Proceedings Paper:
Jan-Christoph, K., Lee, J.-U., Stowe, K. et al. (2023) Lessons learned from a Citizen Science project for Natural Language Processing. In: Vlachos, A. and Augenstein, I., (eds.) Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics. 17th Conference of the European Chapter of the Association for Computational Linguistics, 02-06 May 2023, Dubrovnik, Croatia. Association for Computational Linguistics, pp. 3594-3608. ISBN: 9781959429449.

Lessons Learned from a Citizen Science Project for Natural Language Processing

Jan-Christoph Klie[1] **Ji-Ung Lee**[1] **Kevin Stowe**[1,2] **Gözde Gül Şahin**[1,3]
Nafise Sadat Moosavi[1,4] **Luke Bates**[1] **Dominic Petrak**[1]
Richard Eckart de Castilho[1] **Iryna Gurevych**[1]

[1]Ubiquitous Knowledge Processing Lab (UKP Lab)
Department of Computer Science and Hessian Center for AI (hessian.AI)
Technical University of Darmstadt
www.ukp.tu-darmstadt.de

[2]Educational Testing Service
[3]KUIS AI, Koç University
[4]Department of Computer Science, The University of Sheffield

Abstract

Many Natural Language Processing (NLP) systems use annotated corpora for training and evaluation. However, labeled data is often costly to obtain and scaling annotation projects is difficult, which is why annotation tasks are often outsourced to paid crowdworkers. Citizen Science is an alternative to crowdsourcing that is relatively unexplored in the context of NLP. To investigate whether and how well Citizen Science can be applied in this setting, we conduct an exploratory study into engaging different groups of volunteers in Citizen Science for NLP by re-annotating parts of a pre-existing crowdsourced dataset. Our results show that this can yield high-quality annotations and attract motivated volunteers, but also requires considering factors such as scalability, participation over time, and legal and ethical issues. We summarize lessons learned in the form of guidelines and provide our code and data to aid future work on Citizen Science.[1]

1 Introduction

Data labeling or *annotation* is often a difficult, time-consuming, and therefore expensive task. Annotations are typically drawn from domain experts or are crowdsourced. While experts can produce high-quality annotated data, they are expensive and do not scale well due to their relatively low number (Sorokin and Forsyth, 2008). In contrast, crowdsourcing can be relatively cheap, fast, and scalable, but is potentially less suited for more complicated annotation tasks (Drutsa et al., 2020). Another approach is using Citizen Science, which describes the participation and collaboration of volunteers from the general public with researchers to conduct science (Haklay et al., 2021). Over the past decade, Citizen Science platforms, which rely on unpaid volunteers to solve scientific problems, have been used for a wide variety of tasks requiring human annotation (Hand, 2010), e.g., classifying images of galaxies (Lintott et al., 2008) or for weather observation (Leeper et al., 2015).

While Citizen Science has been shown to produce high-quality annotations in ecological or environmental projects (Kosmala et al., 2016), its potential has so far not been investigated in depth for Natural Language Processing (NLP). Our goal in this work is to assess the practicality of undertaking annotation campaigns for NLP via Citizen Science. We analyze whether volunteers actually react to our calls and participate, how the resulting quality is compared to crowdsourcing, what the benefits and shortcomings are and what needs to be taken into account when conducting such a project. We especially are interested in differences between annotators recruited via different channels, which we investigate by advertising to different social media platforms, NLP-related mailing lists, and university courses. To explore this possibility, we use the PERSPECTRUM dataset (Chen et al., 2019, CC-BY-SA) that focuses on the task of stance detection and can be motivated by fighting misinformation and promoting accurate debate in internet discussions. We replicated a portion of the annotations in this dataset using citizen scientists instead of crowdworkers. To accomplish this goal, we designed an annotation workflow that is suitable for Citizen Science and allows us to recruit volunteers across a variety of platforms.

[1]https://github.com/UKPLab/eacl2023-citizen-science-lessons-learned

Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, pages 3594–3608

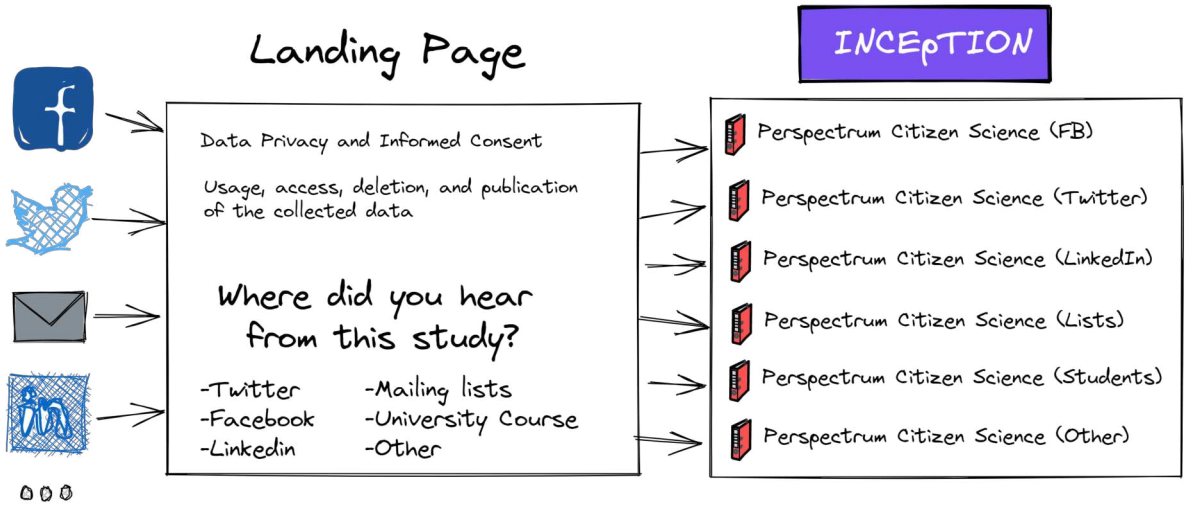


Figure 1: We advertised our project via various social media, mailing lists and university courses. Volunteers then are onboarded via the landing page and donated annotations via INCEpTION.

Our contributions are the following:

1. We provide a systematic study on Citizen Science across different channels and analyze turnout and quality. For this, we re-annotate parts of the PERSPECTRUM dataset using Citizen Science and compare these to the original, crowdsourced annotations.

2. We provide guidelines and recommendations on how to successfully conduct a Citizen Science project for NLP annotation and discuss critical legal and ethical aspects.

3. We provide a platform for future Citizen Science projects that handles onboarding, anonymous access, work assignment and the annotating itself.

Our results show that using Citizen Science for linguistic annotation can result in high-quality annotations, but that attracting and motivating people is critical for its success, especially in the long-term. We were able to attract 98 volunteers when conducting our Citizen Science project which resulted in 1,481 annotations over 2 months, thereby re-annotating around 10% of the original dataset. We find that annotations obtained through mailing lists and university students were of high quality when comparing them to the original, adjudicated crowdsourced data. We thus conclude that Citizen Science projects have the potential to be applied to NLP annotation if they are conceptualized well, but are best suited for creating smaller datasets.

