



Impact of technology-supported collective intelligence on patient-led research

Bastian Greshake Tzovaras, Marc Santolini

► To cite this version:

Bastian Greshake Tzovaras, Marc Santolini. Impact of technology-supported collective intelligence on patient-led research. NESTA Collective Intelligence Grant. 2021. <hal-04167308>

HAL Id: hal-04167308

<https://hal.science/hal-04167308v1>

Submitted on 20 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Impact of technology-supported collective intelligence on patient-led research

Bastian Greshake Tzovaras^{1,2}, Marc Santolini^{2,3}

1. Open Humans Foundation, Sanford, USA

2. Université de Paris, INSERM U1284, Center for Research and Interdisciplinarity (CRI), Paris, France

3. Just One Giant Lab, Paris, France

Introduction

Our health and well-being are influenced by a large web of interacting factors; including biological, environmental, behavioural and social factors. This makes designing interventions and finding optimal strategies for care and treatments a highly complex problem. In addition, patients – especially with rare or chronic diseases – are often better experts at managing their disease, improving their quality of life and well-being, than medical professionals, due to their immediate experience and the impact on their lives. So far, this represents a “dark matter” for public health investigations, underused information that could actually lead to a better understanding of yet unknown complex associations between patient behavior and symptoms. To make use of these resources, there is increasing interest to tap into this tacit knowledge and formalize it through patient-led research (1,2).

Examples of patient-led research are many and diverse. Already in 2011, users of the web platform PatientsLikeMe created studies on amyotrophic lateral sclerosis (ALS) (3). Similarly, patients with cluster headaches – a neurological disease with a big unmet treatment need – developed their own research protocol to self-experiment on whether psilocybin might be a treatment option for them (4), with initial findings giving rise to an ongoing clinical trial (5). It is likely that patient-led research can not only minimize the costs for drug development and can bring them to market sooner (6), but it can also catalyze policy interventions (7).

Despite these promises, there remain significant technical and social challenges to the adoption of patient-led research. In particular, there is an unmet need of connecting patients with complementary skill sets to each other to enable impactful patient-led research. In line with this hypothesis, the use of matchmaking systems has been shown to help guide and foster volunteer contributions to existing citizen science projects on the online citizen science platform SciStarter (8). Yet, such contributions focus on projects *already* formed that are looking for specific, predetermined contributions. Here we consider the more general context of promoting engagement in the early stages of project development, emphasizing complex and open-ended contributions.

Research Objective

The self-organization of patients is currently limited by the inability to connect them with others that have the right skill sets to transform their ideas into conclusive research projects. By allowing patients to explicitly express open research needs in a structured form, we aim to facilitate such connections. We

hypothesized that this self-organization can be improved through targeted notifications that include the required skill sets and the topic of open needs. While the former incentivizes interactions based on declared attributes, the latter represents the role of shared interest based on undeclared, tacit knowledge.

Our objective was to test how these presented features, implemented in a recommendation system, could help the self-organization of patient-led research and lead to an improved effectiveness. In the long-term, we expect that repeated notification-driven engagement will alter the general community interactions on the platform in both structure and frequency, ultimately leading to self-sustained patient-communities.

Experimental Design/Protocol

Experimental design

In our initial design, we aimed to connect two existing platforms, *JOGL* and *Open Humans* to allow patients to collectively plan and implement projects. [JOGL](#) facilitates the mobilisation of contributors for open innovation projects, through a social and collaboration network approach. *Open Humans* focuses on data management/sharing and already engages patient communities that are designing their own PLR projects. As part of this connection, we aimed to implement our intervention, namely an email notification system to monitor the engagement of participants with open needs and relate it to the need characteristics (skill/location/topic).

The Covid pandemic created an unforeseen context that affected the deployment of this plan as it led to an organic integration of both communities that preceded the technological development. At the same time, a lot of traditional patient-led research – which happens within the boundaries of specific diseases and conditions – went on hold for a number of reasons: First of all, patients struggled with being less available both in terms of time and mental capacity, due factors like lockdowns, school closures and pandemic-related unemployment. At the same time, patients with chronic conditions are often particularly at risk of severe COVID infections, leading to a particularly urgent interest in pandemic response. This shift of interest in patient communities coincided with pandemic response interests from DIYBio and Maker communities, leading to the joint OpenCovid19 community.

As a result this community of patients and non-patients formed prior to the implementation of the recommendation system, allowing for longitudinal data collection & learning of barriers faced in the absence of automated matchmaking during this frame. We describe the observations made over this longitudinal investigation for two case studies – the patient-led Quantified Flu and CoughCheck projects.

After the technological implementation of the email recommendation system, four email campaigns were sent to all JOGL users, containing 2 needs each (either with a high skill/need match or randomly selected). We collected and monitored engagement data across these campaigns.

Recruitment

While we had planned to recruit participants for this experiment from the existing *JOGL* community and the *Open Humans* patient communities, due to the COVID-19 pandemic our recruitment strategy was adapted to the changed circumstances and was focused on the OpenCovid19 community, bringing together the JOGL & Open Humans communities, as well as around 4,000 new community members. This recruitment preceded the technical implementation of the OH-JOGL integration and the recommendation system.

