Gender Imbalance and Spatiotemporal Patterns of Contributions to Citizen Science Projects: The Case of Zooniverse

Khairunnisa Ibrahim[1,2], Samuel Khodursky[3,4] and Taha Yasseri[1,5,6,7]*

[1]Oxford Internet Institute, University of Oxford, Oxford, United Kingdom, [2]Faculty of Arts and Social Sciences, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam, [3]Laboratory of Evolutionary Genetics and Genomics, The Rockefeller University, New York, NY, United States, [4]Department of Physics, University of Oxford, Oxford, United Kingdom, [5]School of Sociology, University College Dublin, Dublin, Ireland, [6]Geary Institute for Public Policy, University College Dublin, Dublin, Ireland, [7]Alan Turing Institute for Data Science and AI, London, United Kingdom

Citizen Science is research undertaken by professional scientists and members of the public collaboratively. Despite numerous benefits of citizen science for both the advancement of science and the community of the citizen scientists, there is still no comprehensive knowledge of patterns of contributions, and the demography of contributors to citizen science projects. In this paper we provide a first overview of spatiotemporal and gender distribution of citizen science workforce by analyzing 54 million classifications contributed by more than 340 thousand citizen science volunteers from 198 countries to one of the largest online citizen science platforms, Zooniverse. First we report on the uneven geographical distribution of the citizen scientist and model the variations among countries based on the socio-economic conditions as well as the level of research investment in each country. Analyzing the temporal features of contributions, we report on high "burstiness" of participation instances as well as the leisurely nature of participation suggested by the time of the day that the citizen scientists were the most active. Finally, we discuss the gender imbalance among online citizen scientists (about 30% female) and compare it with other collaborative projects as well as the gender distribution in more formal scientific activities. Online citizen science projects need further attention from outside of the academic community, and our findings can help attract the attention of public and private stakeholders, as well as to inform the design of the platforms and science policy making processes.

Keywords: citizen science, digital divide, circadian pattern, zooniverse, human dynamics

INTRODUCTION

The growth of online collaboration between scientists and the public on research projects in recent years have led to numerous discoveries at unprecedented rates in various fields of science. Among these are projects called *citizen science* which involve non-scientists in research work. Not only does online citizen science facilitate new breakthroughs [1], they do so thanks to voluntary contributions. The collective efforts of citizen scientists, as these volunteers are known, to one of the online citizen science platforms called the Zooniverse[1], was estimated to be wort United States $1.5 million in 2015 by

---

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

Sauermann and Franzoni [2] and represent the equivalent of 34 years' full time service [3] and 50 years of non-stop research [4]. In terms of research outputs, several Zooniverse projects have led several important discoveries, particularly in the field of astronomy. Blickhan, Trouille, and Lintott identify two types of discovery: intended ones which are known targets of certain research projects, and serendipitous ones, which arise from "our innate human ability to notice things that are out of the ordinary" [5].

It is not only the scientific community that benefits from citizen science; research has shown that volunteers gain learning opportunities, positive attitudes toward science and the ability to participate in research [6–9] as well opportunity for socializing and participating in an online community [7, 10]. This has proven especially pertinent in the wake of the Covid-19 pandemic, when the Zooniverse and other similar platform saw a surge in participation as many home-bound individuals turned to the Internet to do educational activities and pass their free time [11, 12]. In addition to making science more open and accessible, online citizen science accelerates research by leveraging human and computing resources [13–15], tapping into rare and diverse pools of expertize [2, 16], providing informal scientific education and training, motivating individuals to learn more about science [17]. Given these advantages, there has been a surge in research that seek to understand the activities and drives of both the scientists and the citizen scientists involved, many of them focused on understanding how best to support volunteer learning [18] and the learning outcomes of volunteers who engage in the Zooniverse whether on their own initiative or as part of their educational exercise [18, 19]. As Rotman et al. note, by "understanding the shared and unique motivations of these two groups" citizen science developers can design "the technical and social infrastructures needed to promote effective partnerships" [10].

This paper contributes toward those efforts by focusing on the geographical and temporal patterns of participation in the multi-project citizen science platform, the Zooniverse. Therefore, it adds to the current literature on online citizen science, which includes previous studies on motivations of citizen scientists [14, 15] and the frequency, productivity and intensity of volunteers' engagement [20]. Other studies have examined volunteers' preference between contributing a classification—defined as a unit of task completed in a citizen science project [4]—and participating in a discussion forum [8, 14]. Still others have looked at factors that promote or hinder participation, such as whether taking part in discussion forums have any influence on classification activities [21] and why volunteers only "drabble" or contribute occasionally, or even drop-out [7]. To date, however, there is little research on where volunteers come from, and when they are active, although Ponciano et al. [20], and Sauermann and Franzoni [2] do briefly examine temporal patterns with regard to frequency of activity over certain time periods.

Access to the technologies and skills to participate in online activities and for data creation and sharing is becoming ever more ubiquitous. Consequently, significant amount of research has been undertaken to discover the spatiotemporal patterns of these activities toward better understanding of human behavior, society and technology. A common finding among geographical studies of online activity is the disproportionate spatial distribution of both participation and data creation. These "stark core-periphery patterns" [22] reaffirm the dominance of resources and power in *the West*, and the lack of both in *the global South*. These are patterns that have been mapped and replicated across various online platforms, such as the photo-sharing site Flickr [23], the online encyclopedia Wikipedia [22], Twitter [24] and Google Maps [25]. Studies that examine where and what data are and are not available online reveal social realities, processes and divisions [21–23].

Temporal patterns of online activity have also been extensively studied in the field of *Human Dynamics*. Recent research include the editing behavior of editors on OpenStreetMap [26], Wikipedia [27, 28], message sending on online dating platforms [29], check-ins on the location-based social network Foursquare [30], phone calls [31], text messages [32], and traffic on the video-sharing platform YouTube [33]. These studies helped discern human tendencies at different time-scales. For instance, Noulas et al., by comparing Four square activity during the weekday and weekends [30], found distinct variance in when and where people check in. Similarly, Kaltenbrunner et al. discovered that people were considerably more active on the technology-news website Slashdot during the weekday compared to weekends [34]. Temporal patterns were also used to infer other information not readily available, such as the geographical location of editors of different language editions of Wikipedia [28].