2 Background

Prior work has developed various means and strategies for annotating large datasets. So far, annotation studies in NLP mainly use domain-experts or crowdworkers, or a mix of both (Nguyen et al., 2015). Crowdsourcing in particular has received increasing attention over the past decade (Wang et al., 2013).

Paid Experts Recruiting domain experts (e.g., linguists) for annotation studies has been a widely accepted method to generate linguistically annotated corpora. Famous examples are the Brown Corpus (Francis and Kucera, 1979) or the Penn Treebank (Marcus et al., 1993). While the resulting datasets are of the highest quality, domain experts are often few, and such annotation studies tend to be slow and expensive (Sorokin and Forsyth, 2008). Although many researchers moved on to annotation studies that recruit crowdworkers, expert annotations are still necessary in various fields, e.g., biomedical annotations (Hobbs et al., 2021).

Crowdsourcing To accelerate the annotation process and reduce costs, researchers have utilized crowdsourcing as a means to annotate large corpora (Snow et al., 2008). The main idea behind crowdsourcing is that annotation tasks that do not require expert knowledge can be assigned to a large group of paid non-expert annotators. This is commonly done via crowdsourcing platforms such as Amazon Mechanical Turk (AMT) or Upwork and has been successfully used to annotate various datasets across different tasks and

domains (Derczynski et al., 2016; Habernal and Gurevych, 2017). Previous work compared the quality between crowdsourcing and expert annotations, showing that many tasks can be given to crowdworkers without major impact on the quality of annotation (Snow et al., 2008; Hovy et al., 2014; De Kuthy et al., 2016).

Although crowdworkers can substantially accelerate annotation, crowdsourcing requires careful task design and is not always guaranteed to result in high quality data (Daniel et al., 2018). Moreover, as annotators are compensated not by the time they spend but rather by the number of annotated instances, they are compelled to work fast to maximize their monetary gain—which can negatively affect annotation quality (Drutsa et al., 2020) or even result in spamming (Hovy et al., 2013). It can also be difficult to find crowdworkers for the task at hand, for instance due to small worker pools for languages other than English (Pavlick et al., 2014; Frommherz and Zarcone, 2021) or because the task requires special qualifications (Tauchmann et al., 2020). Finally, the deployment of crowdsourcing remains ethically questionable due to undervalued payment (Fort et al., 2011; Cohen et al., 2016), privacy breaches, or even psychological harm on crowdworkers (Shmueli et al., 2021).

Games with a Purpose A related but different way to collect annotations from volunteers is *games with a purpose*, i.e., devising a game in which participants annotate data (Chamberlain et al., 2008; Venhuizen et al., 2013). Works propose games for different purposes and languages. For instance, anaphora annotation (PhraseDetectives, Poesio et al. 2013), dependency syntax annotation (Zombilingo, Fort et al. 2014), or collecting idioms (Eryiğit et al., 2022). It has been shown that if a task lends itself to being gamified, then it can attract a wide audience of participants and can be used to create large-scale datasets (von Ahn, 2006). Finally, Lyding et al. (2022) investigate games with a purpose in the context of (second) language learning to simultaneously crowdsource annotaions from learners as well as teachers. One such example is Substituto, a turn-based, teacher-moderated game for learning verb-particle constructions (Araneta et al., 2020). We do not consider gamification in this work, as enriching tasks with game-like elements requires considerable effort and cannot be applied to every task.

Citizen Science Citizen Science broadly describes participation and collaboration of the general public (the citizens) with researchers to conduct science (Haklay et al., 2021). Citizen Science is a popular alternative approach for dataset collection efforts, and has been successfully applied in cases of weather observation (Leeper et al., 2015), counting butterflies (Holmes, 1991) or birds (National Audubon Society, 2020), classifying images of galaxies (Lintott et al., 2008) or monitoring water quality (Addy et al., 2010). Newly-emerging technologies and platforms further allow researchers to conduct increasingly innovative Citizen Science projects, such as the prediction of influenza-like outbreaks (Lee et al., 2021) or the classification of animals from the Serengeti National Park (Swanson et al., 2015). *LanguageARC* is a Citizen Science platform for developing language resources (Fiumara et al., 2020). It is however not open yet to the public to create projects and does not easily allow conducting a Citizen Science meta-study as we do in this work. One work using LanguageARC is by Fort et al. (2022) (LD) who collected resources to evaluate bias in language models. They did not investigate the impact of using different recruitment channels which we do. Other projects using LanguageARC are still running and it is too early to derive recommendations from.

Compared to crowdsourcing, Citizen Science participants are volunteers that do not work for monetary gain. Instead, they are often motivated intrinsically. For instance, they may have a personal interest on positively impacting the environment (West et al., 2021), or in altruism (Rotman et al., 2012). Asking for unpaid work also entails various issues like finding good ways of how to attract volunteers, and ethical considerations (Resnik et al., 2015; Rasmussen and Cooper, 2019) that need to be addressed (cf. §5). Intrinsic motivation also has the potential of resulting in higher-quality annotations compared to crowdsourcing. For instance, Lee et al. (2022) find in their evaluation study with citizen scientists that their participants may have been willing to take more time annotating for the sake of higher annotation accuracy. However, as their main goal was to conduct an evaluation study for their specific setup, this finding cannot be generalized to other Citizen Science scenarios. So far, only Tsueng et al. (2016) provide a direct comparison between crowdsourcing and Cit-

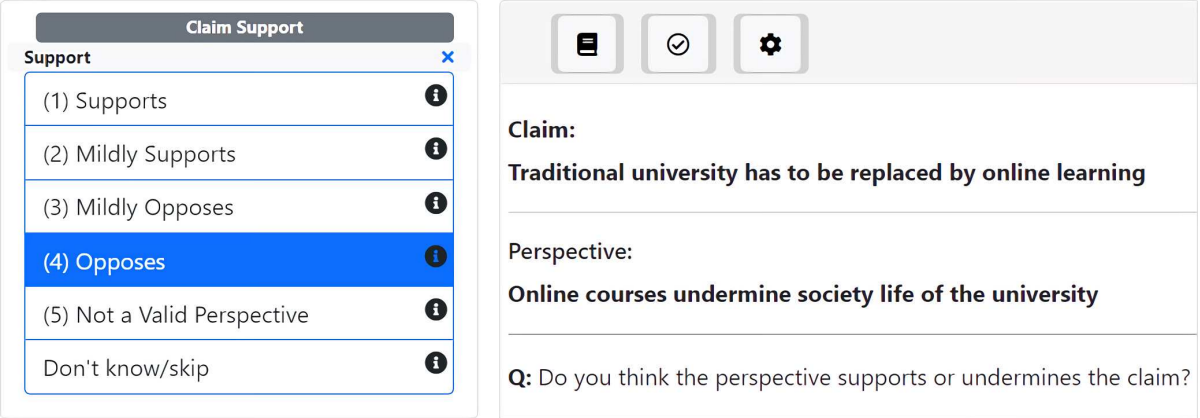


Figure 2: Assigning a label to an instance in the INCEpTION text annotation platform.

izen Science and show that volunteers can achieve similar performance in mining medical entities in scientific texts. They recruit participants through different channels such as newspapers, Twitter, etc., but do not compute channel-specific performance, making it difficult to assess whether the quality of the resulting annotation depends on the recurrent channel. In contrast, in the present work, we explicitly consider the recruitment channel in our evaluation and furthermore provide a discussion and guidelines for future Citizen Science practitioners. Also, it attracts intrinsically (not only fiscally) motivated volunteers that are often skilled in the task and can provide high-quality annotations, thus potentially combining the advantages of expert annotations and crowdsourcing. Relying on unpaid annotators entails several issues, including attracting volunteers and ethical considerations (Resnik et al., 2015; Rasmussen and Cooper, 2019) that need to be taken into account (see §5).

