Deep Sea Spy: an online citizen science annotation platform for science and ocean literacy

Marjolaine Matabos[1], Pierre Cottais[1], Riwan Leroux[1], Yannick Cenatiempo[2], Charlotte Gasne-Destaville[1], Nicolas Roullet[3], Jozée Sarrazin[1], Julie Tourolle[1], Catherine Borremans[1]

[1] Univ Brest, Ifremer, BEEP, F-29280 Plouzané, France

[2] Ikadoc, 1 Grand rue, 31460 Loubens-Lauragais France

3 7 Rue Jean Giono – 31130 Balma France

Corresponding author: marjolaine.matabos@ifremer.fr

Abstract

The recent development of deep-sea observatories has enabled the acquisition of high temporal resolution imagery for the study of deep-sea communities' dynamics from hourly to multi-decadal scales. Camera systems deployed at hydrothermal vents have acquired, since 2010, over 11 Tera bytes of data that cannot be processed by research labs only. While deep learning offers an alternative to human processing, training algorithms requires substantial annotated reference datasets. The project Deep Sea Spy enable citizens to contribute to the annotation of pictures acquired with underwater platforms. The annotation of over 45 000 images supported the development of a multi-participant data validation workflow that can be applied to similar databases. We also present the impact of the platform on the civil society, and how it can serve education and inform managers and policy makers. Deep Sea Spy and the proposed workflow has a strong potential to enhance environmental observation and monitoring.

Key-words

Crowdsourcing, Deep-sea hydrothermal vents, Education, Ocean literacy, Image processing, EMSO-Azores, Ocean Networks Canada

1. Introduction

Increasing threats on the ocean calls for an urgent and comprehensive assessment of deep-sea ecosystems status (Franke et al., 2020; Roberts et al., 2023). Changes in the deep include the indirect effects of anthropogenic climate change (Levin and Le Bris, 2015; Sweetman et al., 2017), pollution (Courtene-Jones et al., 2017; Kane and Clare, 2019), as well as the direct effects of deep-sea fisheries (Morato et al., 2006), oil and gas extraction, and potentially deep-sea mining (Miller et al., 2018; Van Dover, 2014). Our ability to detect long-term trends and changes as a consequence of human activities requires a good knowledge of the natural dynamics of ecosystems and their associated environmental drivers. In the deep sea, hydrothermal vents still constitute a relatively 'pristine' environment but industries are increasingly interested in these metal-rich environments (Boschen et al., 2013). Predicting biological responses to deep-sea mining and informing mining regulations require a good understanding of community responses to changes in environmental conditions, the role of biotic interactions in structuring communities as well as species biology (Van Dover et al., 2020). Recently, the development of deep-sea observatories (Cannat et al., 2011; Favali and Beranzoli, 2006; Juniper et al., 2007; Matabos et al., 2022) and associated instrumentation (Porter et al., 2009) provides unprecedented means to investigate and characterize ecosystems at increasing temporal resolution (Matabos et al., 2016). This is particularly true in heterogeneous and remote environments, where the poor accessibility and limited amount of ship time on-site impedes the fine characterization of the environment and their associated faunal communities. Deep-sea observatories provide power and communication to instruments deployed on the seafloor allowing for long-term time-series of multidisciplinary data (e.g., geological, physical, chemical, ecological) with resolution from seconds to decades. More specifically, monitoring of faunal communities is now enabled by the use of optical imagery deployed on these deep-sea platforms (e.g., Aguzzi et al., 2015; Lelièvre et al., 2017; Robert and Juniper, 2012; Van Audenhaege et al., 2022)

The technology to acquire and process underwater marine imagery has significantly evolved in recent years, including high-definition cameras, illumination systems and analytical software. As a result, *in situ* imaging technology are increasingly used in marine science (review in Durden et al., 2016b) to quantify species abundance and distribution in the water column (Biard and Ohman, 2020) and on the seafloor (Devine et al., 2020), study species biology (Matabos et al., 2015; Zweifler et al., 2017) and map benthic communities and habitats (e.g., Girard et al., 2020; Macedo et al., 2022; Marcon et al., 2014; Van Audenhaege et al., 2021). Image analysis is non-invasive and allows the monitoring of animals in their natural environment over long periods of time. But these advances have led to new challenges including storage, management and annotation relating to 'big data' (Schoening et al., 2018). Particularly, multidisciplinary seafloor observatories generate data that accumulate faster than the processing power of research laboratories. Manual processing of these data is time-consuming, highly labour-intensive, and beyond the human capacity currently available. Their effective exploitation requires more human resources and additional computational solutions.

In this context, the TEMPO(-mini) ecological module equipped with deep-sea lights and a camera, called SMOOVE, was developed in order to monitor vent communities' dynamics from hourly to multi-decadal scales (Auffret et al., 2009; Sarrazin et al., 2007). Two versions of the module were deployed and connected to deep-sea observatories: one (TEMPO-mini) at 2,200 m depth at Main Endeavour Field (MEF) on the Juan de Fuca Ridge (JdFR), connected to the Endeavour node of Ocean Networks Canada observatory (Barnes et al., 2008); the other (TEMPO) at 1,700 m depth at the Lucky Strike vent field (LS) on the Mid-Atlantic Ridge (MAR), connected to the autonomous EMSO-Azores observatory (Cannat et al., 2016). To date, analyses of sub-samples of the images acquired with the modules brought new insights on local community dynamics such as the role of tides and inertial currents on species behaviour (Cuvelier et al., 2017; 2014; Lelièvre et al., 2017) or the role of local variations in hydrothermal venting and the high stability of mussel habitats at the Lucky Strike vent field along the slow-spreading MAR (Sarrazin et al., 2014; Van Audenhaege et al., 2022), significantly advancing our knowledge on vent ecology. However, since their first deployment in 2010, the archive now contains over 7,000 hours of video sequences, representing over 11 Tera bytes (Tb) of imagery data, and is still growing.

While artificial intelligence, more particularly machine and deep-learning, can help the annotation of these large databases, a large set of reference data is required to train and validate algorithms, especially if one wants to classify rare morphotypes or species (Durden et al., 2021). The lack of annotated datasets for marine environments has hampered the use of these approaches, although they are increasingly used for underwater imagery (Han et al., 2020; McEver et al., 2023; Piechaud et al., 2019; Villon et al., 2018). Since the first development of the Zooniverse platform by the NASA to engage citizens in the annotation of galaxies (Lintott et al., 2008), a number of web-based tools have been developed to involve society in image annotations and contribute to building reference databases (Anton et al., 2021; Grol et al., 2020; Matabos et al., 2017; Ra et al., 2022). In this paper, we present the image annotation platform

Deep Sea Spy (DSS) that was developed to help the annotation of video images acquired by the SMOOVE camera of the TEMPO and TEMPO-mini ecological modules deployed at deep-sea hydrothermal vents. The main objective of the project was to build a web-based application for manual imagery processing that will help gather useful information for scientists while also raising awareness among the general public about these remote ecosystems and the threats they face (Boschen et al., 2013). By involving citizens in the scientific process of imagery annotation, we mostly tackled two important aspects: i) raising awareness about scientific research, environmental issues and the deep ocean, and ii) offering new ways for data collection and processing to handle the bottleneck from big data generated by Research Infrastructures (RIs). Overall, we contributed to enhancing interactions between society and RIs. This paper presents a preliminary analysis of citizens' data in the framework of the EMSO-Azores and Ocean Networks Canada observatories that can be used as a reference guideline for future development within other RIs aiming to help with complex data processing through public engagement. More specifically, this paper aims to assess the thresholds defined to optimize data quality with processing efficiency, evaluate citizens performance in annotating complex deep-sea hydrothermal images and establish a data validation workflow.

2. The Deep Sea Spy platform

Together, both SMOOVE cameras acquired 128 minutes of video a day of a siboglinid tubeworm *Ridgeia piscesae* assemblage at MEF (JdFR), and a bed of the bathymodiolin mussel *Bathymodiolus azoricus* at the LS vent field (MAR; Figure 1). While the frequency of images to be annotated depends on the scientific question to be addressed in relation with the phenomena of interest (e.g., tidal signal, seasonal variations, inter-annual variability), we can assume that the analyses of a picture extracted every 10 seconds will cover the full temporal range of variability in community dynamics. Automated detection was one of the first path explored to help in annotating such large datasets, but the first approaches showed that the human eye still performed better than the machine for extracting data from complex imagery (Aguzzi et al., 2009; Aron et al., 2010; Matabos et al., 2017; Purser et al., 2009; Schoening et al., 2012). More recently, deep-learning approaches offer new solutions for automatic classification but in the absence of a large training dataset, these could not yet be applied to our images (Durden et al., 2021).

The astronomical community was the first to use citizen science into data processing methods by asking volunteers to classify galaxies from space imagery (https://www.zooniverse.org/projects, Galaxy Zoo; Fortson et al., 2012; Lintott et al., 2008). Since, the Zooniverse platform has been hosting a growing number of projects in various disciplines and was a great success, leading to large number of scientific publications. They also managed to engage a significant part of the population, highlighting the desire of the public to contribute to tangible scientific work (Fortson et al., 2012; Raddick et al., 2010). Since then, crowdsourcing, where a large number of citizens contribute to research projects through online classification/processing of data with few prerequisite knowledge, became a recognized and popular form of citizen science (Silvertown, 2009). Considering our yearly growing imagery database, we calculated that it would require ~10,000 participants annotating 10 images a day over a month to process a full year of imagery dataset (each image being annotated 10 times), not to mention the number that would be required to process more than 10 years of imagery.

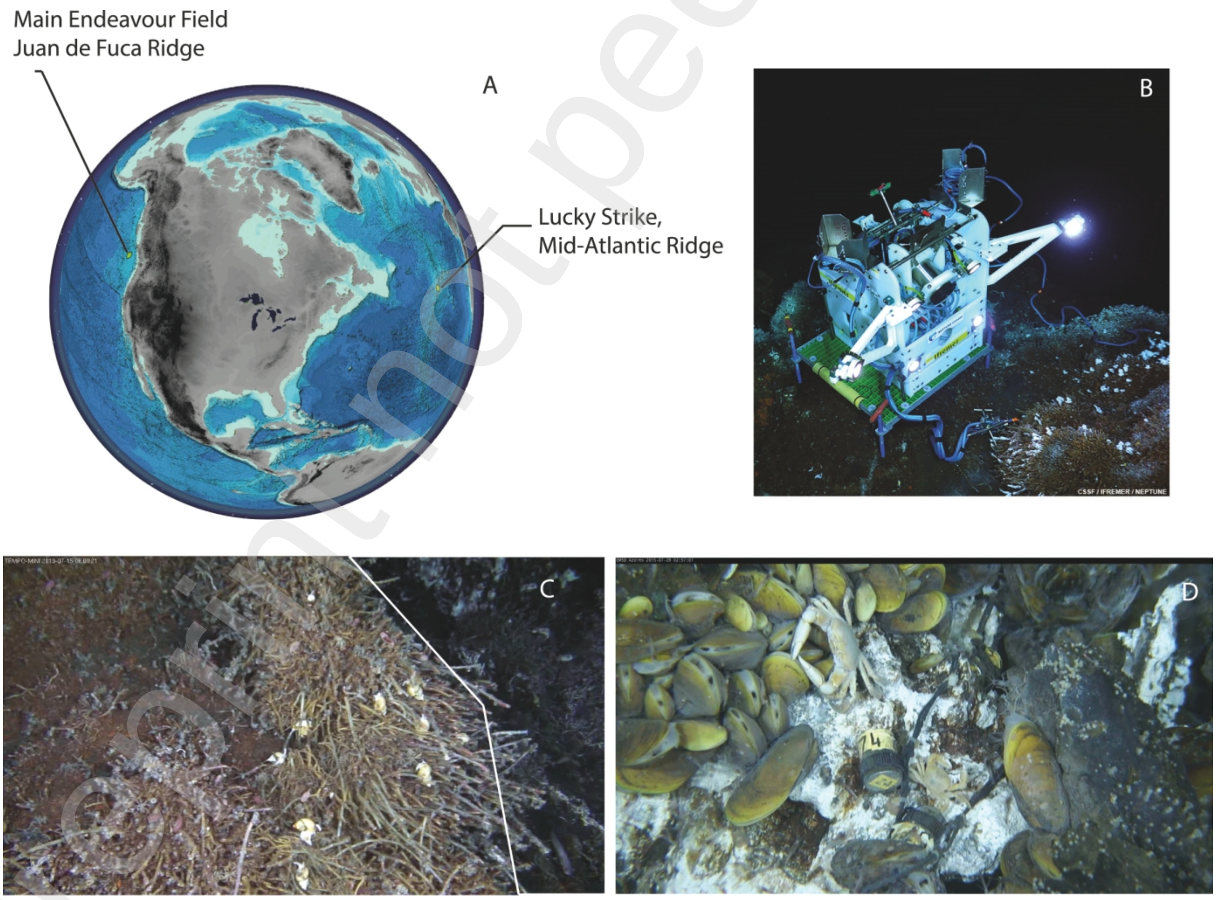

Figure 1. Data acquisition and images. A. Location of the two ecological observatory modules on the mid-Atlantic ridge (MAR) and Juan de Fuca ridge (JdFR). B. TEMPO-mini, equipped

with the SMOOVE camera, monitors siboglinid tubeworms at the Main Endeavour Field of the JdFR. C & D. Field of view filmed by SMOOVE cameras on the JdFR (C) and MAR (D). The white line on C delineates the "background" area (see text).

