Ref. Ares(2017)5346144 - 02/11/2017

ENVRI^plus DELIVERABLE

D14.2
Preliminary results from the online annotation tool Deep Sea Spy

WORK PACKAGE 14 - Citizen Observatories and Participative Science

LEADING BENEFICIARY: IFREMER

Author(s):	Beneficiary/Institution
Marjolaine MATABOS	IFREMER
Catherine BORREMANS	IFREMER
Julie TOUROLLE	IFREMER

Accepted by: Mairi Best (WP 14 leader)

Deliverable type: [REPORT]

Dissemination level: PUBLIC

Deliverable due date: 31.10.2017

Actual Date of Submission: 31.10.2017

A document of ENVRI^plus project - www.envri.eu/envriplus

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

ABSTRACT

Ifremer in association with the SEM Noveltis (Labège, France) developed a web-based application linked with a structured database. The software is built as a game with dedicated missions. The goal of each mission is to annotate a series of images extracted from archived video sequences coming from our deep-sea observatories. The annotation tool has been developed and was officially launched in March 2017 through French media and international scientists network (Interridge, Deep-sea biology society). The first public mission attracted more than four hundred active participants and led to the annotation of more than fifteen thousand images. This document presents the preliminary results.

Project internal reviewer(s):

Project internal reviewer(s):	**Beneficiary/Institution**
Mairi Best	EMSO

Document history:

Date	**Version**
26.10.2017	Draft for comments
29.10.2017	Corrected version
31.10.2015	Accepted by Mairi Best

DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to Marjolaine Matabos _(Marjolaine.Matabos@ifremer.fr)

TERMINOLOGY

Acronyms used in this report are briefly explained in **Appendix A**. In addition, the project glossary is found at https://envriplus.manageprojects.com/s/text-documents/LFCMXHHCwS5hh.

PROJECT SUMMARY

ENVRIplus is a Horizon 2020 project bringing together Environmental and Earth System Research Infrastructures, projects and networks together with technical specialist partners to create a more coherent, interdisciplinary and interoperable cluster of Environmental Research Infrastructures across Europe. It is driven by three overarching goals: 1) promoting cross-fertilization between infrastructures, 2) implementing innovative concepts and devices across RIs, and 3) facilitating research and innovation in the field of environment for an increasing number of users outside the RIs.

ENVRIplus aligns its activities to a core strategic plan where sharing multi-disciplinary expertise will be most effective. The project aims to improve Earth observation monitoring systems and strategies, including actions to improve harmonization and innovation, and generate common solutions to many shared information technology and data related challenges. It also seeks to harmonize policies for access and provide strategies for knowledge transfer amongst RIs. ENVRIplus develops guidelines to enhance transdisciplinary use of data and data-products supported by applied use-cases involving RIs from different domains. The project coordinates actions to improve communication and cooperation, addressing Environmental RIs at all levels, from management to end-users, implementing RI-staff exchange programs, generating material for RI personnel, and proposing common strategic developments and actions for enhancing services to users and evaluating the socio-economic impacts.

ENVRIplus is expected to facilitate structuration and improve quality of services offered both within single RIs and at the pan-RI level. It promotes efficient and multi-disciplinary research offering new opportunities to users, new tools to RI managers and new communication strategies for environmental RI communities. The resulting solutions, services and other project outcomes are made available to all environmental RI initiatives, thus contributing to the development of a coherent European RI ecosystem.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

TABLE OF CONTENTS

A document of ENVRI^plus project - www.envri.eu/envriplus

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

Preliminary results from the online annotation tool Deep Sea Spy

I. INTRODUCTION

1) Scientific background

Most of the current knowledge of deep-sea environments is based on punctual, at best yearly, oceanographic cruises. Since 2006, deep-sea ecosystems are continuously being monitored using video cameras deployed on deep-sea platforms. The acquisition of high-frequency video data from deep-sea observatories like EMSO-Azores (www.emso-fr.org/fr/EMSO-Azores) or Ocean networks Canada/NEPTUNE Canada (NorthEast Pacific Time-series Underwater NEtwork) provide for the first time information on species behaviour, feeding habits, growth, possibly reproduction and organisms' response to changes in environmental conditions. Cameras deployed on those observatories acquire hourly video data representing thousands of hours and Tera Bytes of footage that require 10 times more hours of viewing to extract useful information. Since their first deployment in 2006, more than 5 Tb of video data from both the Atlantic and the Pacific oceans were acquired that cannot possibly be analyzed by a few researchers. Only with the help of citizen scientists will we be able to process the huge archive of imagery.

2) Objective of the global project

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 as well as raise awareness among the general public about deep-sea ecosystems, including the resources they provide and the state-of-the art scientific knowledge of these environments.

In order to meet this goal the specific objectives were to:

- Develop an online image annotation program that will allow participants to simultaneously perform defined tasks on the extracted images.
- Organise the output information in a searchable database compatible with existing Ifremer and EMSO databases following ENVRI + standards.
- Develop a communication plan to share the application and the knowledge

Ifremer, in association with the company Noveltis (Labège, France), developed the web-based application linked with a structured database. The software is built as a game with dedicated missions. The goal of each mission is to annotate a series of images extracted from archived video sequences acquired with deep-sea observatories. The information obtained by annotation is stored and exported in a structured database.

3) Report content

This report presents the first results acquired with the annotation tool. For a description of the tool itself, report to the deliverable 'D14.1 Prototype of a web-based annotation tool ready for user testing' (Matabos et al. 2016). The first test mission was available online from October 2016 to March 2017 to beta-testers, but only a few images were available. The first official mission of the application started in March 2017 and is still on-going.

II. GENERAL STATISTICS ON THE APPLICATION

From the admin webpage of the tool (Fig 1), the game administrators can have an overview of the main statistics of the application and ongoing mission including number of images annotated, number of users per week, month, over the mission and globally. While the application itself 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.

Note: All results presented below are those reported on October 10th, 2017

Statistiques générales

Exporter sous format Excel

Nombre de participants	Global	Mission courante
Total	418	359
Année en cours	401	359
Année dernière	28	0
Mois en cours	0	0
Mois dernier	34	34
Semaine en cours	0	0
Semaine dernière	17	17

Nombre d'images annotées	Global	Mission courante
Total	15022	12906
Année en cours	14120	12906
Année dernière	902	0
Mois en cours	0	0
Mois dernier	859	859
Semaine en cours	0	0
Semaine dernière	266	266

Nombre d'annotations	Global	Mission courante
Total	124639	121835
Crevette alvinocarididae	1929	1893
Pycnogonide	34880	34880
Autre	8	0
Autre poisson	756	219
Crabe araignée	20435	20435
Crabe bythograeidé	2476	1158
Escargot buccinidé	47779	47779
Ophiure	89	89
Poisson Cataetyx	978	307
Poisson chimère	266	64

FIG. 1. VIEW OF THE ADMINISTRATION STATISTIC PAGE (DEEPSEASPY.IFREMER.FR/ADMIN)

1) Users

Out of 848 registered participants only less than half (i.e. 418) have annotated at least one image. Several reasons can be put forward to explain this behaviour:

- People are just curious about the application and the project, but do not go all the way to annotate images because of lack of interest or time.

- Some might find going through the tutorial constraining and/or tedious
- Some might find the application and annotation too complicated

A way of gathering feedback would help better define the behaviours at play in this observation. However, chances are these same persons will not take on their time to fill in the survey.

The number of participants drops after 10 images annotated (Table 1, Fig. 2). It is interesting to note that 10 images corresponds to the threshold to take the first level, but only a few of the participants actually took the quiz it is necessary to pass a quiz to move on to the next level). It is possible that some people find the process too complicated and feel discouraged when they reach the first quiz. Another possibility is that the system of quiz generates pressure (similar to 'school exams'). This is supported by the fact that many participants annotated up to 15 pictures without ever taking the quiz, one participant annotated 25 images without ever passing the quiz.

TABLE 1. USERS' PARTICIPATION IN THE ANNOTATION TOOL (AS OF OCTOBER 10TH, 2017)

Number of registered users	848
Number of active participants (at least one image annotated)	418
Number of participants that annotated > 10 images	189
Number of participants that annotated > 100 images	29
Number of participants that annotated > 500 images	5
Number of participants that annotated > 1000 images	2
Number of participants that annotated > 2000 images	1

Only 29 persons annotated more than 100 images and only 2 more than a thousand. There is an inverse exponential relationship between the number of images annotated and the number of participants (Fig. 2). This result was expected considering how time-consuming, and somehow wearisome, the task is.

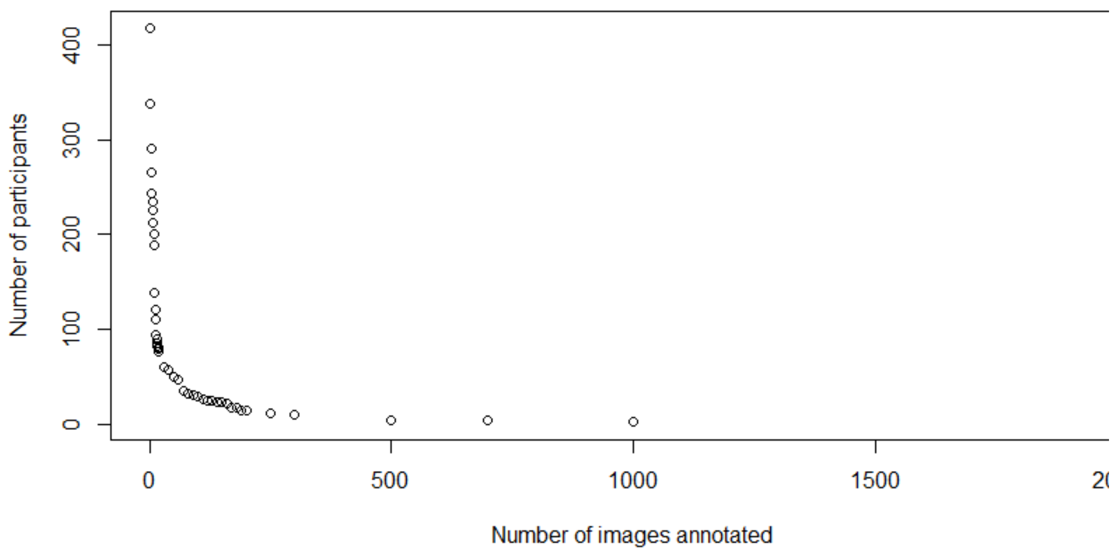


FIG. 2. NUMBER OF PARTICIPANTS BY NUMBER OF IMAGES ANNOTATED.

