Type of Contribution: PAPER

**Title:**

**To crowdsource or not to crowdsource: citizen science as a force of revealing historical evidence from Croatian Glagolitic manuscripts**

Keywords: Digital Humanities, Crowdsourcing, Collaborative transcription system, Citizen Science, Digitization

# Introduction

Historical manuscripts are very often the only evidence of the everyday life in local communities, especially manuscripts written in parishes, like those written in Zadar islands. Local historical manuscripts from Zadar area dating from 15th to 19th centuries are very often written simultaneously in Glagolitic and Latin scripts, as well as in several languages, namely Croatian redaction of Old Slavic, Croatian, Latin and Italian languages. Sometimes *hidden* in GLAM institutions, or in parish archives and private homes, those manuscripts present the only trustful written evidences of everyday life in Zadar area, and as such the interpretation of those manuscripts has to be included in the interpretation of the overall national and European history. Those manuscripts are of special interest for local citizens, as they reveal the history of their ancestors and ex-fellow citizens, of their villages, brotherhoods and parishes. Therefore, the local community is interested in its research.

The GLAM community is involved in a number of projects of digitization of manuscripts, and strives to assist Digital Humanities (DH) scholars in identifying, gathering, organizing, analysing and interpreting data, as much as in identifying available tools and software for data organization, visualization, and other activities involved in conducting research. One of such activities is manuscript text transliteration. Done in traditional environment, transliteration was usually done by a single scholar, then revised by one or several colleagues, and finally published in printed form as an edited edition. Nowadays, manuscript transliteration is done in digital environment, by the use of digital tools for transliteration allowing collaborative work of several scholars working on the same document, and is usually published in digital form. A number of GLAM institutions and research centres are organizing campaignes intended to involve public to participate in identification and transliteration of manuscripts. Their goal is to involve citizens in scientific activities and use the knowledge and motivation of the crowd to transliterate more manuscripts in less time. Involving citizens in identification, digitization, communication, readings and interpretation of local historical manuscripts is an important part of activities and endeavour reassembled under common term *Citizen science*, meaning involving public, usually amateur groups, users of GLAM institutions or civil associations in scholarly research. One of the activities organized with a goal to involve citizens in scholarly research is crowdsourcing.

**Theoretical framework & Research questions & Methodology & Research Results & Discussion & Conclusion**

# *Crowdsourcing*

Crowdscourcing as a concept is related to Citizen Science (CS), called also “crowd science”, “networked science”, or “massively-collaborative science” (Franzoni and Sauerman, 2014; Young, 2010; Wiggins and Crowstone, 2011). The main idea is to enable citizens to participate in the scientific research, “as observers, funders, in identifying images or analysing data, or providing data themselves” (European Commission Website)[[1]](#footnote-1). Citizens Science (CS) is described either as a method, a movement, or a social capacity, but always refers to “scientific work undertaken by members of the general public, often in collaboration with professional scientists and scientific institutions.”[[2]](#footnote-2) CS and crowdsourcing are closely connected to Informal Science Education (ISE), a movement oriented to increasing public science literacy[[3]](#footnote-3).

Crowdsourcing is related to GLAMs (Galleries, Libraries, Archives and Museums) which have a long history of involving citizens deeply in the wide range of activities and engagement with members of the public (Owens, 2012, Holley, 2010). Both *Crowdsourcing* and *Citizen Science* are recognized by European Commission as means of the democratisation of science, and as such are represented in EU Strategy. The Citizen Science movement in Europe is organized through The European Citizen Science Association (ESCA), which is enhancing the participation of citizens in scientific research, as well as supporting and performing citizen science projects.[[4]](#footnote-4)

Crowdsourcing, a process of leveraging public participation in projects and activities (Dunn, Hedges: 1), together with Citizen Science, have become an unavoidable method in a range of academic disciplines. Although some contend that the idea itself dates back in 1880s when so called *gifted amateurs* were engaged to voluntarily supply the Oxford English Dictionary with spellings, etymologies and definitions, Brabham (2014: 9) warns that, although similar to it, crowdsourcing is a new phenomen that relies on the technology of the Internet. It flourished in natural sciences in a project Galaxy Zoo[[5]](#footnote-5) where over 250.000 amateurs are bringing forward scientific discoveries of new classes of galaxies by collecting astronomical data (Hedges, Dunn, 2018: xi). The term *crowdsourcing* was first coined by Jeff Howe, editor of *Wired* magazine, in his article “The Rise of Crowdsourcing”, and in his blog *Crowdsourcing: Tracking the Rise of Amateur*. The term was adopted very quickly by the popular press and blogers. Although over 40 definitions of crowdsourcing are found (Estelles-Arolas and Gonzales-Ladron-de-Guevara, 2012), it is usually defined as “an online, distributed problem-solving and production model that leverages the collective intelligence of online communities, to serve specific organizational goals” (Brabham, 2013).