2.1. User interface and design

The online annotation platform provides a user interface (UI) to allow for the annotation of images. The interface was built with the objective to be aesthetically attractive while providing the minimum information required for a good completion of the tasks. While some studies showed that even for complicated tasks, citizens can perform as well as experts (Butt et al., 2013; Delaney et al., 2008), we tried to keep the annotation task as simple as possible to allow the participation of a wide range of people and ensure the robustness of the acquired data. New registered participants can follow a simple tutorial so each of them can benefit from a minimum of training to accomplish the requested tasks. The tutorial explains how to use the annotation system and provides an overview of all species to be annotated and/or measured (Figure 2). Several functionalities facilitate the task including the display on the screen of a thumbnail that illustrates how to annotate a given species, the ability to zoom on the image and the possibility to request help from another participant. The participant can access the tutorial at any time during a session.

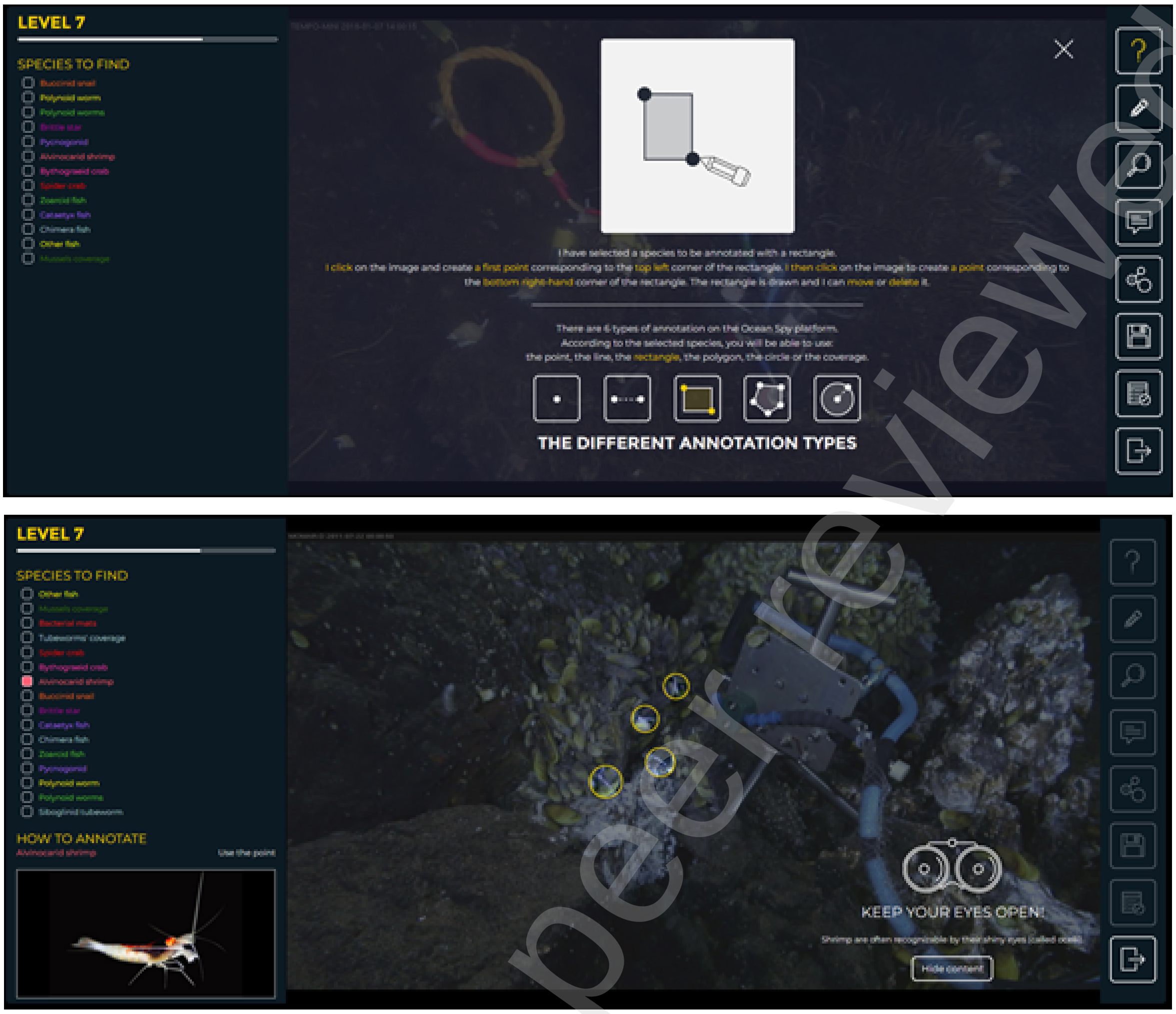

Figure 2. Tutorials providing training for the annotation system (top panel) and species recognition (bottom panel) in the Deep Sea Spy online image annotation tool.

While, to our knowledge, most citizen annotation platforms propose quite diverse and changing images involving spatial sampling (Lintott et al., 2008; Robinson et al., 2017; Van den Bergh et al., 2021), the SMOOVE cameras have been recording the same area of ~1 m² since 2011. The processing of these complex images (Figure 1) can be quickly laborious and repetitive. The full annotation of an image, thereafter referred to as "annotation session", requires few seconds to 20 minutes to an expert depending on the number of animals to be annotated. The question of citizen engagement over time was a real issue during the development process. Designing the UI and application in a gamified way was a solution chosen to enrol a large number of people, and boost the motivation of participants not interested in science in the first place (Apostolopoulos and Potsiou, 2022; Wang et al., 2022). The UI was built to be user-friendly and intuitive. For this, attention was given to different aspects including the graphic design and game mechanisms including specific missions, leader boards, levels and rewards.

Elements of gamification – Several elements were included in the application to maintain citizen engagement (Wang et al., 2022). Because video image acquisition is still ongoing, the dataset to process is infinite which can be discouraging for contributors. A system of missions was developed allowing to inform on their progress, advancement and completeness. To each mission is assigned a goal, a dataset and a set of species to annotate. This allows the display of a progression bar. Other elements such as leader boards (i.e., my progression) per mission and across all missions as well as the level and rewards of the participant are displayed. Level and rewards are specific to each participant and depend on the number of images annotated. Participants have to annotate a given number of images to reach the next level where they receive a virtual reward, being a 3D reconstruction of one of the species, and the possibility to annotate a new species. Species to annotate are of increasing difficulty as the participant reaches higher levels. This level mechanic was designed to ensure that the participant is properly trained and skilled to contribute to more complex annotations.

An administration page, secured by a Central Authentication Service (CAS) identification protocol, allows the configuration of all these custom options, providing some flexibility to adapt the application to other scientific projects. In addition, from the administration page, the administrators get an overview of the main statistics including the number of images annotated, number of users per week/month/over the mission and globally. While the web application and the project website exist both in French and English, the admin page is currently only available in French but can be easily adapted in other languages. This flexibility allowed the development of additional applications, inspired by our project, that officially started in 2023 (e.g., Deep Reef Spy and Shore Spy available on the platform Ocean Spy; https://ocean-spy.ifremer.fr/).

2.2. The database

Annotations and images' metadata are stored in an independent PostgreSQL database associated with the Deep Sea Spy application. The Data Model is available in supp mat 1 & 2. The definition of the data model required that i) all annotations are associated with an image; ii) all annotations are associated with an observer; and iii) all annotations are stored in pixels. Images' information stored in the DSS database allows to track back associated metadata stored in the Ifremer Oracle database (Table 1) and to import data in this central information system.

Table 1. Information included in the Oracle database of the Deep Sea Spy project

Image	Participant	Annotation
Observatory (EMSO-Azores, Ocean Networks Canada) Latitude (degrees decimal) Longitude (degrees decimal) Depth (meters) Camera type/model Zoom value Date of acquisition (from the extraction) Time of acquisition (from the extraction) Still image ID (unique number) 'Annotation Mission' name/ID	ID (unique number) Date of registration Personal information • Pseudo • Email address • Gender* • Age* • Job* Ranking	Date of observation Corresponding participant Corresponding image Date of observation Time of observation Unit of measure (pixel) Taxon (animal) name Position of each animal Measurement of each animal Area types' polygons pixels

* Non mandatory

3. General statistics

In this paper, we focus on data related to the first annotation mission, entitled "Tides at 1,700 m depth?", that consisted in 6 months of video data at both locations (i.e., JdFR and MAR). The objective was to assess the role of tidal variations on species behaviour to confirm previous observations (Cuvelier et al., 2014; Lelièvre et al., 2017; Mat et al., 2020). The contribution of participants allowed for the processing of a significantly higher number of images and inclusion of a higher number of species enabling a more comprehensive description of the faunal communities. The dataset included a picture every 4 hours at the JdFR and 1 picture every 6 hours at the MAR, resulting in 3,978 unique images (thereafter called photos) which had to be annotated 10 times before being removed from the list to increase the confidence in data quality. This threshold was arbitrarily chosen based on expert opinion, researchers in the lab having a long-lasting experience in image annotation and student supervising in the field. The following general statistics only relies on presence (or abundance) data. The mission lasted 3 years from March 2017 to May 2020 and stopped when each photo was annotated 10 times. Over the three years, 1,130 participants annotated 39,255 images from the 3,978 photos of the mission (i.e., 39,780 images), reaching a total of 313,300 annotations of organisms. The discrepancy between the number of annotated photos and their total number can be explained by data cleaning (i.e., removal of "draft" accounts used for tests and demonstrations). Only results for one species per location were considered in this paper (the snail *Buccinum thermophilum* and the crab *Segonzacia mesatlantica* in Pacific and Atlantic respectively) which correspond to accessible annotations for level 0 participants, in order to maximize the number of available photos.

3.1. Participation

The daily participation rate (i.e., the number of active participants per day) ranged from 0 to 64 and depicted three major peaks in March 2017, June 2018 and May 2020, as well as additional intermediate ones (Figure 3). Most of the peaks could be attributed to outreach or media communication events, highlighting their importance to recruit new participants and maintain enthusiasm and motivation among the community (De Vries et al., 2019; Golumbic et al., 2020; Rüfenacht et al., 2021). For instance, the launch in March 2017 with a press release was widely covered in the media (radio, newspapers and local TV). The highest participation occurred at the end of the mission and was fostered by a game contest organised during the COVID-19 lock-down in May 2020. The first three annotators (in terms of the number of images annotated) were rewarded, the first prize being a visit of the Ifremer research institute and free admission to one of three re-known aquaria in France. The incentive of prizes for the best annotator generated enthusiasm among the community. The third major peak in June 2018 resulted from a communication performed by a third party (national radio broadcast). The intermediate peaks mostly occurred upon the launch of the project and resulted from the various media coverage that followed the original press release. Minor increases in participation rate always followed general public conferences or science exhibition and events.

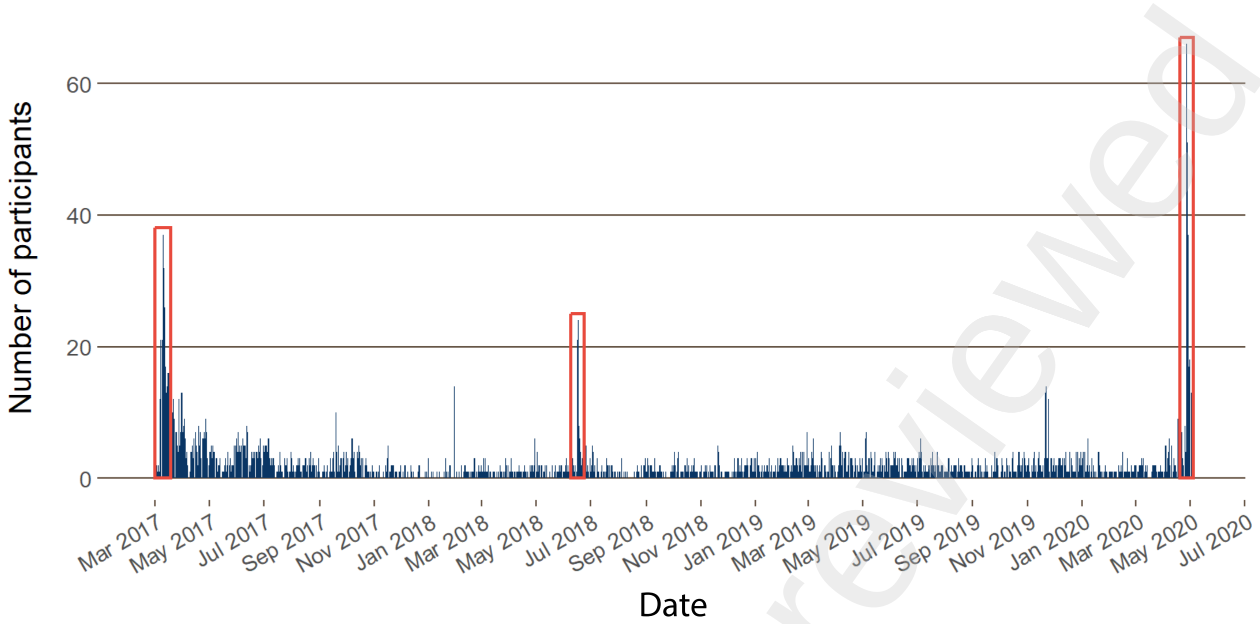

Figure 3. Number of active participants (i.e., who annotated at least an image) per day over time during the first mission. Red rectangles highlight major peaks of participation resulting from specific communication and outreach events (see text).