According to a survey by the Zooniverse, about 30% of their respondents are from the United Kingdom, 35% from the United States and the rest from other parts of the world [35]. Our paper investigates the patterns of activity in the Zooniverse between 2009 and 2013 by analyzing 54 million classifications made in 17 different citizen science projects by more than 340,000 volunteer citizen scientists. Our focus is on the spatial and temporal patterns of activity; we are interested in where volunteers contribute to citizen science from, and during what times of the day and week they are most active. We seek to answer three basic questions:

1) Where do volunteers contribute from, and when?
2) What are some of the reasons for these patterns?
3) What is the gender distribution among the contributors?

MATERIALS AND METHODS

Data Overview

This study utilizes a dataset of 54 million classifications made in 17 different citizen science projects in the Zooniverse made between November 2009 and June 2013 (see Supplementary Information for a list and timeline of the projects). The classifications were produced by more than 340,000 volunteers from 198 different countries. Each classification record includes a unique classification id, the volunteer's distinct user-id and their approximate geographical location, the timestamp of the

TABLE 1 | Socioeconomic variables Adapted from the World Bank[3].

Variable	Description
Population	Includes all residents "regardless of legal status or citizenship" but excludes refugees who are not "permanently settled in the country of asylum." The latter are generally considered part of the population of their country of origin. The values are mid-year estimates of the population
GDP per capita	The gross domestic product divided by the midyear population. GDP is the "sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products." Data are in expressed current United States dollars
Internet users per 100 people	"People with access to the worldwide network"
Gross primary school enrollment (%) (PSE)	Total is the total enrollment in primary education, regardless of age, expressed as a percentage of the population of official primary education age. GER can exceed 100% due to the inclusion of over-aged and under-aged students because of early or late school entrance and grade repetition
Gross secondary school enrollment (%) (SSE)	The total is the total enrollment in secondary education (all programmes), regardless of age, expressed as a percentage of the population of official secondary education age. GER can exceed 100% due to the inclusion of over-aged and under-aged students because of early or late school entrance and grade repetition

TABLE 2 | Scientific culture variables adapted from the World Bank.

Variable	Description
Research and development expenditure (% of GDP)	Figures indicate the "current and capital expenditures (both public and private) on creative work undertaken systematically to increase knowledge, including knowledge of humanity, culture, and society, and the use of knowledge for new applications" and includes "basic research, applied research, and experimental development."
Researchers in R&D (per million people)	Figures indicate the number of professional involved in "the conception or creation of new knowledge, products, processes, methods, or systems and in the management of the projects concerned," including postgraduate PhD students
Scientific and technical journal articles	Figures indicate the number of scientific and engineering articles published in the fields of physics, biology, chemistry, mathematics, clinical medicine, biomedical research, engineering and technology, and earth and space sciences

classification and the project to which the classification is made. The records also include gender information. However, volunteers are not required to identify as either male or female, so gender information in the dataset is derived from a separate analysis of the first names that they used to register using the Gender API[2,3] by Zooniverse team. The complete dataset is available at https://doi.org/10.5281/zenodo.583182.

Geographical Analysis

Country data on citizen science is obtained by extracting the numbers of unique volunteers and the aggregate number of classifications made per country. The dataset uses two-letter codes to identify countries (e.g. GB for United Kingdom, US for United States). Country-level socioeconomic data are obtained from the World Bank. Four types of socioeconomic variables are taken from the World Bank database: population, income (GDP per capita), Internet connectivity (number of Internet users) and education (primary and secondary school enrollment). **Table 1** provides an overview of the socioeconomic variables.

The three variables for scientific culture are also obtained from the World Bank. An overview is provided in **Table 2**.

Considering most of the variables vary orders of magnitude among different countries, and are often best described by fat-tailed distributions such as log-normal or scale-free, we transform

them logarithmically and work with the Log transformed variables in the next section.

Temporal Analysis

The daily cycles of activity are examined for the 20 countries most active in the Zooniverse. The standardized time is taken to be the Greenwich Meridian Time (GMT). All other countries' data were adjusted to the GMT. Several countries, most notably the United States, Canada, Australia, the Russian Federation and Brazil, span several time zones. For countries with large geographical expansions yet a single time zones, such as the Russian Federation and Brazil the decision was taken to adjust the data using an average of +3 and −3 h respectively. For the United States, Canada and Australia, the data for various regions are aggregated and adjusted according to their various time zones. The hourly contributions by region are then added to represent the countries' overall activity per hour.

The circadian rhythms for each country ($Circadian_{\Delta t}$) are obtained by dividing the number of classifications made at any particular hour ($Classifications_{\Delta t}$) with the total number of classifications made within that country ($Classifications$).

$$Circadian_{\Delta t} = \frac{Classifications_{\Delta t}}{Classifications}$$

Meanwhile, the universal curve is the ratio of the total number of classifications per hour made by the 20 countries over the overall number of classifications from the 20 countries made throughout the 24-h period.
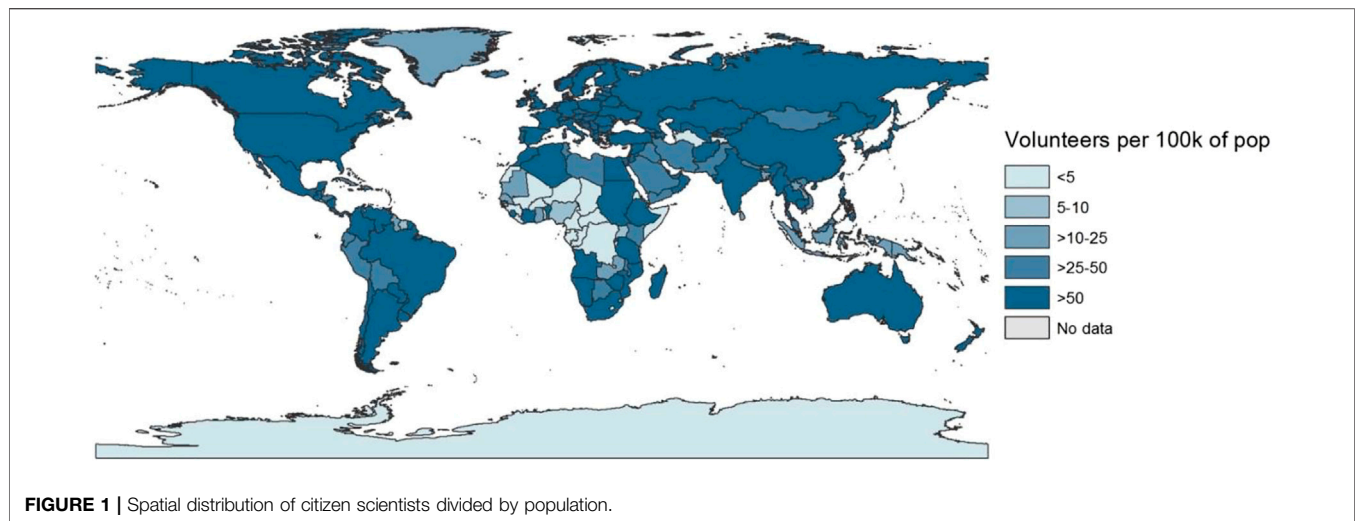
FIGURE 1 | Spatial distribution of citizen scientists divided by population.

