



Collaboration and Performance of Citizen Science Projects Addressing the Sustainable Development Goals

COLLECTION:
CONTRIBUTIONS OF
CITIZEN SCIENCE TO
THE UN SDGS

RESEARCH PAPER

CAMILLE MASSELOT

RATHIN JEYARAM

RAPHAEL TACKX

JOSE LUIS FERNANDEZ-MARQUEZ

FRANÇOIS GREY

MARC SANTOLINI

*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

Measuring the progress towards the Sustainable Development Goals (SDGs) requires the collection of relevant and reliable data. To do so, Citizen Science can provide an essential source of non-traditional data for tracking progress towards the SDGs, as well as generate social innovations that enable such progress. At its core, citizen science relies on participatory processes involving the collaboration of stakeholders with diverse standpoints, skills, and backgrounds. The ability to measure these participatory processes is therefore key for the monitoring and evaluation of citizen science projects and to support the decisions of their coordinators. Here, we show that the monitoring of social interaction networks provides unique insights on the participatory processes and outcomes of citizen science projects. We studied fourteen early-stage citizen science projects that participated in an innovation cycle focused on SDG 13, Climate Action, as part of the Crowd4SDG project. We implemented a monitoring strategy to measure the collaborative profiles of citizen science teams. This allowed us to generate dynamic interaction networks across complementary dimensions, making visible both formal and informal interactions associated with the division of labor, collaborations, advice seeking, and communication processes of the projects during their development. Leveraging jury evaluation data, we showed that while team composition and communication are associated with project quality, measures of collaboration and activity are associated with engagement quality. Overall, monitoring social interaction dynamics helps build a more comprehensive picture of participatory processes, which is of importance for guiding citizen science projects and for designing initiatives leveraging citizen science to address the SDGs.

CORRESPONDING AUTHOR:

Marc Santolini

Université Paris Cité,
Inserm, System Engineering
and Evolution Dynamics,
F-75004 Paris, France

Learning Planet Institute,
F-75004 Paris, France

marc.santolini@cri-paris.org

KEYWORDS:

Citizen science; network
science; monitoring and
evaluation; Sustainable
Development Goals; team
science; collaboration

TO CITE THIS ARTICLE:

Masselot, C, Jeyaram, R, Tackx, R, Fernandez-Marquez, JL, Grey, F and Santolini, M. 2023. Collaboration and Performance of Citizen Science Projects Addressing the Sustainable Development Goals. *Citizen Science: Theory and Practice*, 8(1): 45, pp. 1–16. DOI: <https://doi.org/10.5334/cstp.565>

INTRODUCTION

The United Nations Sustainable Development Goals (SDGs) are a series of development targets designed to address the world's most pressing societal, environmental, and economic challenges by 2030. Measuring their progress requires the obtention of timely, relevant, reliable data across a multitude of stakeholders. By engaging in scientific activities, citizens can foster the progress towards the resolution of the SDGs (Fritz et al. 2019), for example by generating evidence to identify gaps in their monitoring (Franzoni et al. 2021), collecting and analyzing data to support the decisions taken by local and national stakeholders (Ballerini and Bergh 2021; Fraisl et al. 2020), and accelerating the development of solutions (Kokshagina 2021; Masselot et al. 2022).

Compared with traditional scientific work, citizen science requires the definition of processes of engagement and coordination, from simple data collection to co-design strategies (Franzoni and Sauermann 2014; Haklay 2018; SenabreHidalgo et al. 2021). Frameworks have recently been developed to assess the impact of citizen science projects towards the SDGs (Parkinson et al. 2022), to understand modes of co-production for sustainability (Chambers et al. 2021) and to evaluate success of online teams in terms of scientific contribution and public engagement (Cox et al. 2015). Yet those frameworks often apply to large teams or to advanced projects, and organizers of initiatives such as Crowd4SDG (Crowd4SDG consortium 2020) lack supporting evidence to guide their practice in forming and coordinating successful citizen science projects at early stages. The evaluation of participatory processes such as those involved in citizen science emphasizes measures of diversity, engagement, collaboration, and learning (Jaeger et al. 2022; Schaefer et al. 2021). The ability to measure these participatory processes is therefore key for the monitoring and evaluation of citizen science projects.

Mixed methods involving digital traces and questionnaires are traditionally used in social studies describing collaborative activities, for example to understand how social networks shape individual performance in collaborative learning (Poquet et al. 2020), or to describe how team interactions and community organization shape collective performance within open research programs (Kokshagina 2021; Masselot et al. 2022) or open source communities (Gargiulo et al. 2022; Klein et al. 2015; Klug and Bagrow 2016). Complementing digital traces, the collection of self-reported data yields qualitative insights across perceived interactions (Deri et al. 2018). However, building a comprehensive group-scale network requires the engagement of a large proportion of the participants involved in the self-report activity in order to

accurately represent the social network, calling for specific survey instruments that allow for collection of the social ties of a participant while minimizing survey burden. The recent availability of such survey instruments has rendered collaborative network data collection easier and scalable, allowing researchers to capture temporal organizational networks within groups of various sizes (Tackx et al. 2021).

In this study, we tackle the question of how participatory processes shape project performance within early citizen science projects addressing the SDGs. To do so, we focused on Crowd4SDG, a European project that guides young individuals from pitching an idea on a social platform to the design of prototypes of citizen science projects via a one-year cycle of innovation. We developed and implemented a framework to monitor the activity and collaborations across 14 citizen science projects from the Crowd4SDG project. We highlight how this framework generated complementary interaction networks informing on the division of labor, collaborations, advice seeking, and communication processes of the citizen science projects. Finally, we show the usefulness of these measures for monitoring engagement and supporting the evaluation process and discuss how this framework could be used in future programs.

METHODS

DESCRIPTION OF THE GEAR CYCLE

The Crowd4SDG project organizes three one-year cycles of innovation, aimed at coaching teams of young social entrepreneurs through the steps of building a citizen science project. Each project tackles a challenge related to Climate Action or involves crowdsourcing tools that can generate data relevant for tracking progress towards the SDGs. The innovation methodology used follows a “GEAR Methodology” to coach teams through the innovation process required to develop new citizen science projects (Figure 1a). Each GEAR cycle includes several phases of online coaching and in-person support: Gather, Evaluate, Accelerate and Refine. The Gather phase is promoted as a global crowdsourcing of ideas, called the Open17 Challenge (Figure 1b), on the social network Goodwall. Some participants entered the phase with their own team, others were assigned teammates by a teaming algorithm (see Supplemental File 2: Teaming Algorithm for the parameters used in the algorithm). At the start of the Evaluate phase, 30 to 50 participants are selected to enroll in the Open17 weekly coaching during which they learn about developing and pitching their citizen science project. The best teams then benefit from a Challenge-Based Innovation Workshop (CBIW), which focuses on building a working prototype for the project, using crowdsourcing tools developed by the

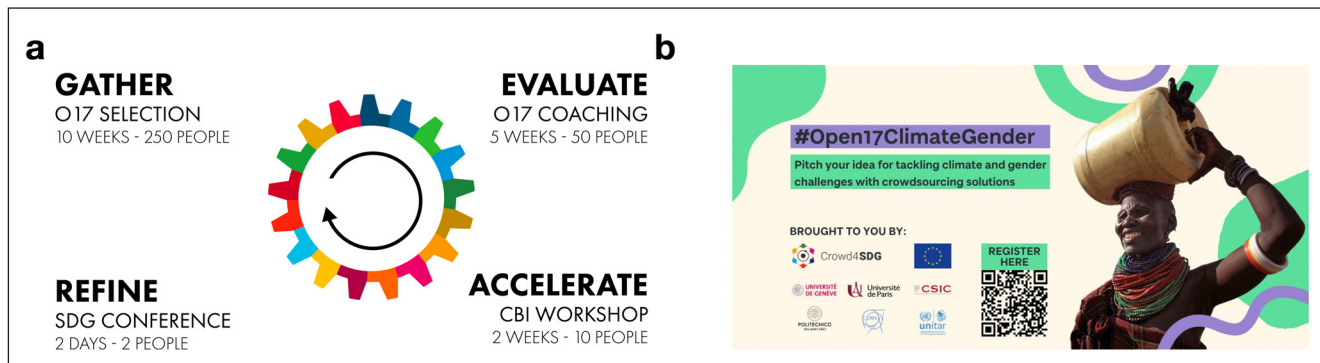


Figure 1 Description of the challenge. (a) Schematics of a GEAR cycle. (b) Visual for the GEAR Cycle 2.