The number of participants increased steadily over time. The official media launch in March 2017 was covered by a large number of media in Brittany and, to a lesser extent, in France (see reference list of journal, radio and television broadcast at the end of the document) and led to a high number of registrations the two following weeks (Fig. 3). A renewed effort in communication at the beginning of the academic year helped increase the activity level on the website by re-engaging existing participants and fostering new registrations (Fig. 3). The communication effort in September was conducted through publishing of web news on the project website (www.deepseaspy.com) and the start of school projects involving high-school students (i.e. 42 registrations in 2 weeks), and through the participation of the team to science events (class, science fest in Paris) (i.e. 22 registrations in 3 days).

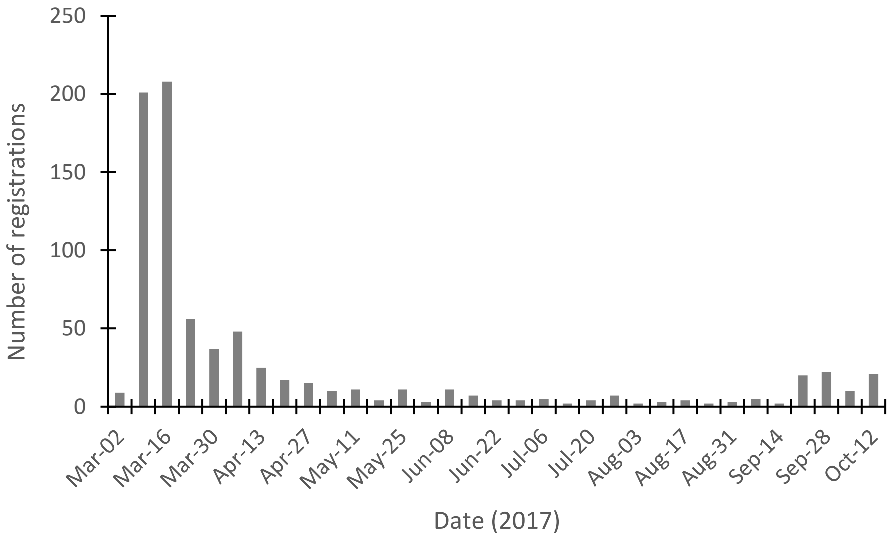


FIG. 3. NUMBER OF REGISTRATIONS SINCE THE OFFICIAL MEDIA LAUNCH ON MARCH 7TH, 2017.

2) Countries

The participants come from 19 different countries but more than 90% live in France. However, one can note that the participant who annotated the highest number of images (> 2000) is from Belgium. This distribution of participant results from the strongly biased communication effort towards France. It will be good in the future to improve communication at the international level, in particular through the ENVRIplusand the Deep-sea community.

TABLE 2. NUMBER OF PARTICIPANTS BY COUNTRY.

Country	Number of participants	Country	Number of participants
France	385	French Polynesia	1
Switzerland	7	Denmark	1
Belgium	6	New Caledonia	1
Canada	3	Japan	1
Ireland	2	Germany	1
United States	2	Mayotte	1
Reunion	2	Israel	1
United Kingdom	1	French Southern Territories	1
Portugal	1	Finland	1
Martinique	1		

3) *Images*

A total of 15,024 images were annotated including 2,216 in the test mission and 12,908 in the ongoing mission. At present, each image must be annotated ten times before it is discarded from the mission. Images are extracted from videos, there is thus no maximum number of pictures. A program developed at Ifremer allows for the extraction of pictures from videos. As such, the frequency of extraction (from seconds to hours) depends on the scientific question. If extracted every second, the number of images is virtually infinite considering we acquire 780 hours of video a year.

III. TEST MISSION: OCTOBER 2016 – MARCH 2017

The test mission '*citizens performance: how do you compare?*' only included 18 pictures and was mostly dedicated to test the application for bugs as well as collecting feedback from different types of participants including scientists, general public and students. The beta-testers included Ifremer employees, classes with kids from 9 to 17 years old, the ENVRIplus community and the general public during the science fest in Brest (October 2016). This phase helped us highlight bugs, and improve the ergonomics of the tool (in particular the zoom system). A new version of the application is now online since September including all bugs fixed and a new design.

IV. PUBLIC MISSION: SINCE MARCH 2017

A new mission including 4,000 images called *'Tides at 1700 m deep?'* was started for the official launch of the application. In this mission participants are asked to annotate 6 months of images from the Atlantic and Pacific in order to monitor the presence of animals at different times of the day. From scientific results that our team have recently highlighted on the influence of tides on the behaviour of vent species (Cuvelier *et al*. 2104, Lelièvre *et al*. 2017), we have reasons to suspect that more of them respond to the tides. Each image has to be annotated 10 times before it is discarded.

1) Progress

The mission reached 32.4% of advancement, corresponding to 12,908 images annotated out of 40,000. So far, each image has been annotated 3 or 4 times.

2) Types of participants

From the preliminary experience we gained the last months, participants can be classified in four types:

- **Classes**: we collaborated with teachers in classes with kids from 9 to 17 years old. Students participated as a class project. Their participation being mandatory, this type of collaboration ensures a minimum number of annotations. The advantage lies in the fact that every academic year sees the arrival of new students, thus mitigating the loss of participants because of boredom.
- **One-time participants**: these participants are part of the general public and probably heard of the application through the media, a friend or by visiting an exhibition. They represent 79% of all participants. Some of them create an account out of curiosity but never annotate an image, others will annotate one or more and never come back.
- **Interested participants**: they annotated 10 images or more, in several sessions. They are different from the above category because they came back to the game after a first try.
- **Involved participants**: they have annotated more than 100 images. These type of participants are rare considering images are monotonous and repetitive. In this mission, only 28 and 4 people annotated more than 500 and 1000 images respectively (in Table 1, one of the participants corresponds to our demonstration and testing account)

All of these categories must be highly considered in the process.

With classes, it is important to maintain long-term collaborations as they represent about 30 new participants each year. Long-term relationships can be insured by exchanging email and giving visio-conferences to the class.

One-time participants, although they seem to poorly contribute, are nevertheless important as they represent the bulk of the participants. With good communication this category can reach high numbers.

The two last categories are obviously important. If some suggest that important contributors should be encouraged with special consideration, communicating with participants is extremely time-consuming, unless a dedicated position is available for the project.

In addition, we record in the data base age groups and jobs. This data is not yet available and need more caution in interpretation as this piece of information is not mandatory when creating an account, but it would be interesting in the future to consider age and job areas within these different categories in communication strategies, and when improving and developing future versions of the application.

3) Scientific results

Note: because of some issues in the extraction from the database, not all data was analyzed and the data presented below only represent a sub-sample of all the data collected. A student in *computer science is currently working on developing a tool to inquire and export data from the application database.*

a. Juan de Fuca/ONC data

A total of 170 participants contributed to the data acquired for the results presented below. These results correspond to 2,251 images analyzed corresponding to 1,416 unique images from July 4th to 22nd 2014 (Table 3, Fig. 4).

TABLE 3. STATISTICS OF THE ONGOING MISSION DETAILED FOR EACH SPECIES.

Species	Min level to annotate	# images annotated	# unique image annotated	# of users	Total # of individuals counted
Buccinid	0	1082	378	161	4993
Polynoid worms	3	289	229	8	763
Pycnogonid	7	563	495	2	8102
Spider crab	0	272	200	72	1504
Zoarcid fish	0	229	133	73	280

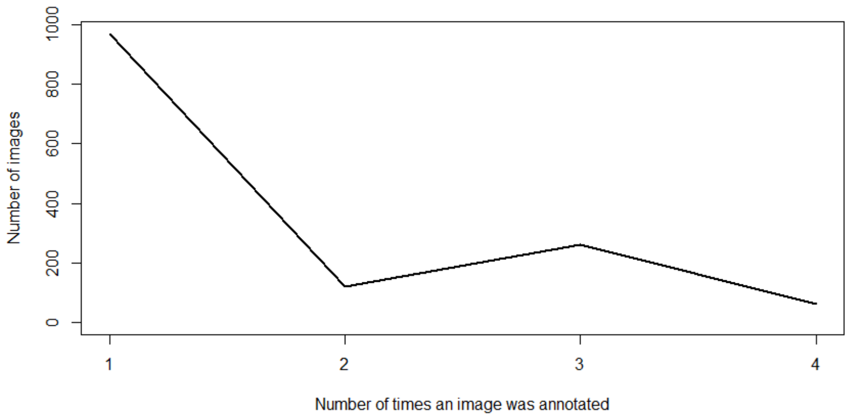
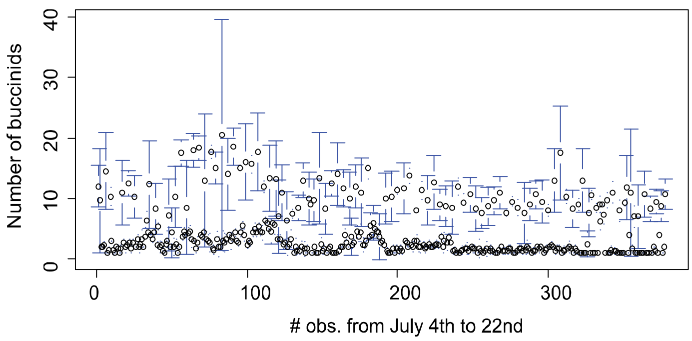


FIG. 4. NUMBER OF TIMES EACH SINGLE IMAGE WAS ANNOTATED IN THE SUB DATASET ANALYSED FOR THE JUAN DE FUCA (ONC) DATA IN THE CURRENT ONGOING MISSION (N = 1,416).

Due to some issues in the extraction as mentioned above, only data from July 4th to 22nd 2014 seemed reliable for single species. We provided below details for each species but the spider crab. Indeed, the spider crab is a rare species and we suspect some participants, despite the tutorial, confused them with the pycnogonids. We would expect over that period one or two individuals of that non-endemic species. This gives us insights on improvement to make for future versions of the application.

❖ Evolution of buccinids

The graphs below show the evolution of buccinid abundances from July 4th to July 22nd 2014 (Figs 5 and 6). Over July 2014, 1,082 images have been annotated, corresponding to 378 unique images. Altogether 4,993 buccinids were counted. Each image was analysed in average 3 times (Fig 5, right panel). We can note that when abundances are higher, errors bars are larger. This pattern was previously observed in a study that analysed how the crowd compared to an expert in counting sablefish from video data (Matabos et al. 2017) . The higher the abundance, the higher the error, highlighting the importance of having a high number of persons annotating the same image. This study showed that if the number of annotations is high enough, the median appears to be a good proxy of the actual number of individuals. This number is dependent on the type of image (more or less complex habitat, i.e. dark pelagic vs hydrothermal vents) and the complexity of the task.