$$Universal = \frac{\sum_{20} Classifications_{\Delta t}}{\sum_{20} Classifications}$$

RESULTS

Spatial Patterns of Citizen Science Activity

The Zooniverse volunteers are distributed worldwide, but the rate and preferences of their engagement vary. The majority of volunteers and activity come from North America and Western Europe. Even when controlled for population, the same pattern holds (**Figure 1**), indicating that the majority of citizen scientists are to be found in Western, developed countries, regardless of the numbers of people within each country. These spatial patterns of engagement and activity reflect other geographical trends in online activity. Similar patterns are observed on Flickr [23], the online encyclopedia Wikipedia [22], Twitter [24] and Google Maps [25]. As with these other online platforms for creating and sharing knowledge, The Zooniverse receives intense bursts of activity in the global North, while other parts of the world, particularly the African continent, remain seemingly disconnected from these platforms.

What are the reasons for the uneven geographies of engagement in citizen science? We considered various factors, which fall into two categories: **socio-economic** and **scientific culture**.

Socioeconomic factors can indicate the propensity of the average individual in a country to engage in citizen science. Citizen science requires the voluntary participation of many people, which is more readily provided if the volunteers are relatively well-off, have easy access to the Internet, and have ample free time. Our analysis of the temporal patterns of activity, shared later in the paper, also point to citizen science as being primarily a leisurely pursuit, taken up after the normal workday is over. Finally, the nature of the citizen science projects indicates at least a certain level of education on the part of the volunteer, in order for them to understand and be interested in the projects.

A country's *scientific culture* can also be a good indicator for its residents' involvement in citizen science. Countries that spend more on research and development, cultivates more researchers, and produce more scientific outputs signal a vibrant scientific community that encourages experimentation and learning. This positive inclination toward knowledge production may also extend to those outside the formal institutions of science, for even when individuals move into non-scientific careers, their early education will most likely have included science subjects.

Using the latest available indicators from the World Bank (http://data.worldbank.org), we analyzed the relationship between engagement in citizen science, defined as the number of contributions made from each country. The socioeconomic variables include population, GDP per capita, Internet connectivity (the number of Internet users per 1,000 people) as well as primary school enrollment. Countries' scientific cultures are measured by the expenditure on research and development initiatives, the number of researchers employed in various sectors and the overall number of publications in science and technology journals and books.

The socioeconomic analyses confirm a positive link between citizen science and socioeconomic indicators; that is, the wealthier and connected a country is and the more educated its residents, the higher the rate of participation and contribution. **Table 3** shows the Pearson correlation results of the socioeconomic analysis (controlling for population). On the whole, most of the variables have a positive relationship with citizen science engagement, although the magnitude of these relationship varies.

Having these primary results in hand, we conducted tests on a number of regression models on socioeconomic indicators. Our analysis of just over 200,000 cases, each representing a volunteer from one of 90 countries (**Table 4**), revealed that over 85% of the variation in classifications made per country can be explained by the combination of all five factors: population, GDP per capita, Internet users per 100 people, and gross school enrollment (Model 1). When the educational variables are excluded (Model 2), the adjusted R^2 is still fairly high at 0.80. However,

TABLE 3 | Bivariate correlations (Pearson correlation) among citizen science and socioeconomic variables among the countries under study. The highest correlation is seen between the number of Internet users and the number of contributions to the Zooniverse projects. The independent variables are the percentage of Internet users, Gross Domestic Product per capita, and the percentage of adults with completed primary education. For more details see the Methods section.

	%Internet users	GDP per capita	%Primary education
Volunteers (log)	0.72	0.58	0.51
Contributions (log)	0.80	0.66	0.41

TABLE 4 | Models of socioeconomic factors for the geographies of citizen science. The dependent variable is the logarithm of the number of contributions from each country. Different models consider different sets of independent variables.

	Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE
Constant	-12.96^a	0.028	-11.60^a	0.021	—	—
Population (log)	0.856^a	0.002	0.935^a	0.002	—	—
GDP per capita	0.328^a	0.007	-0.465^a	0.007	—	—
Internet users	4.944^a	0.018	6.786^a	0.019	7.010^a	0.012
SSE	-0.014^a	0.000	—	—	-0.037^a	0.000
PSE	0.031^a	0.000	—	—	-0.008^a	0.000
Adjusted R^2	0.856		0.801		0.658	

$^a p < 0.01.$

when only Internet connectivity and school enrollment are considered (Model 3), the outcome is still considerable at 0.66. These results indicate that taken separately, each factor has varying rates of influence on citizen science, but in tandem they explain the major bulk of engagement in the Zooniverse. What is notable among these results is the negative sign for secondary education coefficients which indicates that when population and the number of internet users are controlled for, higher level of education is associated with lower participation rate in comparison to the rest of the countries. This is further discussed below in the context of citizen science platforms being empowering tools for groups of society who would otherwise not engage in formal education or scientific research.

Apart from this overall trend, to be able to identify the outliers (overperformers and underperforms) compared to the model prediction, residuals from the Model 1 for each country are shown in **Figure 2**. Where we see under-representation from North-American countries and over-representation from some of the developing countries in Asia and Africa.

Our analysis of the influence of scientific culture on citizen science engagement indicate that more scientifically active countries are more likely to be involved in citizen science. We ran a correlation test of over 21,000 cases from 56 countries with available science data—spending on research and development, the number of researchers and publications—from the World Bank. **Table 5** shows the result of this analysis. All three variables show a positive correlation, with R&D budget and science publications being highly correlated at over 70 and 80% respectively.

A linear combination of the three variables contribute to 73.2% of the variation in citizen science activity (**Table 6**). The results indicate that while research spending and the

availability of human resource in science are necessary, it is the outcomes of these investment and resources that have greatest impact on involvement in citizen science. In other words, researchers and funding agencies should aim for increasing their publication record as this research shows that greater publications can influence more involvement in science even outside academia and industry.