Crowdsourcing in humanities became possible with the growth of DH, with rapid and creative deployment of online platforms and virtual meeting places for humanists to collaborate remotely. Those platforms are nowadays adjusted to crowdsourcing task, to be able to give the citizens a possibility to interact with scholars, to be introduced with their tasks, methods and research results, to contribute to projects with their knowledge, and perform tasks sometimes on corpus far larger than the one scientists themselves could process. Although there is a range of crowdsourcing projects within DH, it is still mostly used as a method of using the power of the crowd to transcribe handwriting (i.e. Transcribe Bentham project).

The term *crowd* occurs to be slightly problematic. Owens warns that the term is misleading, as it refers to anonymous mass of people, while the most successful crowdsourcing project in libraries, archives, and museums have not involved *crowds*, but engaged members of the public (Owens: 1). Crowdsourcing in that sense is not equated with community sourcing as crowdsoursing involves more closed groups, while community sourcing includes open calls for participation (Hedges, Dunn, 2018: 17).

The benefits of involving citizens in scientific research are twofolded and can be divided to the benefits to institutions/projects and to citizens involved. Benefits to libraries, in particular, are listed by Holley (2010) and range from achieving goals the library would never have time of financial support to achieve by itself, to adding value to and improving data, and demonstrating the value of libraries in the community and strengthening and building trust of the users to the library. Holley researched a number of crowdsourcing projects in libraries, and made a *Volunteer Profile*, which brings the idea that volunteers in such projects are loyal, work hard and long, are very interested in the subject, appreciate the possibility of learning and contributing, they find the work personally rewarding and want to achieve the project/group goals, etc., and they are motivated by love, interest, cause, possibility to give something back to the community, etc. (Holley, 2010).

Academic crowdsourcing presents a distinctive type of crowdsourcing, oriented to and conducted within scholarly projects. In this paper, the models of citizen participation in the scholarly projects in the field of research of historical manuscripts will be researched. The benefits gained by the citizens, as well as those gained by the scholarly project will be studied, together with the attitude of humanities’ scholars towards involving citizens in scientific research. Benefits of involving citizens in scholarly project are not fully researched, and with this paper it will be put in a focus of Croatian scholarly projects in humanities, particularly those within Croatian Glagolitic research field.

**Research study**

Within a project *Digitization, bibliographic description and research of texts written on Glagolitic, Croatian Cyrillic and Latin scripts until the end of 19th century in Zadar and Šibenik area* (further *Written heritage*), carried out at the University of Zadar in co-operation with Vestigia Manuscript Research Centre of University of Graz, Austria, manuscripts, predominantly local historical manuscripts from Zadar area, written on Glagolitic cursive and Latin cursive scripts, are digitized, catalogued, presented on scientific portal together with rich and extensive metadata and researched by a group of humanities scholars. One of the aims of the project is to provide digital scholarly platform for research of historical manuscripts, to publish and share research data and communicate written heritage of Zadar and Šibenik area to variety of users and interested communities (Tomić, 2018: 1). Historical manuscripts digitized, catalogued and research within the first phase of the project *Written heritage* are of great local historical importance for Zadar area. They include register books (birth, christenings, matrimony, or death registers), land records, brotherhood registers and matriculas, homily and lecture collections and other manuscripts written in parishes of Zadar archdiocese, including city of Zadar, its hinterland and predominantly Zadar islands. Digitization in general, but digitization of historical manuscripts in particular, provides broad access not only to the contents, but much importantly, to the evidence being concealed in those valuable codeces. Texts of those manuscripts are usually the only written trace of everyday life, customs, personal and family histories, etc. Memory of lots of places are captured and hidden in those precious manuscripts, being nowadays hidden in various archives and private homes. Those local manuscripts are written on cursive Glagolitic script, palaeographicaly described as Glagolitic script of Northen Dalmatian style, developed in scriptorium in Tkone, on the island of Pašman, one of the islands of Zadar archipelagos (Štefanić, 1969: 20). That particular script is readable to a close circle of scholars, but civil associations of amateurs in the area of Glagolitic scripts are working on transliterations of such manuscripts. The idea of this paper is to see the possibilities of and problems with involving those amateurs in specific tasks within scholarly projects.