Participants' involvement also varied greatly, with the three most active participants contributing a fifth of all annotated images (i.e., 8,299 images; Figure 4). The median number of annotated images by participant was 7 and ranged from 1 to 4,444 out of 39,255 images. The most active participant contributed 37% (i.e., 116,754 annotations), and the three most active 43% (i.e., 136,019; see Figure 4) of all annotations.

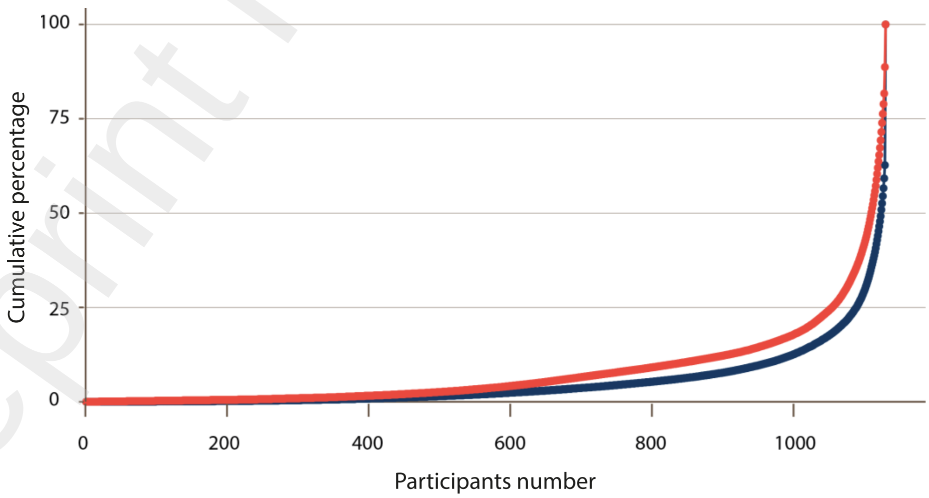

Figure 4. Cumulative proportion of annotations (blue) and annotated images (red) with increasing number of participants (sorted from least to most active).

3.2 Participants' profile

Almost all participants have indicated their country, France being the most frequent (91% of all participants). Participants originated from 26 countries, 8 francophone and 4 anglophone. However, francophone countries represent 97.7% of the participants (among those who have indicated their country) while anglophones represent 0.7%. The lack of participation from other countries most likely resulted from the language barrier and less efficient communication and outreach efforts at the international level. Only 11% of participants indicated their age. Among them, the most represented age range was from 10 to 20 years old (30% of participants that provided this information), highlighting the importance of outreach activities with schools detailed hereinafter. When it comes to participants' jobs, only 4% provided that information, a quarter of them being students in science.

3.3. Participants' behaviour

Most of the participants (87%) remained active within a week at most after registration, and 75% participated only the day they registered. Considering only the time between the first and last annotation, the number of participants that contributed only over one day reaches 83%. Indeed, 91% and 94% of the participants have started annotating within the same day or within three days following their registration, respectively. This behaviour suggest that most people connect by curiosity but do not feel involved enough to further contribute. A fifth of participants have annotated only one image and never annotated again. More interestingly, half of registered users (966) have never annotated any image. This could be due to a lack of interest, a lack of time to go further in the process, or issues in handling the UI.

The annotation time (AT) considers the time used to complete the full annotation of an image (i.e., annotation session). The average AT was 4 minutes and 10 sec, and the median 2 min, with 93% of participants having an AT lower than 10 min. In rare cases, the database recorded unusual time to complete a session of annotation, up to almost three days for the longest one, but most images (97%) were annotated in less than 9 min.

4. Pre-processing of the data: from citizen to scientific data

4.1. Merging multi-participant data: the deeptools R package

The challenge of such crowdsourcing database with repeated annotations from different users is to be able to merge multiple independent classifications into one organism occurrence (i.e., potentially annotated ten times). The detection and classification of organisms hence relies on a "vote" from multiple participant judgments (Fortson et al., 2012). Recent methodology was proposed to aggregate multi-participant classifications (Swanson et al., 2016), but are difficult to apply to our classification where the number of individuals by species and their distribution in the field of view is critical to the research. In this case, we need to take into account the spatial information of annotations, being the position in pixels of each individual organism. The deeptools package (2018, https://github.com/DeepSeaSpy/deeptools) was developed to tackle this issue and provide a method and functions to identify common organisms among participants in three steps (Figure 5). First, Voronoï polygons are drawn around each organism identified by all users. Voronoï diagrams allow to delineate a non-overlapping area around each object (i.e., organism) under the assumption that they all have equidistant separation (Figure 5D). Voronoï polygons are then cropped using a buffer area to define polygons around each single annotation (Figure 5E). A quick exploration suggests a buffer area of twelve pixels was enough to find overlapping annotations and fit the real size of the organisms in our photo database. Voronoï polygons are then transformed into rasters to recover the information about each pixel that constitutes the polygons and determine overlapping polygons. The last step consists in defining groups of polygons under the assumption that overlapping ones across users represent one single organism (Figure 5C). However, some polygons can be assigned to different groups which required developing a decision process to select groups under the constraint that a polygon can only belong to one group. The best combination of polygons was chosen following two rules. If a polygon is found in two groups, the procedure assigns the polygon to i) the most consensual group, i.e., identified by the highest number of participants, and ii) the group that have the greater number of overlapping pixels. These two rules have to be applied in a loop to reduce grouping possibilities until each individual polygon belongs to only one group (Figure 5C; function find_groups_in_image() in deeptools). If a polygon is not assigned to any existing group, a new group is created.

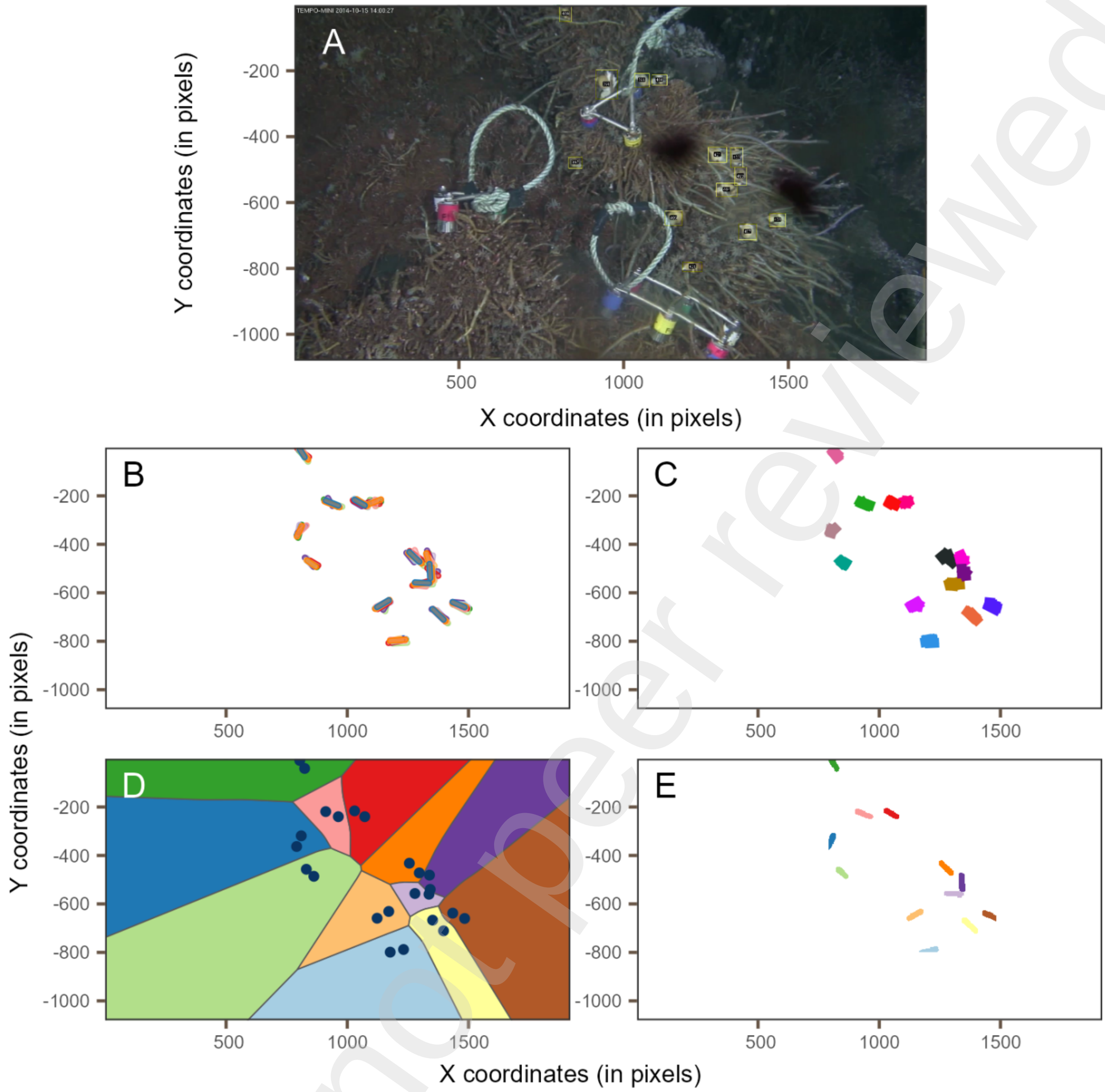

Figure 5. Procedure to merge multiple annotations in one single occurrence in a photo. A. Original photo showing the distribution of 14 real buccinid gastropods. (B, C). Procedure to delineate polygons for each contributor. B. Representation of all annotations performed by ten contributors with overlap between annotations of the same individual. Each colour represents a different participant. C. Groups of polygons that correspond to a single buccinid occurrence after merging all contributor's annotation with the deeptools package. Each colour represents a unique organism. In total 14 groups (i.e. unique buccinid) were detected by the participants. (D, E). Procedure to define polygons for a given contributor. D. Voronoï polygons delineated around each annotation. E. Voronoi polygons cropped using a buffer in pixel around the annotation.

The final output provides a list of unique organisms each associated, among others, with the photo ID (i.e., SMOOVE module, acquisition date and time), pixel coordinates in the image and the number of times the organisms were seen across all users who annotated the image (i.e., hereinafter called Agreement Rate - AR). For the full list of associated metadata, the output is provided in an open dataset (Cottais and Matabos, 2024; 10.5281/zenodo.10984717). Each user is given equal weight and the data can be cleaned through the choice of an agreement threshold.

4.2. Citizen data validation

Agreement Rate (AR) — In order to assess our confidence in citizens' performance, data quality was checked by comparing citizen annotations with those conducted by an expert at the lab using the Image J annotation software (Schneider et al., 2012). For the rest of the analyses we assume that the dataset acquired by the expert represents the real number of animals, i.e. the reference dataset, although this assumption could be biased as image analyses performed by humans is prone to error (Durden et al., 2016a). Prior to data comparison, we removed all annotations performed in the background in JdFR images to only focus on the tubeworm assemblage which has scientific interest (see Figure 1C and 5A), but also because some participants only counted animals in the foreground while others annotated the whole image. In addition, object identification in the back is unreliable because of their distance and low illumination (Figure 5A) and results are expected to have low accuracy. The reference dataset reported 15,571 individuals (i.e., 14,985 buccinids and 586 crabs) on 3,213 photos, when citizens identified a total of 35,168 individuals (33,602 buccinids and 1,566 crabs) on 3,844 photos. These results show a high number of false positives in the citizen data. Note that the difference in the number of annotated photos between expert and participants results in the absence of individuals in some photos, according to the expert. Altogether, 81% of organisms observed by the expert were annotated by at least one participant, while 3,030 individuals (19%) were only observed by the expert. To evaluate citizen's effectiveness in annotating individuals of buccinids and crabs, we analysed the results in relation to the AR that corresponds to the relative number of times a unique individual was detected across all participants. This AR is set to a minimum threshold above which all detected individuals are considered (Figure 6).