Problems/Obstacles

The main obstacle we encountered during the implementation of our experimental design was the start of the COVID-19 pandemic, forcing us to significantly change our initial plan. In particular, as outlined above, the recruitment of patient communities was hindered. Beyond this, the creation of a custom system allowing us to track which recommendations were sent, then clicked by a user on a mailer system while being able to regenerate recommendations was more challenging than expected. Unforeseen incompatibility issues occurred between different software development environments and programming languages. The problem has been resolved by a more complex but isolated system that encapsulates the recommended system engine. Because of the delay generated by this, the first wave of emails could only be sent on Thursday Feb 26 2021.

Given these two problems, we collected only a limited amount of observations. While these can be sufficient to generate basic statistics, they limit our ability to generalize given the general community shift during the COVID-19 pandemic. This creates a confounder with its own user dynamics and a baseline engagement that might not represent a bona fide “control” situation. Further data collection in the ongoing implementation of the system will be needed to overcome this exogenous factor.

Analysis/Results

Case studies

Given the change in experimental design necessitated by the COVID-19 pandemic we decided to follow two projects – *Quantified Flu* & *CoughCheck* – that were created as part of the OpenCovid19 community (Fig 1) as case studies. Both projects included members of the Open Humans communities and patients amongst project participants and members of our team were involved as participant observers.

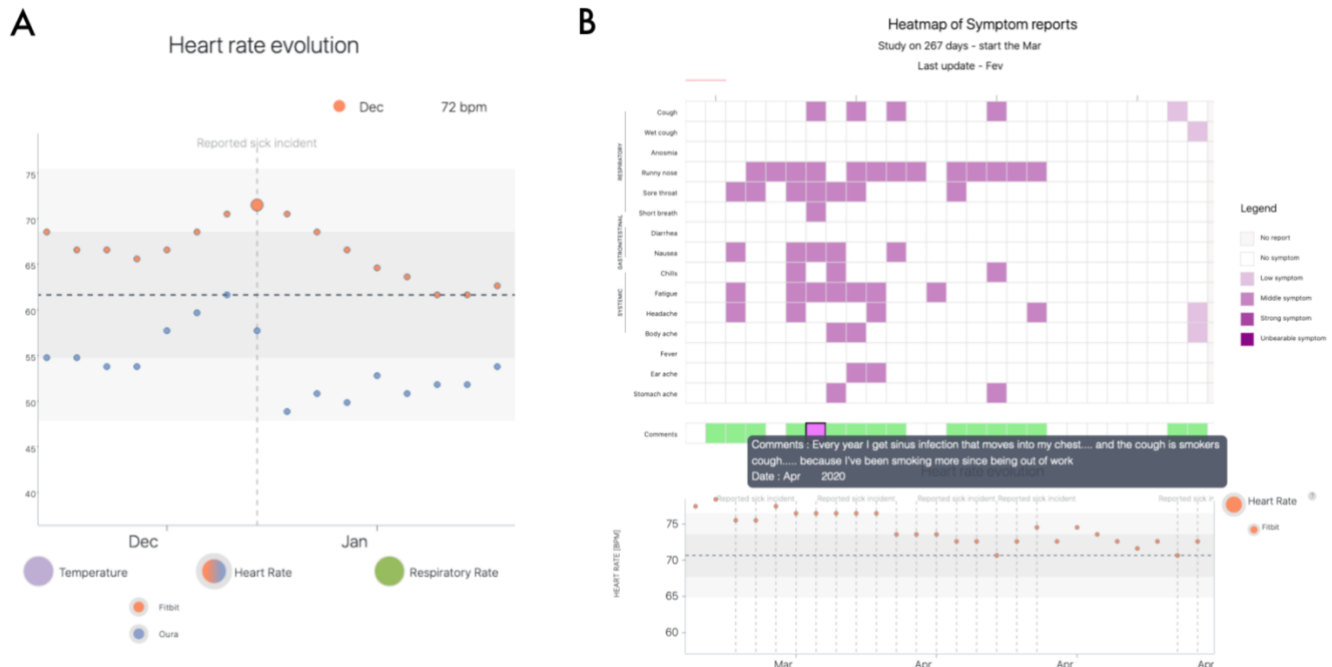


Figure 2: Screenshot of the Quantified Flu web app. A: Retrospective visualizations allow users to make sense of former sickness events B: On-going symptom reporting shows symptoms & physiological data alongside each other

Collective Intelligence during the development

Following the initial brainstorming session during the community call, a Slack channel was created in the OpenCovid19 Slack for the project to manage both the further brainstorming as well as the development of the *Quantified Flu* website. The development process was done collaboratively, using an open source approach with the code being managed on GitHub. Beyond the initial set of community members, the participants also joined the JOGL OpenCovid19 community to advertise open needs and find additional contributors for the software implementation. The Slack channel eventually gathered a total of 146 members with 50 active users over time.

As the automated matchmaking process on JOGL was not implemented at this time yet, Quantified Flu members actively reached out to other JOGL/OpenCovid19 participants to find and recruit more active contributors. In this particular case study this approach worked, yielding additional code contributors that provided expertise in skills lacking from the initial group (e.g. data visualizations in JavaScript and iOS mobile development) Collectively, a total of 844 messages by 50 individuals were sent on the Slack channel and 316 commits were done on GitHub by 7 active contributors between March 10 and September 2020. Importantly, this work collectively engaged patients coming from a variety of different patient communities as well as non-patients together.