Crowd4SDG consortium partners and other relevant ones. The most promising projects are invited to participate in the Geneva Dialogue, an opportunity to meet sponsors and potential partners amongst the international organizations in Geneva. Each phase of the GEAR methodology filters projects based on their novelty, relevance, feasibility, and appropriate use of crowdsourcing tools, and helps participants advance towards practical deployment.

The citizen science projects developed in the three GEAR cycles of Crowd4SDG aim to address the nexus between Climate Action (SDG 13) and several other key SDGs: Sustainable Cities (SDG 11 in 2020), Gender Equality (SDG 5 in 2021), and Peace, Justice and Strong Institutions (SDG 16 in 2022).

DESCRIPTION OF TEAM PROJECTS

In this work, we study the “Evaluate” phase of the GEAR cycle 2, the coaching program where teams ideate their project and engage in interactions with their peers and mentors. This focus allowed us to gather data on a large enough sample of projects corresponding to 14 teams. We document in [Table 1](#) the objectives of the projects, along with team size and the final stage of the challenge they achieved.

COMMUNICATION DATA

A Slack workspace was used by the teams during the GEAR cycle as a means to communicate with other teams and with the organizing team.

The data was extracted in JSON format using the export function available to the owners/admins of the Slack workspace. This allowed us to gather a data frame containing across all public channels the messages (post contents), their time stamp, sender, and channel it was sent in. The raw data was then processed to obtain mentions. A mention occurs when a Slack user types in a message the Slack username of a target user prefixed by “@”. Each recorded mention has information on the source (who wrote the message), target (who is being mentioned),

and the timestamp (when the message was sent). Slack also allows users to broadcast messages by citing all users in a channel or a workspace by using specific commands (@all, @here, @channel_name). These were not included as mentions in order to focus on direct interactions only.

SURVEY DATA

We used two types of surveys: those related to participant attributes (e.g., their background or country of origin), and those related to participant interactions (e.g., who they collaborated with or sought advice from).

The initial survey was related to attributes only and was disseminated using a Google Form at registration to the Evaluate phase.

We then disseminated 4 weekly surveys related to social interactions and activities using the CoSo platform ([Tackx et al., 2021](#)) (Supplemental File 2: Figure S1). The CoSo platform is designed to collect self-reported interaction data with a simple, reactive interface, and an analysis-ready database. To document their interactions, the users could select target users across all other participants and organizers. The interactions spanned prior ties in the first survey (“Which of these people did you know personally before?”), and on a weekly basis their advice-seeking behavior (“Who did you seek advice from last week?”) and collaborations (“Who did you work with last week?”). To document their activity, they could select across 26 activities encompassing routine activities within research teams inspired from the CRediT contribution taxonomy, as well as specific questions regarding Crowd4SDG, for example specific tool usage. Activities encompassed different levels of complexity in their realization. They ranged from tasks that could be performed in a distributed fashion such as preparing the final pitch and analyzing data, to tasks involving higher levels of collaboration such as brainstorming.

The surveys were advertised through Slack, and the organizing team dedicated 10 minutes for participants to fill them during weekly sessions, ensuring a high engagement:

TEAM NAME	TEAM SIZE	PROJECT OBJECTIVES	FURTHEST STAGE ACHIEVED
WOMER	2	Assess the effect of climate change on women in indigenous communities in Colombia. Empower indigenous women via their engagement in citizen science.	Refine
DonateWater	4	Map the operating status of water handpumps in rural communities of Nigeria using crowdsourcing techniques. Discuss changes in water access solutions.	Refine
Andapé Institute	3	Map the accessibility of sidewalks and their conditions in the urban environment in Brazil. Design an ecological sidewalk from recycled materials.	Accelerate
Climate Gender Justice	2	Generate area-specific data related to climate change impact on sexual and reproductive health and rights.	Accelerate
Water Warriors	3	Organize and classify emergencies and the responses adopted to solve them.	Accelerate
Women & Technology Against Climate Change	3	Map women participation in climate change projects connected to technology.	Evaluate
SDesiGn	2	Increase gender equality in architecture and product designs that prioritize environment protection.	Evaluate
UpGet app - CitiCERN Project	2	Build a digital wallet to get eco-friendly rewards.	Evaluate
Eco Winners	3	Build Digital Platform where farmers in Nigeria—of which 70% are women—can rent farming tools and equipment at low cost on a peer-to-peer basis.	Evaluate
Women 4 Sustainable World	2	Set up meteorological data collection kits that send alerts to users on their mobile phones on favorable periods for planting seeds, and guidelines for making treatments with fertilizers or pesticides, or adapting the type of seeds to the soil.	Evaluate
TEAM FOILED	3	Allow affected communities to crowdsource damages caused by floods and access help faster.	Evaluate
PAM	3	Map and classify climate disasters and give particular attention to women’s voices.	Evaluate
Rights of Climate	2	Assess the link between women’s land rights and their vulnerability to climate change in Malawi.	Evaluate
Flood Rangers	2	Generate gender-segregated data related to the impact of flood on men and women in Nigeria	Evaluate

Table 1 GEAR cycle 2 project description.

Except for the team “Flood Rangers,” who answered only one CoSo survey before dropping out from the program, at least one member of each team answered surveys each week (Supplementary File 2: Table 1).

TEAM FEATURES

To document measures of participation, we monitored features related to team composition, communication, collaboration, and activity (features from Figure 7 are in italics below).

For team composition, we built measures of size, diversity, education level, and prior experience. *Team size* was assessed using the number of members of a team. Background diversity was assessed by computing the

background span—that is, the number of unique academic backgrounds in the team as declared in the registration form. The *education level* was computed by taking the average level of education in a team based on the response to the question “What is your current or highest level of education,” to which we attributed the following score based on the answer: 0 for secondary school, 1 for high school, 2 for undergraduate, and 3 for graduate. Finally, *prior experience* was computed as the average answer to the question, “Have you participated in data projects or contributed as citizen scientist to data production before?” (yes = 1 and no = 0) within each team.

For communication, we leveraged the activity and interactions on Slack public channels. The *Slack activity* was

assessed as the total number of messages posted by team members. For interactivity, we measure *Slack interaction intra team* as the number of mentions among members of a team, and *Slack interaction organizing team* as the number of mentions between members of a team with the organizing team. We counted mentions regardless of their directionality.

For collaborations, we focused on the number of collaborations within the teams, as well as the centrality of the teams within the advice network. For the intra-team collaborations (*coso interaction intra-team*), we summed for each team the weights of the intra-team edges in the “work with” collaboration network. For the centrality in the advice network, we computed the Burt constraint (Burt, 2004), a measure of social capital that takes low values when a neighborhood is diverse (ties to separate neighborhoods), and higher values when the neighborhood is constrained (dense ties to the same neighborhood). *Advice diversity* was computed by taking the negative of Burt constraint, with higher values indicating higher levels of diversity (more structural holes). This quantifies the ability of a team to leverage diverse sources of information for advice seeking.

Finally, for the activity, we focused on measures of diversity and engagement of activities performed. For diversity, we computed the *activity span* as the proportion of activities performed by a team among the 26 listed. For engagement, we considered the *activity regularity* by first computing the number of activities reported by a team each week, and then computing the negative of the Gini index on the resulting vector. The Gini index ranges from 0 (perfectly regular) to 1 (perfectly irregular). $1 - \text{Gini}$ is higher if activities are regularly conducted across weeks. Finally, we quantified for each team the *survey engagement* as the proportion of survey responses per team across all CoSo surveys, a measure of engagement to the study.

TEAM PERFORMANCE DATA

To quantify team performance, we used the scores that teams obtained in their assessment by the jury and the Crowd4SDG organizing team (features from Figure 7 are in *italics* below).

Performance was assessed through 4 team outcome metrics (crowdsourcing component, feasibility, relevance, novelty) as judged by a panel of experts selected by the Crowd4SDG consortium, and by 4 process parameters (project documentation, members attendance, commitment, and weekly evaluation) as assessed by the organizing team.