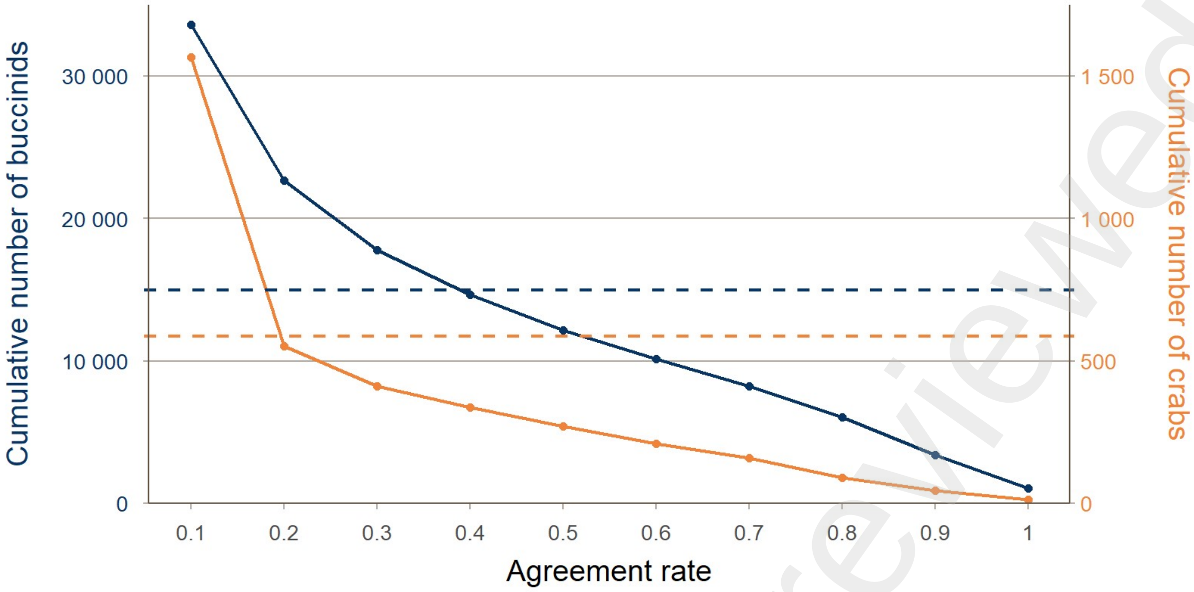

Figure 6. Cumulative number of unique individuals detected depending on the citizen agreement rate threshold for the snail *Buccinum thermophilum* (blue, JdFR) and the crab *Segonzacia mesatlantica* (orange, MAR), with expert counts (dashed lines).

Among citizens, most annotated individuals have been identified by only one participant, corresponding to 32% of the buccinid snails and 65% of the crabs. However, among "consensual" individuals, i.e., that have been detected by at least two participants (AR > 0.1), 54% of the buccinid snails and 49% of the crabs have been identified by at least half of the participants ($0.5 \leq AR \leq 1$). Indeed, we observed a sharp decrease in organism detection when considering an AR between of 0.1 and 0.2, with the loss of about a third of the buccinid snails and two thirds of the crabs (Figure 6). This result supports the principle behind the 'vote' concept, based on the fact that two people will not make the same error (i.e., false positive at the exact same location in the image). As expected, the most active participants, in terms of number of annotations, are also the ones that identified the highest number of individuals assigned to only one group (AR = 0.1), hence contributing to the high number of false positives. Conversely, some of the most active participants did not annotate many organisms, leading to a high number of false negative. This difference can reflect participant enthusiasm in performing well, by fear of missing an individual leading to over-annotation, or to wrongly identify a species leading in this case to a precautionary behaviour and under-annotation. Indeed, some organisms can be partially hidden behind larger ones such as tubeworms or mussels, making them hard to identify. The presence of a white 'patch' could then be interpreted as part of an individual when it is not, or erroneously discarded leading to the major difference in detection level.

The AR which fits the best the expert identification differed depending on the species. Considering buccinid snails, an AR of 0.4 appeared as a good threshold as it leads to the same number of individuals as the one detected by the expert (Figure 6). However, because the total amount of individuals does not inform on differences within a given photo, we calculated a delta between unique individuals annotated by citizens compared to the expert, by photo and in relation to AR threshold (Figure 7). The box plot showed that, despite some outliers, false positives are equally distributed among images. For instance, when considering all unique individuals (i.e., AR = 0.1), the difference between citizens and expert is below ten individuals for most of the photos. Results supported the choice of 0.4 as a proper AR to account for the real number of buccinid snails with the absence of significant difference with the expert (Figure 6 and 7). It is noteworthy that differences with the reference dataset displayed only slight changes between AR of 0.2 and 0.6. Above these thresholds, there is a significant underestimation of individual counts and a sharp drop with increasing AR, reaching next to 0 for an AR of 1 (i.e., all participants detecting a given individual). Surprisingly, AR > 0.1 leads to a non-negligible number of false positives although it is expected that two participants could not find non existing organism in the same location. This result can be attributed to the complexity of the image where several objects can be similar in shape and colour. For instance, some buccinid hidden behind tubeworms are recognisable by colour, hence, a patch of microbial mats/filaments or large pycnogonids (arthropods) partially hidden could be mistaken for a buccinid snail.

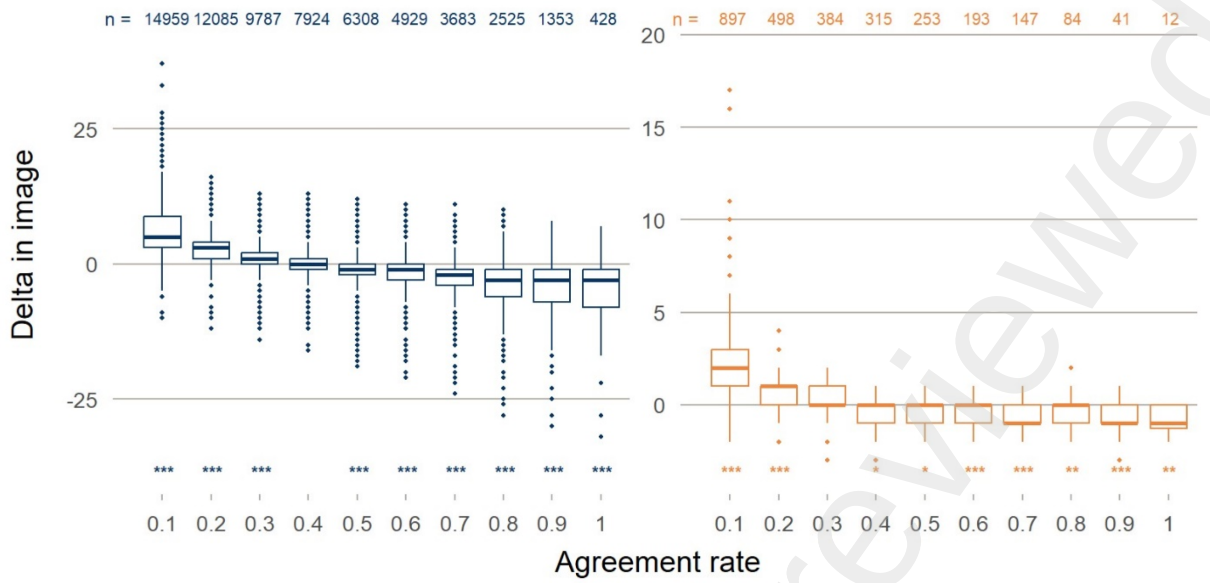

Figure 7. Difference between expert and participants in abundances of the snail *Buccinum thermophilum* (left) and the crab *Sgonzacia mesatlantica* (right) as function of the agreement rate threshold from each annotated image. Note that for crabs, one outlier corresponding to 60 observed organisms considering an AR = 0.1 was removed to ease the readability. Averages of delta distributions per agreement rate were compared to zero with Student t-tests (***: p-value < 0.001 ; **: p-value < 0.01 ; *: p-value < 0.05 ; ·: p-value < 0.1). n: number of images.

Regarding the crabs, the difference with expert counts evolves more rapidly depending on the AR with the best fits occurring for AR of 0.2 and 0.3 (Figure 6). Considering both the total amount of unique individuals across the entire dataset, and by photo, an agreement of 0.2 appeared optimal to minimise the number of false negatives (Figures 6 and 7). Above this threshold, the number of detected animals sharply decreased. This difference with buccinid snails can be attributed to the complexity of the task. Indeed, crabs are a territorial species that inhabits mussel beds (Matabos et al., 2015). They are often partially hidden among mussels and can only be detected through the presence of a claw or piece of carapace among mussel shells making them hard to see. In addition, because of their territorial and aggressive behaviour, only few individuals colonise the field of view, leading to a high number of images with a real absence of crabs. This might lead most active participants to get used to the absence of crabs on the photos, and thus quickly validate an image 'by habit', whether or not it contains a crab.

Subsampling among participants — The number of times an image had to be annotated was set to ten, based on expert opinion. To assess if a reduced number of participants would have been sufficient to correctly detect real individuals, data were sub-sampled to test changes in detection level depending on the number of participants annotating a photo. Three sub-sampling rates were considered (i.e., three, four and five participants) and the number of detected individuals was then compared to data collected with ten participants per photo as well as with the expert reference data, considering several levels of AR. To this end, we randomly sub-sampled three, four and five participants out of the ten who annotated a given photo. For each number of participants, the process was iterated five times and then averaged. Finally, the number of detected individuals was considered for the different AR thresholds across all images (Figure 8) and by photo. The number of detected buccinids showed little variation depending on the number of participants except for low level of AR (i.e., < 0.4), where the number of false positives is higher when considering a reduced number of participants (Figure 8). For an AR of 0.4, the same number of individuals were detected considering five or ten participants. Considering differences between participants and expert by photo confirmed this pattern (results not shown). From these results, considering the annotation of buccinids, the choice of an AR of 0.4 based on five participants per photo appeared as the optimal strategy to obtain the best detection of real individuals while minimising the number of contributions required from the participants. Dividing by two the number of required contributors would double the annotation power of the DSS platform.

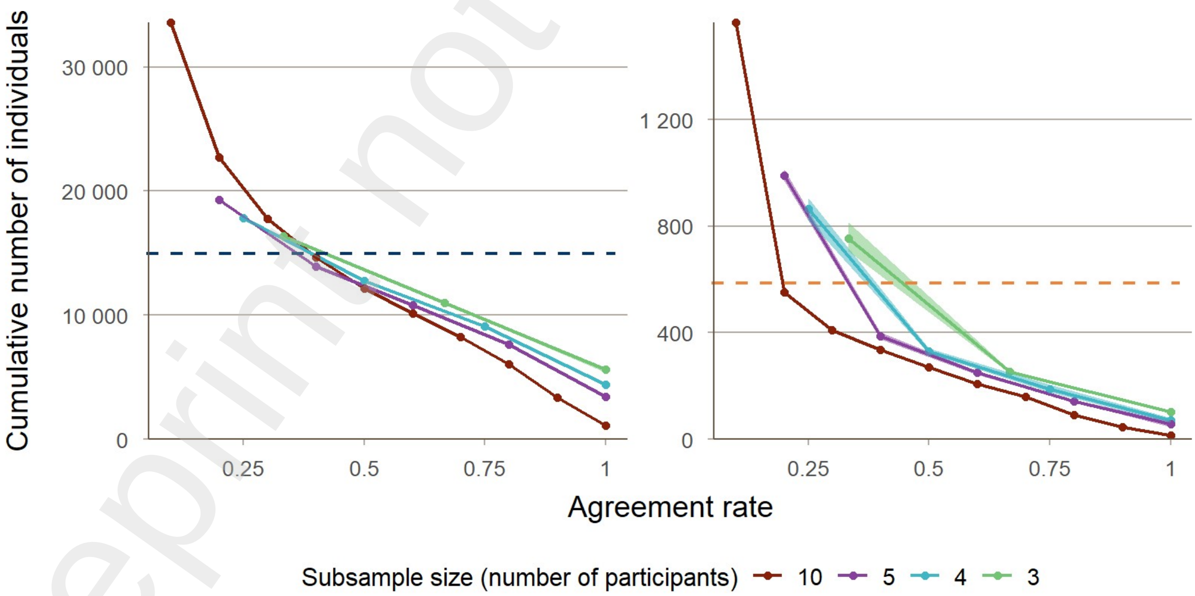

Figure 8. Abundances of the snail *Buccinum thermophilum* (left) and the crab *Segonzacia mesatlantica* (right) across all images as function of the Agreement Rate threshold. Average abundances resulting from the annotation of the ten participants (red) were compared with

averages after sub-sampling with three (green), four (blue) and five participants (purple) per photo and with expert counts (dashed).

Considering crab data, differences in the number of detections were higher with an increasing number of contributors (Figure 8). For all sub-sample sizes, considering all unique individuals leads to a high number of false positives based on the expert data. Conversely, considering individuals annotated by at least 2 participants results in a sharp drop in the number of detected individuals and a high number of false negative. This observation can most likely be attributed to the lower number of individuals within an image. Hence, for this species, it seems more optimal to maintain the number of contributors by photo to 10, in order to obtain the best detection of real individuals (Figure 8).