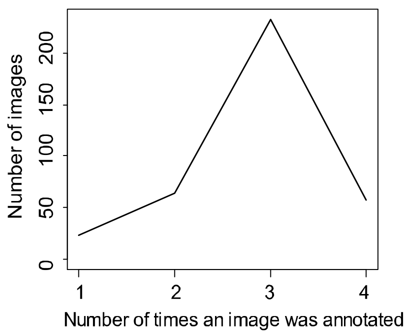


FIG. 5. LEFT PANEL. EVOLUTION OF THE BUCCINID ABUNDANCE SHOWING THE MEAN (BLACK DOT) AND STANDARD DEVIATION (BLUE LINES) AROUND THE MEAN. RIGHT PANEL. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANALYSED.

Mean and median were quite different on certain images (Fig. 6), this observation is often related to one outlier in the data because the participant may have simply inadvertently hit the wrong key or knowingly entered biased results (Matabos *et al*. 2017). These preliminary results thus confirmed that median, by mitigating the influence of outliers, is a better approach to summarize citizen science data.

Periodogram analyses revealed significant periods at 8h, 16h, 32h, 48h, 64h, 80h, and 96h (Fig. 6. bottom panel). A periodicity of 16h (corresponding to inertial currents) and its harmonics have already been highlighted in the variations of abundance of the buccinid *Buccinmum thermophilum* from that same site (Cuvelier *et al*. 2014; Lelièvre *et al.* 2017). These results are thus in accordance with what has already been observed and confirms previous studies. Indeed, this time-series is the first to include so many individuals of buccinids allowing for more robust statistical analyses. These results reveal the potential of the application and citizen science to help in the processing of big imagery datasets.

A document of ENVRIplus project - www.envri.eu/envriplus

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

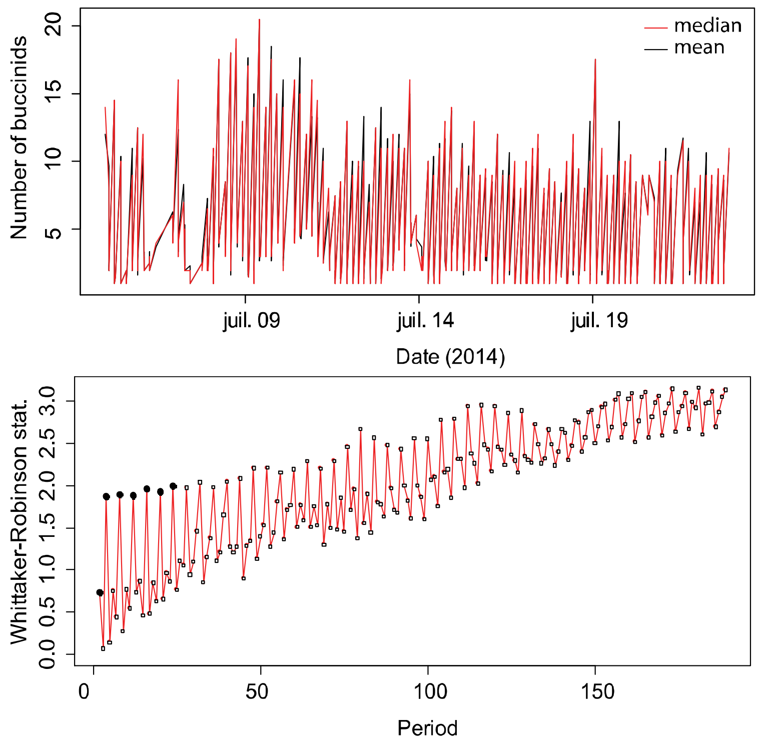


FIG 6.

TOP PANEL. EVOLUTION OF THE NUMBER OF BUCCINIDS OVER THE STUDY PERIOD. MEAN AND MEDIAN WERE CALCULATED FOR EACH IMAGE.

BOTTOM PANEL. PERIODOGRAM COMPUTED ON THE BUCCINID ABUNDANCE DATA. THE BLACK ROUND DOTS REPRESENT SIGNIFICANT PERIODICITIES.

While the data hasn't been fully processed yet, all coordinates of individuals in pixels are available. This will allow us to generate maps to study the spatial distribution of species in the field of view. Below an example of maps that can be generated.

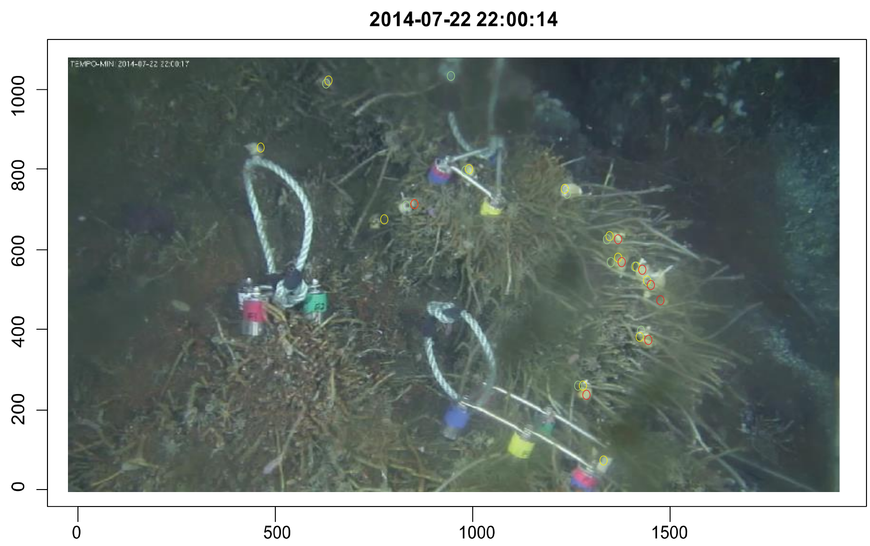
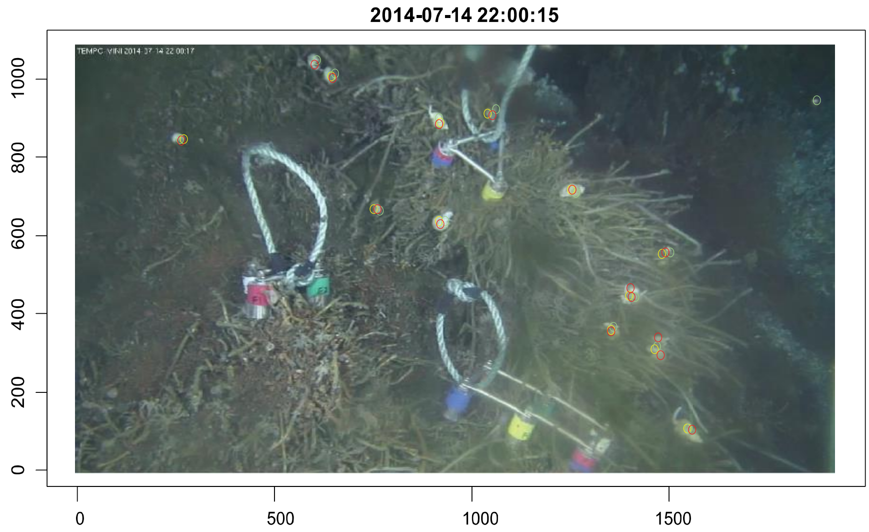
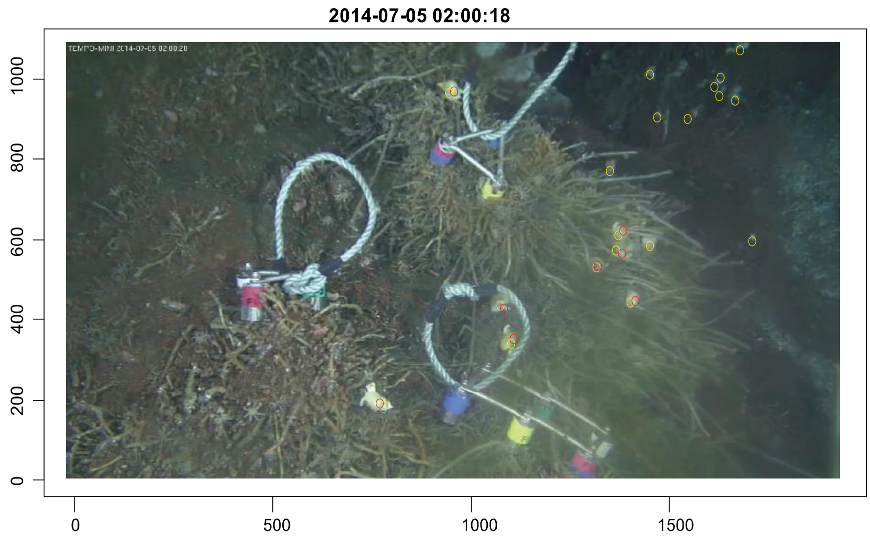


FIG. 7.

MAPS OF THE DISTRIBUTION OF BUCCINIDS IN THE FIELD OF VIEW. THE THREE DIFFERENT COLORS (GREEN, YELLOW, RED) CORRESPOND TO THREE DIFFERENT PARTICIPANTS/ANNOTATORS. EXAMPLE OF MAPS DONE FOR THREE DIFFERENT DATES AND TIMES.





These maps show concordance among observers for obvious individuals present in the front of the field of view. Some users though appear more enthusiastic and might over annotate in presence of a resembling shape/colour. On the bottom map, we can see that the 'yellow' participant found a number of buccinids in the back of the field of view, unnoticed/ignored by the two others.

❖ Evolution of polynoids

The graphs below show the evolution of the polynoid abundance from July 4th to 22nd 2014 (Fig 8). Over the month of July 2014, 289 images have been annotated, corresponding to 229 unique images. Altogether 763 polynoids were counted. Annotating polynoids is only available starting at level 3, resulting in a lower number of annotated images. In consequence most of the images annotated for polynoids were only analysed once (Fig 8C). Results of the periodogram computed on the polynoid abundances did not show any periodicities (graph not shown), while a previous study highlighted a periodicity of 12h, corresponding to the tidal signal, and of 16h, corresponding to inertial currents (Lelièvre et al. 2017). More data is needed to compare data among participants, and with previous knowledge.