More precisely, *crowdsourcing* was assessed using the mean score attributed to the question “Is there an effective crowdsourcing component?” (yes = 1 and no =

0). We measured the *feasibility*, *relevance*, and *novelty* by computing the mean score attributed by the jury on a scale from 0 to 5 to the questions “Feasibility: Is the project implementable with reasonable time and effort from the team?,” “Novelty: Is the pitch based on a new idea or concept or using existing concepts in a new context?,” and “Relevance: Is the solution proposed relevant to the challenge or potentially impactful?”

In terms of process, all variables were integer values with scores ranging from 0 to 5 for deliverables and attendance, 0 and 1 for commitment. For *weekly evaluation* the score was a continuous value ranging from 0 to 10 scoring the overall quality of their weekly pitch sessions. *Deliverables* measured the total number of deliverables submitted and documented on the platform Innprogress (<https://innprogresstest.unige.ch/>) among the expected ones. *Attendance* was estimated by the proportion of sessions attended by team members. *Commitment* was scored 1 if teams were willing to continue their project after the end of Evaluate, or 0 otherwise.

NETWORK CONSTRUCTION

The SDG and background networks in Figure 2 are built using the co-occurrence of answers in multiple choice questions across participants in the first survey. In these networks, two nodes are linked if they are co-cited in an answer, and the weight of the link is the number of participants who reported this co-occurrence. The SDGs were declared in questions related to past projects: “Have you contributed to projects on SDGs before,” and “Which of the SDGs was the project addressing? Select all that apply.” The backgrounds were selected across a multiple-choice question asking “What are your main fields of work or study? Select all that apply.”

The Slack mention network links a user A to a user B if they mention them, with a weight corresponding to the number of times A has mentioned B. When aggregating at the team level, intra-team mentions are encoded as self-loops with an edge weight given by the sum of the intra-team links weights.

The CoSo networks are directly inferred from the surveys. In Figures 5 and 6, we aggregated the networks over all timepoints collected, yielding weighted interaction networks where edge weights correspond to the number of times an interaction was reported. Figure 6 further aggregates the individual networks at the team level.

In Figure 6c, centralization is computed as the Gini coefficient of the degree distribution. More precisely, we compute for each team its undirected, total degree (number of neighbors). We then compute the Gini coefficient of the degrees. Its value ranges from 0 (all degrees equal) to 1 (one team dominates the degree distribution), and indicates the

degree to which interactions are concentrated towards one “hub” node in the network.

Network centrality measures were computed using the igraph library in R. Network visualizations were produced with Gephi 0.9.7 with a force layout.

STATISTICAL ANALYSES

Data were analyzed using the R software. The association between performance and team features was assessed using Pearson’s correlations. The p value for the correlation is calculated by computing the t statistic (cor.test function in R), with the null hypothesis that the correlation between the dependent and independent variable is 0. The level of significance was set at $p = 0.1$.

RESULTS

In this study, we provide an analysis of teamwork and collaborations during the second GEAR cycle of innovation from the Crowd4SDG project. The citizen science projects developed in this GEAR cycle aim to address the nexus between Climate Action (SDG 13) and Gender Equality (SDG 5). To study the process of generation of citizen science projects in this context, we focus on the “Evaluate” phase (see Methods), a coaching program in which teams build their project and engage in interactions with their peers and mentors.

COHORT DESCRIPTION

The cohort was made up of a total of 38 participants covering 17 nationalities (Supplemental File 2: Figure S2).

A total of 14 teams were formed with sizes varying from 1 to 4 members and showing diversity in terms of age and gender (Supplemental File 2: Figure S3b). All teams comprised students, mostly at the university level (12 out of 14), with some high school-level teams (2/14). Overall, 68% of participants were younger than 25 years old, with an age range of 16 to 32 years old (Supplemental File 2: Figure S3a).

Some participants had prior experience with citizen science, with 26% (10/38) answering positively to the question “Have you participated in data projects or contributed as citizen scientist to data production before?” Moreover, the prior experience of participants with SDGs covered most goals, with a primary focus on Climate Action and Gender Equality, as expected from the topic of the GEAR cycle (Figure 2a). Interestingly, the participants (and the teams themselves) displayed a high level of interdisciplinarity, with backgrounds spanning natural sciences, technology, and humanities (Figure 2b and S2).

TEAM COMMUNICATION

During the Evaluate phase, teams used a Slack workspace to discuss with other team members from their own or from another team, as well as with the organizing team. Since the challenge was fully conducted online, this workspace was a central repository for communications at the cohort level. We analyzed the data from the public channels of the Slack workspace to study the patterns of engagement of participants within and across teams, as well as with organizers.

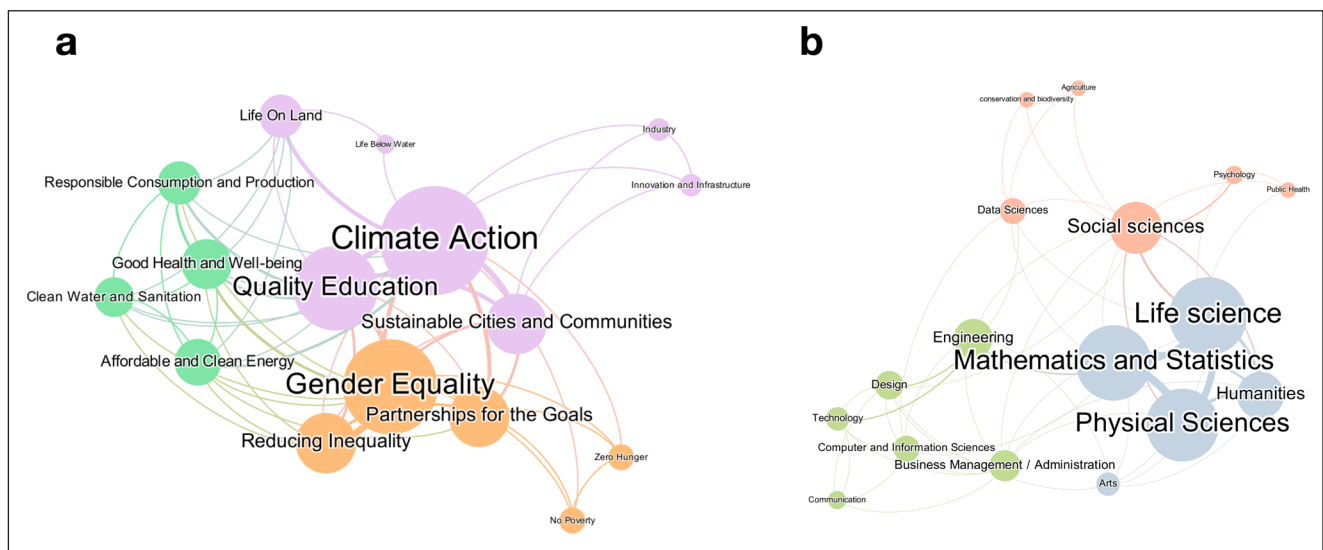


Figure 2 Description of the cohort. Co-occurrence networks across participants highlighting (a) their previous experience with Sustainable Development Goals and (b) their main background or field of study. In these networks nodes are linked by the number of times they are co-reported by a participant, and colors correspond to denser subnetworks as determined by the modularity algorithm. Participants had prior experience with Climate Action and Gender Equality, and came from interdisciplinary backgrounds.

We first observed that the activity of the Slack workspace, measured by the number of posts per week, closely follows the phases of the GEAR cycle, with low activity outside of the phases (Figure 3a). This might be because teams would work solely during the program, or because they would synchronize on other communication channels outside of these phases, such as Whatsapp, e-mails, or private Slack conversations (Supplemental File 2: Figure S4).

To examine the interaction dynamics between participants, we used a network approach. This allows for representation of the flow of information characterizing this

phase, in particular highlighting the interactions with the organization team. We computed the number of mentions of a “target” participant B from a “source” participant A as an indication of a directed interaction from A to B. Mapping participants to their respective teams, we derived a directed, weighted network indicating the interaction strength between and within teams (Figure 3b).