**The main goal** of this paper is to research the possible models of citizen participation in the scholarly projects, for a start in the field of transcription of historical manuscripts. The other goals include the insight into the benefits of academic crowdsource/civil science projects gained by the citizens, as well as those gained by the scholarly project itself, together with the attitude of humanities’ scholars towards involving citizens in scientific research. Additionally, possibilities of increasing public science literacy through crowdsourcing projects will be drafted.

The methodology used in this research is a case-study, including interviews (with scholars and citizens), and a simulation of controlled crowdsourcing campagne *Crowdsourcing Glagolitic historical manuscripts from Zadar area*. The sets of interviews will be conducted to gain the insight into scholars’ and public attitude towards using crowdsourcing for transliteration of Glagolitic manuscripts within scholarly projects and to arouse the interest of both sides into the civil participation in scholarly projects. The Glagolitic crowdsourcing campaigne *Crowdsourcing Glagolitic historical manuscripts from Zadar area* scenario will be drafted, including the determination of openness of crowdsourcing (crowdsourcing vs. community sourcing), description of selected tasks intended to be crowdsourced and foreseen outcomes and visualization of results.Also, a specification for adjustment of manuscript transcription online tool (Indigo information system) for crowdsourcing will be done within the case study.

**The expected result** is a model of public participation of citizens in scholarly projects involving transcription of Glagolitic manuscripts. The model should include the level of openness, targeted audience, outline of possible tasks for citizens and the form of those tasks in regards to project goals and to CS features, form of communication of results, a projection of potential benefits for citizens and for a scholarly project. Additionally, a specification of a tool for crowdsourcing transcription of historical Glagolitic manuscripts within scholarly project will be conducted.

# REFERENCES

Suzan A. Alteri and Daniel D. Golodner (2009). Using Web-based Exhibits to Enhance Interactive Arhcival Scholarship. In *Digital scholarship* / edited by Marta Mestrovic Deyrup. New York, London : Routledge, 132-144.

Brabham, D. (2013). *Crowdsourcing*. Massachusetts Institute of Technology, Cambridge.

Estellés-Arolas, E. and González-Ladrón-de-Guevara (2012). Towards an Integrated Crowdsourcing Definition. *Journal of Information Science* 38 (2): 189-200.

Franzoni, C., Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. Research Policy 43,1-20.

Holley, R. (2010) Crowdsourcing: How and Why Should Libraries Do It?. D-Lib Magazine 16, no. 3/4. <http://www.dlib.org/dlib/march10/holley/03holley.html>

Nielsen, M., (2011). Reinventing Discovery: The New Era of Networked Science. Princeton University Press, Princeton, NJ.

Owens, T. (2012). The Crowd and the Library. <http://www.trevorowens.org/2012/05/the-crowd-and-the-library/>

Owens, T. (2013) Digital Cultural Heritage and the Crowd. Curator: The Museum Journal. Vol 56. Iss: 1.

Štefanić, Vjekoslav (1969). Glagoljski rukopisi I. Zagreb: JAZU.

Tomić, Marijana (2018). Project Digitization, Bibliographic Description and Research of Texts Written in Glagolitic, Croatian Cyrillic and Latin Scripts Until the End of the 19th Century in the Zadar and Šibenik Area (Written Heritage) *// Cataloging & Classification Quarterly,* **56**, 4; 1-5 doi:10.1080/01639374.2018.1491438

Young, J. 2010. Crowd science reaches new heights. The Chronicle of Higher Education May 28, 2010. <https://www.chronicle.com/article/The-Rise-of-Crowd-Science/65707/>

1. <https://ec.europa.eu/digital-single-market/en/citizen-science> [↑](#footnote-ref-1)
2. <http://www.sisnetwork.eu/media/sisnet/Policy_brief_Citizen_Science_SiSnet.pdf> [↑](#footnote-ref-2)
3. <http://www.birds.cornell.edu/citscitoolkit/publications/CAISE-PPSR-report-2009.pdf> [↑](#footnote-ref-3)
4. <https://ecsa.citizen-science.net/about-us> [↑](#footnote-ref-4)
5. Galaxy Zoo accessed at http://www.galaxyzoo.org/. [↑](#footnote-ref-5)