most of the centroids representing the Tagliamento River are placed outside the buffer with a radius of 500 meters when the starting geometry is the line (88.8%). Furthermore, the fact of passing from a buffer with a radius of 100 to one with 500 meters only slightly increases the number of centroids falling within the reference geometry (from 3 to 10). These results indicate the difficulty of respondents in tracing the geometry of the river accurately. However, when the lagoon is considered, the centroids of the polygons fall almost entirely within the benchmark geometry (94.4%).

TABLE 4 – Data quality when the original geometry used by respondents is the line or the polygon

Miramare Castle				
GEOMETRIC OBJECT	WITHIN BUILDING	WITHIN AREA	OUTSIDE AREA	TOTAL
POLYGON	8 (26.7%)	23 (76.7%)	7 (23.3%)	30 (100.0%)
Tagliamento River				
GEOMETRIC OBJECT	WITHIN BUFFER 100 M	WITHIN BUFFER 500 M	OUTSIDE AREA	TOTAL
LINE	3 (3.4%)	10 (11.2%)	79 (88.8%)	89 (100.0%)
POLYGON	1 (8.3%)	1 (8.3%)	11 (91.7%)	12 (100.0%)
Marano-Grado Lagoon				
GEOMETRIC OBJECT	WITHIN	-	OUTSIDE AREA	TOTAL
LINE	3 (100.0%)	-	0 (0.0%)	3 (100.0%)
POLYGON	84 (94.4%)	-	5 (5.6%)	89 (100.0%)

SOURCE: Own elaboration, 2020

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

3.4. The quantitative analysis

As mentioned, the sample was constructed in such a way that the six different groups of students can be considered as independent samples. For this reason, it is possible to employ the t-test for equality of means. When the indicator on GIS knowledge is considered, the value of this indicator is on average extremely low for the first and second level (0.08 and 0.11, respectively). It is quite evident that the difference in the average number of GIS skills reported could not be significant. However, when the first two levels are aggregated and this value is compared with the third, the difference is of 0.96 (the values of the indicator are 0.10 and 1.06, respectively). This difference is decidedly significant, as shown in row T1 of Table 5. This result empirically validates the assumption on the importance of formal GIS training in the preparation of the sampling plan. In fact, the experiment focuses on the practical ability of participants to use GIS tools and, for this reason, the first two levels do not appear that different, while the third distinguishes itself from the others.

There is also a certain difference between the first two levels considered together and the third (0.34 against 0.13) when the indicator on IT skills is considered. In this case, the lowest value is associated with the third level. This result can be interpreted as the fact that those who specialize in GIS do not focus on other activities such as the construction of websites or writing a blog. This finding is confirmed by the t-test, which is significant as shown in row T2 of Table 5. This outcome, like the first, is not surprising because the sample includes respondents attending the university course of Public Relations (to which three of the courses involved in the experiment pertain). These students have only a theoretical knowledge on GIS, but, on the other hand, they develop advanced computer skills for web communication.

As anticipated, due to the Covid-19 pandemic, the conduction of the survey was supervised in person for some courses, while it was done remotely for others. It is therefore necessary to understand whether the physical presence of the researcher affected the participation of the students in the completion of the georeferencing exercises. The simple observation of the means of the completed exercises leaves no doubt, because when the supervision occurred in presence, there is a mean of almost 3 completed exercises out of 5 (there are five, because the two optional mapping exercises that students could do on landscape perceptions are also considered). On the other hand, when the supervision is online, the participation drops to 2.2. A further aspect that should be considered is that the percentage of students who completely skipped the spatial part of the geo-questionnaire (34 students out of 217, that is 15.7%) is only 8% in the face-to-face supervision while it is 22.2% in the online modality. This finding is not too surprising because in the online supervision there was no possibility

TABLE 5 – Independent sample tests for equality of variances and means

Label test	F	p-value	Variances	Difference	t	p-value	Means
T1	9.29	0.00	not equals	0.96	10.93	0.00	not equals
T2	28.86	0.00	not equals	-0.21	-3.18	0.00	not equals
T3	5.55	0.02	not equals	0.63	2.71	0.01	not equals
T4	0.04	0.85	equals	0.33	1.62	0.11	equals
T5	0.08	0.77	equals	0.34	2.27	0.02	not equals
T6	0.07	0.80	equals	0.34	2.70	0.01	not equals

SOURCE: Own elaboration, 2020. The significance level for the test is set at 5%

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

to control whether students were really completing the spatial section of the survey in their homes. A similar outcome, despite being quite trivial, highlights that the production of spatial data is strongly affected, above all, by the desire to participate in the survey as well as by any intrinsic difficulties of doing it alone.