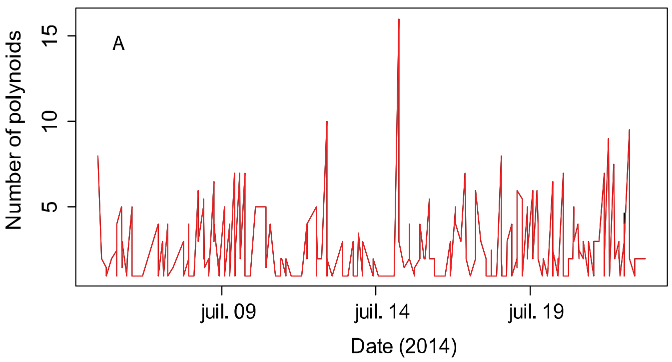
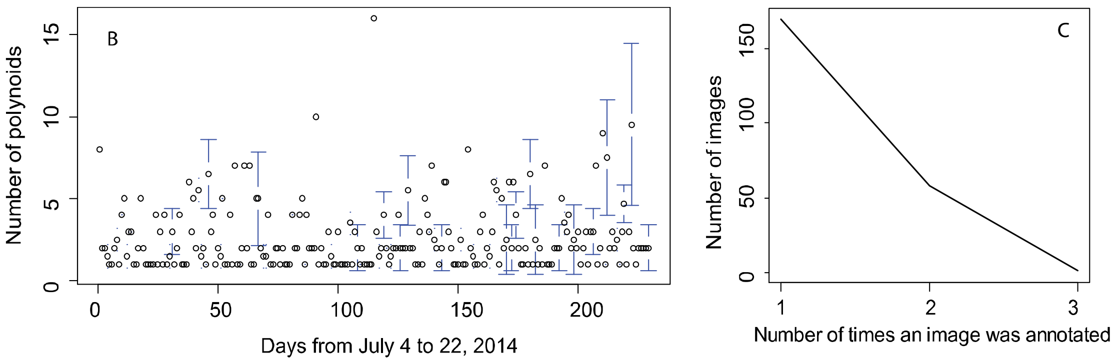


FIG. 8. EVOLUTION OF THE ABUNDANCE OF POLYNOIDS FROM JULY 4TH TO 22ND, 2014. A. EVOLUTION OF THE MEAN (BLACK) AND MEDIAN (RED) OF POLYNOID ABUNDANCE. B. EVOLUTION OF THE MEAN (BLACK DOTS) AND THE STANDARD DEVIATION (BLUE BARS). C. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANALYSED.



❖ Evolution of pycnogonids

Similarly to polynoids, the participant needs to reach a certain level (i.e. 7) to be able to annotate pycnogonids. The level has only been reached by 3 people which is currently not sufficient to insure reliable results on pycnogonid abundance and mapping. We nevertheless present results obtained so far from these participants. Between July and October 2014, 563 images have been annotated, corresponding to 495 unique images. Altogether 8,102 pycnogonids were counted. Over the 495 images annotated, only 65 were annotated twice and 2 annotated three times.

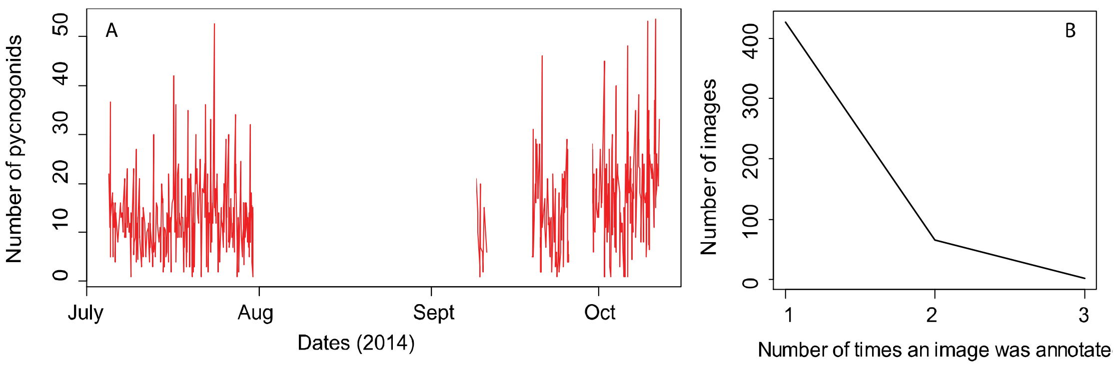
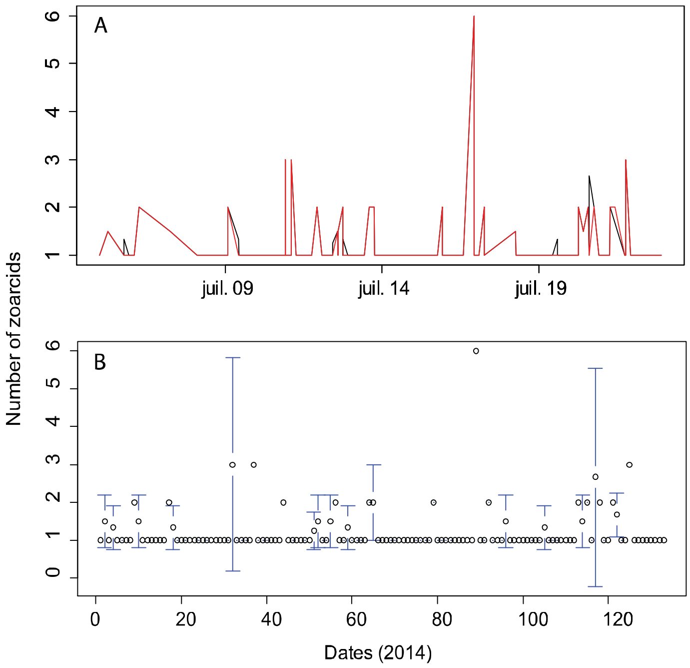


FIG. 9. EVOLUTION OF THE ABUNDANCE OF PYCNOGONIDS IN 2014. A. EVOLUTION OF THE MEDIAN. B. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANALYSED.

Again, while a previous study showed the occurrence of a periodicity of 12h, corresponding to the tidal signal, in the variations of pycnogonids abundances (Lelièvre *et al.* 2017), the periodogram computed on this dataset (only for July 2014) did not reveal any significant periodicity.

❖ Evolution of zoarcid fish abundances

The graphs below show the evolution of zoarcid abundances from July 4th to July 22nd 2014 (Fig. 10). Over July 2014, 229 images have been annotated, corresponding to 133 unique images. Altogether 280 zoarcids were counted.



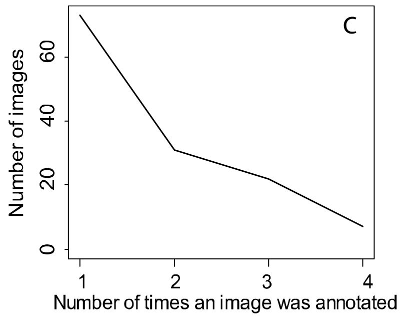


FIG. 10. EVOLUTION OF THE ABUNDANCE OF ZOARCIDS FROM JULY 4TH TO 22ND, 2014. A. EVOLUTION OF THE MEAN (BLACK) AND MEDIAN (RED) OF ZOARCID ABUNDANCE. B. EVOLUTION OF THE MEAN (BLACK DOTS) AND THE STANDARD DEVIATION (BLUE BARS). C. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANALYSED.

Even though, zoarcids seem quite obvious in the image, we observed a big standard deviation when there are more than 3 zoarcids. However, more data is needed to make conclusions.

b. MAR data (Mid-Atlantic ridge)

A total of 151 participants contributed to the data acquired for the results presented below. These results correspond to 1,136 images analysed, representing 528 unique images (91% of the total number of images from the Atlantic in the mission) of the entire study period from July 27th 2014 to January 31st 2015 (Table 4). Most of the images were annotated 1 to 3 times (Fig. 11).

TABLE 4. STATISTICS OF THE ONGOING MISSION DETAILED FOR EACH SPECIES FROM THE MID-ATLANTIC RIDGE (EMSO-AZORES DATA).

Species	Min level to annotate	# images annotated	# unique image annotated	# of users	Total # of individuals counted
Bythogreid crab	0	757	383	124	1151
Cataetyx fish	0	169	147	69	304
Chimera fish	0	46	43	25	61
Other fish	0	164	139	55	216
Polynoid worm	2	274	241	18	426
Alvinocarid shrimp	3	364	305	14	1893
Ophiuorid	5	64	60	3	89

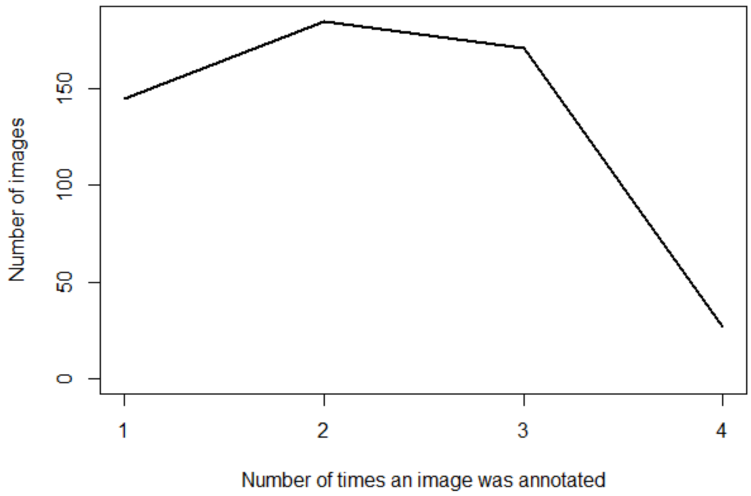


FIG 11. NUMBER OF TIMES EACH SINGLE IMAGE WAS ANNOTATED IN THE SUB DATASET ANALYSED FOR THE MID-ATLANTIC RIDGE (EMSO-AZORES) DATA IN THE CURRENT ONGOING MISSION (N = 528).

Not all species are detailed below. For the species not displayed in this document (i.e. other fish, polynoid worm, ophiuroid, and Chimera fish), there is only one replicate annotation for most images, making it difficult to draw conclusions on consensus across participants. However, one thing that was commonly observed for each species is the occurrence of absurd data. It would be interesting to compare 'User ids' among these absurd data to be able to identify if there are all linked to a reduced number of observers, in which case, the given observers should be disregarded for future analyses. A first analysis of absurd data revealed that they are not attributed consistently to the same observer. We observed two types of absurd data: i) one-time participant who misannotated on purpose; ii) regular serious participants who were a bit enthusiastic and will believe simply made mistakes.