Temporal trend in the evolution of buccinid abundance — Depending on the scientific question, the absolute number of individuals in a photo might not be the most instructive piece of information. For instance, in this mission, the objective was to explore the role of tides in species behaviour. Variations in abundance over time was thus more informative. While we lacked data to make robust temporal analyses on crab abundances, we investigated temporal trend of the buccinid snail population according to the AR and expert reference data. Interestingly, while increasing the AR among participants leads to the loss of buccinids, the relative trend remained similar for thresholds between 0.2 and 0.6 (Figure 9). This robustness in data distribution can be explained by the fact that false positives and false negatives are equally distributed among the ca. 3,000 photos which contributes to smooth annotation errors. Above these threshold values, the curve from citizen data tends to flatten as a result of a higher number of false negatives. This result highlights the power of citizen contribution in monitoring species abundances across time.

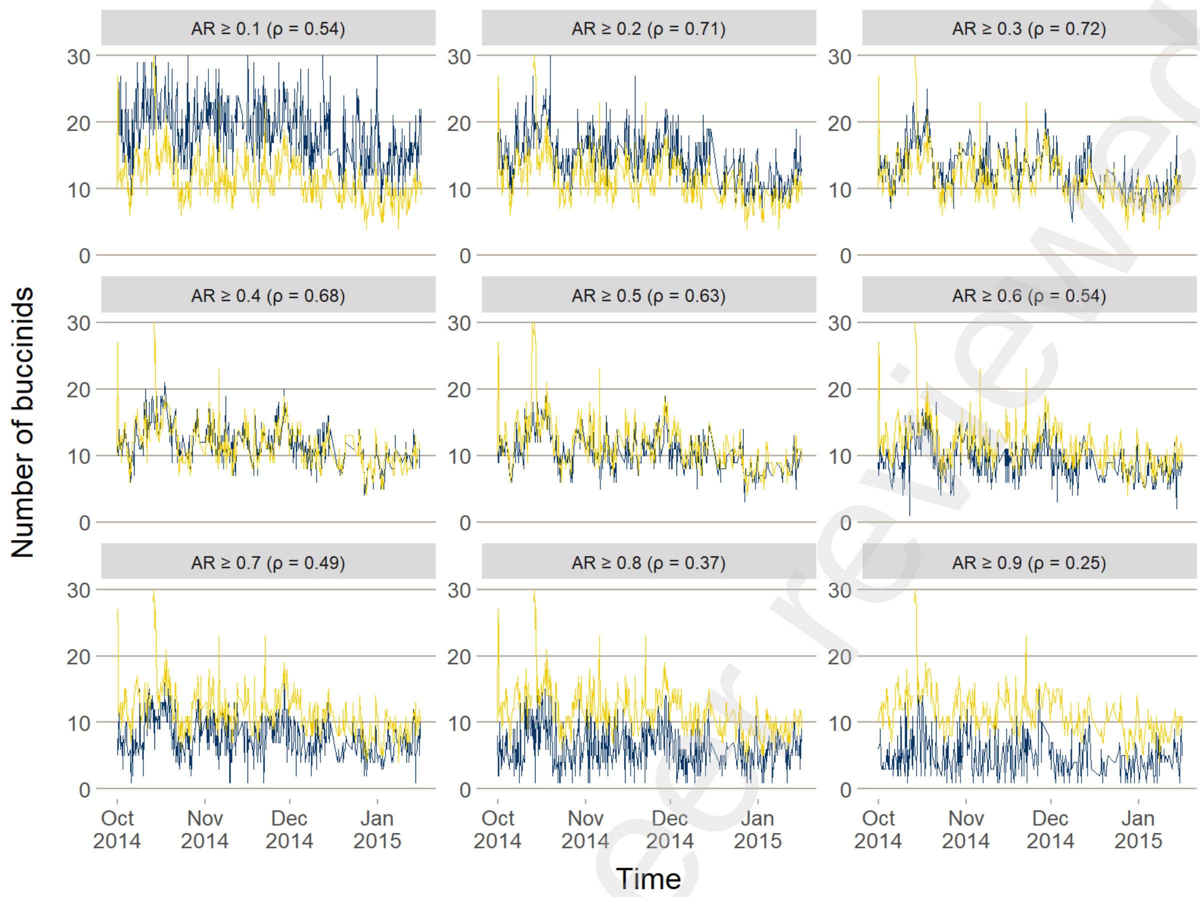

Figure 9. Evolution of the snail *Buccinum thermophilum* abundances depending on participants (blue) agreement rate threshold compared with expert (yellow). Pearson correlation coefficients (ρ) between expert and citizen data are shown on top of the graphs and are all significant (p value < 0.001).

5. Citizen science as a tool for ocean literacy

In this project, we contributed to enhancing interactions between society and research, including deep-sea infrastructures. Involving citizens in the scientific process of imagery annotation is mutually beneficial for scientists and society by i) raising awareness about scientific research, environmental issues and the deep ocean, and ii) offering new ways for data collection and processing to handle the bottleneck generated by imagery big data.

5.1. Public outreach

A number of actions and material was developed in parallel to support communication around the project, raise awareness about the deep sea, its wonders and the threats it is facing, and make the platform an educational support available to teachers.

First, a project website "behind the scene" (www.deepseaspy.com) was developed to provide participants with background information. It provides the project description, science background on hydrothermal vent ecosystems, their associated biodiversity, and deep-sea observatories. Short video sequences, photos and introductory texts enable citizens to take ownership of the topic.

To enhance DSS visibility in the long term, we set up a partnership with the Océanopolis aquarium based in Brest, France. A multimedia computer terminal (Figure 10A) was installed in the deep-sea section of the 'Pavillon Bretagne' exhibition area with a specific "visitor" account. To date, the public has annotated 3,175 images ranking them 4th contributor globally. During the 1st mission analysed in this paper, they annotated 2,747 images corresponding to 2,102 photos making them the 2nd best contributor. We suspect that most visitors tried annotating without validating their images such that the above number are under-estimated. The visitors annotated 409 unique buccinid snails from which only 146 are true positives (i.e., found by the expert). Similarly, among the 55 crabs they annotated, only 2 are common with those of the expert. These results highlight the lack of implication from visitors. This could result from the environmental context, where visitors tend to just « play » with interactive set up and equipment compared to citizens contributing willingly from their own computer. Hence, while the computer terminal and the associated exhibition material displayed at the aquarium might contribute to communicate around the project and questions about the deep sea, we conclude that these annotations should not be considered for further analyses or as reference dataset to train machine algorithms.

To increase the impact of the project, the research team participated in a number of public events including public conferences in many cities and regions, science events including the iconic science festival (Fête de la Science) organised every fall across the country (Figure 10B). For these occasions, visitors are invited to test the application using a computer set up at the Ifremer booth. In Paris, the application was demonstrated to the French Minister of Higher Education, Research and Innovation (https://youtu.be/EHcHzs_vqXI?t=2996), to the public on site but also to all virtual participants who watched the YouTube channel of the science dissemination program "L'Esprit Sorcier".

Figure 10. Example of outreach events related to the Deep Sea Spy (DSS) project. A. Computer terminal with the DSS application at the Océanopolis aquarium in Brest (photo credit N. Roullet). B. Demonstration to kids at the science festival in Brest, France (photo credit Ifremer). C. School kids with the DSS educational booklet (photo credit A. Bianic, école Sainte Anne, Saint-Thonan, France). D. Video conference from the ship to introduce the ship crew to kids on land (photo credit F. Le-Moigne, école Mouez Ar Mor, Ploumoguer, France).

5.2. Educational resources

A number of collaborations with schools were developed along the course of the project from kinder garden to high school as well as university through professor training. Material was adapted to kids' age. For kids from 3 to 12 years old, we developed educational booklets dedicated to teachers to turn the application into a school project (Figure 10C). They propose information and exercises adapted to the French National Academic program to introduce deep-sea exploration, biodiversity and scientific methodology. They support the participation of classrooms to online annotation by helping teachers explaining the science behind the project, its importance for society, and the scientific approaches used. Educational booklets are freely available online (https://www.deepseaspy.com/en/Educational-material2). School interactions were extended in 2019 to a project called "Plouarnautes" that was designed to engage kids from several classes into the experience of an oceanographic cruise from material preparation to operations at sea. The project involved a visit at the Ifremer centre in Brest, video conferences between the scientific team and crew onboard and the classes on land, and the publication of a blog over the lifetime of the project. In 2019, two schools in Ploumoguer and Plougastel-Daoulas (Britanny, France) were involved (Figure 10D).

Resources targeting high-school teachers were also developed in collaboration with Océanopolis and the national education network "Réseau Canopé". Every year, the region of Brittany organises a week of science immersion for high school students selected from schools across the entire region. Since 2013, DSS is part of the program which includes a general conference on deep sea and hydrothermal vents followed by a practical 'lab' to analyse images. We also initiated a collaboration with the regional school academy to integrate the project into schools through the French education special program entitled Culture, Society and Information Technology (CSTI). CSTI provides means for the teachers to develop middle and high schools' students scientific culture through partnership with research institutions. The academy supports these projects by providing an online space that lists potential resources and partnerships, but also by fostering contacts with local researchers. In this context, DSS was selected and officially supported by the academy.

The advantage in working with schools lies in the fact that students change every year, bringing new participants. Not only these projects contribute to increasing the number of participants, but constitute a powerful tool to raise awareness about the deep sea among the young generation. The choice of tool is of utmost importance, and we emphasise here the importance of developing these resources in collaboration with educational professionals and in accordance with national programs. The tools developed also proved useful to make accessible, popularize and explain the scientific approach, demonstrating that everyone can contribute to research, thus removing barriers between science and society. Citizen science approach can thus help improve the quality, credibility, or relevance of research projects (Winickoff et al., 2016).

6. Discussion and conclusions

This paper presents the first results of the Deep Sea Spy project since its launch in 2017 and highlights the tremendous potential of citizens to support research and contribute to building large databases of annotated images (e.g., Anton et al., 2021; Kuminski et al., 2014; Lintott et al., 2008). Comparing annotations by a marine biologist expert from the lab with citizen participants showed some limitations in finding accurate abundances in our set of images. Discrepancies between citizen and expert data can be related to video quality, abundance of species of interest, participant experience, and task complexity as observed in similar studies (Langenkämper et al., 2019; Wick et al., 2020). For instance, video quality is variable and can be affected by illumination with a degradation of intensity over time, or by biofouling on the camera lens, that might mask part of the field of view, thus making it hard to distinguish organisms (Cowling et al., 1998). These issues are common at deep-sea observatories where sensors are deployed over long periods, from several month up to several years. The presence on the camera of a microchlorination system to protect the optical sensor and projector from biofouling (Delauney et al., 2010) helped mitigate this problem.

We have showed that the experience of participants can have several effects, accounting for most of the false positives, but also false negatives. Indeed, because of the complexity of the images in terms of objects, texture and homogeneous colours, targeted species can be hard to classify. Organisms can be partially hidden behind large engineer species (e.g., mussels, tubeworms), making it hard to detect them only based on their shapes. As a result, the best clue of the presence of an organism is the location in the field of view (i.e., its habitat) combined with the colour. At vents, many organisms, including microbial mats and filaments, display colours in a gradient from white to light brown. It is thus easy to misidentify patches in areas where organisms are expected. The occurrence of false negatives or positives may also result from the participant's behaviour. Independently of their annotating experience, some participants tend to under-estimate the number of individuals in the image by fear of wrongly annotating objects while others appear more concerned about missing an organism and tend to over-annotate. Hence, participant experience does not appear as an important factor of accuracy. This is contradictory to a number of studies that highlighted higher performance on trained participants (e.g. Delaney et al., 2008; Matabos et al., 2017; Wick et al., 2020). However, several studies also showed that citizens can perform almost as well as professionals (Crall et al., 2011; Holt et al., 2013). In our case, because the field of view remains constant over time, a participant can quickly learn how to recognise an organism as long as a tutorial picturing the targeted species in its environment is available. In addition, an expert can also make mistakes as annotating thousands of images is repetitive and can lead to fatigue and a drop in attention (Durden et al., 2016a; Swanson et al., 2016). This could explain part of the discrepancies in classification between the expert and participants, and probably accounts for some of the false positives for agreement rates above 0.1 in buccinid snails counts. Swanson et al. (2016) showed that the aggregated participant answers were more accurate (97.9%) than those of individual experts (96.6%) when compared with the consensus expert assessments. It would have been interesting in our study to multiply the number of experts, but this could not be achieved due to limited human resources and time.

Finally, while some studies showed that even for more complicated tasks, participants can perform as well as experts (Butt et al., 2013; Delaney et al., 2008), task complexity appeared as an important factor to take into account. Indeed, the level of optimal agreement rate strongly differed between the two considered species. This suggests that the validation procedure should be adapted and reconsidered for each analysed species. These differences can be explained by a combination of factors including, organisms' sizes, abundances and behaviour (i.e., more or less hidden in mussels or tubeworms). Considering this outcome, future missions could be targeted on only one species to facilitate the detection from non-trained citizens by making the task more manageable (Langenkämper et al., 2019).