3 Study Design

To study the feasibility of Citizen Science for NLP annotation, we asked volunteers recruited via various channels to re-annotate an existing, crowdsourced dataset. The general setup is described in Fig. 1. To conduct a systematic study, we identified the following four necessary steps: 1) Identifying a suitable dataset (§3.1); 2) Selecting suitable recruitment channels to advertise our project on (§3.2); 3) Building a landing page for onboarding participants that asks for informed consent and the channel from which they originated (§3.3); 4) Setting up the annotation editor to which participants are forwarded after the onboarding (§3.4).

3.1 Dataset selection

We first conducted a literature review of relevant crowdsourced NLP datasets to identify the ones that could be accurately reproduced via Citizen Science. We assessed datasets for the following two criteria: 1) **Availability**: the dataset must be publicly available to make proper comparisons in terms of annotator agreement; 2) **Reproducibility**: the annotation setup including annotation guidelines needs to be reproducible to ensure similar conditions between citizen scientists and crowdworkers. We focused on datasets that are targeted towards contributing to social good to encourage volunteers to participate. Unfortunately, many inspected datasets did not fulfill both of these requirements. Overall, we identified two main issues while screening over 20 candidate datasets. First, many datasets used Tweets which impacted reproducibility as Twitter only allows researchers to publish the tweet identifiers. This leads to irrecoverable instances when tweets were deleted. Second was the lack of precise guidelines. For instance, many considered datasets about societal biases lack explicit descriptions of what is considered a stereotype. As such biases are often also impacted by the respective cultural background of annotators, they are difficult to reproduce without specific guidelines.

In the end, we decided on the stance detection task of the PERSPECTRUM dataset (Chen et al., 2019). The task provides clear instructions, publicly available data, and is motivated by social good (fighting misinformation/promoting accurate debate in internet discussions). Each instance consists of a claim–perspective pair (cf. Fig. 2) and annotators are asked if the claim *supports*, *opposes*, *mildly-supports*, *mildly-opposes*, or is *not a valid*

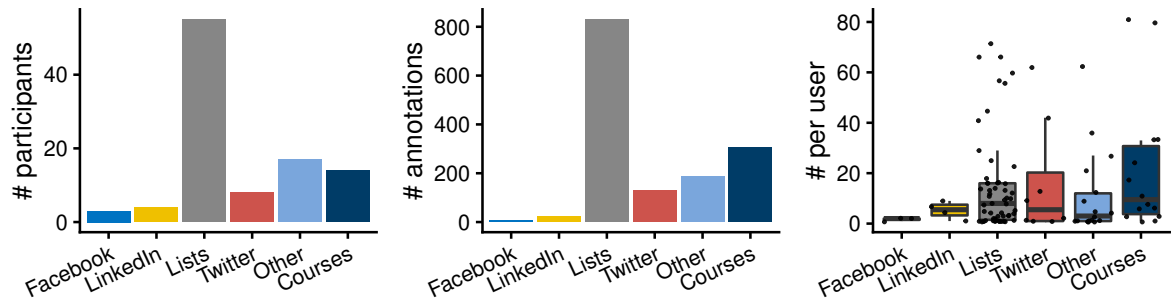


Figure 3: Participants, annotations and annotations grouped by the channel via they were recruited. It can be seen that overall, most participants and annotations were contributed by annotators recruited via mailing lists. Annotators from mailing lists and courses yielded the volunteers who contributed the most individually.

perspective. Following the original work, we also evaluated the annotations on a coarser tagset that only contains the categories for *support*, *oppose* and *not a valid perspective*. Overall, the dataset consists of 907 claims and 8,370 different perspectives which yield 11,805 annotated instances. In preliminary studies, we received further feedback that forcing annotators to provide an explicit label for each instance could lead to increasing frustration, especially for ambiguous or complicated instances. To lessen the burden for our voluntary annotators and keep them motivated in the annotation task, we allowed them to skip instances (*Don't know/skip*) which was not present in the original annotation editor for PERSPECTRUM.

3.2 Recruitment channels

To recruit annotators, we advertised our project on three social media platforms, namely, Twitter, LinkedIn and Facebook. Unfortunately, after creating the Facebook organization and advertising the project, the account was banned due to "violating their community standards" and has so far remained banned. One of our team members then promoted our annotation study on their personal Facebook to attract participation from this social media platform. In addition, the team members advertised the work on Twitter and in relevant LinkedIn groups such as COMPUTATIONAL LINGUISTICS and MACHINE LEARNING AND DATA SCIENCE.

We further promoted the study via two external mailing lists (i.e., CORPORA-LIST, ML-NEWS). Late in the project, we received interest from other faculty to advertise the task in their courses—an offer that we gladly accepted. For this, participation was completely voluntary and anonymous, students' grades were not affected by participation, and authors were not among the instructors. To evaluate different recruitment channels separately, we asked participants on the landing page to answer the question: "Where did you hear from this study?". We also allowed volunteers to not disclose how they found out about the study, this is referred to as "Other" or "Undisclosed" in this paper. Final participation counts are given in Fig. 3. We deliberately limited our outreach, e.g. we did not use university social media accounts or colleagues with large follower bases. Also, we made sure to not exhaust channels by posting too many calls for participation.

3.3 Landing page

We implemented a customizable landing page web application catering to the needs of Citizen Science projects. The link to such a landing page was shared via the respective recruitment channels. The landing page contained information about the study itself, its purpose, its organizers, which data we collected, and its intended use. This landing page toolbox is designed so that it can easily be adapted to future Citizen Science projects. To allow project creators to use an annotation editor of their choice, we designed the toolbox to act as an intermediary that collects a participant's consent for the actual annotation study. This ensures that only participants that have been properly informed and have explicitly provided their consent are given access to the study. For future Citizen Science projects, the tool further assists organizers through the landing page creation process to foster an ethical collection of data by asking several questions, that are listed in the appendix.

3.4 Annotation editor

INCEpTION (Klie et al., 2018) offers a configurable, web-based platform for annotating text documents at span, relation and document levels. To make it usable in Citizen Science scenarios, we extended the platform with three features, namely, (1) the ability to join a project through a link, (2) support for anonymous guest annotators, and (3) a dynamic workload manager. Allowing citizen scientists to participate in the project anonymously as guests without any sign-up process substantially reduced the entry barrier and made it easier for us to satisfy data protection policies. The same is true for the ability of joining a project through an invite link. Upon opening the link, annotators were greeted with the annotation guidelines and were directly able to start annotating. Finally, we implemented a dynamic workload manager that takes as input the desired number of annotators per document and then automatically forwards annotators directly to the document instances requiring annotation. Upon finishing annotating an instance, INCEpTION was configured to automatically load and display the next instance for annotation, similar to popular crowdsourcing platforms. We also included rules for handling other issues that may occur with voluntary annotations such as recovering instances that annotators have started to work on but then abandoned. Additionally, we modified the existing user interface to improve the annotation workflow. This mainly included implementing a dedicated labeling interface that allows users to select a single label for an instance via a radio button group. Annotation of an instance thus required two user actions: first, selecting the document label, and second, confirming the annotation, thereby moving on to the next document.