- Evolution of crabs

The graphs below show the evolution of crab abundances from July, 27, 2014 to January, 31, 2015 (Figs 12). Over the time study, 757 images have been annotated, corresponding to 383 unique images. Altogether 1,151 crabs were counted. Each image was analysed on average 1 to 3 times (Fig 12, right panel).

A document of ENVRIplus project - www.envri.eu/envriplus

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

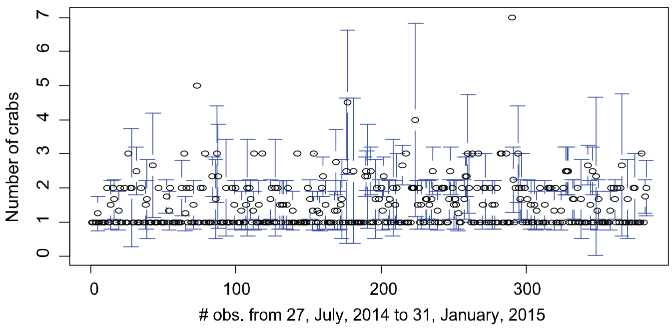
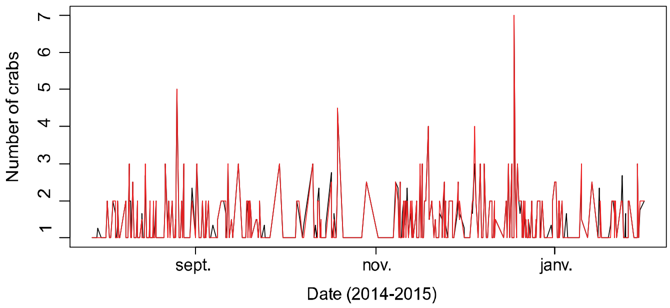
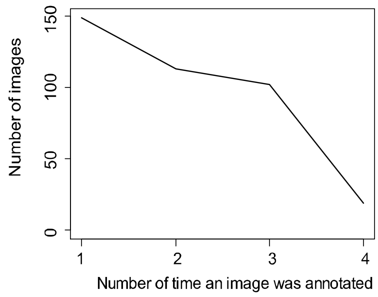


FIG 12. TOP LEFT. EVOLUTION OF CRAB ABUNDANCE SHOWING THE MEAN (BLACK DOT) AND STANDARD DEVIATION (BLUE LINES) AROUND THE MEAN. BOTTOM LEFT. EVOLUTION OF THE NUMBER OF CRABS OVER THE STUDY PERIOD: MEAN (BLACK) AND MEDIAN (RED). BOTTOM RIGHT. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANNOTATED.





Mean and median were quite similar (Fig. 12). Crabs are usually easy to observe in the image. Looking closer at the data, the median was lower than the mean, due to outliers overcounting crabs. Periodogram analyses did not reveal any periodicities in crab abundance variation. Considering the small amount of crabs, these results might be due to the low statistical power in relation to low abundance. In this case, observer bias might be too important to detect any signal.

Similarly to what is presented for the buccinids in the Pacific data, coordinates of individuals in pixels are available and will be used to generate maps to study the spatial distribution of crabs in the field of view. Below is an example of maps that can be generated (Fig. 13). In this case, distribution could reveal territoriality behavior.

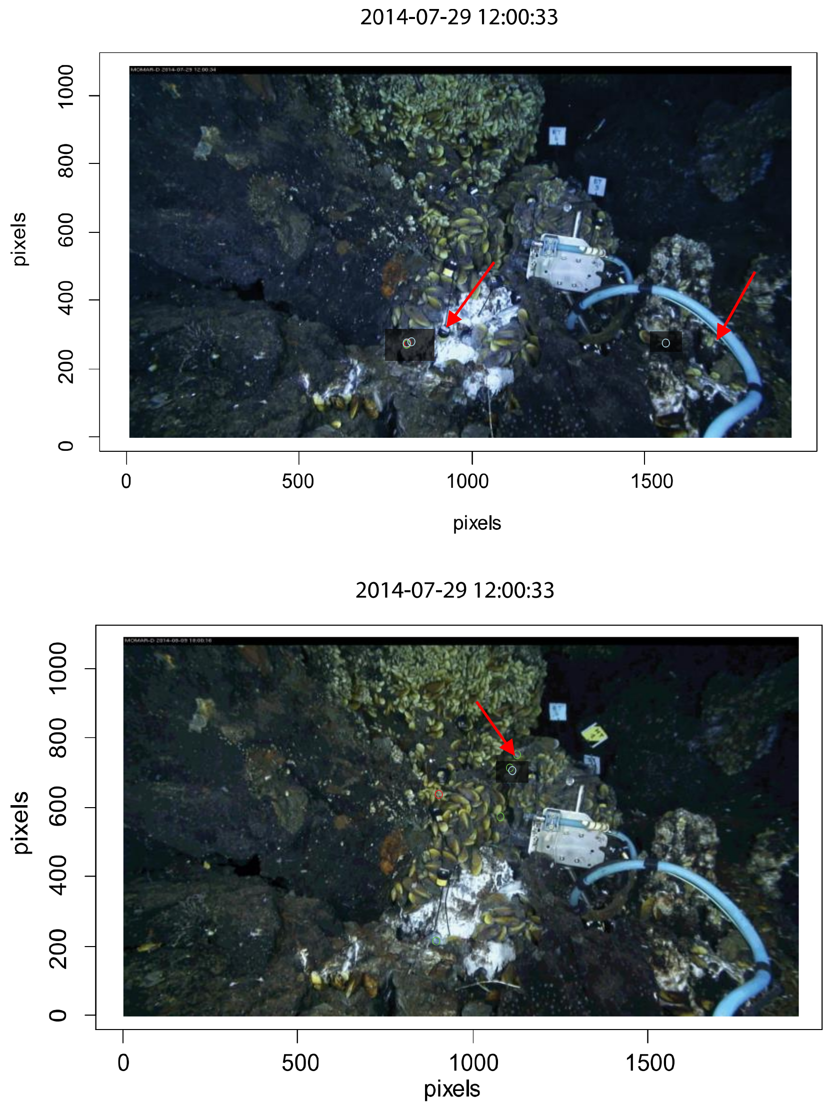


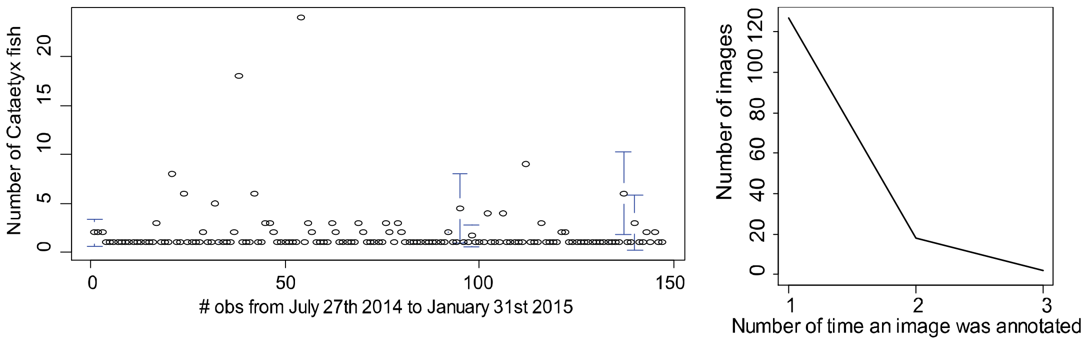
FIG. 13.

MAPS OF THE DISTRIBUTION OF CRABS IN THE FIELD OF VIEW. THE DIFFERENT COLORS CORRESPOND TO DIFFERENT PARTICIPANTS/ANNOTATORS. THE SQUARES AND ARROWS SHOW THE POSITION OF THE CRABS. EXAMPLE OF MAPS DONE FOR TWO DIFFERENT DATES AND TIMES.

These maps show concordance among observers for obvious individuals present in the front of the field of view. Some users appeared more enthusiastic and over-counted crabs.

❖ Evolution of Cataetyx fish

The graphs below show the evolution of abundances for the fish *Cataetyx* from 27, July, 2014 to 31, January, 2015 (Fig. 14). Over the time study, 169 images have been annotated, corresponding to 147 unique images. Altogether 304 Cataetyx fish were counted. Most images were annotated only once (Fig 14 right panel).



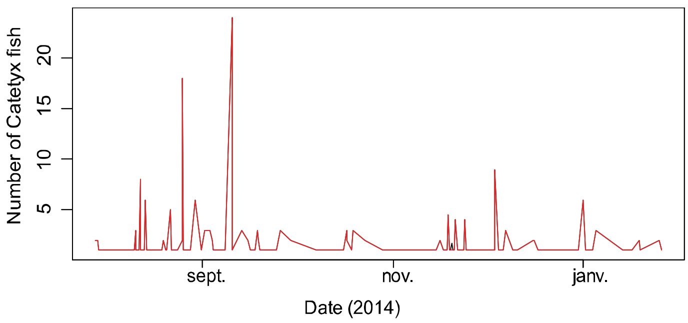


FIG 14. TOP LEFT. EVOLUTION OF CATAETYX ABUNDANCE SHOWING THE MEAN (BLACK DOT) AND STANDARD DEVIATION (BLUE LINES) AROUND THE MEAN. BOTTOM LEFT. EVOLUTION OF THE MEAN (BLACK) AND (MEDIAN) NUMBER OF CATAETYX. TOP RIGHT. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANALYSED.

Regarding *Cataetyx* abudances, images containing *Cataetyx* were only annotated twice for 20 images. It is thus currently impossible to draw any conclusions on these results. However, *Cataetyx* fish are easily identifiable in the image and we would expect citizen data to be reliable. However Fig. 14 show some absurd data with two dates reporting 18 and 24 fish respectively. It is extremely rare to observe more than one *Cataetyx* in the field of view, especially considering the size of the animal compared to the surface covered (one fish will occupy in length half the field of view). It might be reasonable to ignore all *Cataetyx* abundance above two individuals in the future, unless there is an agreement among a high number of observers, which would suggest the occurrence of an event.

The periodogram did not reveal any periodicity in the abundance evolution of *Cataetyx* fish, but similarly to crabs, this species is characterized by low abundances lowering the statistical power of the analysis.