We observe a high centralization of interactions to the organizing team (Figure 3c), both in terms of incoming (teams reaching out to organizers) and outgoing links (organizers reaching out to teams). While the workspace

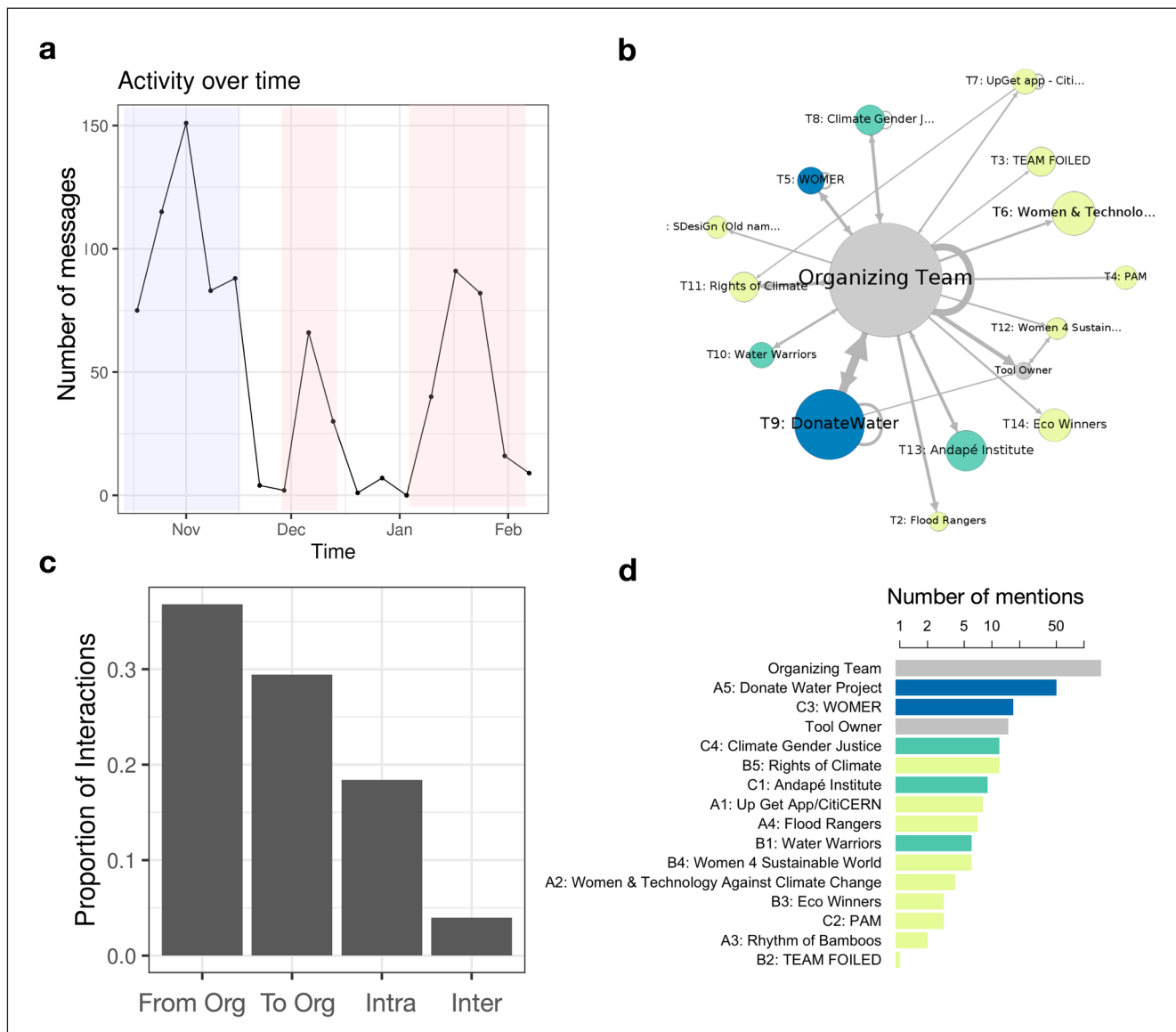


Figure 3 Communication activity. (a) Total number of posts on Slack per week. The Evaluate phase is highlighted in blue, and the Accelerate phase, consisting of two periods, is highlighted in red. (b) Mention network extracted from Slack during the Evaluate phase. Nodes represent aggregated individuals at the team level. Teams are linked by weighted edges quantifying the number of times an individual from one team mentions an individual from another team. Self-loops denote intra-team interactions. Grey color denotes the organizing team, and the color denotes the stage achieved in the program: in order, Evaluate (yellow), Accelerate (green), Refine (blue). (c) Proportion of Slack mentions that are from the organization team, towards the organization team, intra-team or inter-team. (d) Number of mentions (sent or received) per team, following the color code from b.

was also used for within-teams interactions, there were very few inter-team interactions (Figure 3c), confirming that the workspace was mostly used as a means to interact with organizers.

Beyond the organizing team, we found that the two teams that were eventually selected as finalists of the program, Donate Water Project and WOMER, had the highest network centrality teams when considering their weighted degree (i.e., the total number of incoming and outgoing mentions they partake in), suggesting that team level of engagement early in the program is important for project success.

TEAM ACTIVITIES

While Slack informs on participant engagement and their interactions with organizers, it does not provide information on what activities teams perform, or what type of (informal and formal) interactions occur. Such information can help guide coordinators in managing citizen science communities. To gather deeper insights into team dynamics, we performed weekly surveys on activities performed and on collaborations during the four weeks of the Evaluate phase preceding the presentations to the jury.

The activities most performed were consistent with the purpose of the Evaluate phase: coaching teams into generating a feasible, novel citizen science project. As such, the main activity performed across the 4 weeks was the preparation of the final pitch (Figure 4a). The early weeks were enriched in activities related to brainstorming and ideation, task planning, team building and literature review, while later weeks showed activities related to the preparation of documentation material and result interpretation. Moreover, it is interesting to note a significant number of participants declared “Meeting with people affected by the problem you are trying to solve” during the 4 weeks, a marker of engagement with stakeholders. The number of activities and their regularity varied widely across teams (Supplemental File 2: Figure S5), with an overall stronger push at the last week, suggesting a deadline effect (Figure 4b).

COLLABORATION DYNAMICS

Beyond activities performed, the surveys enquired about formal (“who did you work with?”) and informal (“who did you know before?” and “who did you seek advice from?”) interactions (Figure 5a,b,c). These surveys were aimed at investigating the collaborative dynamics during the GEAR, its evolution in time, and eventual impact on team performance.

In the GEAR cycle, participants could join as a team, or as individuals. The latter were assigned to a team using a matching algorithm (see Supplemental File 2: Teaming

Algorithm). The existence of pre-formed teams is revealed in the “prior ties” network (Figure 5a). Yet, beyond intra-team links, we found that several participants acted as bridges between teams in the prior ties network. This is probably because the Gather phase was able to tap into already existing communities, in particular through the social platform Goodwall.

Work collaborations occurred mostly within teams, as well as with organizers (Figure 4b,d), while only few inter-team interactions were observed. However, advice seeking interactions, in which participants report having asked for advice from another participant, showed more inter-team interactions, with around 10% of them being inter-team ties (Figure 4c,e). Moreover, while participants sought primarily advice within their own team in the first week, they gradually increased their outreach to the organizers, eventually constituting 55% of interactions. In both networks, organizers occupied the most central position, acting as bridges between teams.

COMPARISON OF THE INTERACTION NETWORKS

The collected data allowed us to infer 4 interaction networks: communications from Slack mentions, and prior ties, collaborations, and advice seeking from surveys. When aggregated at the team level, these constitute a “multiplex” network, with the same nodes (the teams) having different types of links. Here, we question whether these networks provide similar or complementary information to inform on team behavior.

We show in Figure 6a the networks at the team level. One can observe that the networks have similar densities (Figure 6c), but different structures: The Slack mentions network is much more centralized than the surveyed interaction networks (Figure 6c), indicating that Slack usage was mostly used to exchange with the organizing team who acted as a strong hub. When aggregating the networks, one obtains a more comprehensive interaction network (Figure 6b), doubling the density of links compared with any single network (Figure 6c).