A netnographic analysis – which bins messages into subcategories of *Ideas*, *Concept*, *Prototyping* and *Social* – of the *Quantified Flu* Slack shows the use of collective intelligence throughout the development cycle (see Fig 3). While topics like *Protocol Design* and *Feature suggestions* are frequent at the early stages of development, we see them appear in bursts throughout the whole development period, indicating

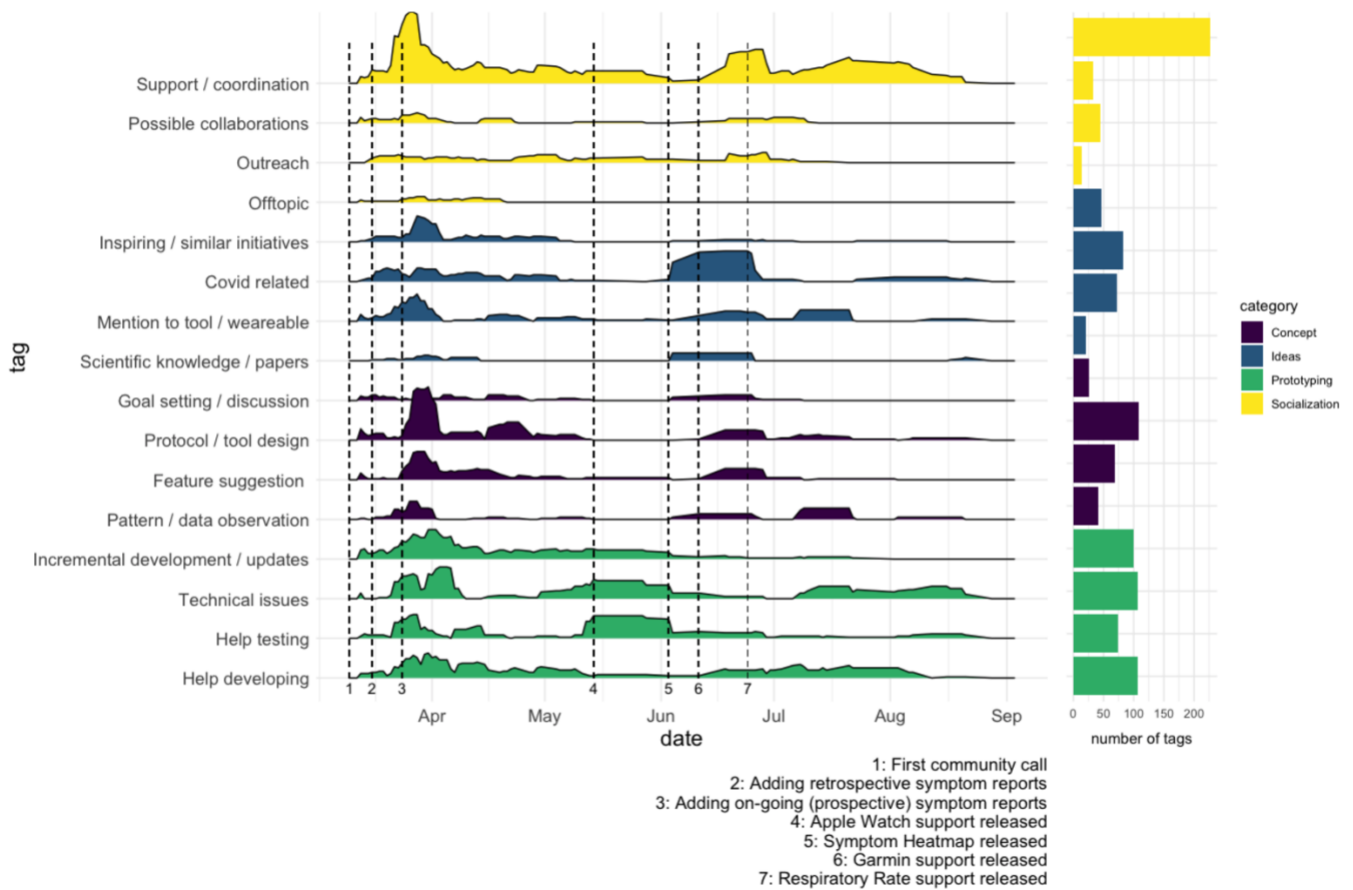


Figure 3 Netnographic analysis showing the evolution of the activity of the Quantified Flu Slack channels across an array of qualitative dimensions.

how new feature releases spark new feature ideas and vice versa. We also observe that *Help developing/testing* is actively used throughout the process.

Importantly, some of the features most used by the community only developed through the co-creation process the community engaged in.

Consequences of the collective intelligence-enabled co-design

In the latest prototype stage, *Quantified Flu* presents a working *mobile Health* (mHealth) prototype that was designed by and with the future users of it. The application engaged 190 individuals, with 92 of them using symptom tracking. Typically, mHealth applications show participation inequalities similar to those found in other online communities, where user activity decays according to power-law functions. In contrast, for *Quantified Flu* we unexpectedly observe a linear decay, showing that users are more engaged than typical. Furthermore, most mHealth applications struggle with high rates of user drop out, with as little as 2% of users showing sustained use. In contrast, over 50% of *QF* users exhibit daily use for a period of more than 3 months (see Fig 4). We attribute this high user-engagement over a prolonged period of time to the co-creation approach that led to the prototype: Prior studies have found that users are more