Temporal Patterns of Citizen Science Activity

Distribution of Time Between Contributions to Zooniverse

Many social and complex systems variables exhibit fat-tailed distributions. The wealth of people [36], the populations of cities [37], and the number of citations to papers [38], are all quantities that are distributed very far from a normal distribution, with few instances of very large values an many instances of small values. The distribution of time between successive contributions by each citizen scientist to Zooniverse is shown in **Figure 3**. The fat-tailed distribution of the time intervals is in line with previous reports on similar patterns among Wikipedia edits [27], emails sent/received [39], phone calls made [40], and many more examples of *unscheduled* human activities. This type of distribution of time-intervals is a fingerprint of *bursty* behavior, in which sessions of many activities are often followed by long waiting times between sessions [41].

Circadian Patterns of Activity

We extracted the number of classifications made by volunteers in 1-h windows throughout the day, divided the values for each window to the total number of contributions throughout the 24-h period, and plotted the circadian charts of activity. **Figure 4** shows the circadian patterns of the 20 most active countries in Zooniverse. We repeated the above analysis for the day of week, however, the fluctuations among days of the week do not seem to be following a significant pattern (**Supplementary Material**).

We also calculated the average temporal patterns for all the selected countries to produce the universal curve shown in **Figure 5**. This curve represents the general pattern of activities over the 24-h period. It shows that overall, activity in the Zooniverse starts to pick up gradually after 5 am. The number of classifications continue to increase throughout the day, with small dips in activity in the afternoon (around 2 pm) and early evening (around 6 pm). The rate of activity peaks at 9 pm, and falls in volume in the following hours.

This universal circadian pattern for the 20 selected countries indicates that while classifications were made at all hours of the
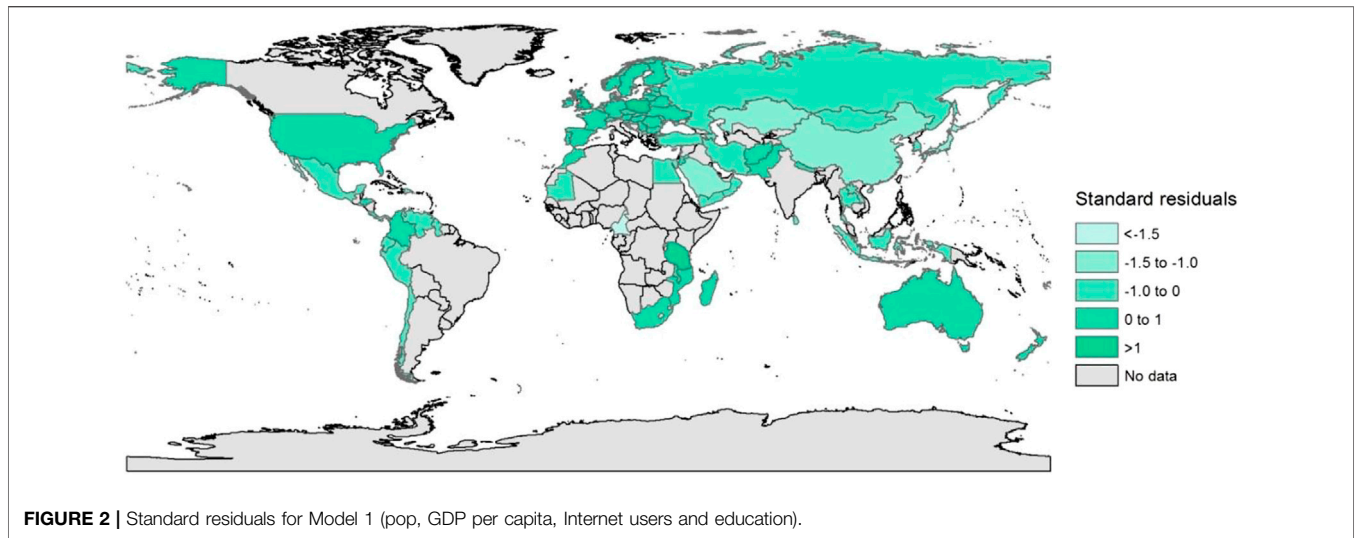
FIGURE 2 | Standard residuals for Model 1 (pop, GDP per capita, Internet users and education).

TABLE 5 | Pearson-moment correlation among citizen science variables and science indicators. The highest correlation is between the number of contributions and the number of S&T publications.

	R&D budget (log)	# Of researchers (log)	S&T publications (log)
Volunteers (log)	0.75	0.48	0.80
Contributions (log)	0.78	0.56	0.82

TABLE 6 | Model of Scientific culture factors for the geographies of citizen science. The dependent variable is the number of contributions from each country.

	B	SE	95% CI	
Intercept	1.545[a]	0.033	1.481	1.609
Budget (log)	-0.298[a]	0.005	1.124	1.140
Researchers (log)	0.832[a]	0.004	-0.308	−0.288
S&T publications (log)	1.132[a]	0.004	0.824	0.840
Adjusted R²	0.732			

[a]$p < 0.01$.

day, most of the activity took place at night. The peak hour of 9 pm is typically leisure time for most households. Assuming that most people only have about an hour or less to contribute to the Zooniverse, this pattern corresponds with findings from a survey on the Zooniverse community, in which the majority of respondents selected the option "Only when I have spare time" in answer to the query on when they tend to classify [35]. It also reflects the tendency of people to collectively contribute to a pursuit *en masse* when they are in possession of free time and the resources to do so, a phenomenon that Shirky calls "cognitive surplus" [42].

Time of First Contributions

Two-thirds of volunteers to the Zooniverse make only one classification and do not return [20]. Given that such a large proportion of volunteers do not continue to contribute, it is important to know when most people first come to the project, so
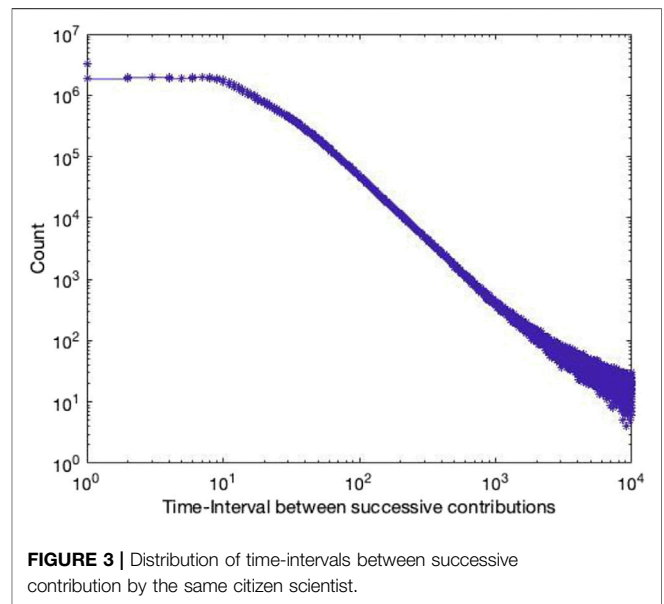
FIGURE 3 | Distribution of time-intervals between successive contribution by the same citizen scientist.