To further assess the topological similarity between the networks obtained, we computed the Jaccard similarity between any pair of networks, that is, the ratio of the number of links in common (intersection) to the total number of links present in both networks (union). Completely dissimilar networks would have a Jaccard of $J = 0$, while identical networks have $J = 1$. We find that the collaboration (“work with”) and advice seeking networks are the most similar ($J = 0.74$), while their similarity with the Slack mention network is much smaller ($J \sim 0.2$). Prior ties are predictive of collaboration and advice seeking ($J = 0.2$) but not of Slack mentions ($J \sim 0$), which is probably due to the fact that most interactions on Slack were with the



Figure 4 Activities during the Evaluate phase. **(a)** Total number of reports of an activity per week. We see a switch from brainstorming/ planning activities to the documentation and preparation of the final presentation. **(b)** Total number of activities reported per team per week.

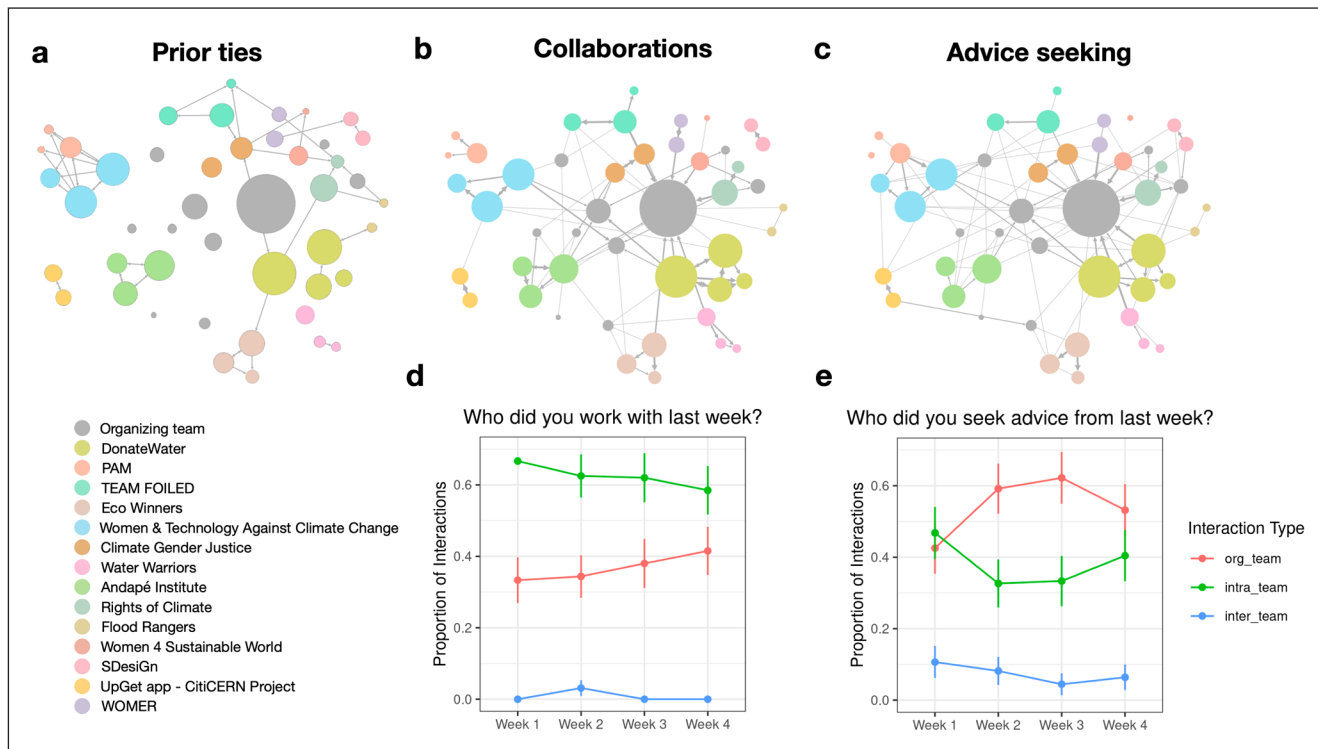


Figure 5 Collaboration activity. (a–c) Participant interaction networks constructed from self-report data from CoSo, using prompts: (a) “Which of these people did you know personally before?” (b) “Who did you work with last week?” and (c) “Who did you seek advice from last week?” The size of a node is proportional to the total number of interactions of a node across the 3 networks. (d) Proportion of interactions in the collaboration network that involve the organization team (red), or that are intra- (green) or inter- (blue) team, in time. Error bars denote the standard error of the estimate given a number of interactions observed, assuming a binomial statistics. (e) Same than d, for the advice seeking network.

organizer team. Finally, the networks show a similarity to the full, aggregated networks ranging between $J = 0.4–0.5$, indicating that a network measured with a single method encapsulates less than half of the information about formal and informal social interactions.

Overall, we find that the collected interaction profiles from digital traces and from surveys highlight different aspects of the social interactions, providing complementary insights to inform community management.

TEAM PERFORMANCE

Finally, we analyzed whether features of team composition, communication, teamwork and collaboration were associated with team performance at the Evaluate phase. The performance was measured using various features that can be grouped into two overarching categories: outcome, that is, the evaluation of the project itself; and process, that is, the assessment of engagement within the program (see Methods and Supplemental File 2: Figure S6a). Given the small number of teams from which we can compute an association with performance (14 data points), we use a correlation analysis with a soft significance threshold at $p = 0.1$. We present the results of this analysis in Figure 7, where we highlight that the quality

of the outcome is generally associated with team profiles and communication activity from Slack, while the level of engagement in the program (process) as judged by the organizing team is associated with self-reported measures of collaborations and activity.

More precisely, for the team composition, we find that team size is associated with the use of at least a tool in the Citizen Science Solution Kit (crowdsourcing), suggesting a need for human power to set up a crowdsourcing infrastructure. The diversity of backgrounds in the team is associated with the novelty of the project, supporting findings that interdisciplinarity begets innovative work (Singh et al. 2022). Prior experience with citizen science is important for the relevance and novelty of the project, indicating the importance of past work in related areas to achieve well-defined, innovative projects in this short time span. Similarly, we find that the average education level in a team is associated with the novelty, feasibility, and relevance of the project.

In the case of communication activity, we find that the overall Slack activity (which is very correlated with the number of interactions with the organizing team, see Supplemental File 2: Figure S6b) is associated with the relevance of the project, highlighting the role of mentoring

