



Working with Volunteered and Crowdsourced Geographic Information (V&CGI)

Instructors:

Dr. Peter Mooney (Maynooth University, Ireland) peter.mooney@mu.ie ,

Dr. Levente Juhász (Florida International University GIS Center, United States of America)
ljuhasz@fiu.edu

Target audience: Staff of national mapping agencies, public authorities, early career researchers in geographic related disciplines, other organisations and private companies dealing with or interested in Volunteered or Crowdsourced Geographic Information. Some basic knowledge of, and experience working with, geospatial data (mostly point and polygon-based data) will be helpful. Participants are informed that a basic knowledge of a programming language is required in order to complete some of the practical exercises and assessments.

Course objectives: Up to recently geographic data was exclusively available from authoritative sources such as National Mapping and Cadastral Agencies, professional companies, etc. The rapid emergence of Volunteered Geographic Information (VGI) and Crowdsourced Geographic Information (CGI) (referred to for brevity as V&CGI in this document) has challenged and changed this situation. V&CGI have transformed from being considered 'disruptive' and poor quality to well-known mainstream data sources used widely in industry, research, and other applications. The goal of this course is to introduce participants to V&CGI, the current state-of-the-art research in these areas, methods for obtaining V&CGI data (API sources, processing GeoJSON, etc), and advanced topics such as assessment of the quality of these data. Various examples from practical applications and hands-on practical work will help illustrate the key characteristics of V&CGI data. These practical examples will also help participants to explore the opportunities provided in terms of research, geospatial data integration, etc. by data sources from V&CGI. The course will also equip participants with a suitable blend of the current theory and practice around understanding the quality of V&CGI data and assessing its fitness for different purposes and applications. Free and open source software and openly accessible V&CGI will be used without any need for special hardware or software installations.

Topics tackled: What are V&CGI? Understanding the state-of-the-art literature and how V&CGI is different to traditional sources of geospatial data; Key readings in both V&CGI. Tagging, ontologies and folksonomies; Practical demonstrations of how to access, use, analyse and visualise Citizen and user-generated content online; Data quality in V&CGI? Advances topics such as data licensing, data ethics and reproducibility with V&CGI.