4 Results

We conducted our study between January and March 2022 and promoted the task in successive rounds across all recruitment channels. In total, we were able to recruit 98 participants who provided 1481 annotations resulting in 906 fully annotated instances. Each instance with at least one annotation has received on average 1.63 annotations. Detailed statistics are provided in the appendix.

Participation To identify promising channels for future Citizen Science studies, we report the number of annotators per channel, the total number of annotations per channel and per user (cf. Fig. 3). Overall, we find that the most effective channel for public outreach are mailing lists (55 participants). Asking students in university courses to participate was the second most effective with 14 participants. Facebook, LinkedIn, and Twitter only yielded three, four, and eight participants respectively. We further find a highly skewed distribution of annotations per user, as many annotators only provide a few annotations while a few annotators provide many annotations. For instance, the most active annotators were two students who provided ∼80 annotations as well as six participants from mailing lists who provided ∼60–80 annotations each. For Twitter and "undisclosed", only a single annotator made over 60 annotations. We also find that on average, participants from university courses provided the most annotations per person. When looking at participation over time (see Fig. 5), we observe increased activity in annotations made after the call for participation has been posted to the respective channel. For many channels, the count quickly flattens. Interestingly, Twitter sees a second spike long after the post was made. We attribute it to people sharing the post in our community quite a while after the initial release. We did not track whether individual volunteers came back for another round of annotations after their initial participation.

Coverage Overall, our 98 volunteers have provided 1,481 annotations to 906 unique instances (approximately 8% of the original dataset) over two months. This is comparable to other Citizen Science projects like Fort et al. (2022), which had 102 participants in total. They annotated three tasks and collected 2347, 2904 and 220 submissions over eight months. Table 1 shows the resulting coverage of our Citizen Science annotation study. While this still leaves room for improvement, the number of annotations collected nonetheless shows that Citizen Science can be viable in real life settings and is a promising direction to investigate in further studies, especially for creating focused and smaller-scale resources.

Quality In terms of annotation quality, we find that most channels yield annotations that highly agree with the gold labels (cf. Table 2), even though our annotations are not adjudicated yet. We further find that volunteers from university courses and mailing list show the highest accuracy, followed by Twitter and "undisclosed". Only LinkedIn yields

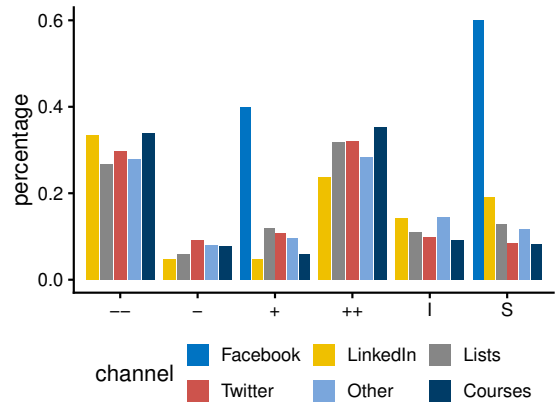


Figure 4: Label distribution grouped by channel. Labels are *supports* (++), *mildly-supports* (+), *mildly-opposes* (-), *opposes* (--), *not a valid perspective* (I) and *Skip* (S).

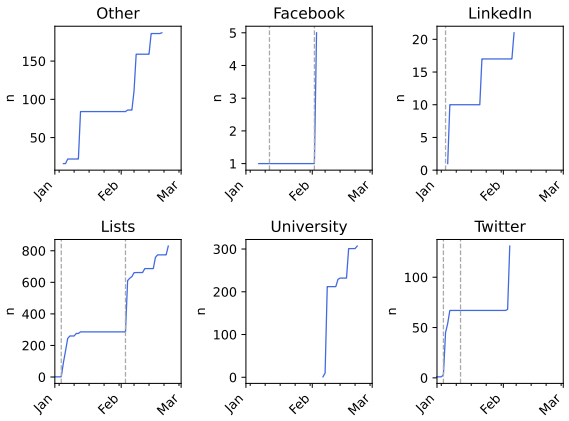


Figure 5: Annotations made over time. Vertical lines indicate when calls on the respective channels have been posted.

lower accuracy than 70% on the coarse label set.

For the majority of channels (with the exception of Facebook and LinkedIn), we only see a skip percentage of ∼10% (cf. Fig. 4). This indicates our volunteers are actually willing to spend time and effort to solve the task at hand, as adding a "Don't know/skip" option in crowdsourcing usually is an invitation for workers to speed through the tasks and not provide useful annotations. The exception is Facebook, where we find that a majority of the annotations from Facebook were labeled as *I don't know/skip* (3 out of 5). Further analysis of the label distribution grouped by channel (cf. Fig. 4) shows that all channels except for Facebook display a similar distribution in terms of annotated labels. This indicates that we can expect a rather stable annotation performance across citizen scientists recruited from different channels.

5 Discussion and Takeaways

Here we present lessons learned, discuss legal challenges and ethical considerations, as well as provide guidelines for future Citizen Science projects.

Table 1: Claims, claim clusters, and individual claim-perspective pairs that have been annotated at least once. We call the set of a claim and a perspective together with its paraphrases a claim cluster.

Name	# Annotated	# Total	% Annotated
Claims	388	907	42.78
Clusters	739	5092	14.51
Total	906	11805	7.67

Table 2: Annotation accuracy compared to the crowd-sourced and adjudicated data from PERSPECTRUM. The five annotations from Facebook (three of them were skipped) and *Don't know/skip* annotations are omitted.

Channel	Coarse	Fine
University	0.92	0.82
LinkedIn	0.69	0.62
Mailing Lists	0.90	0.82
Undisclosed	0.84	0.75
Twitter	0.85	0.73

Channel-dependent differences Our results clearly differ across recruiting channels. We find that overall, Facebook and LinkedIn have the lowest turnout and accuracy when compared to the gold labels, followed by Twitter. Our assumption for the overall low participation is that our network for these channels was not large enough. Advertising our study to NLP-related and university-internal mailing lists and university courses yielded the highest number of participants who also provided the most and best-quality annotations. Although our results show that students may outperform participants from other channels, we also acknowledge that this may not always be a viable option to recruit citizen scientists. Overall, our findings indicate that it is important to address the respective target groups that may be interested in a specific study. However, we also note that continuously advertising Citizen Science studies to the same channels may have a negative impact, as it can cause participation fatigue and lead to fewer volunteers participating. One possible solution could be the use of LanguageARC (Fiumara et al., 2020)

from the LDC and centralize calls for participation.

Motivating volunteers In contrast to crowdsourcing, there is no monetary or other extrinsic motivation that could be used to attract Citizen Science annotators. Thus, annotator motivation is a crucial question for Citizen Science studies. As Fig. 5 shows, citizen scientists can be quickly motivated to participate, but can also quickly lose interest in a given annotation study. This can become an issue with a low number of participants, yet our results also indicate that we were able to find highly-motivated participants (8 out of 98 in our results).