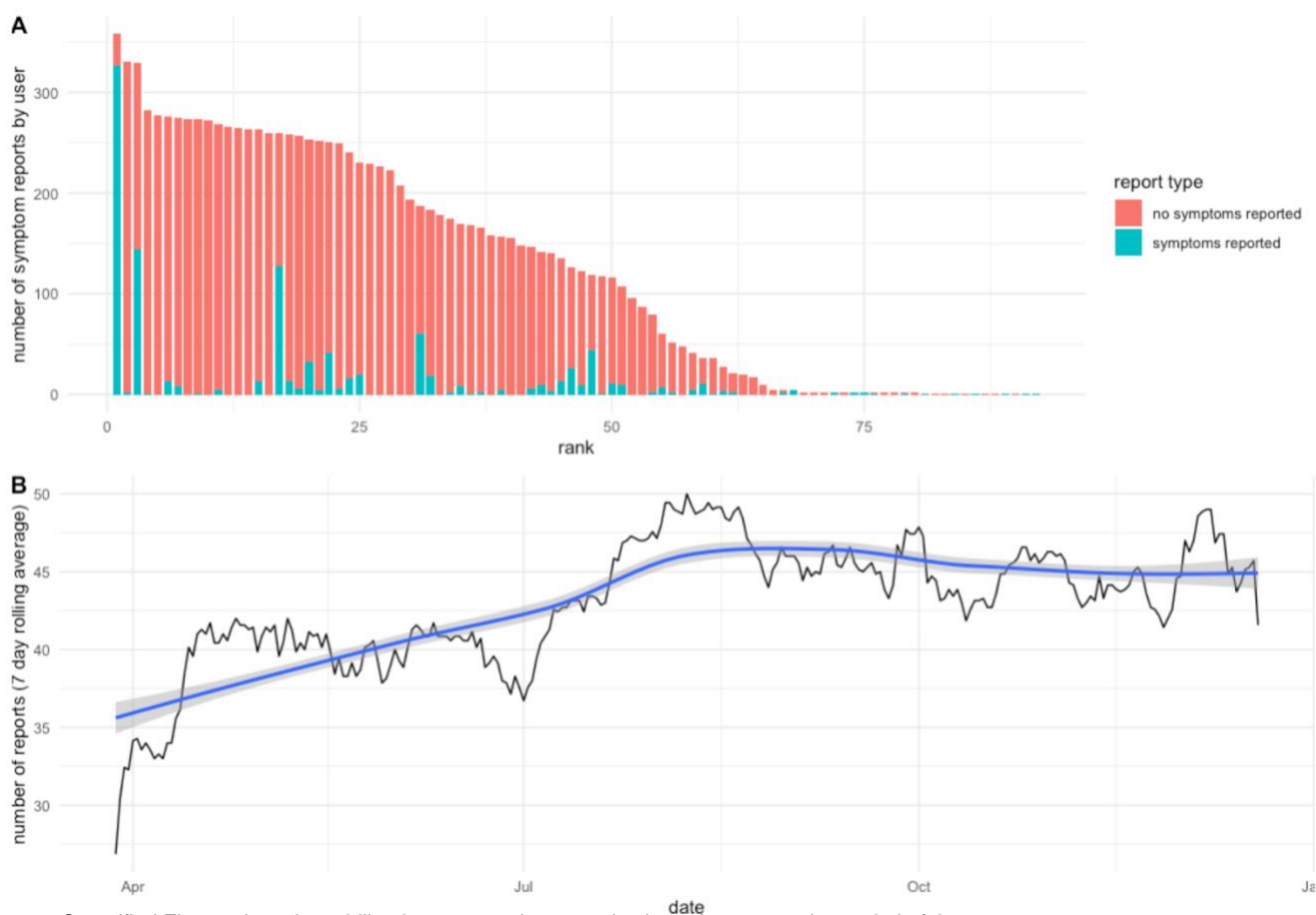


Figure 4 Quantified Flu consistently mobilized an engaged community through an extensive period of time.

likely to continue using mHealth apps if there is a good fit between user and application (9), and a co-creation process with the future users – as applied for QF – can be a key way of achieving this fit.

CoughCheck

Similarly to the *Quantified Flu* case study, the *CoughCheck* project emerged at the early days of the pandemic in March 2020. The project's goal was to release a mobile app that would make use of AI to detect the presence of COVID-19 based on the sound of a person's cough with the intermediate goal to collect enough cough audio recordings for the initial training of the machine learning algorithms. The project was spearheaded by participants from Latin America with a focus on Argentina in particular, as the accessibility of traditional COVID tests was extremely limited there. Started by individuals (patients- and non-patients) that weren't initially neither part of the JOGL or Open Humans communities, they joined the OpenCovid19 program and Open Humans in March. In particular, the *CoughCheck* project engaged with the patient communities of Open Humans for expertise in how to responsibly handle patient-data and getting patient insights on their lived experiences.

Unlike the *Quantified Flu* project, this project did encounter significant problems during the prototyping and implementation phase, leading to no prototype being developed at the end. Despite being able to attract 64 participants to their active Slack channels (1,144 messages sent) and holding regular, virtual

community meetings similar to QF, the *CoughCheck* project failed to attract the “right” kind of contributors to take their project forward. As the project aimed to develop a mobile app, *CoughCheck* was looking for contributors with expertise in UI/UX design, mobile development, and machine learning. And while the project managed to attract a large number of members through JOGL, most of them did not bring the skills that were actually needed by the project. Instead we observed an excess of participants with skills in marketing and project management that were not able to contribute to the development that was needed at the time. As a result the project came to an end in May 2020 without the actual implementation of a prototype.