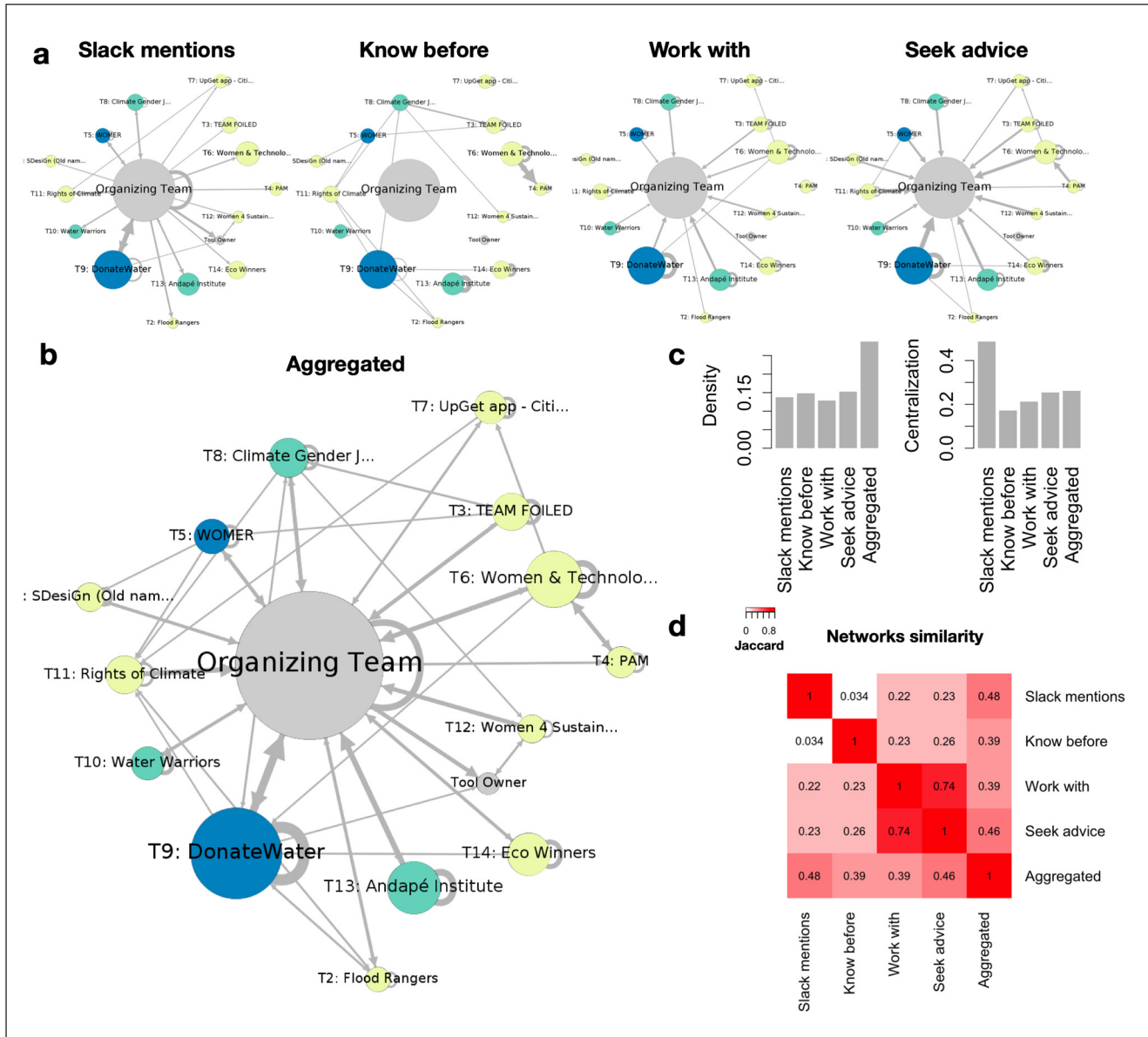


Figure 6 Comparison of the interaction networks. (a) Team-level networks for the different interaction networks collected. (b) Corresponding aggregate network, where edge weights are the sum over weights across the individual networks shown in a. (c) Network density and centralization (see Methods) across the 4 considered networks. (d) Jaccard similarity between the networks in a. The similarity measures the number of edges shared between any two networks (intersection), divided by the total number of edges present in both networks (union), and ranges from 0 (most dissimilar) to 1 (most similar).

for helping teams craft a relevant project. Intra-team interactions from Slack mentions are associated with *relevance*, *novelty*, and *crowdsourcing* aspects of the project, as well as with the quality of *deliverables*. Interestingly, we find similar results when measuring the intra-team collaborations with CoSo surveys, indicating that the digital traces do capture relevant qualitative information about team interactions.

In contrast, we find that team engagement in activities and advice seeking is associated with the quality of the process, as judged by the organizing team, encompassing team commitment, attendance, weekly evaluation, and

their ability to produce qualitative deliverables. Beyond engaging in diverse activities in a regular manner, survey engagement was found to be a strong predictor of program engagement. Moreover, we note the importance of the ability of teams to engage in advice seeking from diverse network neighborhoods, as measured by (lower) Burt constraint (Burt 2004) in the advice-seeking network. These results may indicate that the organizing team, who was responsible for judging these criteria, was particularly sensitive to the ability of teams to engage and collaborate throughout the cycle, information that was not readily available to other experts.

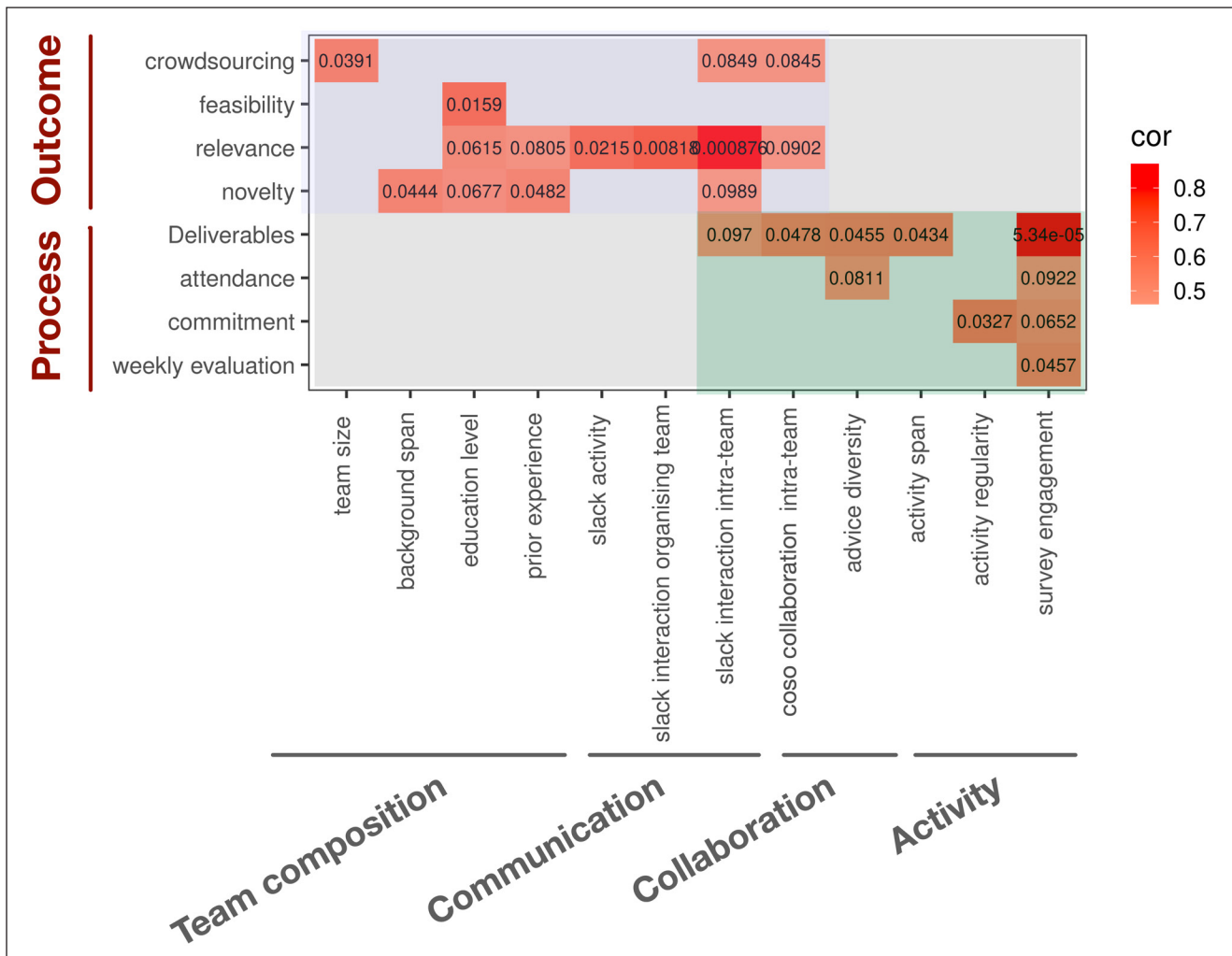


Figure 7 Association with performance. Correlations between performance assessment (rows) and team features (columns). The correlation value is indicated when the correlation is significant at the $p = 0.1$ level. Shaded areas correspond to sets of features associated with metrics related to outcome (blue) or to process (green).

DISCUSSION

Processes of engagement and coordination are fundamental to citizen science projects (Jaeger et al. 2022; Schaefer et al. 2021). Here, we showcase a framework to measure indicators of participation, contribution, and collaboration during the elaboration of citizen science projects. We show that surveys of social interactions collected at several points in time provide information otherwise invisible from digital traces obtained from a Slack workspace that can be leveraged by practitioners who guide citizen science projects at their early stage of development.

Given the nature of the program, time could be set aside by the organizers for engaging participants in surveys on a weekly basis, as part of the curriculum. As such, the engagement with the survey instrument was particularly good, allowing to obtain a near-complete coverage. In other contexts where regular meetups with participants

would not be feasible, the method could be adapted to incentivize participants to build and analyze their collaboration and stakeholder network and learn from it, for example by providing a dashboard for visual feedback (Tackx et al. 2021).