Compared to other groups, university students in particular provided a high amount of quality annotations. Considering the findings by Phillips et al. (2018), who do not find statistical differences in terms of quality between students participating for course credit vs. no extrinsic reward—asking students to participate in such projects as part of their coursework might be another good option, but needs to ensure an ethical data collection. For instance, such an approach has been used to annotate the Georgetown University Multilayer Corpus (Zeldes, 2017). Nonetheless, one remaining question is how to keep participants motivated and participate in several sessions as our results indicate that a vast majority of our volunteers only participated in a single session and that participation quickly stops shortly after a call has been posted to the respective channel.

Finally, we want to emphasize the inclusion of a *Don't know/skip* option for Citizen Science annotators. Whereas in crowdsourcing studies, annotators may exploit such an option to increase their gain (Hovy et al., 2013), from the feedback we got during our pilot study, it is crucial to keep volunteers motivated for Citizen Science. For this work, we did not provide a survey that asks about the motivation, as we thought that this might deter potential participants We however suggest that future studies provide such a survey that is as unintrusive as possible to further analyze why participants take part in the respective annotation project.

Legal challenges One substantial challenge in implementing Citizen Science studies is the potentially wide outreach they can have and, consequently, the varying kinds of data protection regulations they have to oblige. To preempt any potential issues that can arise—especially when data that can be used to identify a person (personal data, e.g. obtained during a survey or login credentials) is involved—we recommend researchers who plan to implement a Citizen Science study consider the most strict regulations that are widely accepted.

For the GDPR (European Parliament, 2016), currently one of the strictest data protection regulations, we recommend researchers to explicitly ask voluntary participants for their informed consent when collecting personal information. This includes informing participants beforehand about (1) the purpose of the data collection, (2) the kind of personal and non-personal data collected, (3) the planned use of the data, (4) any planned anonymization processes for publication, and finally, (5) how participants can request access, change, and deletion of the data. We further recommend assigning one specific contact person for any questions and requests for access, change, or deletion of the data. This may seem like additional work when compared to crowdsourcing, but transparent and open communication is one of the key factors to build trust—which is necessary for voluntary participants to consider such studies and provide high-quality annotations. Finally, participants should be informed and agree to the annotations donated being published under a permissive license.

Ethical and economical considerations Although Citizen Science can substantially reduce annotation costs, we emphasize the importance of considering an ethical deployment that does not compromise the trust of the participants. Moreover, given increasing concerns regarding the ownership and use of collected data (Arrieta-Ibarra et al., 2018), one should grant participants full rights to access, change, delete, and share their own personal data (Jones and Tonetti, 2020). This ensures that participants are not exploited for "free labor"—in contrast to approaches like reCAPTCHA (von Ahn et al., 2008), where humans are asked to solve a task in order to gain access to services. Whereas CAPTCHAs were initially intended to block malicious bots, they are becoming increasingly problematic due to their deployment and use by monopolizing companies which raises ethical concerns (Avanesi and Teurlings, 2022) . It is especially important to take the data itself into consideration; exposing volunteers to toxic, hateful, or otherwise sensitive speech should be avoided if they are not informed about it beforehand.

Recommendations Overall, we derive the following recommendations for future Citizen Science studies. 1) our call for annotations resonated the most with the target group that is likely to benefit the most from contributing to it: NLP researchers coming from mailing lists and university students. Therefore, the target audience should be carefully selected, for instance by identifying topic-specific mailing lists or respective university courses. This further means that the purpose of data collection should be made clear and that the results should be made publicly available. 2) the research question of the study should conform to the respective ethical and legal guidelines of the potential target group which should clearly be communicated to make the project accountable. 3) participation should be easy with clearly formulated annotation guidelines and, moreover, the annotation itself should be thoroughly tested beforehand to ensure that participants do not get frustrated due to design errors or choices. For instance, in our preliminary study, we got the feedback that some instances are frustrating to annotate and hence added an option to skip. 4) analyzing participation over time shows that a Citizen Science project has to be continuously advertised in order to stay relevant and achieve high participation. Otherwise, it will be forgotten quickly. This can be done by sharing status updates or creating preliminary results. Fifth, we recommend asking about user motivation before, during or after the annotation with a survey to better understand the participants and their demographics.

6 Conclusion

In this work, we presented an exploratory annotation study for utilizing Citizen Science for NLP annotation. We developed an onboarding process that can easily be adapted to similar projects and evaluated Citizen Science annotations for re-annotating an existing dataset. Furthermore, we extended the INCEpTION platform, a well-known open-source semantic annotation platform, with a dynamic workload manager and functionality for granting access to external users without registration. This enables its usage for Citizen Science projects. We advertised the study via Twitter, Facebook, LinkedIn, mailing lists, and university courses and found that participants from mailing lists and university courses are especially capable of providing high-quality annotations. We further discuss legal and ethical challenges that need to be addressed when conducting Citizen Science projects and provide general guidelines for conducting future projects that we would like to have known before starting. Overall, we conclude that Citizen Science can be a viable and affordable alternative to crowdsourcing, but is limited by successfully keeping annotators motivated. We will make our code and data publicly available to foster more research on Citizen Science for NLP and other disciplines.

Future Work We see the following directions for further research and evaluation to better understand in which settings Citizen Science can be applicable and how to use it best. Here, we used PERSPECTRUM as the dataset to annotate and mentioned in the participation calls that it benefits the social good. Therefore, it would be interesting to conduct more projects and see which datasets are suitable as well as whether volunteers participate, even if there is no extrinsic motivation. Then, it can also be tested how annotator retention develops, especially when project are running longer. The call for participation itself could also be investigated for the impact it has on turnout, motivation and quality.

7 Limitations

Throughout this article, we analyzed whether Citizen Science applies to linguistic annotation and showed that we can attract volunteers that donate a sizeable number of high-quality annotations. This work, however, comes with limitations that should be taken into account and tackled in future work. First, we based our analysis on a single annotation campaign and dataset that we advertised as being relevant for the social good. Therefore we suggest conducting more such annotation projects, also with different kinds of tasks. Second, we did not perform a user survey that for instance asked for user motivation. This is why we can only speculate about the motivation of our participants and suggest future works to explicitly prepare such a survey. Third, using Facebook as a channel might be viable, but we were not able to properly analyze it, as our account was blocked shortly after creation and never was reinstantiated. Finally, based on participation and annotation numbers, we see Citizen Science as more of an option for annotating smaller datasets, or longer-term projects that are more actively advertised than in our study which took place over two months and for which we deliberately limited the outreach.

Acknowledgments

We thank our anonymous reviewers, Michael Bugert and Max Glockner for their detailed and helpful comments to improve this manuscript. We are especially grateful for the discussions with Michael and Anne-Kathrin Bugert regarding our study setup. Finally, we thank the many volunteers that donated annotations for this project, as this work would not have been possible without their generous participation.

This work has been funded by the German Research Foundation (DFG) as part of the Evidence project (GU 798/27-1), UKP-SQuARE (GU 798/29-1), INCEpTION (GU 798/21-1) and PEER (GU 798/28-1), and within the project "The Third Wave of AI" funded by the Hessian Ministry of Higher Education, Research, Science and the Arts (HWMK). Further, it has been funded by the German Federal Ministry of Education and Research and HMWK within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

References

K Addy, L Green, E Herron, and K Stepenuck. 2010. Why volunteer water quality monitoring makes sense. usdanifa volunteer water quality monitoring national facilitation project, factsheet ii. *US Department of Agriculture, Washington, DC*.