Case studies: Lessons learned

The two different outcomes of the *Quantified Flu* & *CoughCheck* case studies – which started off from similar points of departure (engaged and motivated communities, significant growth in contributors) – highlight the importance of matching and connecting the “right” people to enable collective intelligence. In the absence of an automated matchmaking and notification system, both projects had to manually recruit contributors or hope that the needed contributors would find the project on their own with dramatically different outcomes: In the case of *Quantified Flu*, the identification of needed volunteers succeeded, leading to the iterative development of a project that ended up in a mature state that has already been written up in a scientific publication that’s already available as a preprint (10) and is currently in revision with an academic journal. In the case of *CoughCheck* this match between project needs and contributor skills could not be achieved, leading to the project faltering after a few weeks of initial activity.

Implementation of a recommender system

Since we could not recruit the patient communities as initially planned *after* the development of the recommendation technology, we decided to focus on the whole JOGL community to provide a proof-of-concept of how such a recommender system could engage users in needs formulated by other projects. In order to facilitate interactions between community members across the vast number of projects, we developed a mailer that sends 2 recent needs (not more than 3 months before the recommendation) to

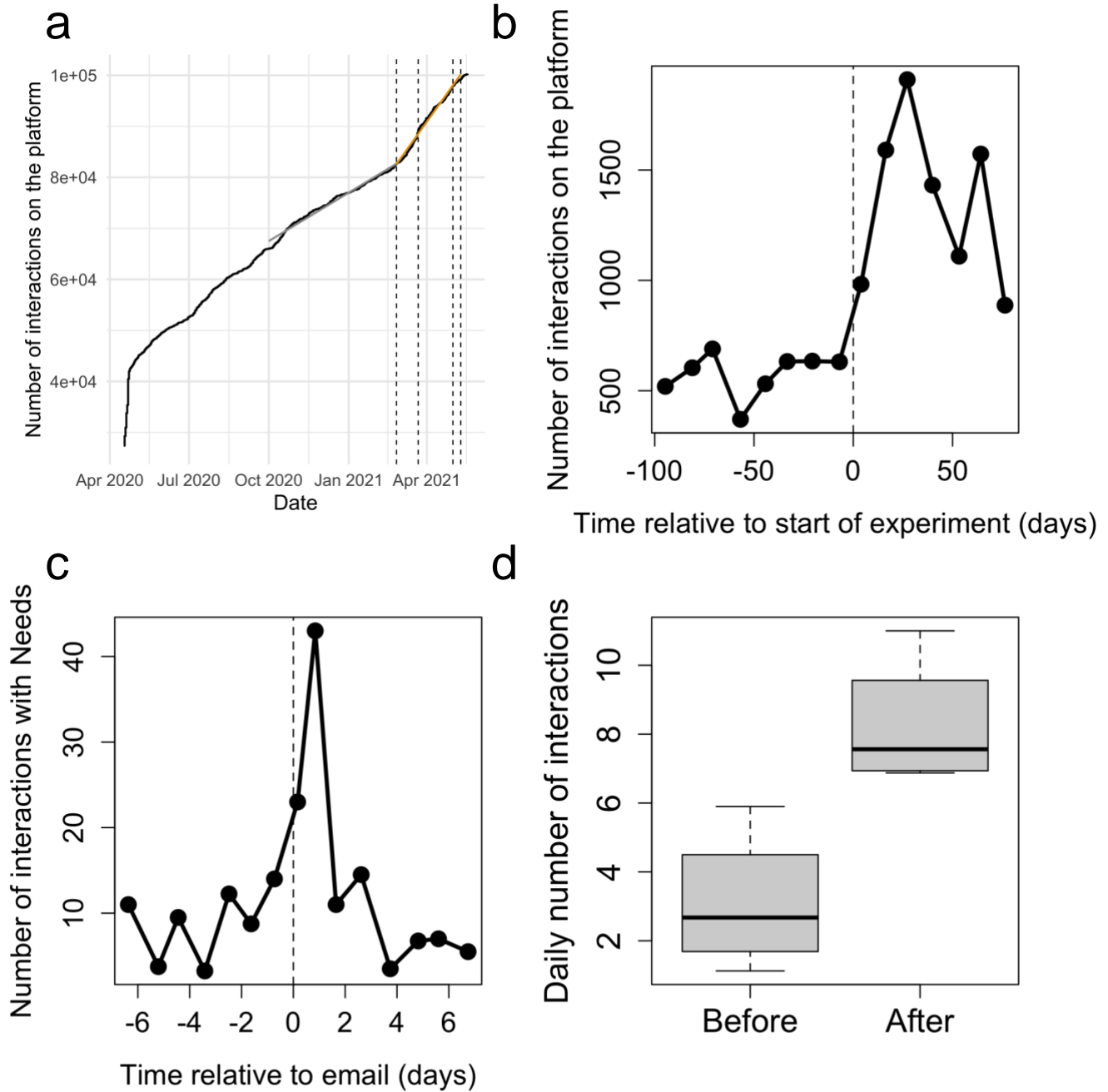


Figure 5 a) Cumulative evolution of the number of interactions on the platforms. Interactions denote actions that a user performs over a project, post, need, or another user (e.g follow, clap, or save). Dashed lines indicate email campaigns, and gray and yellow lines indicate best fits over a 6 months period before and 2.5 months period after the first email. b) Number of interactions on the platform relative to the start of the experiment, showing the impact of the first email campaign on platform activity. c) Number of interactions with needs relative to an email campaign (1 week before and after), showing a specific increase in activity for the 2 days following an email. d) Average daily number of interactions during the 2 days after (right) and the 7 days before (left) for the four campaigns.

the users of the platform. Among the two needs, the first one is recommended based on network proximity using the method of (11), while the other one is taken at random in the recent needs. The recommender system was deployed in February 2021, and four email campaigns were sent afterwards (Fig 5a).

test for the presence of SARS-CoV-2 using multiple approaches." or *"CoronaDetective: A project to develop an open-source detection method for SARS-CoV-2, analogous to the GMO Detective."*) or education (e.g. *"edgame: PC Serious Game : learning about COVID-19 and sanitary measures to stay healthy and limit the spread."* or *"studentwebinars: help students all over the world understand what viruses are, why/how they make us sick and how we can fight them"*). Several of these projects have benefited from a micro-grant program, and they had to document their project properly for reviewers to assess them. In addition, these projects benefited from a community interest and regular community calls.