Finally, it is necessary to understand whether the level of GIS knowledge influences the ability to participate in the experiment. The students belonging to the first two levels completed about 2.4 exercises each, while those in the third slightly more than 3 (we recall again that the maximum number of exercises in the survey is 5). The difference of 0.63 is considered significantly different, as shown in T3 of Table 5. In other words, we have verified how the specific knowledge acquired in university courses influences the ability to complete at least one of the required exercises. If we now exclude from the analysis the students who did not complete any exercises, we can investigate the effect of the level of GIS knowledge on the number of exercises completed: the students of the first two levels have completed 2.95 exercises each, while those of the third 3.28. The difference of 0.33 in this case is smaller than the previous one and is not considered statistically significant, as shown in row T4 of Table 5. It seems that the simplicity of the exercises was such that, once the mechanism of interactive mapping was understood, the total number of exercises completed was not substantially different between "experts" and those who did not have a formal training in GIS.

Beside the participation of the students in the spatial component of the survey, it is relevant to analyze the quality of the data they provide. In this paper, we focus in particular on the ability of students to identify the correct position of the Miramare Castle, the Tagliamento River and the Marano-Grado lagoon. In this case, a new measure, different from the previous ones, must be used. Therefore, in the final part of this analysis we want to study how the indicator that counts the number of correct exercises (ranging from 0 to 3) behaves according to the level (first two levels versus the third) and the mode of supervision (in person versus remotely). In the first case, as it is possible to read from line T5 of Table 5, the first two levels have an average of 1.09 of correct exercises against 1.43 of the third level (the difference is equal to 0.34). As for the impact of the type of supervision, we can refer to the results presented in row T6 of Table 5. The online supervision shows an average 1.02 of correct exercises against 1.36 in the case of physical presence (the difference is still of 0.34). Looking at the values of the tests, the two cases behave in basically the same way: the Levene's test leads us to say that the variances are not different, and the t-test allows us to affirm that the differences found are both significant, so the respective means are not equal. In other words, both the level of practical GIS knowledge and the mode of supervision led to significant differences in the ability to complete the exercises correctly, and therefore in the quality of the data obtained from the experiment.

4. Concluding remarks

The paper tries to call attention to the quality of user-producer contents. The results mainly highlighted two aspects. The first relates to the fact that not everyone is comfortable dealing with interactive cartography, given that there is a large number of students who do not access the maps in the experiment, either because they do not know where the requested places are or because they do not know how to find them on the map. The second important result is that those who claim to be able to map do not always produce high-quality data. In fact, the statistical analysis showed that there were two factors that favored the collection of accurate map-based contributions: formal training in the GIS field and the method of supervision.

Our results highlight that probably not everyone is able to produce spatial data, as having a formal GIS training (i.e. belonging to the third level) increases the possibility of performing the exercises correctly. These considerations therefore raise concerns about the fact that, despite the Internet being in principle open to the contribution of everyone, only a few people can provide accurate geographic contributions. Since university students are those who, by deciding to invest time in their training, demonstrate that they want to increase their skills and abilities, the results show that the ability to provide new high-quality geographical information is not for all.

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

The other aspect that emerged is on the importance on the conditions of the experiment, since when the supervision occurs online, participants seem to be less involved. It is important to remember that when we planned the experiment, we did not consider the possibility of students conducting the survey independently, at their home. The need to resort to the online supervision was a consequence imposed by the coronavirus lockdown. However, the participation "from home" is the one that comes closest to the reality of the VGI, which is of the citizen who wants to contribute to the production of new geographical information through online cartographic tools alone. The fact that even university students are not comfortable with online maps must make us reflect on the barriers that any citizen may encounter when using these types of online tools.

This paper seeks to contribute to the VGI domain, but there are many issues that remain unexplored, such as the relationship between the respondent's location and participation. In Bressan et al. (2020) it was highlighted that when volunteers are asked to indicate sites perceived as being degraded and of exceptional beauty within the region, they do so by indicating places close to their home. In this new survey the issue is to verify whether the knowledge of the (correct) geographical position of the place depends on the participant's familiarity with the municipality in which the element to be mapped is located. Since in the questionnaire participants were asked for both their domicile and the municipalities with which they are familiar, it is possible in principle to explore this issue with a high level of detail. Furthermore, another priority appears to be deepening the study of data quality, by using other measures of positional accuracy and employing other reference geometries for the evaluation of correctness of the mapping activity.