Marianne Grace Araneta, Gülşen Eryiğit, Alexander König, Ji-Ung Lee, Ana Luís, Verena Lyding, Lionel Nicolas, Christos Rodosthenous, and Federico Sangati. 2020. Substituto – A Synchronous Educational Language Game for Simultaneous Teaching and Crowdsourcing. In *Proceedings of the 9th Workshop on NLP for Computer Assisted Language Learning*, pages 1–9, Gothenburg, Sweden.

Imanol Arrieta-Ibarra, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier, and E. Glen Weyl. 2018. Should We Treat Data as Labor? Moving Beyond "Free". *AEA Papers and Proceedings*, 108:38–42.

Vino Avanesi and Jan Teurlings. 2022. "I'm not a robot," or am I?: Micro-labor and the immanent subsumption of the social in the human computation of ReCAPTCHAs. *International Journal of Communication*, 16(0):1–19.

Jon Chamberlain, Massimo Poesio, and Udo Kruschwitz. 2008. Addressing the resource bottleneck to create large-scale annotated texts. In *Semantics in Text Processing. STEP 2008 Conference Proceedings*, pages 375–380.

Sihao Chen, Daniel Khashabi, Wenpeng Yin, Chris Callison-Burch, and Dan Roth. 2019. Seeing Things from a Different Angle: Discovering Diverse Perspectives about Claims. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 542–557, Minneapolis, Minnesota.

K. Bretonnel Cohen, Karën Fort, Gilles Adda, Sophia Zhou, and Dimeji Farri. 2016. Ethical Issues in Corpus Linguistics And Annotation: Pay Per Hit Does Not Affect Effective Hourly Rate For Linguistic Resource Development On Amazon Mechanical Turk. *Proceedings of the tenth international conference on language resources and evaluation (LREC 2016)*, 2016(W40):8–12.

Florian Daniel, Pavel Kucherbaev, Cinzia Cappiello, Boualem Benatallah, and Mohammad Allahbakhsh. 2018. Quality Control in Crowdsourcing: A Survey of Quality Attributes, Assessment Techniques, and Assurance Actions. *ACM Computing Surveys*, 51(1):1–40.

Kordula De Kuthy, Ramon Ziai, and Detmar Meurers. 2016. Focus annotation of task-based data: A comparison of expert and crowd-sourced annotation in a reading comprehension corpus. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 3928–3935, Portorož, Slovenia.

Leon Derczynski, Kalina Bontcheva, and Ian Roberts. 2016. Broad Twitter corpus: A diverse named entity recognition resource. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 1169–1179, Osaka, Japan.

Alexey Drutsa, Valentina Fedorova, Dmitry Ustalov, Olga Megorskaya, Evfrosiniya Zerminova, and Daria Baidakova. 2020. Crowdsourcing Practice for Efficient Data Labeling: Aggregation, Incremental Relabeling, and Pricing. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, pages 2623–2627, Portland, Oregon, USA.

GülŞen Eryiğit, Ali Şentaş, and Johanna Monti. 2022. Gamified crowdsourcing for idiom corpora construction. *Natural Language Engineering*, 1(1):1–33.

European Parliament. 2016. Consolidated text: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance).

James Fiumara, Christopher Cieri, Jonathan Wright, and Mark Liberman. 2020. LanguageARC: Developing Language Resources Through Citizen Linguistics. In *Proceedings of the LREC 2020 Workshop on "Citizen Linguistics in Language Resource Development"*, pages 1–6, Marseille, France.

Karën Fort, Gilles Adda, and K. Bretonnel Cohen. 2011. Amazon Mechanical Turk: Gold Mine or Coal Mine? *Computational Linguistics*, 37(2):413–420.

Karën Fort, Bruno Guillaume, and Hadrien Chastant. 2014. Creating *Zombilingo* , a game with a purpose for dependency syntax annotation. In *Proceedings of the First International Workshop on Gamification for Information Retrieval*, pages 2–6, Amsterdam, The Netherlands.

Karën Fort, Aurélie Névéol, Yoann Dupont, and Julien Bezançon. 2022. Use of a citizen science platform for the creation of a language resource to study bias in language models for French: A case study. In *Proceedings of the 2nd Workshop on Novel Incentives in Data Collection from People: Models, Implementations, Challenges and Results within LREC 2022*, pages 8–13, Marseille, France.

W. N. Francis and H. Kucera. 1979. Brown corpus manual. Technical report, Department of Linguistics, Brown University, Providence, Rhode Island, US.

Yannick Frommherz and Alessandra Zarcone. 2021. Crowdsourcing Ecologically-Valid Dialogue Data for German. *Frontiers in Computer Science*, 3:1–21.

Ivan Habernal and Iryna Gurevych. 2017. Argumentation Mining in User-Generated Web Discourse. *Computational Linguistics*, 43(1):125–179.

Mordechai Haklay, Daniel Dörler, Florian Heigl, Marina Manzoni, Susanne Hecker, and Katrin Vohland. 2021. What Is Citizen Science? The Challenges of Definition. In Katrin Vohland, Anne Land-Zandstra, Luigi Ceccaroni, Rob Lemmens, Josep Perelló, Marisa Ponti, Roeland Samson, and Katherin Wagenknecht, editors, *The Science of Citizen Science*, pages 13–33. Springer International Publishing, Cham.

Eric Hand. 2010. People power: Networks of human minds are taking citizen science to a new level. *Nature*, 466(7307):685—687.

Elizabeth T. Hobbs, Stephen M. Goralski, Ashley Mitchell, Andrew Simpson, Dorjan Leka, Emmanuel Kotey, Matt Sekira, James B. Munro, Suvarna Nadendla, Rebecca Jackson, Aitor Gonzalez-Aguirre, Martin Krallinger, Michelle Giglio, and Ivan Erill. 2021. ECO-CollecTF: A Corpus of Annotated Evidence-Based Assertions in Biomedical Manuscripts. *Frontiers in Research Metrics and Analytics*, 6:1–13.

Anthony M Holmes. 1991. *The Ontario Butterfly Atlas*. Entomologists' Association, Toronto.

Dirk Hovy, Taylor Berg-Kirkpatrick, Ashish Vaswani, and Eduard Hovy. 2013. Learning whom to trust with MACE. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1120–1130, Atlanta, Georgia.

Dirk Hovy, Barbara Plank, and Anders Søgaard. 2014. Experiments with crowdsourced re-annotation of a POS tagging data set. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 377–382, Baltimore, Maryland.

Charles I. Jones and Christopher Tonetti. 2020. Nonrivalry and the Economics of Data. *American Economic Review*, 110(9):2819–2858.

Jan-Christoph Klie, Michael Bugert, Beto Boullosa, Richard Eckart de Castilho, and Iryna Gurevych. 2018. The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation. In *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, pages 5–9.

Margaret Kosmala, Andrea Wiggins, Alexandra Swanson, and Brooke Simmons. 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10):551–560.

Ji-Ung Lee, Jan-Christoph Klie, and Iryna Gurevych. 2022. Annotation Curricula to Implicitly Train Non-Expert Annotators. *Computational Linguistics*, pages 1–22.