Finally, we observe a small impact of geographical proximity on engagement: needs that are part of projects which are less than 100km to the user show a stronger engagement.

Overall, this indicates that skills similarity, project popularity and activity are drivers of engagement, while proximity is a milder indicator. The top two predictors in our case are project popularity (project saves, $p=1.9e-4$) and shared skills with the project ($p=1.1e-2$), closely followed by project activity (number of needs in project $p=1.3e-2$) and number of shared skills with the need ($p=1.6e-2$). This highlights the fact that engagement is driven by general interest in active projects. As such, the help that can be provided for a given need might not be direct, but rather indirect through the opportunities that the engaged user can provide to the project (their own connections towards relevant people who could help). We note however that the small amount of data collected does not allow final conclusions. Indeed, the skills required for recent needs might have been irrelevant to the wider community, or might have been too precise to be matched correctly to existing participants. A future refinement of the algorithm should focus on creating a smaller, coarse-grained database of skills and integrate dimensions other than skills that could connect a user to the need, for example whether projects are part of a shared community of interest (part of the same JOGL program or subgroup).

Conclusions/Reflection

While our experimental plan had to adapt quite dramatically, we observe that matchmaking can be key to drive the growth and success of nascent open, patient-driven projects. Due to the pandemic circumstances our initial experimental design was flipped on its head, with a large enrollment of new participants preceding the implementation of automated recommendations. In order to adapt to this new reality, we first engaged in a wizard-of-oz like experiment, in which the matchmaking between unfilled project needs and newly joined community members was done manually by community managers as well as by individual patient-driven projects themselves to attract appropriate new contributors to complement their existing skill sets. The recommender system was eventually released after the specific community engagement on patient-led research projects, allowing to gather preliminary results on the role of project activity, popularity, and to some extent skill similarity and geographical proximity on engagement. Importantly, the launch of the recommender system triggered a consistent uptake in community activity, showing how it can help foster interactions. Given the felt urgency of the pandemic, our first manual matchmaking was not accompanied by a detailed data collection. If a similar situation should arise in the future, we would pay particular attention to generate more rigorous learnings at this early stage, through the distribution of surveys amongst community members to both understand their needs as well as the effectiveness of the manual matchmaking.

Beyond this, our two highlighted case studies show that this manual approach to matchmaking is not without its problems, as it is unlikely to scale beyond a small set of participants. This is particularly true given our observation that most of the community activity during the start of the OpenCovid19 program was by motivated individuals & projects that looked for particular skill sets among the other contributors, while these efforts were hindered by the lack of automating this matchmaking.

Furthermore, our case studies show that growing the number of contributors alone is not enough for projects to successfully implement collective intelligence. Rather, projects require a set of contributors with diverse skills to cover the broad spectrum of needs each project has. While both *Quantified Flu* and *CoughCheck* attracted similar numbers of contributors overall, only the former project succeeded in covering the required expertise to a level that a successful prototyping could begin, while the latter suffered from a lack of diversity in participant expertise.

Overall, these points all highlight the importance of smart systems to facilitate the automated matchmaking between existing or nascent patient-led projects with contributors that can bring necessary expertise. As such we are excited to continue our work following this report, to further explore the impact of the now-implemented matchmaking and how it can be optimized to support further patient-led research.

Wider significance and lay summary

Patient-led research often focuses on the creation of otherwise ‘undone science’ with the aim of improving the quality of life of patients that are affected by a given disease or condition. These projects – in particular if they involve chronic disease – are powered by the expertise and lived experience of the patients themselves. But typically, patient-led research requires a complex network of skills that goes beyond the first-hand patient-experiences, such as software development, data analysis, community engagement and many more. Matching interested participants/patients that have the right skills with patient-led research projects is thus key to ensure their success. In this way, patient-led research is emblematic of open innovation projects at large, which all face similar issues. As such our findings on the importance of automating relevant matchmaking are applicable as well for the larger open innovation field.

Just as patient-led research, open innovation approaches implemented in hackathons, online challenges, and even open source software development rely on the creation of self-organized teams. At present the success of this self-organization frequently relies on luck – whether they are able to attract and find contributors with the needed, complementary skills. Our preliminary findings and case studies show both how matching participants can work out successfully and unsuccessfully, highlighting the importance of relevant matchmaking for improving success rates in open innovation settings. Beyond the importance of matchmaking based on an exact matching of skill sets, we find that the quality of project documentation, characterized by project activity (number of posts and needs) and popularity (number of “claps”, a publicly visible marker of popularity, as well as “saves”, a nonpublic marker of interest), is a strong driver of engagement for potential new contributors.