❖ Evolution of Alvinocarid shrimp

The graphs below show the evolution of shrimp abundances from 27, July, 2014 to 31, January, 2015 (Fig. 15). Over the time study, 364 images have been annotated, corresponding to 305 unique images. Altogether 1,893 shrimp were counted. Most images were annotated only once (Fig 15 right panel).

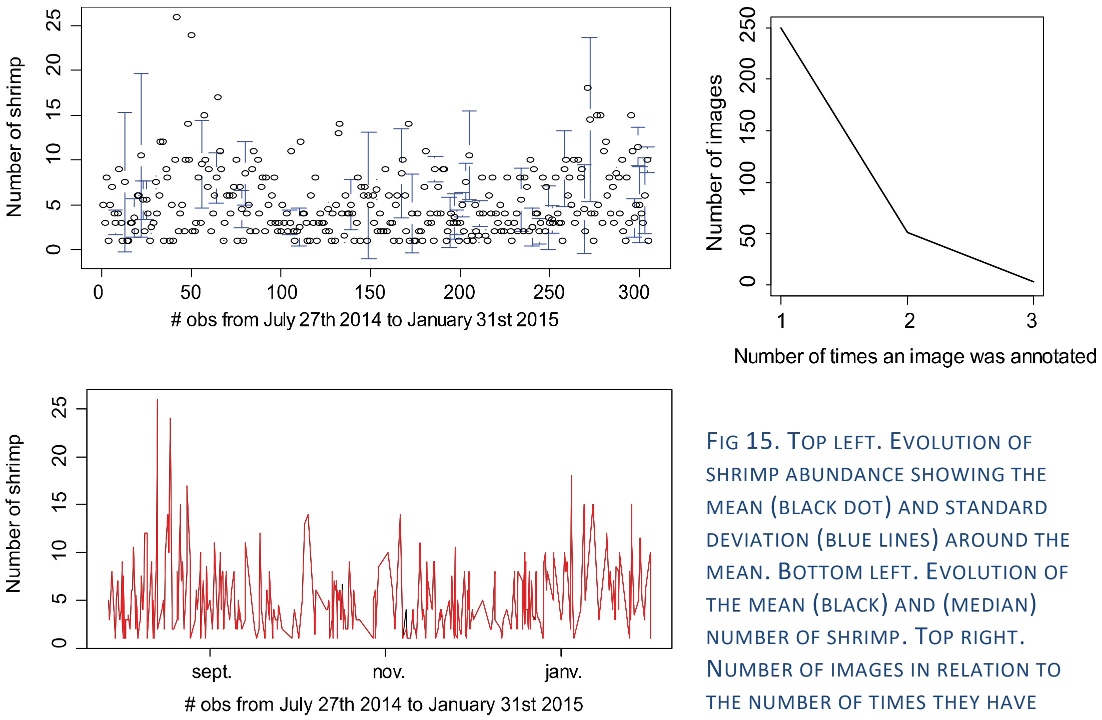


FIG 15. TOP LEFT. EVOLUTION OF SHRIMP ABUNDANCE SHOWING THE MEAN (BLACK DOT) AND STANDARD DEVIATION (BLUE LINES) AROUND THE MEAN. BOTTOM LEFT. EVOLUTION OF THE MEAN (BLACK) AND (MEDIAN) NUMBER OF SHRIMP. TOP RIGHT. NUMBER OF IMAGES IN RELATION TO THE NUMBER OF TIMES THEY HAVE BEEN ANALYSED.

It is premature to draw conclusions on consistency among observers as most images were only annotated once. Shrimp are hard to detect in the camera field of view when the image is not zoomed and we expect high variations among participants. Standard deviation, when images were annotated more than once, highlighted this variation (Fig 14). It is likely that only zoomed video sequences will be used to study shrimp abundance variation. Indeed, the periodogram did not reveal any periodicity in abundance variations, while a student in the lab detected a 14 day periodicity corresponding to circa-lunar signal.

V. PUBLIC EVENTS AND COMMUNICATION

Since the launch in March 2017, a number of communication actions were undertaken as part of exhibitions, pedagogic actions and conferences.

1) Design and project website

We worked with a graphic designer to develop visual guidelines associated with the tool, including a logo, color codes and a dedicated font (Fig. 16). A last version of the tool was deployed at the end of the summer 2017 including all graphic changes.

A project website (Fig. 16, www.deepseaspy.com) was developed parallel to the annotation tool in order to provide participants with background information. The project website, entitle *'behind the scene'* provides information on the project, the science around hydrothermal vents, the fauna, observatories and provide regular updates and news. When news are published on the website, the project's icon on the annotation application flashes.

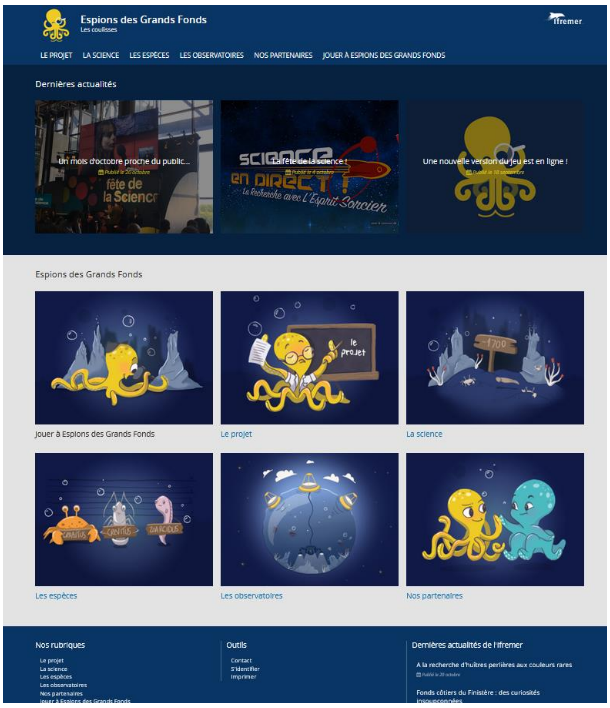


FIG. 16. ABOVE THE LOGO WITH OKTO THE OCTOPUS. ON THE RIGHT, THE LANDING PAGE OF THE ASSOCIATED PROJECT WEBSITE.

In collaboration with the communication and audiovisual departments we also produced a number of little video sequences to introduce several aspects of hydrothermal vent ecosystems to the public. The first one was published in September (https://www.youtube.com/watch?v=DqIs0iJNqIM). The second one will be available soon. We plan to publish a news/video every two weeks.

2) The center for ocean discovery: Océanopolis

A multimedia computer terminal (Fig. 17) was installed in the deep-sea area of the 'Pavillon Bretagne' at the aquarium Océanopolis in Brest. A special account was created specifically for the exhibition. To date, the visitors have annotated 353 images ranking them 7th on the current mission and 8th globally. We suspect that most visitors try annotating without validating the image. It will be interesting to assess how the visitors performed compared to school students and volunteer citizens.

Fig. 17. The DeepSeaSpy computer terminal at the center for ocean discovery Océanopolis in Brest, France (Pavillon Bretagne).

3) Actions in science events

❖ Science Fest

Our team participated to the science fest in Brest in 2016 and at the 'Cité de la Science et de l'Industrie in Paris in 2017'. For these occasions, visitors are invited to test the application using a computer set up at the Ifremer booth. The booth usually involves a computer, monitor, an accompanying poster and some samples of animals collected in the deep sea. For the Science Fest 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" (Fig. 18).

Fig. 18. A. Sandra Fuchs presenting deep sea vent and Deep Sea Spy to the Frédérique Vidal, the French Minister for higher education, Research and Innovation at the Cité de la science in Paris (October 2017). On the right: Ifremer booth with Deep Sea Spy at the Cité de la science in Paris (October 2017).

❖ Ocean Fest

In June 2017, our team was present at the 'Aquarium de la Porte Dorée' in Paris, France (Fig. 19). The day was dedicated to elementary classes. Each class had a 20 minutes explanation on hydrothermal vents, observatories, and how citizen science can help researchers. We also introduced them to the deep sea in general and the fauna inhabiting it.

Fig. 19. First Ocean Fest to be organised at the aquarium of the Porte Dorée, in Paris.

❖ Immersion science

Every year, the region of Brittany organises a week of science immersion for high school kids selected from schools across the entire region. Since 2013, our lab contributes with a presentation on the deep sea. In addition, we propose since 2014 an imagery workshop or 'lab', where students are invited to annotate our images (Fig. 20). For the 2016 and 2017 editions, the application Deep Sea Spy was made available during the lab.

Fig. 20. Immersion science 2017, held in Loctudy, Britanny, France. On the right, students working on the application

4) Conferences

The application was presented in the framework of several public conferences through 2016 and 2017, others are planned in 2018. Below is the list since the launch of the game:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

- Matabos M. Lumière sur les abysses. Conférence publique à l'Université du Temps Libre, Quimper, 06 avril 2017, Quimper.
- Sarrazin J., Matabos M, Sarradin P-M. En direct des grands fonds marins! 'Immersion Sciences' (Académie de Rennes), March 28, 2017, Ile Tudy.
- Matabos M. et Sarrazin J. Lumière sur les abysses. Conférence publique pour l'association du musée du sable. Mercredi 15 mars 2017, Château d'Olonne, France.
- M. Matabos & A. Abreu. 'L'océan, un monde à sauvegarder', Festival du livre et de la Presse d'écologie, October 8th 2016, Paris, France.

5) Media diffusion

The media release emitted in March 2016 for the launch of the application generated great interest from the media, particularly regional media. They include written articles in paper and web journals, radio chronicles and interviews, and live television broadcasts. The list of the main published stories can be found at the end of this document. We are still regularly contacted by journalists about the project.

6) Classes collaboration/school

We initiated a collaboration with the regional school academy to integrate the project into schools. The French education has a 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, the project *Deep Sea Spy* was selected as a project officially supported by the academy. The project leader will receive training by educational personnel in order to develop a two hours course that will be delivered to a group of middle and high school teachers from schools across Brittany.

Additionally, in collaboration with Océanopolis, an equivalent course will be built and targeted to elementary school teachers in order to bring the projects into classes for kids between 5 and 11 years old. This age range appeared to be the most enthusiastic public from our experience.

The advantage in working with schools lies in the fact that students change every year, bringing new participants. One can hope that every year few students will keep contributing to the annotation, slowly building our regular participant list.