that project managers can identify measures to retain participation, and schedule them at the appropriate times. We investigated time of first contribution for the 20 countries and found that show that on average there is a marked concentration of volunteers making an initial contribution at 9 pm (**Figure 6**). This corresponds with the universal
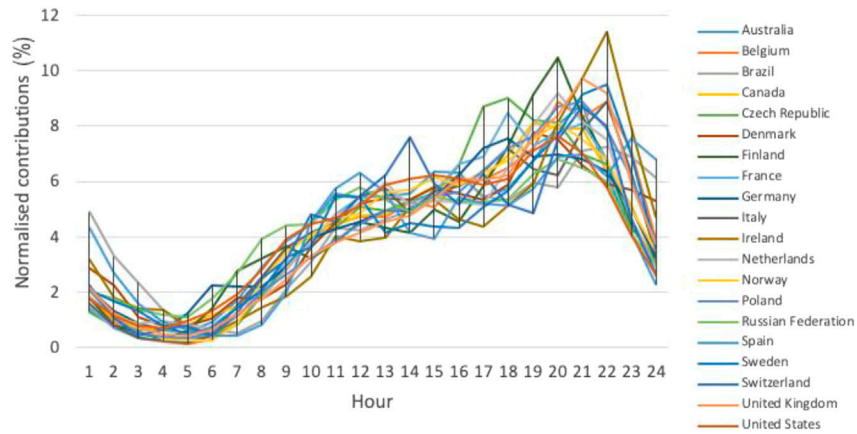
FIGURE 4 | Circadian rhythms of contributions from 20 countries to the Zooniverse.
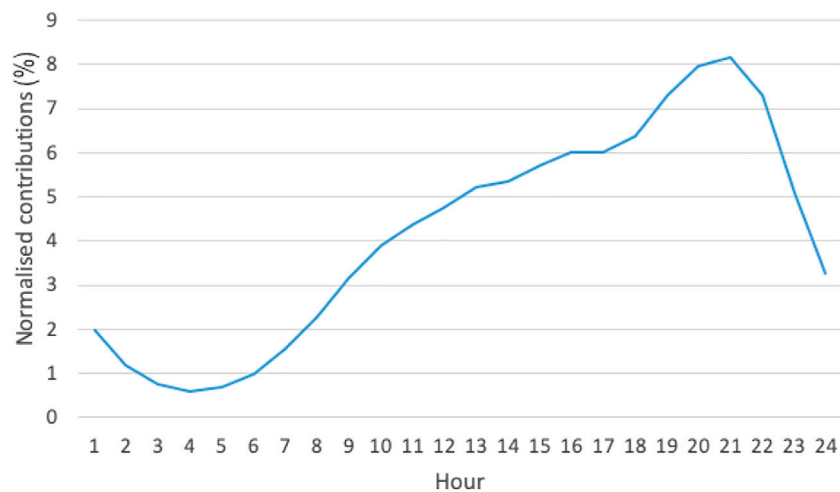

FIGURE 5 | The universal curve showing the average contributions per hour for the 20 countries.

circadian pattern above, where 9 pm is shown to experience the highest activity rates in the 20 countries. Here we see a sharper peak around 9 pm compared to the overall activity curve. Citizen science in general seems to be a leisure activity, undertaken when individuals have time to spare. The prevalence of nighttime activity might also be due to individuals first learning about the various projects through social media, whether directly from the Zooniverse or others [2], when they become connected to these informational outlets after work and dinner.

Gender Balance Across Projects

Overall the gender imbalance among the contributors to the Zooniverse projects is large. Among all the contributions, where the gender of the contributor could be determined with high confidence, only about 30 per cent are committed by female contributors. Whilst the share of female citizen scientists' contributions is considerably larger than similar collaborative

projects such a Wikipedia [43] and free/libre/open source software development [44], there is still a large gap of about 40% between the two main genders. **Figure 7** shows the percentage of female volunteers for each country (see Supplementary Information for the full list).

Female volunteers are underrepresented in most countries. In many countries, women make up less than one-third of number of volunteers whose gender is known. The female ratio of participation in the United Kingdom and Australia, for instance, is 25%, while the figures for United States, Canada and Germany are between 27 and 30%. What is notable here is we see no clear correlation between these ratios and the estimated percentage of female authors in each country [45] with India (44% female), Iran (40% female), and Ukraine (38% female) among the countries with the larger share of female contributors in Zooniverse.

An explanation for this deviation can be that in countries where opportunities for formal participation in knowledge
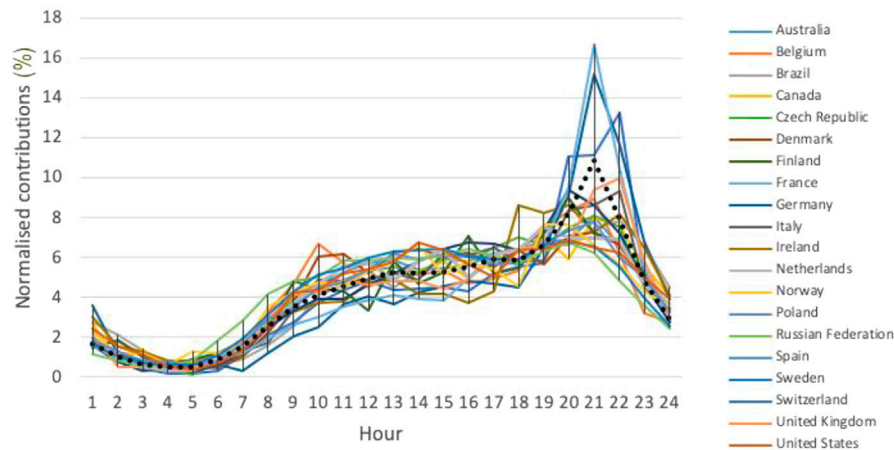
FIGURE 6 | Time of first contribution, when volunteers make their first, and for many their only, classification.
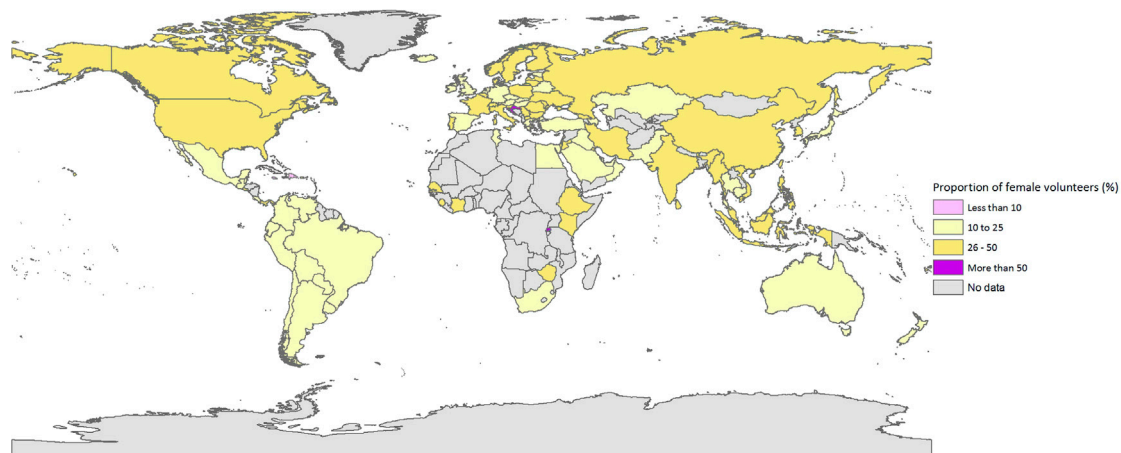