These findings highlight the fact that contributions can come from serendipitous interactions with generally interested individuals who can create further bridges towards relevant resources. We expect these insights to be relevant for a number of open innovation communities that rely on volunteer contributors. In

particular, the use of regular notifications promoting simple yet useful actions is important to build an “architecture of attention” for large communities. In addition, we found that project activity leads to a stronger engagement. As such, in addition to notifications, it is important to promote guidelines for projects to regularly document their activity in the form of posts and needs. Such an activity is in turn rewarded as such projects will attract more attention, generating further activity (posts, needs) and increasing their popularity (claps, saves).

In order to communicate our learnings from the OpenCovid19 community organisation challenges, we have developed the [community playbook](#) “How to Build and Manage Open Communities: the top 47 do’s and don’ts for building and engaging open communities”. This playbook is aimed at practitioners and has been made available as an open project on JOGL. In addition, the development of the recommender system has generated interest at the Center for Research and Interdisciplinarity (CRI Paris), with whom we have applied to several public funding agencies (EU, ANR) to leverage and improve recommender systems to propose relevant similar projects to newly developed projects. More generally, we are interested to continue developing and monitoring the recommender system in JOGL for the various objects on the platform, adapting to the needs of the communities that partake in challenges and programs.

Finally, we have made our framework openly available by sharing and documenting the [open source code](#) used in the implementation of the recommender system. This will allow other practitioners to apply the lessons learned during this project and replicate them in other contexts.

Programme impact

While JOGL and Open Humans have been leveraging collective intelligence in its communities for a number of years already, the urgency of the COVID-19 pandemic along with the jointly formed new OpenCovid19 community made us appreciate very new challenges in relation to collective intelligence. In particular, the rapid growth of the community with over 1,000 new contributors joining within the span of only a few weeks made us appreciate the limitations of human-facilitated collective intelligence.

Before the start of the program both Open Humans and JOGL grew at a pace where manual matchmaking and community management was possible. Within the pandemic-context this quickly came to its limits as the demand for community management and matchmaking rapidly outgrew our capacities. As such we feel more confident about the need to plan ahead for the future and investigate more closely how machine-facilitated collective intelligence can be implemented in our individual communities as well as in the cross-cutting OpenCovid19 community to improve the success rates of patient-led research projects.

Furthermore, as part of the Nesta programme, we got the chance to interact with a number of other collective intelligence practitioners and learned a lot from them. In particular, we saw how valuable it is to start collecting structured data and qualitative insights early on to better understand and shape both our communities as well as the experimental design itself.

Beyond the other grant awardees, the method used by Nesta itself to organize the collective intelligence of the awardees on a purely online setup through regular workshops and meetings with the other projects

was a valuable learning experience. We could see firsthand how to organize and structure a program that includes a diverse set of participants with very different backgrounds and projects.

As a next phase of our experiment, we plan to explore in-depth how automated recommendations can improve patient-led projects by implementing our initially planned experiments: indeed, the necessary technology is now in place and the community has been integrated, providing the ground for a continuous monitoring and improvement of our system. Which means that once the pandemic situation improves across the globe, we will be able to engage in more ‘traditional’, disease-specific patient-led research as initially planned. To support our future work in that direction, both Marc Santolini and Bastian Greshake Tzovaras have recently submitted independent proposals to the French national research funding agency (ANR): “COPPA: Co-learning & knowledge practices of personal science communities” and “CORES: Quantitative assessment of open collaborations in science and engineering”, aimed at continuing our work on collective intelligence and patient-led research. These proposals are now in second round of review for a final decision in July.

Moreover, we have in the past year developed a micro-grant system based on community open peer review that provides fine-grain measures of project success along a diversity of axes, from feasibility to originality and impact. By associating measures of community interactions on the platform with such measures of project success, we will be able to obtain a more granular understanding of organization principles that guide collective intelligence in community-based research.

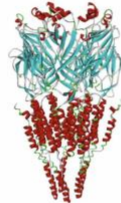
Appendix

Example recommendation email sent to the JOGL users, highlighting two recent needs from the platform:



Hello 🙋
We've been thinking about you!

We believe these projects could use your help. Have a look at them and don't hesitate to leave a comment.
This is a new JOGL beta feature. These recommendations are based on your profile and interactions on the JOGL platform and also highlight recent needs expressed on the platform.



COMBINATION ENTRY BLOCKER (RHN-ACE2) - ATPase INHIBITOR

What we do

Review of literature and RD research of ATPase inhibitor or building new molecules of interest, maybe it is the solution to this crisis !

What we need

DATA COLLECTION ON SARS-COV2 CLINICAL CASES
I need data collection skills to answer a few questions of this project. If you think you can help in this area !

What we're looking for

Data collection
Data science

[READ MORE](#)



📄 136

OPEN DATA FOR COMBATING ANTIMICROBIAL RESISTANCE IN AFRICA

What we do

MicroBIS uses AI to identify bacteria and determine antibiotics sensitivity-resistance patterns of antibiotics.

What we need

RESOURCES FOR BUILDING MICROBIS

The resources needed for building a fully functional web platform for displaying the AI models that interprets the core functions of MicroBIS are divided into two parts. First one: technical/human resource needs and developer needs. The current project team can provide the complete technical and human resource needs for developing the MicroBIS platform. The deficiency is with the developer needs which includes: Cloud storage and hosting service, high performance computers, etc.