7) Educative booklets

We are currently developing educative booklets for kids between 6 and 11 years old. These booklets are targeted to the younger public and will serve as support for the elementary teachers bringing the application into their class. Two booklets will be created, one for kids from 6 to 8, and a second for kids from 9 to 11 (Fig. 21). In the future, we hope to create one for younger kids from 3 to 5, aiming to just raise awareness on the deep sea, as kids that age are not able to use the application online.

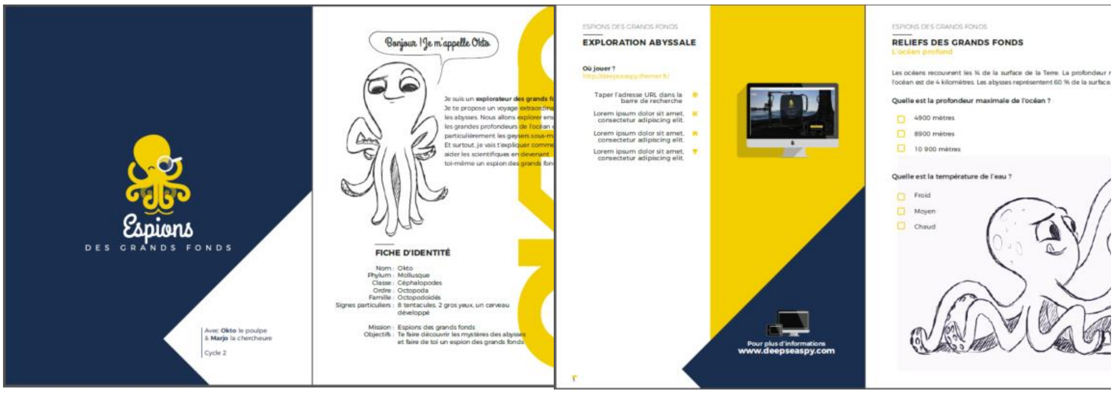


FIG. 21. EXAMPLES OF THE FUTURE BOOKLETS. THE ABOVE ONE IS TARGETED TO KIDS FROM 6 TO 8 YEARS OLD.

VI. CONCLUSIONS

All the initial objectives were fulfilled. The three objectives stated in the initial proposal document were:

- Design an online interface for video annotation
- Design a data model and database to archive information
- Develop a communication and educational plan

The online interface, including a tool to automatically extract images from video (developed in-house), the online interface, and programmed tasks, is now finished, functional and available (http://deepseaspy.ifremer.fr). Among the 800+ participants that registered with the application, ca. 450 are actively participating in the annotations. In 7 months more than 15,000 images (some identical) have been annotated.

The data model has been developed and all information is archived in the database center SISMER at Ifremer. An IT engineering student is currently working on developing a tool to generate requests to the database and export data in a form usable for biologists.

Finally, we have been working with our communications department to develop a communication plan. A first media release generated a lot of interest from the French media. This initial action was followed by a number of participations to public events through conferences or booth demonstrations. In parallel, we initiated a collaboration with colleagues at the center for ocean discovery Océanopolis in Brest, and the life science inspector and professors at the regional education academy to develop an educational strategy. This includes a training for professors in middle and high school to include Deep Sea Spy in their teaching program. For elementary school, we are currently developing educational booklets as well as a training workshop for teachers for them to use Deep Sea Spy as an educational tool in class and raise awareness of the deep ocean among the young generation.

The next challenge will be the processing of such a 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 (Wiggins et al. 2011; Bonter & Cooper 2012; Bird et al. 2014; Kosmala et al. 2016) and will provide directions on future analyses. Another approach will be to citizen data will to implement deep learning algorithms: annotations will help train computer programs for the automatic detection of animal species in the image (Kuminski et al. 2014). However, it will be important to maintain a high interest among the public and schools to ensure a high participation rate and thus the acquirement of valuable data.

IMPACT ON PROJECT

In this project we contributed to enhancing interactions between society and RIs. By involving citizens in the scientific process of imagery annotation, we particularly 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 generated by RIs big data. This document presents a preliminary analyses of citizens' data in the framework of the EMSO-Azores observatory project, and can be used as reference guidelines for future development within other RIs aiming to help with complex data processing through public engagement.

IMPACT ON STAKEHOLDERS

This project produced a certain number of deliverables that can applied to other RIs. First the system architecture of the application and database are available and can easily be adapted to other sets of images. In addition, we believe it is important that future databases are built on the one that was developed in this project in order to ensure standardisation of image databases across RIs to ultimately build a common archive and query system.

Finally, our team developed a certain experience in citizen engagement (both general public and schools) through the large number of public events and actions that were undertaken. This experience can serve other RIs within the ENVRIplus community in developing educational and communication plans.

AKNOWLEDGEMENTS

We would like to thank all the people at IFREMER that help in the project. Guillaume Clodic designed the program to automatically extract pictures from the video. Patrick Bossard provided it support to run all applications on the IFREMER infrastructure. Johanna Martin and Jade Burdallet from the communication team assisted for all the public diffusion. Stéphane Lesbats designed and realized the short videos. Finally, we are grateful to Sandra Fuchs and Jozée Sarrazin for their involvement and support.

REFERENCES

Bird TJ, Bates AE, Lefcheck JS, Hill NA, Thomson RJ, Edgar GJ, Stuart-Smith RD, Wotherspoon S, Krkosek M, Stuart-Smith JF, et al. 2014. Statistical solutions for error and bias in global citizen science datasets. Biol Conserv. 173:144–154.

Bonter DN, Cooper CB. 2012. Data validation in citizen science: A case study from Project FeederWatch. Front Ecol Environ. 10:305–307.

Cuvelier D, Legendre P, Laes 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:e96924.

Kosmala M, Wiggins A, Swanson A, Simmons B. 2016. Assessing data quality in citizen science. Front Ecol Environ [Internet]. [cited 2016 Dec 14]; 14:551–560. Available from: http://doi.wiley.com/10.1002/fee.1436

Kuminski E, George J, Wallin J, Shamir L. 2014. Combining Human and Machine Learning for Morphological Analysis of Galaxy Images. Publ Astron Soc Pacific [Internet]. [cited 2015 Dec 15]; 126:959–967. Available from: http://www.jstor.org/stable/info/10.1086/678977

Lelièvre Y, Legendre P, Matabos M, Mihály S, Lee RW, Sarradin P, Arango CP, Sarrazin J. 2017. Astronomical and atmospheric impacts on deep-sea hydrothermal vent invertebrates. Proc R Soc London B Biol Sci [Internet]. [cited 2017 May 11]; 284. Available from: http://rspb.royalsocietypublishing.org/content/284/1852/20162123.article-info

Matabos M, Sarrazin J, Borremans C, Tourolle J, Decker C 2016. D14.1 Prototype of a web-based annotation tool ready for user testing. Available from: http://www.envriplus.eu/wp-content/uploads/2015/08/D14.1.pdf

Matabos M, Hoeberechts M, Doya C, Aguzzi J, Nephin J, Reimchen TE, Leaver S, Marx RM, Branzan Albu A, Fier R, et al. 2017. Expert, Crowd, Students or Algorithm: who holds the key to deep-sea imagery “big data” processing? Methods Ecol Evol [Internet]. [cited 2017 Mar 20]; 8:996–1004. Available from: https://authorservices.wiley.com/api/pdf/fullArticle/13951974

Wiggins A, Newman G, Stevenson RD, Crowston K. 2011. Mechanisms for data quality and validation in citizen science. In: Proc - 7th IEEE Int Conf e-Science Work eScienceW 2011. [place unknown]; p. 14–19.

LIST OF COMMUNICATIONS ACTIONS

A. Web journals

Jouer pour percer les mystères des fonds marins. (31/03/2017). *up-inspirer.fr*. http://ct.moreover.com/?a=30081689928&p=20q&v=1&x=0aMV8IB3A2VBWXzxsXGy-A

Lumière sur les abysses. (05.03.2017). *culturesciences.fr*. http://culturesciences.fr/agenda/lumiere-sur-abysses

Sciences participatives : devenez un espion des grands fonds ! (07.03.2017) *Valeursvertes*. http://www.valeursvertes.com/sciences-participatives-devenez-un-espion-des-grands-fonds/

Un jeu pour découvrir les sources hydrothermales (08.03.2017). *Sciences et Avenir web*. http://ct.moreover.com/?a=29829825891&p=20q&v=1&x=YQpBdw80anPbq4QkP-7qmQ

L’Ifremer recherche des espions des grands fonds (08.03.2017). *Le Marin*. http://www.lemarin.fr/secteurs-activites/environnement/27951-devenez-espion-des-grands-fonds-pour-lifremer

Deep Sea Spy: aider la science en jouant les espions (15.03.2017). *plongez.fr*. http://www.plongez.fr/deep-sea-spy-aider-la-science-en-jouant-les-espions/

A découvrir ces images formidables d’une Tour Eiffel à 1700 mètres de profondeur en 3D ! (18.01.16). *Seableue*. http://www.seableue.fr/a-decouvrir-ces-images-formidables-dune-tour-eiffel-a-1700-metres-de-profondeur-en-3d/

B. Printed journals

Les marées donnent le rythme, Sciences Ouest n° 353-juin 2017.

Devenir un espion des grands fonds. *Sciences Avenir*, p24, 01/04/2017.

Plongez, jouez, communiquez, *brief*, p20, 01/04/2017.

Les comptages sont encore en cours, lire Un jeu pour aider les chercheurs, Sciences Ouest n° 350-mars 2017.

Devenez un espion des grands fonds, *Le Télégramme*, p44, 08/03/2017.

Conférence « Lumière sur les abysses », *Le journal des Sables*, p57, 09/03/2017.

Devenez un espion des grands fonds, *Le journal des enfants*, p13, 16/03/2017.

Devenez un espion des grands fonds, *L'Alsace*, p14, 16/03/2017.

C. Television broadcast

WebTV L'Esprit sorcier: https://youtu.be/EHcHzs_vqXI?t=2996, live with the French Minister of Higher Education, Research and Innovation during the Science Fest at the 'Cité de la Science et de l'Industrie' in Paris, 6-8 October 2017.