Our framework is particularly suited to investigate measures related to teamwork. The organizational literature shows that the effectiveness of traditional teams depends on their composition, the collaboration of their members, the task allocation, and the activity level (Hackman 1987). Here, we showed that we could monitor proxies for these features, and that they were associated with the ability of teams to produce well-defined deliverables, an indicator of team performance to a standardized task. Beyond small-scale teamwork, the proposed framework can be interesting for quantifying contributions within larger projects. This would allow fine-grain recognition of the different activities achieved, acting both as an incentivization mechanism for

monitoring, as well as a reward system for the (usually volunteer) work done.

Beyond team dynamics of early-stage projects, leveraging social networks measurements within citizen science programs offers opportunities to document and understand the build-up of a community around a citizen science project, the engagement patterns of participants, and the contribution to different tasks. This is particularly useful to facilitate the coordination processes of potentially large communities (Kokshagina 2021; Santolini 2020), allowing the core team to react and assess whether certain individuals or sub-projects would need help.

Yet this work has limitations. First, the case study could offer only a small sample size, and more data will be needed in further studies to validate the associations with performance. Moreover, during the Evaluate phase, the citizen science projects are at a very early stage of ideation, which did not allow to investigate interactions between teams and citizens. Future work could investigate more mature projects. Finally, the activities performed have different levels of complexity, and do not require the same levels of engagement or collaboration. To highlight these fine-grained aspects would require more qualitative insights from participants, as well as an adaptative strategy to integrate these insights as new (sub)tasks. Beyond the specificity of our project, such an adaptative co-design strategy to account for the varied activities performed is also an important step to be conducted by citizen science projects, as it renders visible the contribution structure and affordances of engagement.

CONCLUSION

One challenge that organizers of programs like Crowd4SDG face is to support with evidence their decisions related to the formation of citizen science teams and their management, as well as the directions they give to participants to maximize the relevance of the data they generate, their ability to develop innovative solutions, and eventually their impact on the problem they are addressing.

Here, we implemented a monitoring framework leveraging digital traces as well as self-reports to gather compositional and social interaction data during the makeup of citizen science projects. This approach complements traditional outcome-driven metrics in the evaluation of science (Fortunato et al. 2018) by emphasizing the importance of the participation process (Jaeger et al. 2022; Schaefer et al. 2021). We reconstructed a multi-layer social network with interactions of various types, from informal social ties to formal collaborations. We showed that these layers

obtained from various means (passive digital traces and active self-reports) cover multiple complementary facets of the interaction dynamics, informing both on interactions with coordinators from the organizing team, as well as intra-team and inter-team interactions. We showed that network centrality measures can be leveraged to quantitatively assess the relative centralization within a given layer, informing on the reliance over a few central nodes. In particular, we found that the ability of a team to manage their social capital by forming interactions across diverse neighborhoods in the network is important for the success of their project, a finding in line with the literature on innovation (Burt 2004). Furthermore, we showed that measures of team composition, intra-team collaborations, and communication with the organizers are associated with the quality of the projects, in particular the relevance and novelty of their solutions to the SDGs. Measures of engagement in activities and advice seeking are on their end associated with the elaboration process, in particular the ability of teams to provide timely deliverables.

Overall, we introduced a framework to monitor the evolution of participatory processes in citizen science projects. The obtained interaction networks reveal both formal and informal relational networks that underlie the collective learning and performing, making visible structural patterns that are otherwise invisible to coordinating teams. Network measures of centrality, peripherality, or diversity can then be leveraged to quantify the embeddedness (or lack thereof) of participants in the ecosystem, informing on concrete interventions to improve engagement and project outcomes. Such insights can therefore prove useful to support practitioners in the design and coordination of programs aiming at fostering engagement, inclusion, and diversity in citizen science projects.

DATA ACCESSIBILITY STATEMENT

Anonymised data can be found in the Supplemental Files 3, 4, 5, and 6.

SUPPLEMENTARY FILES

The supplementary files for this article can be found as follows:

- **Supplemental File 1.** Consent form. DOI: <https://doi.org/10.5334/cstp.565.s1>
- **Supplemental File 2.** Supplementary Tables, Methods, and Figures. DOI: <https://doi.org/10.5334/cstp.565.s2>
- **Supplemental File 3.** Slack mention network. DOI: <https://doi.org/10.5334/cstp.565.s3>

- **Supplemental File 4.** CoSo self-reported interactions. DOI: <https://doi.org/10.5334/cstp.565.s4>
- **Supplemental File 5.** CoSo self-reported activities. DOI: <https://doi.org/10.5334/cstp.565.s5>
- **Supplemental File 6.** Team features and performance. DOI: <https://doi.org/10.5334/cstp.565.s6>

ETHICS AND CONSENT

The surveys and research questions received the ethical approval of the Inserm committee (IRB00003888). Participants gave their consent to the collection of data as they registered to the Evaluate phase (see Supplemental File 1). Data was pseudo-anonymized before analysis.

ACKNOWLEDGEMENTS

We acknowledge Barbara Pernici and Jesus Cerquides for their comments on an early version of this draft, and Amudha Ravi Shankar for her important help in disseminating surveys.

FUNDING INFORMATION

This work was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 872944. This work was also partly supported by the French Agence Nationale de la Recherche (ANR), under grant agreement ANR-21-CE38-0002-01, as well as by Fondation Botnar and Generation Unlimited through the Youth Agency Marketplace (Yoma) program.


COMPETING INTERESTS


The authors are part of the Crowd4SDG consortium.


AUTHOR CONTRIBUTIONS

FG, JLF, and MS acquired funding for the study. FG conceptualized the GEAR cycle, FG and JLF coordinated the implementation of GEAR 2. MS conceptualized and supervised the study. MS and CM designed the monitoring methodology. RT developed the CoSo interface for survey collection. CM, RJ, and RT collected data. RJ curated data. RJ and MS analyzed and visualized the data. CM and MS wrote the manuscript.


AUTHOR AFFILIATIONS


Camille Masselot  orcid.org/0000-0002-0728-673X
Université Paris Cité, Inserm, System Engineering and Evolution Dynamics, F-75004 Paris, France; Learning Planet Institute, F-75004 Paris, France

Rathin Jeyaram  orcid.org/0000-0002-1304-1229
Université Paris Cité, Inserm, System Engineering and Evolution Dynamics, F-75004 Paris, France; Learning Planet Institute, F-75004 Paris, France

Raphael Tackx  orcid.org/0000-0002-7940-6833
Université Paris Cité, Inserm, System Engineering and Evolution Dynamics, F-75004 Paris, France; Learning Planet Institute, F-75004 Paris, France

Jose Luis Fernandez-Marquez  orcid.org/0000-0002-0254-0794
Centre Universitaire d'Informatique, University of Geneva, CH

François Grey  orcid.org/0000-0003-3471-5445
Centre Universitaire d'Informatique, Geneva School of Economics and Management, CH

Marc Santolini  orcid.org/0000-0003-1491-0120
Université Paris Cité, Inserm, System Engineering and Evolution Dynamics, F-75004 Paris, France; Learning Planet Institute, F-75004 Paris, France

REFERENCES

- Ballerini, L** and **Bergh, SI**. 2021. Using citizen science data to monitor the Sustainable Development Goals: a bottom-up analysis. *Sustainability Science*. DOI: <https://doi.org/10.1007/s11625-021-01001-1>
- Burt, RS**. 2004. Structural Holes and Good Ideas. *American Journal of Sociology*, 110: 349–399. DOI: <https://doi.org/10.1086/421787>
- Chambers, JM, Wyborn, C, Ryan, ME, Reid, RS, Riechers, M, Serban, A, Bennett, NJ, Cvitanovic, C, Fernández-Giménez, ME, Galvin, KA, Goldstein, BE, Klenk, NL, Tengö, M, Brennan, R, Cockburn, JJ, Hill, R, Munera, C, Nel, JL, Österblom, H, Bednarek, AT, Bennett, EM, Brandeis, A, Charli-Joseph, L, Chatterton, P, Curran, K, Dumrongrojwattana, P, Durán, AP, Fada, SJ, Gerber, J-D, Green, JMH, Guerrero, AM, Haller, T, Horcea-Milcu, A-I, Leimona, B, Montana, J, Rondeau, R, Spierenburg, M, Steyaert, P, Zaehring, JG, Gruby, R, Hutton, J and Pickering, T**. 2021. Six modes of co-production for sustainability. *Nature Sustainability*, 4: 983–996. DOI: <https://doi.org/10.1038/s41893-021-00755-x>
- Cox, J, Oh, EY, Simmons, B, Lintott, C, Masters, K, Greenhill, A, Graham, G and Holmes, K**. 2015. Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects. *Computing in Science & Engineering*, 17: 28–41. DOI: <https://doi.org/10.1109/MCSE.2015.65>
- Crowd4SDG consortium**. 2020. Citizen Science for Monitoring Climate Impacts and Achieving Climate Resilience. *CORDIS Eur. Comm.* URL <https://cordis.europa.eu/project/id/872944/fr>