The next challenge will be the processing of such complex big data set to reach conclusions on the ecology of vent communities. A growing literature is now available on methods to process citizen science data (Bird et al., 2014; Bonter and Cooper, 2012; Kosmala et al., 2016; Wiggins et al., 2011) and will provide directions on future analyses. However, because of the wide variety of citizen science data, even in the specific case of imagery in terms of annotation types, it remains difficult to offer a common standardised validation approach. An interesting result in this study is the robustness in the detection of the temporal variation in abundance evolution, independently of the agreement rate considered. This result is of utmost importance in the context of monitoring, which aims to disentangle natural rhythms from long-term trends. A big challenge in deep-sea research, which is associated with technological constraints, limited access and high costs, is our capability to assess status and evolution of ecosystems (Aguzzi et al., 2011; Danovaro et al., 2017). Disentangling natural cycles from long-term changes caused by natural processes, climate or human activities, requires a long-term monitoring of these environments. Such a robust performance in citizen to detect trends in time series holds great potential for the processing of observatory data and will contribute to unlock the bottleneck associated with the exponential growth of imagery databases.

Another envisioned approach for the future is to use citizen data to train deep learning algorithms to detect organisms in the photos (Kuminski et al., 2014). Indeed, reaching good detection rate and performance using machine learning requires large reference datasets that are today not available in marine environments (Durden et al., 2021). Citizens can produce large datasets, that if properly validated, have great potential to advance machine learning applications (Anton et al., 2021; Langenkämper et al., 2019; Van den Bergh et al., 2021). However, it will be important to maintain a strong interest among the public and schools to ensure a high participation rate and thus the acquirement of valuable data. While a number of studies proposes new methodologies for the validation of citizen data (e.g., this study; Bird et al., 2014; Wick et al., 2020), training algorithms require clean datasets to ensure a good and reliable learning process. A solution could be to enlist a community of participants to review thumbnail images produced by cropping the photo based on annotation coordinates (Sullivan et al., 2009). This requires an evolution of the platform where citizens can validate and correct existing annotations to help clean reference databases. This would contribute to reduce false positives and correct wrong classifications. Optimising the efficiency of such an approach and ensuring proper validation requires the selection of trained and engaged participants, although coordinating such a community demands significant human and financial resources (Bonter and Cooper, 2012; Sullivan et al., 2009).

By comparing the first mission with expert knowledge, this paper aimed at defining a workflow of citizen science validation for the processing of large observatory imagery database. Indeed, images included in the project are acquired by cameras deployed on deep-sea observatories which aims at monitoring benthic ecosystems over decades to gain a better knowledge on ecosystem natural dynamics, essential to predict and detect variations related to anthropogenic activities and global change (Matabos et al., 2022). As a result, the database is still growing and global deep-sea benthic imagery will exponentially increase worldwide in the future as other observatories develop and new technologies emerge (e.g., Drones; Autonomous Underwater Vehicles, cameras; e.g. Aguzzi et al., 2019; Danovaro et al., 2017). Our workflow completes the set of validation standards proposed in the last decade (e.g. Kosmala et al., 2016; Swanson et al., 2016) and can be applied to any citizen dataset generated from single-point long-term video camera systems, or any image annotation that gives importance to the spatial organisation of objects in the image.

Finally, DSS is more than a citizen science project but rather constitutes a full program allowing for the processing of large volume of imagery data (i.e., crowdsourcing; Silvertown, 2009), and provides a platform to raise awareness about the deep sea through media, public events and conferences, as well as educational resources for kids and teachers. Motivation factors of participants include an interest in science, or more specifically in the project's topic, or the will to learn something (De Vries et al., 2019; Raddick et al., 2013; 2010). While providing participants accessibility to the data they collected and sharing scientific findings are essential to ensure continued participation (De Vries et al., 2019), efforts in communicating the science behind the project and supporting education professionals through free learning and outreach material proved extremely useful to maintain collaborations in the long term. The team dedicated a large effort to communicate the science in schools and local scientific events. Our collaborations with high-school classrooms involved providing data collected by the students that they can manipulate, which harboured a great educational power (Bonney et al., 2009a). However, these activities are time-consuming and hard to maintain in time by scientists only, without a dedicated team. Communicating scientific output to the full community of participants requires a dedicated science outreach program, human resources for data curation and preparation, and a communication plan. Ensuring such aspects can be difficult for individual laboratories with limited human resources. But our experience showed that providing associated educational material and developing collaborations with local stakeholders can strongly contribute to maintain a local engaged community. This new way of "making science" can benefit both citizens and researchers by accelerating the processing of large imagery datasets for researchers, and by learning and engage in science for participants (Bonney et al., 2009b). Citizen science contribute to breaking down barriers between academic research, raises awareness on environmental issues and conservation, and contribute to citizens' engagement and empowerment. Recently, the annotation platform was extended to other ecosystem compartments into a single digital infrastructure, Ocean Spy (https://ocean-spy.ifremer.fr), using a common web-based portal and a unique database hosted at Ifremer. Oceans are changing fast and are increasingly impacted by human activities. Acquiring the necessary knowledge to properly inform environmental management requires technological developments to increase our observation and monitoring capacities, but also new means to accelerate data processing and analyses. Citizens represent a great reservoir of scientists, and citizen science has tremendous potential to enhance scientific knowledge in time and space and increase ocean literacy for the benefit of all (Garcia-Soto et al., 2017).

Acknowledgements

We first warmly thank the 1,130 deep-sea spies who contributed to this first mission. We also thank all the schools that participated to the project: the middle school Collège Dom Michel in

Le Conquet, elementary schools Mouez Ar Mor in Ploumoguer, Sainte Anne in Saint Thonan, Petit bois in Plouguin, école des 4 moulins in Brest, école Prévert in Guipavas, école de Kerisbian in Brest, and the high-school, Lycée Assomption Rennes with a special thanks to Murielle Waendendries for her continued participation and engagement with her students since the beginning. We also thank the crew members and pilots who accepted, during the Momarsat 2019 cruise, to take part in the Plouarnautes project to exchange with classrooms from the ship and share their work. We are grateful to Sébastien Rochette for his contribution in the development of the R deeptools package, and Patrick Bossard for the integration of the web application on the IFREMER IT infrastructure. We also thank the captain and crew of the RVs Pourquoi Pas?, the pilots of the ROV Victor6000 and chiefs scientists, Pierre-Marie Sarradin, Mathilde Cannat and Jérôme Blandin, of the Momarast cruises 2013 and 2014 that enabled the deployment and recovery of the TEMPO module. We also thank the captain and crew of the CCGS John P. Tully that enabled the deployment of the TEMPO-mini module at the Juan de Fuca ridge. We are grateful to all the engineering team from Ocean Networks Canada and the Technological Research and Development (RDT) department at Ifremer for the development and deployment of the ecological modules hosting the SMOOVE cameras. Thank you to Atelier Canopé, and the Rennes French National Academy for opportunities to connect with schools and include the project in national educational array. Finally, the authors would like to dedicate this paper to Anne Rognant, previous curator in charge of scientific and cultural outreach at the Océanopolis aquarium (Brest) who recently passed away. We are forever grateful for her enthusiasm, support and help in developing the educational resources, building a large network of schools and teachers across Brittany, and providing many opportunities to highlight the project through public events and school intervention.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme ENVRIplus under grant agreement No 654182. This work also benefited from State aid managed by the National Research Agency under France 2030: ANR-22-POCE-0007.

References

Aguzzi, J., Chatzievangelou, D., Marini, S., Fanelli, E., Danovaro, R., Floegel, S., Le Bris, N., Juanes, F., De Leo, F., del Rio, J., Thomsen, L., Costa, C., Riccobene, G., Tamburini, C., Lefevre, D., Gojak, C., Poulain, P.-M., Favali, P., Griffa, A., Rountree, R., 2019. New High-Tech Flexible Networks for the

Monitoring of Deep-Sea Ecosystems. Environ Sci Technol 53. (12), 6616-6631. https://doi.org/10.1021/acs.est.9b00409

Aguzzi, J., Company, J.B., Menesatti, P., Garcia, J.A., Bahamon, N., Puig, P., Sardà, F., 2011. Activity rhythms in the deep-sea crustacean: chronobiological challenges and potential technological scenarios. Frontiers in bioscience-Landmark 16, 131–150.

Aguzzi, J., Costa, C., Fujiwara, Y., Iwase, R., Ramirez-Llodra, E., Menesatti, P., 2009. A novel morphometry-based protocol of automated video-image analysis for species recognition and activity rhythms monitoring in deep-sea fauna. Sensors (Basel) 9(11), 8438–8455. https://doi.org/10.3390/s91108438

Aguzzi, J., Doya, C., Tecchio, S., De Leo, F.C., Azzurro, E., Costa, C., Sbragaglia, V., Del Río, J., Navarro, J., Ruhl, H.A., Company, J.B., Favali, P., Purser, A., Thomsen, L., Catalán, I.A., 2015. Coastal observatories for monitoring of fish behaviour and their responses to environmental changes. Rev Fish Biol Fish 25, 463–483. https://doi.org/10.1007/s11160-015-9387-9

Anton, V., Germishuys, J., Bergström, P., Lindegarth, M., Obst, M., 2021. An open-source, citizen science and machine learning approach to analyse subsea movies. Biodivers Data J 9, 1–14. https://doi.org/10.3897/BDJ.9.e60548

Apostolopoulos, K., Potsiou, C., 2022. Consideration on how to introduce gamification tools to enhance citizen engagement in crowdsourced cadastral surveys. Survey Review 54(383), 142-152. https://doi.org/10.1080/00396265.2021.1888027

Aron, M., Sarrazin, J., Sarradin, P., Mercier, G., 2010. Elaboration of a video processing platform to analyze the temporal dynamics of hydrothermal ecosystems, in: AGU Fall Meeting Abstracts. San Francisco, p. 4.

Auffret, Y., Coail, J.-Y., Delauney, L., Legrand, J., Dupont, J., Dussud, L., Guyader, G., Ferrant, A., Barbot, S., Laes-Huon, A., 2009. Tempo-Mini: A Custom-designed instrument for real-time monitoring of hydrothermal vent ecosystems. Instrumentation viewpoint 17.

Barnes, C.R., Best, M., Zielinski, A., 2008. The NEPTUNE Canada And the Design of the Science Experiments. Sea Technology 49, 10–14.

Biard, T., Ohman, M.D., 2020. Vertical niche definition of test-bearing protists (Rhizaria) into the twilight zone revealed by in situ imaging. Limnol Oceanogr 65(11), 2583–2602. https://doi.org/10.1002/lno.11472

Bird, T.J., Bates, A.E., Lefcheck, J.S., Hill, N.A., Thomson, R.J., Edgar, G.J., Stuart-Smith, R.D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J.F., Pecl, G.T., Barrett, N., Frusher, S., 2014. Statistical solutions for error and bias in global citizen science datasets. Biol Conserv 173, 144–154. https://doi.org/10.1016/j.biocon.2013.07.037

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., Shirk, J., 2009a. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. Bioscience. 59(11), 9759(11), 977-9847-984. https://doi.org/10.1525/bio.2009.59.11.9

Bonney, R, Ballard, H, Jordan, R, McCallie, E, Phillips, T, Shirk, J and Wilderman, CC. 2009b. Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education. A CAISE Inquiry Group Report. Washington, DC: Center for Advancement of Informal Science Education (CAISE).

Bonter, D.N., Cooper, C.B., 2012. Data validation in citizen science: A case study from Project FeederWatch. Front Ecol Environ. https://doi.org/10.1890/110273

Boschen, R.E., Rowden, A.A., Clark, M.R., Gardner, J.P.A., 2013. Mining of deep-sea seafloor massive sulfides: A review of the deposits, their benthic communities, impacts from mining, regulatory frameworks and management strategies. Ocean Coast Manag 84, 54–67. https://doi.org/10.1016/j.ocecoaman.2013.07.005

Butt, N., Slade, E., Thompson, J., Malhi, Y., Riutta, T., 2013. Quantifying the sampling error in tree census measurements by volunteers and its effect on carbon stock estimates. Ecological Applications 23(4), 936–943. https://doi.org/10.1890/11-2059.1

Cannat, M., Sarradin, P., Blandin, J., Ballu, V., Barreyre, T., Chavagnac, V., Colaço, A., Crawford, W.C., Daniel, R., Escartı́n, J., Legrand, J., Matabos, M., Rommevaux, C., Roullet, G., Reverdin, G., Sarrazin, J., 2016. EMSO-Azores : Monitoring seafloor and water column processes at the Mid-Atlantic Ridge. Fix03 - Project Newsletter. Service Activities Special 3.

Cannat, M., Sarradin, P., Blandin, J., Escartı́n, J., Colaço, A., 2011. MoMar-Demo at Lucky Strike. A near-real time multidisciplinary observatory of hydrothermal processes and ecosystems at the Mid-Atlantic Ridge, in: AGU Fall Meeting, Abstract OS22A-05. San Francisco.