Espion des grands fonds sur L'instant T. Tebeo – 21/03/2017

D. Radio

IFREMER sur FRANCE INFO. 10:53:11 C'est ma planète - Anne-Laure Barral. On peut devenir espion des grands fonds en participant à un projet de science participative de l'IFREMER, il s'agit de compter les espèces sous-marines et parfois d'en découvrir de nouvelles. 10:56:27

APPENDIX A. ACRONYMS

A complete ENVRIplus Terminology/Glossary can be found online at :
https://wiki.envri.eu/pages/viewpage.action?pageId=14452608

Project acronyms

BEERi: Board of European Environmental Research Infrastructures - is an internal advisory board representing the needs of environmental Research Infrastructures

CA: Consortium Agreement - Legal contract between the ENVRIplus beneficiaries

DL: deliverable

DoW: Description of Work

GA:

1) Grant Agreement - Contract between Coordinator and Commission

2) General Assembly - GA is the ultimate decision-making body of the consortium

EB: Executive Board - supervisory body for the execution of the Project

EC: European Commission - is the executive body of the European Union responsible for proposing legislation, implementing decisions, upholding the EU treaties and managing the day-to-day business of the EU

ESFRI: the European Strategy Forum on Research Infrastructures

PM: Person Month

RI: Research Infrastructure

WP: Work Package

Organisational Acronyms

ACTRIS: Aerosols, Clouds, and Trace gases Research InfraStructure network

BEERI: Board of European Environmental Infrastructures

CEA: Commissariat a l Energie Atomique et aux Energies Alternatives

CINECA: Consorzio Interuniversitario CNR: Consiglio Nazionale Delle Richerche

CNRS: Centre National de la Recherche Scientifique

CODATA: Committee on data for Science and Technology

ConnectinGEO: Coordinating an Observation Network of Networks EnCompassing saTellite and IN-situ to fill the Gaps in European Observations

COOPEUS: Strengthening the cooperation between the US and the EU in the field of environmental research infrastructures

COPERNICUS: previously known as GMES (Global Monitoring for Environment and Security), is the European Programme for the establishment of a European capacity for Earth Observation

CSC: CSC - IT Center for Science

CU: Cardiff University

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182

D4Science: is an organisation offering a Hybrid Data Infrastructure service and a number of Virtual Research Environments

DASSH: Data archive for seabed species (a UK marine biology resource centre)

DIRAC : Distributed Infrastructure with Remote Agent Control

DKRZ: Deutsches Klimarechenzentrum GmBH

EAA : Umweltbundesamt GmbH - Environment Agency Austria

EduGAIN: is an international interfederation service interconnecting research and education identity federations

EEA: European Environment Agency

EGI : European Grid Infrastructure

EGLEU:

EINFRA-1-2014:H2020 Call for e-infrastructures (Managing, preserving and computing with big research data)

EISCAT: EISCAT Scientific Association

EMBL: European Molecular Biology Laboratory

EMBRC: European Marine Biological Resource Centre a consortium of research organisations interested in marine biology

EMODNET: The European Marine Observation and Data Network

EMRP: European Metrology Research Programme

EMSC: Euro-Mediterranean Seismological Centre

EMSO: European Multidisciplinary Seafloor and Water Column Observatory

ENVRI : FP7 project on Implementation of common solutions for a cluster of ESFRI infrastructures in the field of environmental Sciences

EPOS: The European Plate Observing System

EUDAT : H2020 project on Research Data Services, Expertise & Technology Solutions (previously funded by FP7)

EURO-ARGO: European ARGO programme (ARGO are a type of marine survey device)

EUROFLEETS:New operational steps towards an alliance of European research fleets

EUROGOOS: European Global Ocean Survey System

EuroSITES: European Ocean Observatory Network

ERIS: Environmental Research Infrastructure Strategy 2030

ESONET Vi: European Seafloor Observatory NETwork

ETHZ: Eidgenoessische Technische Hochschule Zurich

ESFRI: European Strategy Forum on Research Infrastructures

FIM4R: Federated Identity Management for Research collaborations

FMI: Ilmatieteen Laitos (Finnish Meteorological Institute)

FZJ: Forschungszentrum Juelich GmbH

FixO3: Fix point open ocean observatories (survey programme)

GBIF: Global Biodiversity Information Facility

gCube: is an open-source software toolkit used for building and operating Hybrid Data Infrastructures enabling the dynamic deployment of Virtual Research Environments by favouring the realisation of reuse oriented policies

GEO : The Group on Earth Observations coordinates international efforts to build a Global Earth Observation System of Systems (GEOSS)

GEOMAR: Helmholtz Zentrum Für Ozeanforschung Kiel

GEOSS : Global Earth Observation System of Systems coordinated by GEO (The Group on Earth Observations)

GROOM: Gliders for research ocean observation and management

H2020: Horizon 2020, European level research funding scheme

HELIX Nebula: partnership between big science and big business in Europe that is charting the course towards the sustainable provision of cloud computing - the Science Cloud

IAGOS - In-service Aircraft for a Global Observing System

ICOS : Integrated Carbon Observation System

ICSU: The International Council for Science

INFREMER : Institute Francais de Recherche Pour l'Exploitation de la Mer

INGV: Instituto Nazionale di Geofisica e Vulcanologia

INSPIRE : Integrated Sustainable Pan-European Infrastructure for Researchers in Europe

INRA: Institut National de la Recherche Agronomique

IS-ENES: RI for the European Network for Earth System Modelling

INTERACT: International Network for Terrestrial Research and Monitoring in the Arctic

IPBES: Intergovernmental Platform on Biodiversity & Ecosystem Services

I3: Integrated Infrastructures Initiative (I3) combines several activities essential to reinforce research infrastructures and to provide an integrated service at the European level

JERICO: Towards a joint European research infrastructure network for coastal observatories

LifeWatch:European e-Science infrastructure for biodiversity and ecosystem research

LU: Lund University

LTER: The Long-term Ecological Research Network

LTER-EUROPE : European Long-term Ecosystem Research network of 21 national LTER networks

MBA: Marine Biological Association of the United Kingdom

NERC: Natural Environment Research Council

NILU: Norsk Institutt for Luftforskning (Norwegian Institute of Air Research)

NMI: National Metrological Institutes

PANGAEA: Data Publisher for Earth & Environmental Science (Open Access library aimed at archiving, publishing and distributing georeferenced data from earth system research)

PLOCAN : Consorcio Para el Diseno, Construccion, Equipamiento y Explotacion de la Plataforma Oceanica de Canarias

RCN: Norges Forskningsrad (Research Council of Norway)

RDA: Research Data Alliance

RI: Research Infrastructures – facilities, resources and related services used by the scientific community to conduct top-level research in their respective fields, ranging from social sciences to astronomy, genomics to nanotechnologies.

SCAPE: SCAlable Preservation Environments (FP7 project)

SeaDataNet: Pan-European infrastructure for ocean & marine data management

SIOS: Svalbard Integrated Arctic Earth Observing System

SME: small and medium-sized enterprises

UCPH: Kobenhavns Universitet (Copenhagen University)

UEDIN: University of Edinburgh

UGOT: Goeteborgs Universitet (University of Gothenburg)

UHEL: Helsingin Yliopisto (University of Helsinki)

UiT: Universitetet i Tromsoe (University of Tromso)

UniHB: Universitaet Bremen (University of Bremen)

UNILE: Universita del Salento (University of Salento)

UNITUS: Universita Degli Studi della Tuscia

USTAN : The University Court of the University of St. Andrews (University of St Andrews)

UvA : Universiteit van Amsterdam (University of Amsterdam)

Important Technical Terms/Acronyms

API: Application Program Interface, is a set of routines, protocols, and tools for building software applications

Biodiversity: is the variety of different types of life found on earth

Biodiversity metrics: measurements of the number of species and how they are distributed

CERIF: Common European Research Information Format

CIARD RING: A global directory of information services and datasets in agriculture

Data stream: is a sequence of digitally encoded coherent signals used to transmit or receive information that is in the process of being transmitted

Data pipeline: In computing, a pipeline is a set of data processing elements connected in series, where the output of one element is the input of the next one.

DCAT: is a resource description format vocabulary designed to facilitate interoperability between data catalogues

DOI: Digital Object Identifier

E-infrastructure: can be defined as networked tools, data and resources that support a community of researchers, broadly including all those who participate in and benefit from research

HPC: High Performance Computing

HTC: High Throughput Computing

IoT: The Internet of Things - is a scenario in which objects, animals or people are provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction.

ICT: Information and Communications technology

INFRADEV-4: Subcall of H2020 INFRADEV call for Implementation and operation of cross-cutting services and solutions for clusters of ESFRI and other relevant research infrastructure initiatives

IPR: Intellectual Property Rights

KOS: Knowledge Organization Systems - is a generic term used in Knowledge organization about authority lists, classification systems, thesauri, topic maps, ontologies etc.

LOD: Linked open data is linked data that is open content

LOV: Linked Open Vocabularies

Metadata : is data that describes other data. Metadata summarizes basic information about data, which can make finding and working with particular instances of data easier

NGI: National Grid Initiative

NREN: National Research and Education Network

NRT: Near Real Time - refers to the time delay introduced, by automated data processing or network transmission, between the occurrence of an event and the use of the processed data (For example, a near-real-time display depicts an event or situation as it existed at the current time minus the processing time, as nearly the time of the live event)

OASIS: Advancing Open Standards for the Information Society (non-profit consortium)

ODP: Online Data Processing

OIL-E: The Open Information Linking model for Environmental science - is a semantic linking framework

Ontology: (In computer science and information science) an ontology is a formal naming and definition of the types, properties, and interrelationships of the entities that really or fundamentally exist for a particular domain of discourse

QoE: Quality of user experience

over dispersion : a statistical characteristic of data such that the data have more clusters than compared to what might be expected if the data were distributed randomly in proportion to the time/space available.

NetCDF: a file format.

OceanSITES: s a worldwide system of long-term, open-ocean reference stations measuring dozens of variables and monitoring the full depth of the ocean from air-sea interactions down to the seafloor

OOI: Ocean Observatories Initiative

RDA: Resource Description and Access, a standard for descriptive cataloguing

RM: Reference Model - is an abstract framework or domain-specific ontology consisting of an interlinked set of clearly defined concepts produced by an expert or body of experts in order to encourage clear communication

SensorML - The primary focus of the Sensor Model Language is to provide a robust and semantically-tied means of defining processes and processing components associated with the measurement and post-measurement transformation of observations

Semantics : is the study of meaning

Syntax: In computer science, the syntax of a computer language is the set of rules that defines the combinations of symbols that are considered to be a correctly structured document or fragment in that language

SLA: Service Level Agreement

UV: unmanned vehicles

VL: Virtual Laboratory

VLDATA: this was the name of the failed project proposal so I think it can be deleted

VRE: Virtual Research Environments, web based package tailored to a specific community