What we're looking for

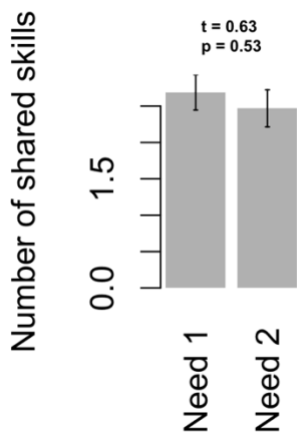
Data science
Devop
Web dev

[READ MORE](#)

This is an experimental feature. We'd love to hear your feedback. If so, you can join us to this group.

Want to learn more?
[Join us on social media!](#)

Impact of the order of needs in user behavior



In the experiment, the first need in the email was based on the proximity between the user and the need in the JOGL network (including users, skills, projects, needs, programs, interests and their relations). However, due to the focus of the program, the recent (“random”) needs, sent as the second need in the email, also contained relevant skills for the users. As a result, as can be observed in the figure on the left, we observed no significant difference in the number of shared skills between Needs 1 and 2. Therefore, this experiment simulates a context where the “relevant” needs are shown in a random order to the user (in terms of rank in the email), allowing to infer the conditions that best engage individuals to interact with needs.

State of the art and novelty

In their seminal 2010 paper, Woolley et al (12) provided evidence that groups exhibit a collective intelligence that cannot be reduced to any individual, but rather builds on group diversity and interactions. They concluded by asking: *“Could a group’s collective intelligence be increased by, for example, better electronic collaboration tools?”*. A similar insight was promoted by Michael Nielsen in the context of large collaborations in citizen science, arguing that “Once the collaboration gets large enough participants cannot possibly pay attention to everything that’s going on [...]. Ideally, the architecture of attention will direct participants to places where their particular talents are best suited to take the next step.” (13). The development of such an “architecture of attention” was experimented in the context of the Citizen Science platform SciStarter, showing how personalized recommendations enhance engagement of contributors towards projects that better reflect their interest as determined by their past contributions (8). Yet, in the context of these studies the contributions are simple and pre-determined. In the present work, we tackle the challenge of open-ended, complex contributions at the early phase of open research projects. Using the rise of patient-led research as a test case (4), and an open-science community as a broader context (14), we examine how a clear signaling of needs that is paired with a notification system in a social platform can influence collective intelligence in domains where group diversity (e.g. in skills and knowledge) is required.

References

1. Mader LB, Harris T, Kläger S, Wilkinson IB, Hiemstra TF. Inverting the patient involvement paradigm: defining patient led research. *Res Involv Engagem*. 2018 Jul 10;4(1):21.
2. Vayena E, Brownsword R, Edwards SJ, Greshake B, Kahn JP, Ladher N, et al. Research led by participants: a new social contract for a new kind of research. *J Med Ethics*. 2016 Apr 1;42(4):216–9.
3. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat Biotechnol*. 2011 May;29(5):411–4.
4. Kempner J, Bailey J. Collective self-experimentation in patient-led research: How online health communities

foster innovation. *Soc Sci Med* 1982. 2019 Oct;238:112366.

5. D'Souza DC. Safety and Efficacy of Psilocybin for the Treatment of Headache Disorders [Internet]. *clinicaltrials.gov*; 2021 Jan [cited 2021 May 27]. Report No.: NCT02981173. Available from: <https://clinicaltrials.gov/ct2/show/NCT02981173>
6. Levitan B, Getz K, Eisenstein EL, Goldberg M, Harker M, Hesterlee S, et al. Assessing the Financial Value of Patient Engagement: A Quantitative Approach from CTTI's Patient Groups and Clinical Trials Project. *Ther Innov Regul Sci*. 2018 Mar;52(2):220–9.
7. Katapally TR, Bhawra J, Leatherdale ST, Ferguson L, Longo J, Rainham D, et al. The SMART Study, a Mobile Health and Citizen Science Methodological Platform for Active Living Surveillance, Integrated Knowledge Translation, and Policy Interventions: Longitudinal Study. *JMIR Public Health Surveill*. 2018 Mar 27;4(1):e31.
8. Dayan N, Gal K, Segal A, Shani G, Cavalier D. Intelligent recommendations for citizen science. In: *CEUR Workshop Proceedings* [Internet]. CEUR-WS; 2020 [cited 2021 May 19]. Available from: <https://asu.pure.elsevier.com/en/publications/intelligent-recommendations-for-citizen-science>
9. Vaghefi I, Tulu B. The Continued Use of Mobile Health Apps: Insights From a Longitudinal Study. *JMIR MHealth UHealth* [Internet]. 2019 Aug 29 [cited 2021 May 31];7(8). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6740166/>
10. Tzovaras BG, Hidalgo ES, Alexiou K, Baldy L, Morane B, Bussod I, et al. Quantified Flu: an individual-centered approach to gaining sickness-related insights from wearable data. *medRxiv*. 2021 Mar 12;2021.03.10.21252242.
11. Ramaciotti Morales P, Lamarche-Perrin R, Fournier-S'niehotta R, Poulain R, Tabourier L, Tarissan F. Measuring Diversity in Heterogeneous Information Networks. *ArXiv E-Prints*. 2020 Jan 1;2001:arXiv:2001.01296.
12. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*. 2010 Oct 29;330(6004):686–8.
13. Nielsen MA. *Reinventing Discovery: The New Era of Networked Science*. Princeton, NJ, USA: Princeton University Press; 2011. 280 p.
14. Kokshagina O. Open Covid-19: Organizing an extreme crowdsourcing campaign to tackle grand challenges. *RD Manag*. 2021 Mar 23;radm.12470.