FIGURE 7 | Ratio of female contributors to Zooniverse projects.

generation activities are sparse for women, there is a larger interest among female citizen scientists to participate through the informal and more accessible environment of citizen science projects.

Another pattern reported for the gender imbalance among formal scientists is the differences between different disciplines with life sciences and social sciences hosting more female scientists compared with physical and mathematical sciences [46]. To test if we see a similar pattern among citizen scientists, we calculated the share of female contributors per project. The top three projects each with more than 50% female contributors are *Snapshot Serengeti*, *Whale FM*, and *Take notes from Nature*, all related to wildlife and nature. On the other hand, the two project with the least number of female scientists (less than 20%) are *Planet Hunters* and *Galaxy Zoo*, both in the field of astronomy. It is notable that whilst the overall pattern of gender imbalance among formal scientists across countries is not preserved among

citizen scientists, the uneven distribution of genders among different fields of science are replicated. According to an NSF report, the share of women in engineering is 13%, computer and mathematical sciences 25%, but they are well-represented in the social sciences 58%, and biological and medical sciences 48%.

DISCUSSION AND CONCLUSION

The growing popularity of citizen science online has a number of implications. Online citizen science acts as a bridge between the science community and the public, as scientific institutions have typically kept a distance from the larger society and scientific processes are obscured behind black boxes and ivory walls. It provides volunteers, who normally will not have access to scientific data and research projects, opportunity to become involved in knowledge production as long as they have access

to computers and Internet connection. In the process, online citizen science helps participants increase their scientific literacy and develop positive attitudes toward science [9]. Indeed, "carefully designed citizen-science projects can be successful environments for increasing adult knowledge of factual science" [6]. Additionally, by leveraging on knowledge creation and sharing models of crowdsourcing [47] and peer-production [48], online citizen science bypasses the problems of funding and human resource shortages. It is by harnessing crowd power and engaging in these new forms of scientific collaboration that online citizen science can help facilitate and accelerate scientific discovery.

Despite the constraints of the study, using data from only one specific online citizen science platform taken from the Zooniverse's early years of activity, and the fact that English is the main medium of communication, this paper provides several important takeaways. These may be of interest and use to online citizen science facilitators, educators and even those who participate in the practice of online citizen science, by leveraging on information about where, when and how volunteers take part in this type of scientific research.

Our research examined the spatiotemporal patterns of activity in online citizen science on the Zooniverse platform. We found that volunteers are unevenly distributed around the world. They are found predominantly in North America and Europe, and in small numbers in the rest of the world. These variations can be explained by socio-economic factors. Over 80 per cent of the variations are due to the combined effects of a country's population, wealth (GDP per capita), Internet connectivity and rate of school enrollment. National emphasis on science, as represented by spending on research and development, the number of researchers employed and publication records, also have an effect on citizen science, accounting for 73% of variation in citizen science activity. By identifying these trends and the factors that produce them, policy makers, scientific institutions as well as citizen science developers can pinpoint strategies to increase involvement in science, one of them through promoting participation in informal science projects such as citizen science.

We also studied the temporal patterns of activity on the Zooniverse, and the results demonstrate that volunteers tend to be most active during the evening, which corresponds with a typical household's leisure time. First time volunteers are also most likely to start classifying during the same time intervals of the day. This effect might slightly change since the pandemic had normalized working from home, blurring the division between work and leisure. This situation also increases time spent online and this has led to increased traffic to and activity on online citizen science platforms, as many have turned to doing research for educational and leisure purposes. Early analysis report that the Zooniverse alone received over 200,000 volunteers in a week early on in the pandemic, citing educators and parents among those who turned to the platform to provide educational activities for home-bound children [11, 12].

Studying the gender imbalance among contributors, we see a higher participation from female citizen scientists compared to Wikipedia and open software development projects, however, still a large gap between the number of male and female contributors is present. We observe smaller gaps in countries with larger gender imbalance in more formal research professions, which suggest citizen science projects can practically play the role of an informal channel for females with strong interest in science where more formal channels are less accessible. This is an encouraging finding, because while online citizen science platforms cannot replace professional training, it can provide allow entry into spaces formerly restricted to a male-dominated scientific elite, and encourage more female involvement in the field. The participation of the children of volunteers and schoolchildren in the Zooniverse projects also serve to provide an welcoming atmosphere and initiation into scientific research for both boys and girls.

The work reported here presents its own limitations. Most notably, the study is limited to the projects all hosted on one citizen science platform, Zooniverse. Even though Zooniverse is among the most prominent citizen science platforms, there are several other examples to which certain features of Zooniverse might not generalize. In addition, the dataset studied here is limited to the "early stage" of Zooniverse life-time and the current state of the art and some of the reported patterns might be different to what we report. Nevertheless, we believe certain tendencies and patterns are universally robust. The selected timeframe is partially determined by the fact that we did want to avoid non-stationary effects and drastic changes that came about at later stages. The other limitation of the current study comes from the omission of language analysis. Several Zooniverse projects have been available only in English and this might have been a barrier against participation for many volunteers. A future work could provide a more detailed analysis in which the language proficiency of users from different countries as well as the availability of the projects in languages beyond English would be considered. Finally, our assumption here is that the main motivations for participation in citizen science projects are 1) contribution to generation of knowledge, and 2) self-education. Whilst the reviewed literature supports these assumptions, one should keep in mind that various motivational patterns have been introduced and discussed in certain case studies ([49]).