Cottais, P., & Matabos, M. (2024). Data from: Deep Sea Spy: an online citizen science annotation platform for science and ocean literacy [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10984717

Courtene-Jones, W., Quinn, B., Gary, S.F., Mogg, A.O.M., Narayanaswamy, B.E., 2017. Microplastic pollution identified in deep-sea water and ingested by benthic invertebrates in the Rockall Trough, North Atlantic Ocean. Environmental Pollution 231, 271–280. https://doi.org/10.1016/j.envpol.2017.08.026

Cowling, M.J., Hodgkiess, T., Parr, A.C.S., Smith, M.J. , Kerr, A., Beveridge, C.M., Clegg, M. and Menlove, R. (1998) The effects of biofouling on imaging underwater, including possible remedies. In: Grattan, K.T.V. (ed.) Applied Optics and Opto-electronics 1998: Proceedings of the Applied Optics Divisional Conference of the Institute of Physics, held at Brighton, 16-19 March 1998 (Institute of Physics Conference). IOP Publishing Ltd., pp. 245-254.

Crall, A.W., Newman, G.J., Stohlgren, T.J., Holfelder, K.A., Graham, J., Waller, D.M., 2011. Assessing citizen science data quality: An invasive species case study. Conserv Lett. 4(6), 433-442. https://doi.org/10.1111/j.1755-263X.2011.00196.x

Cuvelier, D., Legendre, P., Laës-Huon, A., Sarradin, P., Sarrazin, J., 2017. Biological and environmental rhythms in (dark) deep-sea hydrothermal ecosystems. Biogeosciences 14, 2955–2977. https://doi.org/10.5194/bg-14-2955-2017

Cuvelier, D., Legendre, P., Laes-Huon, A., Sarradin, P., Sarrazin, J., 2014. Rhythms and Community Dynamics of a Hydrothermal Tubeworm Assemblage at Main Endeavour Field–A Multidisciplinary Deep-Sea Observatory Approach. PLoS One 9(5), e96924.

Danovaro, R., Aguzzi, J., Fanelli, E., Billett, D., Gjerde, K., Jamieson, A., Ramirez-Llodra, E., Smith, C.R., Snelgrove, P.V.R, Thomsen, L., Van Dover, C., 2017. An ecosystem-based deep-ocean strategy. Science 355(6324), 452-454.

De Vries, M., Land-Zandstra, A., Smeets, I., 2019. Citizen Scientists' Preferences for Communication of Scientific Output: A Literature Review. Citiz Sci 4(1). https://doi.org/10.5334/cstp.136

Delaney, D.G., Sperling, C.D., Adams, C.S., Leung, B., 2008a. Marine invasive species: Validation of citizen science and implications for national monitoring networks. Biol Invasions 10, 117–128. https://doi.org/10.1007/s10530-007-9114-0

Delauney, L., Compère, C., Lehaitre, M., 2010. Biofouling protection for marine environmental sensors. Ocean Science 6(2), 503–511. https://doi.org/10.5194/os-6-503-2010

Devine, B.M., Baker, K.D., Edinger, E.N., Fisher, J.A.D., 2020. Habitat associations and assemblage structure of demersal deep-sea fishes on the eastern Flemish Cap and Orphan Seamount. Deep Sea Res 1 Oceanogr Res Pap 157, 103210. https://doi.org/10.1016/j.dsr.2019.103210

Durden, J.M., Bett, B.J., Schoening, T., Morris, K.J., Nattkemper, T.W., Ruhl, H.A., 2016a. Comparison of image annotation data generated by multiple investigators for benthic ecology. Mar Ecol Prog Ser 552, 61–70. https://doi.org/10.3354/meps11775

Durden, J.M., Hosking, B., Bett, B.J., Cline, D., Ruhl, H.A., 2021. Automated classification of fauna in seabed photographs: The impact of training and validation dataset size, with considerations for the class imbalance. Prog Oceanogr 196, 102612. https://doi.org/10.1016/j.pocean.2021.102612

Durden, J.M., Schoening, T., Althaus, F., Friedman, A., Garcia, R., Glover, A.G., Greinert, J., Stout, N.J., Jones, D.O.B., Jordt, A., Kaeli, J.W., Köser, K., Kuhnz, L.A., Lindsay, D., Morris, K.J., Nattkemper, T.W., Osterloff, J., Ruhl, H.A., Singh, H., Tran, M., Bett, B.J., 2016b. Perspectives in visual imaging for marine biology and ecology: From acquisition to understanding. Oceanography and Marine Biology: An Annual Review 54, 1–72. https://doi.org/10.1201/9781315368597

Favali, P., Beranzoli, L., 2006. Seafloor observatory science: A review. Annals of Geophysics. https://doi.org/10.4401/ag-3125

Fortson, L., Masters, K., Nichol, R., Borne, K. D., Edmondson, E. M., Lintott, C., Raddick, J., Schawinski, K., & Wallin, J. (2012). Galaxy Zoo: Morphological Classification and Citizen Science. In Advances in Machine Learning and Data Mining for Astronomy (pp. 213-236). CRC Press.

Franke, A., Blenckner, T., Duarte, C.M., Ott, K., Fleming, L.E., Antia, A., Reusch, T.B.H., Bertram, C., Hein, J., Kronfeld-Goharani, U., Dierking, J., Kuhn, A., Sato, C., van Doorn, E., Wall, M., Schartau, M., Karez, R., Crowder, L., Keller, D., Engel, A., Hentschel, U., Prigge, E., 2020. Operationalizing Ocean Health: Toward Integrated Research on Ocean Health and Recovery to Achieve Ocean Sustainability. One Earth 2(6), 557–565. https://doi.org/10.1016/j.oneear.2020.05.013

Garcia-Soto, C., van der Meeren, G., Busch, J., Delany, J., Domegan, C., Dubsky, K., Fauville, G., Gorsky, G., K., von, Malfatti, F., Mannaerts, G., Mchugh, P., Monestiez, P., J., S., Weslawski, J., Zielinski, O., 2017. Advancing Citizen Science for Coastal and Ocean Research. Position Paper 23 of the European Marine Board, Ostend, Belgium, V., Kellett, P., Delany, J., McDonough, N. [Eds.], ISBN: 978-94-92043-30-6 112pp.

Girard, F., Sarrazin, J., Arnaubec, A., Cannat, M., Sarradin, P., Wheeler, B., Matabos, M., 2020. Currents and topography drive assemblage distribution on an active hydrothermal edifice. Prog Oceanogr 187, 102397. https://doi.org/10.1016/j.pocean.2020.102397

Golumbic, Y.N., Baram-Tsabari, A., Koichu, B., 2020. Engagement and Communication Features of Scientifically Successful Citizen Science Projects. Environ Commun 14, 465–480. https://doi.org/10.1080/17524032.2019.1687101

Grol, M.G.G., Vercelloni, J., Kenyon, T.M., Bayraktarov, E., Van Den Berg, C.P., Harris, D., Loder, J.A., Mihaljević, M., Rowland, P.I., Roelfsema, C.M., 2020. Conservation value of a subtropical reef in south-eastern Queensland, Australia, highlighted by citizen-science efforts. Mar Freshw Res 72(1), 1–13. https://doi.org/10.1071/MF19170

Han, F., Yao, J., Zhu, H., Wang, C., 2020. Underwater Image Processing and Object Detection Based on Deep CNN Method. J Sens 2020. https://doi.org/10.1155/2020/6707328

Holt, B.G., Rioja-Nieto, R., Aaron Macneil, M., Lupton, J., Rahbek, C., 2013. Comparing diversity data collected using a protocol designed for volunteers with results from a professional alternative. Methods Ecol Evol 4(4), 383–392. https://doi.org/10.1111/2041-210X.12031

Juniper, S.K., Escartín, J., Cannat, M., 2007. Monitoring and Observatories: Multidisciplinary, Time-Series Observations at Mid-Ocean Ridges. Oceanography 20(1), 128–137.

Kane, I.A., Clare, M.A., 2019. Dispersion, accumulation, and the ultimate fate of microplastics in deep-marine environments: A review and future directions. Front Earth Sci (Lausanne) 7, 80. https://doi.org/10.3389/feart.2019.00080

Kosmala, M., Wiggins, A., Swanson, A., Simmons, B., 2016. Assessing data quality in citizen science. Front Ecol Environ 14(10), 551–560. https://doi.org/10.1002/fee.1436

Kuminski, E., George, J., Wallin, J., Shamir, L., 2014. Combining Human and Machine Learning for Morphological Analysis of Galaxy Images. Publications of the Astronomical Society of the Pacific 126(944), 959–967. https://doi.org/10.1086/678977

Langenkämper, D., Simon-Lledó, E., Hosking, B., Jones, D.O.B., Nattkemper, T.W., 2019. On the impact of Citizen Science-derived data quality on deep learning based classification in marine images. PLoS One 14(6), e0218086. https://doi.org/10.1371/journal.pone.0218086

Lelièvre, Y., Legendre, P., Matabos, M., Mihaly, S., Lee, R.W., Sarradin, P., Arango, C.P., Sarrazin, J., 2017. Astronomical and atmospheric impacts on deep-sea hydrothermal vent invertebrates. Proceedings of the Royal Society of London B: Biological Sciences 284(1852), 20162123.

Levin, L.A., Le Bris, N., 2015. The deep ocean under climate change. Science 350 (6262), 766-768. https://doi.org/10.1126/science.aad0126

Lintott, C.J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M.J., Nichol, R.C., Szalay, A., Andreescu, D., Murray, P., Vandenberg, J., 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. Mon Not R Astron Soc 389(3), 1179–1189. https://doi.org/10.1111/j.1365-2966.2008.13689.

Macedo, L., Oliveira, C. De, Lim, A., Conti, L.A., Wheeler, A.J., 2022. High-resolution 3D mapping of cold-water coral reefs using machine learning. Front Environ Sci 10, 1–20. https://doi.org/10.3389/fenvs.2022.1044706

Marcon, Y., Sahling, H., Allais, A.G., Bohrmann, G., Olu, K., 2014. Distribution and temporal variation of mega-fauna at the Regab pockmark (Northern Congo Fan), based on a comparison of videomosaics and geographic information systems analyses. Marine Ecology 35(1), 77–95. https://doi.org/10.1111/maec.12056

Mat, A.M., Sarrazin, J., Markov, G. V., Apremont, V., Dubreuil, C., Eché, C., Fabioux, C., Klopp, C., Sarradin, P., Tanguy, A., Huvet, A., Matabos, M., 2020. Biological rhythms in the deep-sea hydrothermal mussel Bathymodiolus azoricus. Nat Commun 11(1°, 1–12. https://doi.org/10.1038/s41467-020-17284-4

Matabos, M., Barreyre, T., Juniper, S.K., Cannat, M., Kelley, D., Alfaro-Lucas, J.M., Chavagnac, V., Colaço, A., Escartin, J., Escobar, E., Fornari, D., Hasenclever, J., Huber, J.A., Laës-Huon, A., Lantéri, N., Levin, L.A., Mihaly, S., Mittelstaedt, E., Pradillon, F., Sarradin, P., Sarrazin, J., Tomasi, B., Venkatesan, R., Vic, C., IMOVE InterRidge working group, 2022. Integrating Multidisciplinary Observations in Vent Environments (IMOVE): Decadal Progress in Deep-Sea Observatories at Hydrothermal Vents. Front Mar Sci 9, 1–32. https://doi.org/10.3389/fmars.2022.866422

Matabos, M., Best, M., Blandin, J., Hoeberechts, M., Juniper, S.K., Pirenne, B., Robert, K., Ruhl, H.A., Sarrazin, J., Vardaro, M., 2016. Seafloor Observatories, in: Clark, M.R., Consalvey, M., Rowden, A.A. (Eds.), Biological Sampling in the Deep Sea. Wiley, pp. 306–337. https://doi.org/10.1002/9781118332535.ch14

Matabos, M., Cuvelier, D., Brouard, J., Shillito, B., Ravaux, J., Zbinden, M., Barthelemy, D., Sarradin, P., Sarrazin, J., 2015. Behavioural study of two Hydrothermal crustacean decapods: Mirocaris fortunata and Segonzacia mesatlantica, from the lucky strike vent field (mid-Atlantic ridge). Deep Sea Research Part II: Topical Studies in Oceanography 121, 146–158. https://doi.org/http://dx.doi.org/10.1016/j.dsr2.2015.04.008

Matabos, M., Hoeberechts, M., Doya, C., Aguzzi, J., Nephin, J., Reimchen, T.E., Leaver, S., Marx, R.M., Branzan Albu, A., Fier, R., Fernandez-Arcaya, U., Juniper, S.K., 2017. Expert, Crowd, Students or Algorithm: who holds the key to deep-sea imagery 'big data' processing? Methods Ecol Evol 8(8), 996–1004. https://doi.org/10.1111/2041-210X.12746

McEver, R.A., Zhang, B., Levenson, C., Iftekhar, A.S.M., Manjunath, B.S., 2023. Context-driven detection of invertebrate species in deep-sea video. International Journal of Computer Vision 131(6), 1367-1388.