Liza Lee, Mireille Desroches, Shamir Mukhi, and Christina Bancej. 2021. FluWatchers: Evaluation of a crowdsourced influenza-like illness surveillance application for Canadian influenza seasons 2015–2016 to 2018–2019. *Canada Communicable Disease Report*, 47(09):357–363.

Ronald D. Leeper, Jared Rennie, and Michael A. Palecki. 2015. Observational Perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: Temperature and Precipitation Comparison. *Journal of Atmospheric and Oceanic Technology*, 32(4):703–721.

Chris J. Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M. Jordan Raddick, Robert C. Nichol, Alex Szalay, Dan Andreescu, Phil Murray, and Jan Vandenberg. 2008. Galaxy Zoo : Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 389(3):1179–1189.

Verena Lyding, Lionel Nicolas, and Alexander König. 2022. About the applicability of combining implicit crowdsourcing and language learning for the collection of NLP datasets. In *Proceedings of the 2nd Workshop on Novel Incentives in Data Collection from People: Models, Implementations, Challenges and Results within LREC 2022*, pages 46–57, Marseille, France.

Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330.

National Audubon Society. 2020. The Christmas bird count historical results.

An Nguyen, Byron Wallace, and Matthew Lease. 2015. Combining Crowd and Expert Labels Using Decision Theoretic Active Learning. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 3(1):120–129.

Ellie Pavlick, Matt Post, Ann Irvine, Dmitry Kachaev, and Chris Callison-Burch. 2014. The Language Demographics of Amazon Mechanical Turk. *Transactions of the Association for Computational Linguistics*, 2:79–92.

Christopher Phillips, Dylan Walshe, Karen O'Regan, Ken Strong, Christopher Hennon, Ken Knapp, Conor Murphy, and Peter Thorne. 2018. Assessing Citizen Science Participation Skill for Altruism or University Course Credit: A Case Study Analysis Using Cyclone Center. *Citizen Science: Theory and Practice*, 3(1):6.

Massimo Poesio, Jon Chamberlain, Udo Kruschwitz, Livio Robaldo, and Luca Ducceschi. 2013. Phrase detectives: Utilizing collective intelligence for internet-scale language resource creation. *ACM Transactions on Interactive Intelligent Systems*, 3(1):1–44.

Lisa M. Rasmussen and Caren Cooper. 2019. Citizen Science Ethics. *Citizen Science: Theory and Practice*, 4(1):5.

David B. Resnik, Kevin C. Elliott, and Aubrey K. Miller. 2015. A framework for addressing ethical issues in citizen science. *Environmental Science & Policy*, 54:475–481.

Dana Rotman, Jenny Preece, Jen Hammock, Kezee Procita, Derek Hansen, Cynthia Parr, Darcy Lewis, and David Jacobs. 2012. Dynamic changes in motivation in collaborative citizen-science projects. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work - CSCW '12*, page 217, Seattle, Washington, USA.

Boaz Shmueli, Jan Fell, Soumya Ray, and Lun-Wei Ku. 2021. Beyond Fair Pay: Ethical Implications of NLP Crowdsourcing. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3758–3769, Online.

Rion Snow, Brendan O'Connor, Daniel Jurafsky, and Andrew Ng. 2008. Cheap and Fast – But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 254–263, Honolulu, Hawaii.

Alexander Sorokin and David Forsyth. 2008. Utility data annotation with Amazon Mechanical Turk. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–8, Anchorage, AK, USA.

Alexandra Swanson, Margaret Kosmala, Chris Lintott, Robert Simpson, Arfon Smith, and Craig Packer. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data*, 2(1):1–14.

Christopher Tauchmann, Johannes Daxenberger, and Margot Mieskes. 2020. The Influence of Input Data Complexity on Crowdsourcing Quality. In *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion*, pages 71–72, Cagliari, Italy.

Ginger Tsueng, Steven M. Nanis, Jennifer Fouquier, Benjamin M. Good, and Andrew I. Su. 2016. Citizen Science for Mining the Biomedical Literature. *Citizen Science: Theory and Practice*, 1(2):14.

Noortje J. Venhuizen, Valerio Basile, Kilian Evang, and Johan Bos. 2013. Gamification for word sense labeling. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Short Papers*, pages 397–403, Potsdam, Germany.

L. von Ahn. 2006. Games with a Purpose. *Computer*, 39(6):92–94.

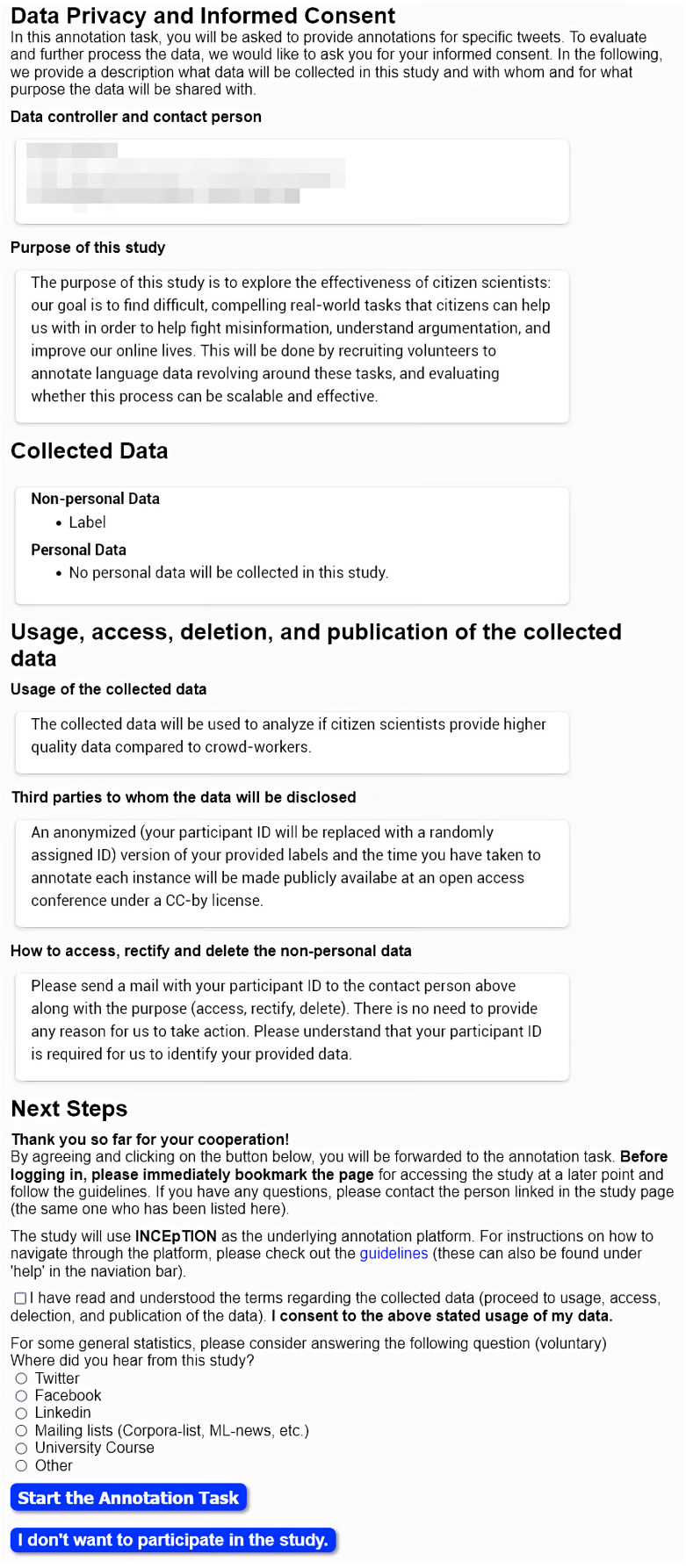
Luis von Ahn, Benjamin Maurer, Colin McMillen, David Abraham, and Manuel Blum. 2008. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. *Science*, 321(5895):1465–1468.