Our findings represent initial forays into research to understand the influence of time and place on involvement in scientific knowledge production. Further studies are needed to discover not just the national-level factors for varying levels of online citizen science involvement, but also variables that affect participation on a micro-scale. Doing so would help identify ways to best tap into the vast reserves of interest, time and effort and channel them toward scientific pursuits that the public can contribute to.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found online at: https://doi.org/10.5281/zenodo.583182.

ETHICS STATEMENT

AUTHOR CONTRIBUTIONS

KI, SK, and TY Analyzed the data. TY designed the study and secured the funding. KI and TY drafted the manuscript. All authors contributed to writing the manuscript and gave final approval for publication.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2021.650720/full#supplementary-material

REFERENCES

1. Nielsen M. *Reinventing Discovery: The New Era of Networked Science*. USA: Princeton University Press (2011). doi:10.1515/9780691202853

2. Sauermann H, and Franzoni C. Crowd Science User Contribution Patterns and Their Implications. *Proc Natl Acad Sci USA* (2015) 112:679–84. doi:10.1073/pnas.1408907112

3. Cox J, Oh EY, Simmons B, Lintott C, Masters K, Greenhill A, et al.Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects. *Comput Sci Eng* (2015) 17:28–41. doi:10.1109/MCSE.2015.65

4. Simpson R, Page KR, and De Roure D. Zooniverse. In: WWW 2014 Companion - Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, April 4–7, 2014. New York, USA: Association for Computing Machinery, Inc (2014). p. 1054. doi:10.1145/2567948.2579215–

5. Blickhan S, Trouille L, and Lintott CJ. Transforming Research (And Public Engagement) through Citizen Science. *Proc IAU* (2018) 14:518–23. doi:10.1017/S174392131900526X

6. Brossard D, Lewenstein B, and Bonney R. Scientific Knowledge and Attitude Change: The Impact of a Citizen Science Project. *Int J Sci Educ* (2005) 27: 1099–121. doi:10.1080/09500690500069483

7. Eveleigh A, Jennett C, Blandford A, Brohan P, and Cox AL. Designing for Dabblers and Deterring Drop-Outs in Citizen Science. In: Conference on Human Factors in Computing Systems - Proceedings, Toronto, Canada, April 26–May 1, 2014. New York, NY, USA: Association for Computing Machinery (2014). p. 2985–94. doi:10.1145/2556288.2557262

8. Luczak-Roesch M, Tinati R, Simperl E, Kleek MV, Shadbolt N, and Simpson R. Why Won't Aliens Talk to Us? Content and Community Dynamics in Online Citizen Science. In: Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM, Ann Arbor, MI, June 1–4, 2014. Palo Alto, CA, USA: The AAAI Press (2014). p. 315–24.

9. Price CA, and Lee H-S. Changes in Participants' Scientific Attitudes and Epistemological Beliefs during an Astronomical Citizen Science Project. *J Res Sci Teach* (2013) 50:773–801. doi:10.1002/tea.21090

10. Kloetzer L, Schneider D, Jennett C, Iacovides I, Eveleigh A, Cox A, Gold M, Kloetzer L, Schneider D, Jennett C, et al. Learning by Volunteer Computing, Thinking and Gaming: What and How Are Volunteers Learning by Participating in Virtual Citizen Science? Proc 2013 Eur Res Conf Netw Access, Learn Careers Identities (pp 73–92.). Eur Soc Res Educ Adults Linköping, Sweden (2013) Available at: https://liu.se/esrea2013?l=en (Accessed January 5, 2021).

11. Cohen K. Citizen Science in a Pandemic: A Fleeting Moment or New Normal? – Humanities & Social Change (2020). Available at: https://hscif.org/citizen-science-in-a-pandemic-a-fleeting-moment-or-new-normal/ (Accessed April 11, 2021). doi:10.1126/science.abg6035

12. Covid-19 Can't Stop Citizen Science (2020). Available at: https://undark.org/2020/04/17/covid-19-citizen-science/ (Accessed April 11, 2021).

13. Hand E. Citizen Science: People Power. *Nature* (2010) 466:685–7. doi:10.1038/466685a

14. Tsvetkova M, Yasseri T, Meyer ET, Pickering JB, Engen V, Walland P, et al.Understanding Human-Machine Networks. *ACM Comput Surv* (2017) 50:1–35. doi:10.1145/3039868

15. Eide AW, Pickering JB, Yasseri T, Bravos G, Følstad A, Engen V, et al.Human-machine Networks: Towards a Typology and Profiling Framework. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Cham, Switzerland: Springer International Publishing (2016). p. 11–22. doi:10.1007/978-3-319-39510-4_2

16. Yasseri T, and Kertész J. Value Production in a Collaborative Environment. *J Stat Phys* (2013) 151:414–39. doi:10.1007/s10955-013-0728-6

17. Straub VJ, Tsvetkova M, and Yasseri T. The Cost of Coordination Can Exceed the Benefit of Collaboration in Performing Complex Tasks. (2020). Available at: http://arxiv.org/abs/2009.11038 (Accessed January 7, 2021).

18. Jackson CB, Østerlund C, Crowston K, Harandi M, and Trouille L. Shifting Forms of Engagement: Volunteer Learning in Online Citizen Science. *Proc ACM Hum.-Comput Interact* (2020) 4:1–19. doi:10.1145/3392841

19. Herodotou C, Aristeidou M, Miller G, Ballard H, and Robinson L. What Do We Know about Young Volunteers? an Exploratory Study of Participation in Zooniverse. *Citiz Sci Theor Pract* (2020) 5:1–14. doi:10.5334/cstp.248

20. Ponciano L, and Brasileiro F. Finding Volunteers' Engagement Profiles in Human Computation for Citizen Science Projects. *HCj* (2014) 1:1. doi:10.15346/hc.v1i2.12

21. Tinati R, Simperl E, Luczak-Roesch M, Kleek MV, and Shadbolt N. Collective Intelligence in Citizen Science – A Study of Performers and Talkers. In, Proceedings of the 2014 ACM conference on Web science, Bloomington, June 23–26, 2014. New York, NY, USA: ACM (2014). p. 295–296. doi:10.1145/2615569.2615651