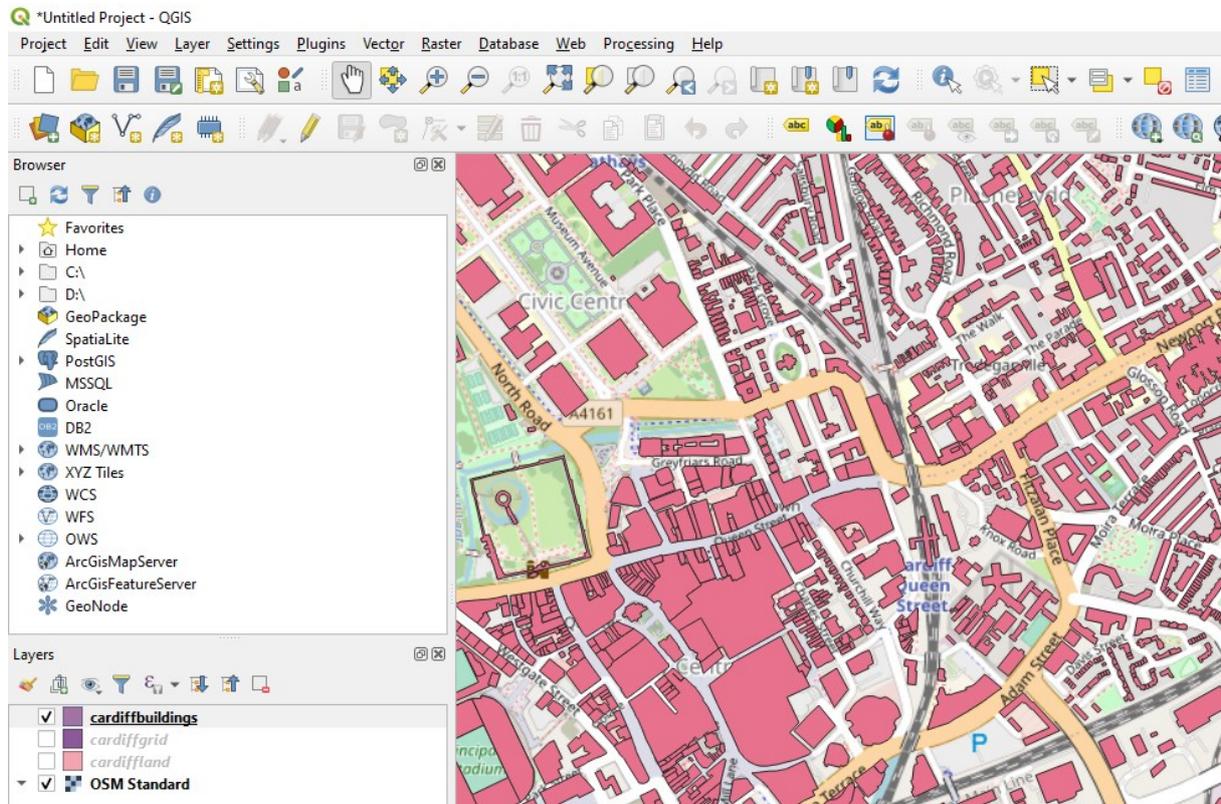


Figure 1. An example within QGIS of OpenStreetMap data (from an ESRI Shapefile) overlaid on an OpenStreetMap map layer supplied by Web Map Services

Module 1: Defining V&CGI (Peter Mooney/ Levente Juhász - 9 hours)

This module will introduce participants to the domain of V&CGI. To properly appreciate and grasp the concepts of V&CGI it is necessary to understand the context within which V&CGI is collected, generated, managed and accessed/disseminated. What do we mean by the term Volunteered Geographic Information and Crowdsourced Volunteered Information and what does it mean in relation to authoritative data sources such as that from National Mapping Agencies or commercial companies? What is the role of citizens in V&CGI? Participants will learn about the different modes of participation (active, passive, volunteered) and the types of tools and systems used by V&CGI projects with emphasis on the role played by open source software and open data. This module will also provide an overview of the state-of-the-art research literature in V&CGI. Participants will be introduced to key readings in this domain in order to understand the evolution of V&CGI as a research topic.

Module 2: V&CGI -Practical examples of working with the data (Peter Mooney - 9 hours)

One of the major challenges encountered by scientists, researchers and practitioners outside of the V&CGI domain is related to accessing the geospatial data itself. In this module participants will be introduced to techniques for accessing data from V&CGI. There will be several technical issues discussed including: working with APIs (application programming interfaces), JSON and GeoJSON, Web Map Services, JavaScript libraries such as Leaflet,



etc. This module will have a practical focus and participants will follow exercises allowing them to explore different ways to access V&CGI data from their own computers/systems. Some programming experience will be required here. However, examples will be provided to help all participants regardless of previous programming experience.

Module 3: V&CGI – Case Study Examples (Levente Juhász - 9 hours)

Based on the content from Module 2 this module will involve a more complex case study that involves spatial analysis and visualization using V&CGI. The core concept behind this module is to demonstrate how V&CGI data is capable of supplying the data and information required for advanced spatial analysis. Participants will also consider the important scientific of reproducibility and there will be a practical examination of the process of computationally recreating an already published study from V&CGI. Programming experience will be required and be very beneficial in this module.

Module 4: V&CGI – Current Context and Future Directions (Peter Mooney/ Levente Juhász - 3 hours)

Module 4 takes a more theoretical approach and considers the current socio-technological landscape V&CGI finds itself in. We consider what the current state-of-the-art questions are in this area and how and if they are likely to be resolved. We also consider some of the future directions of research and industry in this area including data licensing, data ethics, privacy and reproducibility with V&CGI. This brings about discussions about the long-term sustainability of V&CGI. How will the citizens of tomorrow contribute and use V&CGI? As V&CGI becomes more widespread and in some cases higher quality than commercially available data what does this mean for NMCA's?

Notes: This EduServe course will provide as much content as possible as open content. All study materials videos, presentations, lecture notes etc. will be made available freely and openly. It is intended to record all of the on-line webinars and make these recordings available. As stated above some programming ability is necessary in order to undertake this module completely. Python will be used for the programming exercises. For participants not familiar with Python there are many helpful resources available online which can provide introductory lessons which we advise these participants to take before this course. Sites such as <https://www.learnpython.org/> are useful for beginners to programming. All software used in this course will be Free and Open Source Software. This software can be installed easily on most computer architectures and operating systems.