Aobo Wang, Cong Duy Vu Hoang, and Min-Yen Kan. 2013. Perspectives on crowdsourcing annotations for natural language processing. *Language Resources and Evaluation*, 47(1):9–31.

Sarah West, Alison Dyke, and Rachel Pateman. 2021. Variations in the Motivations of Environmental Citizen Scientists. *Citizen Science: Theory and Practice*, 6(1):14.

Amir Zeldes. 2017. The GUM corpus: Creating multilayer resources in the classroom. *Language Resources and Evaluation*, 51(3):581–612.

Appendix A Landing Page



Data Privacy and Informed Consent

In this annotation task, you will be asked to provide annotations for specific tweets. To evaluate and further process the data, we would like to ask you for your informed consent. In the following, we provide a description what data will be collected in this study and with whom and for what purpose the data will be shared with.

Data controller and contact person

Purpose of this study

The purpose of this study is to explore the effectiveness of citizen scientists: our goal is to find difficult, compelling real-world tasks that citizens can help us with in order to help fight misinformation, understand argumentation, and improve our online lives. This will be done by recruiting volunteers to annotate language data revolving around these tasks, and evaluating whether this process can be scalable and effective.

Collected Data

Non-personal Data

- Label

Personal Data

- No personal data will be collected in this study.

Usage, access, deletion, and publication of the collected data

Usage of the collected data

The collected data will be used to analyze if citizen scientists provide higher quality data compared to crowd-workers.

Third parties to whom the data will be disclosed

An anonymized (your participant ID will be replaced with a randomly assigned ID) version of your provided labels and the time you have taken to annotate each instance will be made publicly availabe at an open access conference under a CC-by license.

How to access, rectify and delete the non-personal data

Please send a mail with your participant ID to the contact person above along with the purpose (access, rectify, delete). There is no need to provide any reason for us to take action. Please understand that your participant ID is required for us to identify your provided data.

Next Steps

Thank you so far for your cooperation!
By agreeing and clicking on the button below, you will be forwarded to the annotation task. **Before logging in, please immediately bookmark the page** for accessing the study at a later point and follow the guidelines. If you have any questions, please contact the person linked in the study page (the same one who has been listed here).

The study will use **INCEpTION** as the underlying annotation platform. For instructions on how to navigate through the platform, please check out the guidelines (these can also be found under 'help' in the naviation bar).

☐ I have read and understood the terms regarding the collected data (proceed to usage, access, delection, and publication of the data). **I consent to the above stated usage of my data.**

For some general statistics, please consider answering the following question (voluntary)
Where did you hear from this study?

- ○ Twitter
- ○ Facebook
- ○ Linkedin
- ○ Mailing lists (Corpora-list, ML-news, etc.)
- ○ University Course
- ○ Other

Start the Annotation Task

I don't want to participate in the study.

Appendix B Annotation Guidelines

Perspectrum Citizen Science Annotation

For help with INCEpTION, please see our Quick Tutorial

Welcome to our citizen science annotation project! Here are some useful links:

- For information on how we use your data, Read This. In short, we don't collect any identifying personal data.
- If you have feedback, please leave us some comments via this google form or via email to
- If you are ready to annotate, click the "Annotation" button in the upper left!

Annotation Task

For each of the following tasks check if the perspective provides a view about the given claim. Feel free to annotate as much or as little as you like: there's no limit, and each additional claim to annotate will help us build better models for identifying fake news, building knowledge, and helping fight misinformation on the internet!

Note that in this task we are NOT asking for your personal opinions; instead our aim is to discover perspectives that could possibly be convincing for those with different world view. If you don't understand the claim or perspective, or otherwise find the text un-interpretable, choose the "Don't know/skip" option.

Below are some examples, with the correct response in **bold**:

Claim: The West should invade Syria

Perspective: Sovereign countries should never be invaded.

Q: Do you think the perspective supports or opposes the claim?

- Supports
- Leaning Supports
- Leaning Opposes
- **Opposes**
- Not a Valid Perspective

Claim: The West should invade Syria

Perspective: If the United States does not intervene, the moral responsibility of those dying will be on us.

Q: Do you think the perspective supports or opposes the claim?

- **Supports**
- Leaning Supports
- Leaning Opposes
- Opposes
- Not a Valid Perspective

Claim: The West should invade Syria

Perspective: The Syrian currency has significantly lost its value compared to the Western money, since the end of the World War II.

Q: Do you think the perspective supports or opposes the claim?

- Supports
- Leaning Supports
- Leaning Opposes
- Opposes
- **Not a Valid Perspective**

If you are ready to annotate, click the "Annotation" button in the upper left!

Appendix C Questions to keep in mind for a citizen science project

- What is the purpose of the study?
- What kind of personal and non-personal data will be collected?[2]
- If there is a questionnaire involved, what questions will it involve?
- How will the data be used?
- Is a publication of the data planned and if so, which data will be published and will it be anonymized?
- How can participants request access, change, or deletion of their data?

Appendix D Project Statistics

D.1 Number of participants

In addition to the plots visualizing the number of participants (c.f. Fig. 3), we also list the raw numbers in Table 3.

Channel	Participants
Courses	14
Facebook	3
LinkedIn	4
Lists	55
Twitter	8
Undisclosed	17

Table 3: Number of participants per channel.

D.2 Annotation statistics

In addition to the plots visualizing the annotation counts and label distribution (c.f. Fig. 4), we also list the raw numbers in Table 4.

Table 4: Label distribution grouped by channel. Labels are *supports* (++), *mildly-supports* (+), *mildly-opposes* (-), *opposes* (--), *not a valid perspective* (I) and *Skip* (S).

Channel	Total	Counts						Percentage					
		+	++	-	--	I	S	+	++	-	--	I	S
Courses	307	18	108	24	104	28	25	5.86	35.18	7.82	33.88	9.12	8.14
Facebook	5	2	0	0	0	0	3	40.00	0.00	0.00	0.00	0.00	60.00
LinkedIn	21	1	5	1	7	3	4	4.76	23.81	4.76	33.33	14.29	19.05
Lists	830	98	264	48	222	92	106	11.81	31.81	5.78	26.75	11.08	12.77
Twitter	131	14	42	12	39	13	11	10.69	32.06	9.16	29.77	9.92	8.40
Undisclosed	187	18	53	15	52	27	22	9.63	28.34	8.02	27.81	14.44	11.76

[2]We provided some pre-defined suggestions such as *Name* or *IP* for personal data and *Label* for non-personal data with the possibility to add more in our landingpage module.