22. Graham M, Hogan B, Straumann RK, and Medhat A. Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty. *Ann Assoc Am Geogr* (2014) 104:746–64. doi:10.1080/00045608.2014.910087

23. Graham M, Hale S, and Stephens M. *Geographies of the World's Knowledge* (2011).

24. Graham M, Hale SA, and Gaffney D. Where in the World Are You? Geolocation and Language Identification in Twitter. *The Prof Geographer* (2014) 66:568–78. doi:10.1080/00330124.2014.907699

25. Graham M, and Zook M. Augmented Realities and Uneven Geographies: Exploring the Geolinguistic Contours of the Web. *Environ Plan A* (2013) 45: 77–99. doi:10.1068/a44674

26. Yasseri T, Quattrone G, and Mashhadi A. Temporal Analysis of Activity Patterns of Editors in Collaborative Mapping Project of Openstreetmap. In:

Proceedings of the 9th International Symposium on Open Collaboration, WikiSym + OpenSym 2013, Hong Kong, China, August 5–7, 2013. New York, NY, USA: ACM (2013). p. 1–4. doi:10.1145/2491055.2491068

27. Yasseri T, Sumi R, Rung A, Kornai A, and Kertész J. Dynamics of Conflicts in Wikipedia. *PLoS One* (2012) 7:e38869. doi:10.1371/journal.pone.0038869

28. Yasseri T, Sumi R, and Kertész J. Circadian Patterns of Wikipedia Editorial Activity: A Demographic Analysis. *PLoS One* (2012) 7:e30091. doi:10.1371/journal.pone.0030091

29. Dinh R, Gildersleve P, and Yasseri T. Computational Courtship: Understanding the Evolution of Online Dating through Large-Scale Data Analysis. 2018).

30. Noulas A, Scellato S, Mascolo C, and Pontil M. An Empirical Study of Geographic User Activity Patterns in Foursquare. In: Proc Fifth Int AAAI Conf Weblogs Soc Media, 11 (2011). p. 570–3. papers3://publication/uuid/557455DB-AC4A-4C73-968A-31E7A663BC4E.

31. Aledavood T, López E, Roberts SGB, Reed-Tsochas F, Moro E, Dunbar RIM, et al.Daily Rhythms in Mobile Telephone Communication. *PLoS One* (2015) 10:e0138098. doi:10.1371/journal.pone.0138098

32. Aledavood T, Lehmann S, and Saramäki J. Digital Daily Cycles of Individuals. *Front Phys* (2015) 3:73. doi:10.3389/fphy.2015.00073

33. Gill P, Arlitt M, Li Z, and Mahanti A. Youtube Traffic Characterization. In: Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC. New York, New York, USA: . ACM Press (2007). p. 15–28. doi:10.1145/1298306.1298310

34. Kaltenbrunner A, Gómez V, Moghnieh A, Meza R, Blat J, and López V. Homogeneous Temporal Activity Patterns in a Large Online Communication Space. In: CEUR Workshop Proceedings (2007). Available at:. doi:10.1109/laweb.2007.4383159http://arxiv.org/abs/0708.1579 (Accessed January 5, 2021).

35. TTFNROB. Who Are the Zooniverse Community? We Asked Them... | Zooniverse. *Zooniverse Blog* (2015). Available at: https://blog.zooniverse.org/2015/03/05/who-are-the-zooniverse-community-we-asked-them/ (Accessed January 5, 2021).

36. Davies JB, and Shorrocks AF. Chapter 11 the Distribution of Wealth. *Handb Income Distrib* (2000) 1:605–75. doi:10.1016/S1574-0056(00)80014-7

37. Eeckhout J. Gibrat's Law for (All) Cities. *Am Econ Rev* (2004) 94:1429–51. doi:10.1257/0002828043052303

38. Goldberg SR, Anthony H, and Evans TS. Modelling Citation Networks. *Scientometrics* (2015) 105:1577–604. doi:10.1007/s11192-015-1737-9

39. Goh K-I, and Barabási A-L. Burstiness and Memory in Complex Systems. *Europhys Lett* (2008) 81:48002. doi:10.1209/0295-5075/81/48002

40. Jo H-H, Karsai M, Kertész J, and Kaski K. Circadian Pattern and Burstiness in mobile Phone Communication. *New J Phys* (2012) 14:013055. doi:10.1088/1367-2630/14/1/013055

41. Barabási A-L. The Origin of Bursts and Heavy Tails in Human Dynamics. *Nature* (2005) 435:207–11. doi:10.1038/nature03459

42. Harvey K. Cognitive Surplus. in: *Encyclopedia Of Social Media and Politics*. London, UK: SAGE Publications, Inc. (2014). doi:10.4135/9781452244723.n113

43. Antin J, Yee R, Cheshire C, and Nov O. Gender Differences in Wikipedia Editing. In: Proceedings of the 7th International Symposium on Wikis and Open Collaboration - WikiSym '11, Mountain View, CA, October 3–5, 2011, 11. New York, New York, USA: ACM Press (2011). doi:10.1145/2038558.2038561

44. Barcomb A, Grottke M, Stauffert J-P, Riehle D, and Jahn S. How Developers Acquire FLOSS Skills. *IFIP Advances in Information and Communication Technology*. Cham, Switzerland: Springer (2015). p. 23–32. doi:10.1007/978-3-319-17837-0_3

45. Holman L, Stuart-Fox D, and Hauser CE. The Gender gap in Science: How Long until Women Are Equally Represented?. *PLOS Biol* (2018) 16:e2004956. doi:10.1371/journal.pbio.2004956

46. Huang J, Gates AJ, Sinatra R, and Barabási A-L. Historical Comparison of Gender Inequality in Scientific Careers across Countries and Disciplines. *Proc Natl Acad Sci USA* (2020) 117:4609–16. doi:10.1073/pnas.1914221117

47. Brabham DC. *Crowdsourcing*. Cambridge, MA: Mit Press (2013). doi:10.7551/mitpress/9693.001.0001

48. CropfBenkler RAY, and Benkler Y (2006). The Wealth of Networks: How Social Production Transforms Markets and Freedom. New Haven and London: Yale University Press. 528 Pp. $40.00 (Papercloth). *Soc Sci Comput Rev* (2008) 26:259–61. doi:10.1177/1084713807301373

49. Vohland K, Land-Zandstra A, Ceccaroni L, Lemmens R, Perelló J, Ponti M, et al.The Science of Citizen Science (2020). doi:10.2139/ssrn.3511088 CrossRef Full Text