- Deri, S, Rappaz, J, Aiello, LM and Quercia, D.** 2018. Coloring in the Links: Capturing Social Ties as They are Perceived. *Proceedings of the ACM on Human-Computer Interaction*, 2: 1–18. DOI: <https://doi.org/10.1145/3274312>
- Fortunato, S, Bergstrom, CT, Börner, K, Evans, JA, Helbing, D, Milojević, S, Petersen, AM, Radicchi, F, Sinatra, R, Uzzi, B, Vespignani, A, Waltman, L, Wang, D and Barabási, A-L.** 2018. Science of science. *Science*, 359: eaao0185. DOI: <https://doi.org/10.1126/science.aao0185>
- Fraisl, D, Campbell, J, See, L, Wehn, U, Wardlaw, J, Gold, M, Moorthy, I, Arias, R, Piera, J, Oliver, JL, Masó, J, Penker, M and Fritz, S.** 2020. Mapping citizen science contributions to the UN sustainable development goals. *Sustainability Science*, 15: 1735–1751. DOI: <https://doi.org/10.1007/s11625-020-00833-7>
- Franzoni, C, Poetz, M and Sauermann, H.** 2021. Crowds, citizens, and science: a multi-dimensional framework and agenda for future research. *Industry and Innovation*, 1–34. DOI: <https://doi.org/10.1080/13662716.2021.1976627>
- Franzoni, C and Sauermann, H.** 2014. Crowd science: The organization of scientific research in open collaborative projects. *Research Policy*, 43: 1–20. DOI: <https://doi.org/10.1016/j.respol.2013.07.005>
- Fritz, S, See, L, Carlson, T, Haklay, M (Muki), Oliver, JL, Fraisl, D, Mondardini, R, Brocklehurst, M, Shanley, LA, Schade, S, Wehn, U, Abrate, T, Anstee, J, Arnold, S, Billot, M, Campbell, J, Espey, J, Gold, M, Hager, G, He, S, Hepburn, L, Hsu, A, Long, D, Masó, J, McCallum, I, Muniafu, M, Moorthy, I, Obersteiner, M, Parker, AJ, Weisspflug, M and West, S.** 2019. Citizen science and the United Nations Sustainable Development Goals. *Nature Sustainability*, 2: 922–930. DOI: <https://doi.org/10.1038/s41893-019-0390-3>
- Gargiulo, F, Castaldo, M, Venturini, T and Frasca, P.** 2022. Distribution of labor, productivity and innovation in collaborative science. *Applied Network Science*, 7(1): Article 1. DOI: <https://doi.org/10.1007/s41109-022-00456-0>
- Hackman, JR.** 1987. The design of work teams. In: Lorsch, J. (ed.), *Handbook of Organizational Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Haklay, M.** 2018. Participatory Citizen Science. In: Hecker, S, Haklay, M, Bowser, A, Makuch, Z, Vogel, J and Bonn, A (eds.), *Citizen Science: Innovation in Open Science, Society and Policy*. London, UK: UCL Press, pp. 52–62. DOI: <https://doi.org/10.2307/j.ctv550cf2.11>
- Jaeger, J, Masselot, C, Greshake Tzovaras, B, Senabre Hidalgo, E, Haklay, M and Santolini, M.** 2022. An Epistemology for Democratic Citizen Science [Preprint]. *Open Science Framework*. DOI: <https://doi.org/10.31219/osf.io/j62sb>
- Klein, M, Maillart, T and Chuang, J.** 2015. The Virtuous Circle of Wikipedia: Recursive Measures of Collaboration Structures. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '15*. Association for Computing Machinery, New York, NY, USA, pp. 1106–1115. DOI: <https://doi.org/10.1145/2675133.2675286>
- Klug, M and Bagrow, JP.** 2016. Understanding the group dynamics and success of teams. *Royal Society Open Science*, 3: 160007. DOI: <https://doi.org/10.1098/rsos.160007>
- Kokshagina, O.** 2021. Open Covid-19: Organizing an extreme crowdsourcing campaign to tackle grand challenges. *R&D Management*. DOI: <https://doi.org/10.1111/radm.12470>
- Masselot, C, Tzovaras, BG, Graham, CLB, Finnegan, G, Jeyaram, R, Vitali, I, Landrain, T and Santolini, M.** 2022. Implementing the Co-Immune Open Innovation Program to Address Vaccination Hesitancy and Access to Vaccines: Retrospective Study. *Journal of Participatory Medicine*, 14: e32125. DOI: <https://doi.org/10.2196/32125>
- Parkinson, S, Woods, SM, Sprinks, J and Ceccaroni, L.** 2022. A Practical Approach to Assessing the Impact of Citizen Science towards the Sustainable Development Goals. *Sustainability*, 14: 4676. DOI: <https://doi.org/10.3390/su14084676>
- Poquet, O, Tupikina, L and Santolini, M.** 2020. Are forum networks social networks? A methodological perspective. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, LAK '20*. Association for Computing Machinery, New York, NY, USA, pp. 366–375. DOI: <https://doi.org/10.1145/3375462.3375531>
- Santolini, M.** 2020. Covid-19: the rise of a global collective intelligence? *The Conversation*. <http://theconversation.com/covid-19-the-rise-of-a-global-collective-intelligence-135738>.
- Schaefer, T, Kieslinger, B, Brandt, M and van den Bogaert, V.** 2021. Evaluation in Citizen Science: The Art of Tracing a Moving Target. In: Vohland, K, Land-Zandstra, A, Ceccaroni, L, Lemmens, R, Perelló, J, Ponti, M, Samson, R and Wagenknecht, K (eds.), *The Science of Citizen Science*. Cham: Springer International Publishing, pp. 495–514. DOI: https://doi.org/10.1007/978-3-030-58278-4_25
- Senabre Hidalgo, E, Perelló, J, Becker, F, Bonhoure, I, Legris, M and Cigarini, A.** 2021. Participation and Co-creation in Citizen Science. In: Vohland, K, Land-Zandstra, A, Ceccaroni, L, Lemmens, R, Perelló, J, Ponti, M, Samson, R and Wagenknecht, K (eds.), *The Science of Citizen Science*. Cham: Springer International Publishing, pp. 199–218. DOI: https://doi.org/10.1007/978-3-030-58278-4_11
- Singh, CK, Barme, E, Ward, R, Tupikina, L and Santolini, M.** 2022. Quantifying the rise and fall of scientific fields. *PLOS ONE*, 17(6): e0270131. DOI: <https://doi.org/10.1371/journal.pone.0270131>
- Tackx, R, Blondel, L and Santolini, M.** 2021. Quantified Us: a group-in-the-loop approach to team network reconstruction. In: *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers, UbiComp '21*. Association for Computing Machinery, New York, NY, USA, pp. 502–507. DOI: <https://doi.org/10.1145/3460418.3479363>

TO CITE THIS ARTICLE:

Masselot, C, Jeyaram, R, Tackx, R, Fernandez-Marquez, JL, Grey, F and Santolini, M. 2023. Collaboration and Performance of Citizen Science Projects Addressing the Sustainable Development Goals. *Citizen Science: Theory and Practice*, 8(1): 45, pp. 1–16. DOI: <https://doi.org/10.5334/cstp.565>

Submitted: 21 September 2022 **Accepted:** 05 April 2023 **Published:** 27 June 2023

COPYRIGHT:

© 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Citizen Science: Theory and Practice is a peer-reviewed open access journal published by Ubiquity Press.