Miller, K.A., Thompson, K.F., Johnston, P., Santillo, D., 2018. An overview of seabed mining including the current state of development, environmental impacts, and knowledge gaps. Front Mar Sci 4. https://doi.org/10.3389/fmars.2017.00418

Morato, T., Watson, R., Pitcher, T.J., Pauly, D., 2006. Fishing down the deep. Fish and Fisheries 7(1), 24–34. https://doi.org/10.1111/j.1467-2979.2006.00205.x

Piechaud, N., Hunt, C., Culverhouse, P.F., Foster, N.L., Howell, K.L., 2019. Automated identification of benthic epifauna with computer vision. Mar Ecol Prog Ser 615, 15–30. https://doi.org/10.3354/meps12925

Porter, J.H., Nagy, E., Kratz, T.K., Hanson, P., Collins, S.L., Arzberger, P., 2009. New Eyes on the World: Advanced Sensors for Ecology. Bioscience 59(5), 385–397. https://doi.org/10.1525/bio.2009.59.5.6

Purser, A., Bergmann, M., Lundälv, T., Ontrup, J., Nattkemper, T.W., 2009. Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study. Mar Ecol Prog Ser 397, 241-251. https://doi.org/10.3354/meps08154

Ra, H., Richards, B.L., Rollo, A., Miller-Greene, D., Taylor, J., 2022. Keeping Track of Hawaii's Bottomfish Populations With the Help of Citizen Scientists. Fisheries 47(11), 510–515. https://doi.org/10.1002/fsh.10812

Raddick, J., Bracey, G., Gay, P., Lintott, C., Cardamone, C., Murray, P., Schawinski, K., Szalay, A., Vandenberg, J., 2013. Galaxy Zoo: Motivations of Citizen Scientists. Astronomy Education Review 12, 106. https://doi.org/10.3847/AER2011021

Raddick, M.J., Bracey, G., Gay, P.L., Lintott, C.J., Murray, P., Schawinski, K., Szalay, A.S., Vandenberg, J., 2010. Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. Astronomy Education Review 9. https://doi.org/10.3847/aer2009036

Robert, K., Juniper, S.K., 2012. Surface-sediment bioturbation quantified with cameras on the NEPTUNE Canada cabled observatory. Mar Ecol Prog Ser 453, 137–149. https://doi.org/10.3354/meps09623

Roberts, J.M., Devey, C.W., Biastoch, A., Carreiro-Silva, M., Dohna, T., Dorschel, B., Gunn, V., Huvenne, V.A.I., Johnson, D., Jollivet, D., Kenchington, E., Larkin, K., Matabos, M., Morato, T., Naumann, M.S., Orejas, C., Perez, J.A.A., Ragnarsson, S.Á., Smit, A.J., Sweetman, A., Unger, S., Boteler, B., Henry, L.-A., 2023. A blueprint for integrating scientific approaches and international communities to assess basin-wide ocean ecosystem status. Commun Earth Environ 4(1), 12. https://doi.org/10.1038/s43247-022-00645-w

Robinson, K.L., Luo, J.Y., Sponaugle, S., Guigand, C., Cowen, R.K., 2017. A tale of two crowds: Public engagement in plankton classification. Front Mar Sci 4, 82. https://doi.org/10.3389/fmars.2017.00082

Rüfenacht, S., Woods, T., Agnello, G., Gold, M., Hummer, P., Land-Zandstra, A., Sieber, A., 2021. Communication and dissemination in citizen science, in: Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perello, J., Ponti, M., Samson, R., Wagenknecht, K. (Eds.), The Science of Citizen Science. Springer Nature Switzerland AG, pp. 475–494.

Sarrazin, J., Blandin, J., Delauney, L., Dentrecolas, S., Dorval, P., Dupont, J., Legrand, J., Leroux, D., Leon, P., Lévèque, J.P., 2007. TEMPO: a new ecological module for studying deep-sea community dynamics at hydrothermal vents, in: OCEANS 2007-Europe. Ieee, pp. 1–4.

Sarrazin, J., Cuvelier, D., Peton, L., Legendre, P., Sarradin, P., 2014. High-resolution dynamics of a deep-sea hydrothermal mussel assemblage monitored by the EMSO-Açores MoMAR observatory. Deep Sea Research Part I: Oceanographic Research Papers 90, 62–75. https://doi.org/10.1016/j.dsr.2014.04.004

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9(7), 671-675.

Schoening, T., Bergmann, M., Ontrup, J., Taylor, J., Dannheim, J., Gutt, J., Purser, A., Nattkemper, T.W., 2012. Semi-automated image analysis for the assessment of megafaunal densities at the Arctic deep-sea observatory HAUSGARTEN. PLoS One 7(6), e38179.

Schoening, T., Köser, K., Greinert, J., 2018. An acquisition, curation and management workflow for sustainable, terabyte-scale marine image analysis. Sci Data 5(1), 1–12. https://doi.org/10.1038/sdata.2018.181

Silvertown, J., 2009. A new dawn for citizen science. Trends Ecol Evol 24(9), 467-471. https://doi.org/10.1016/j.tree.2009.03.017

Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., Kelling, S., 2009. eBird: A citizen-based bird observation network in the biological sciences. Biol Conserv 142(10), 2282–2292. https://doi.org/10.1016/j.biocon.2009.05.006

Swanson, A., Kosmala, M., Lintott, C., Packer, C., 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. Conservation Biology 30(3), 520–531. https://doi.org/10.1111/cobi.12695

Sweetman, A.K., Thurber, A.R., Smith, C.R., Levin, L.A., Mora, C., Wei, C.-L., Gooday, A.J., Jones, D.O.B., Rex, M., Yasuhara, M., Ingels, J., Ruhl, H.A., Frieder, C.A., Danovaro, R., Würzberg, L., Baco, A., Grupe, B.M., Pasulka, A., Meyer, K.S., Dunlop, K.M., Henry, L.-A., Roberts, J.M., 2017. Major impacts of climate change on deep-sea benthic ecosystems. Elem Sci Anth 5, 4. https://doi.org/10.1525/elementa.203

Van Audenhaege, L., Broad, E., Hendry, K.R., Huvenne, V.A.I., 2021. High-Resolution Vertical Habitat Mapping of a Deep-Sea Cliff Offshore Greenland. Front Mar Sci 8, 1–18. https://doi.org/10.3389/fmars.2021.669372

Van Audenhaege, L., Matabos, M., Brind'Amour, A., Drugmand, J., Laës-Huon, A., Sarradin, P., Sarrazin, J., 2022. Long-term monitoring reveals unprecedented stability of a vent mussel assemblage on the Mid-Atlantic Ridge. Prog Oceanogr 204. https://doi.org/10.1016/j.pocean.2022.102791

Van den Bergh, J., Chirayath, V., Li, A., Torres-Pérez, J.L., Segal-Rozenhaimer, M., 2021. NeMO-Net – Gamifying 3D Labeling of Multi-Modal Reference Datasets to Support Automated Marine Habitat Mapping. Front Mar Sci 8, 1–17. https://doi.org/10.3389/fmars.2021.645408

Van Dover, C.L., 2014. Impacts of anthropogenic disturbances at deep-sea hydrothermal vent ecosystems: A review. Mar Environ Res 102, 59-72. https://doi.org/10.1016/j.marenvres.2014.03.008

Van Dover, C.L., Colaço, A., Collins, P.C., Croot, P., Metaxas, A., Murton, B.J., Swaddling, A., Boschen-Rose, R.E., Carlsson, J., Cuyvers, L., Fukushima, T., Gartman, A., Kennedy, R., Kriete, C., Mestre, N.C., Molodtsova, T., Myhrvold, A., Pelleter, E., Popoola, S.O., Qian, P.Y., Sarrazin, J., Sharma, R., Suh, Y.J., Sylvan, J.B., Tao, C., Tomczak, M., Vermilye, J., 2020. Research is needed to inform environmental management of hydrothermally inactive and extinct polymetallic sulfide (PMS) deposits. Mar Policy 121, 104183. https://doi.org/10.1016/j.marpol.2020.104183

Villon, S., Mouillot, D., Chaumont, M., Darling, E.S., Subsol, G., Claverie, T., Villéger, S., 2018. A Deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecol Inform 48, 238–244. https://doi.org/10.1016/j.ecoinf.2018.09.007

Wang, Y.F., Hsu, Y.F., Fang, K., 2022. The key elements of gamification in corporate training – The Delphi method. Entertain Comput 40, 100463. https://doi.org/10.1016/j.entcom.2021.100463

Wick, M.J., Angradi, T.R., Pawlowski, M.B., Bolgrien, D., Debbout, R., Launspach, J., Nord, M., 2020. Deep Lake Explorer: A web application for crowdsourcing the classification of benthic underwater video from the Laurentian Great Lakes. J Great Lakes Res 46(5), 1469–1478. https://doi.org/https://doi.org/10.1016/j.jglr.2020.07.009

Wiggins, A., Newman, G., Stevenson, R.D., Crowston, K., 2011. Mechanisms for data quality and validation in citizen science, in: Proceedings - 7th IEEE International Conference on e-Science Workshops, EScienceW 2011. pp. 14–19. https://doi.org/10.1109/eScienceW.2011.27

Winickoff, D.E., Jamal, L., Anderson, N.R., 2016. New modes of engagement for big data research. J Responsible Innov 3(2), 169–177. https://doi.org/10.1080/23299460.2016.1190443

Zweifler, A., Akkaynak, D., Mass, T., Treibitz, T., 2017. In situ analysis of coral recruits using fluorescence imaging. Front Mar Sci 4, 1–12. https://doi.org/10.3389/fmars.2017.00273

Figures legends

Figure 1. Data acquisition and images. A. Location of the two ecological observatory modules on the mid-Atlantic ridge (MAR) and Juan de Fuca ridge (JdFR). B. TEMPO-mini, equipped with the SMOOVE camera, monitors siboglinid tubeworms at the Main Endeavour Field of the JdFR. C & D. Field of view filmed by SMOOVE cameras on the JdFR (C) and MAR (D). The white line on C delineates the "background" area (see text).

Figure 2. Tutorials providing training for the annotation system (top panel) and species recognition (bottom panel) in the Deep Sea Spy online image annotation tool.

Figure 3. Number of active participants (i.e., who annotated at least an image) per day over time during the first mission. Red rectangles highlight major peaks of participation resulting from specific communication and outreach events (see text).

Figure 4. Cumulative proportion of annotations (blue) and annotated images (red) with increasing number of participants (sorted from least to most active).

Figure 5. Procedure to merge multiple annotations in one single occurrence in a photo. A. Original photo showing the distribution of 14 real buccinid gastropods. (B, C). Procedure to delineate polygons for each contributor. B. Representation of all annotations performed by ten contributors with overlap between annotations of the same individual. Each colour represents a different participant. C. Groups of polygons that correspond to a single buccinid occurrence after merging all contributor's annotation with the deeptools package. Each colour represents a unique organism. In total 14 groups (i.e. unique buccinid) were detected by the participants. (D, E). Procedure to define polygons for a given contributor. D. Voronoï polygons delineated around each annotation. E. Voronoi polygons cropped using a buffer in pixel around the annotation.

Figure 6. Cumulative number of unique individuals detected depending on the citizen agreement rate threshold for the snail Buccinum thermophilum (blue, JdFR) and the crab Segonzacia mesatlantica (orange, MAR), with expert counts (dashed lines).

Figure 7. Difference between expert and participants in abundances of the snail Buccinum thermophilum (left) and the crab Sgonzacia mesatlantica (right) as function of the agreement rate threshold from each annotated image. Note that for crabs, one outlier corresponding to 60 observed organisms considering an AR = 0.1 was removed to ease the readability. Averages of

delta distributions per agreement rate were compared to zero with Student t-tests (***: p-value < 0.001 ; **: p-value < 0.01 ; *: p-value < 0.05 ; ·: p-value < 0.1). n: number of images.

Figure 8. Abundances of the snail Buccinum thermophilum (left) and the crab Segonzacia mesatlantica (right) across all images as function of the Agreement Rate threshold. Average abundances resulting from the annotation of the ten participants (red) were compared with averages after sub-sampling with three (green), four (blue) and five participants (purple) per photo and with expert counts (dashed).

Figure 9. Evolution of the snail Buccinum thermophilum abundances depending on participants (blue) agreement rate threshold compared with expert (yellow). Pearson correlation coefficients (ρ) between expert and citizen data are shown on top of the graphs and are all significant (p value < 0.001).

Figure 10. Example of outreach events related to the Deep Sea Spy (DSS) project. A. Computer terminal with the DSS application at the Océanopolis aquarium in Brest (photo credit N. Roullet). B. Demonstration to kids at the science festival in Brest, France (photo credit Ifremer). C. School kids with the DSS educational booklet (photo credit A. Bianic, école Sainte Anne, Saint-Thonan, France). D. Video conference from the ship to introduce the ship crew to kids on land (photo credit F. Le-Moigne, école Mouez Ar Mor, Ploumoguer, France).

Supplementary Material 1. Deep Sea Spy data model showing the logical structure of the PostgreSQL database.

Supplementary Material 2. Dictionnary of terms used in the Deep Sea Spy PostgreSQL database.