As VGI is evolving, a better understanding of the parameters that affect data quality will help enhance the utility of its products as alternative information source.

ISSN 2282-572X (online)
ISSN 0044-9733 (print)

References

Amaduzzi S., Bressan G., Guaran A., Pascolini M., Zaccomer G.P. (2019), "I paesaggi che cambiano: quali risposte contro il degrado?", in: Cerutti S., Tadini M. (a cura di), *Mosaico/Mosaic. Memorie geografiche* NS17, Società di Studi Geografici, Florence, pp. 493-502.

Borruso G. (2010), "La 'nuova cartografia' creata dagli utenti. problemi, prospettive, scenari – The user-created 'new cartography'. Problems, perspectives and scenarios", *Bollettino dell'Associazione Italiana di Cartografia*, 138, pp. 241-252.

Bressan G., Pascolini M. (2019), "Dalle percezioni della popolazione al dato georiferito: studio quali-quantitativo del paesaggio del Friuli V.G.", in: *Atti della Conferenza Nazionale di Geomatica e Informazione Territoriale ASITA 2019*, ASITA, pp. 115-122.

Bressan G., Zaccomer G.P., Grassetti L. (2020), "Studying the Spatial Distribution of Volunteered Geographic Data Through a Non-parametric Approach", in: Gervasi O., Murgante B., Misra S., Garau C., Blečić I., Taniar D., Apduhan B.O., Rocha A.M.A.C., Tarantino E., Torre C.M., Karaca Y. (a cura di) *Computational Science and Its Applications – ICCSA 2020. ICCSA 2020. Lecture Notes in Computer Science*, vol. 12252. Springer, Cham. DOI: 10.1007/978-3-030-58811-3_36.

Bressan G., Amaduzzi, S. (2020), "Map-based surveys for mapping high-quality and degraded sites", in: Gallia A. (Ed.) *Territorio: rischio/risorsa*. Labgeo Caraci, Rome, pp. 197-213.

Fonte C. C., Antoniou V., Bastin L., Estima J., Arsanjani J. J., Bayas, J.C. L., See L., Vatseva R. (2017). *Assessing VGI Data Quality*, in Foody G., See L., Fritz S., Mooney P., Olteanu-Raimond A.M., Fonte C.C., Antoniou V. (a cura di), Mapping and the Citizen Sensor, pp. 137-163, Ubiquity Press: London, UK. DOI: 10.5334/bbf.g.

Goodchild M. (2007), "Citizens as Sensors: The World of Volunteered Geography", *GeoJournal*, 69(4), 211-221. DOI: 10.1007/s10708-007-9111-y.

Guaran A., Pascolini M. (2019, a cura di), *Pianificazione e governo del paesaggio: analisi, strategie, strumenti. L'apporto pluridisciplinare dell'Università di Udine al Piano Paesaggistico Regionale del Friuli Venezia Giulia*, Forum, Udine.

Maiulini E., Cadez L. (2019), *Pianificazione paesaggistica e partecipazione: l'analisi dei dati*. In: Guaran A., Pascolini M. (a cura di), *Pianificazione e governo del paesaggio: analisi, strategie, strumenti. L'apporto pluridisciplinare dell'Università di Udine al Piano Paesaggistico Regionale del Friuli Venezia Giulia*, Forum, Udine, pp. 87-103.

Poplin A. (2015), "How user-friendly are online interactive maps? Survey based on experiments with heterogeneous user", *Cartography and Geographic Information Science*, 42(4), 358-376. DOI: 10.1080/15230406.2014.991427.

Yan Y., Feng C.C., Huang W., Fan H., Wang Y.C., Zipf A. (2020), "Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience", *International Journal of Geographical Information Science*, 34(9), 1765-1791. DOI: 10.1080/13658816.2020.1730848.

Zaccomer G.P. (2018), *L'analisi territoriale socio-economica in ambito paesaggistico. Gli indicatori compositi per la zonizzazione territoriale del Friuli Venezia Giulia*, Forum, Udine.

Zaccomer G.P. (2019), "L'analisi territoriale socio-economica in ambito paesaggistico. Metodi multivariati e indicatori compositi a confronto", in: *Atti della Conferenza Nazionale di Geomatica e Informazione Territoriale ASITA 2019*, ASITA, pp. 1033-1040.

ISSN 2282-572X (online)
ISSN 0044-9733 (print)