www.aquacosm.eu

Workshop on AI in imaging recognition techniques

Deliverable No: 7.3

PROJECT TITLE:

AQUACOSM-plus

Network of Leading Ecosystem Scale Experimental Aquatic Mesocosm Facilities Connecting Rivers, Lakes, Estuaries and Oceans in Europe and Beyond

PROJECT NUMBER:

871081

PROJECT TYPE:

Research and Innovation Action

WORK PROGRAM TOPICS ADDRESSED:

H2020-INFRAIA-2018-2020 / H2020-INFRAIA-2019-1

Deliverable title	Report on Workshop on AI in imaging recognition techniques
Deliverable number	D7.3
Deliverable version	1
Contractual date of delivery	31 July 2021 (M16)
Actual date of delivery	21 December 2021 (M21)
Dissemination level	**PU**
Nature of deliverable	**R**
Workpackage	WP7
Lead Beneficiary	SYKE
Author(s)	Katri Kuuppo (SYKE), Maiju Lehtiniemi (SYKE), Jens Nejstgaard (FVB-IGB)
Editor	Katerina Symiakaki (FVB-IGB), Ayoub El Ghadraoui (FVB-IGB), Jens Nejstgaard (FVB-IGB)
EC Project Officer	Pierre QUERTENMONT

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

Table of Content

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

1. Abstract

The Workshop on AI in imaging recognition techniques (D7.3), was originally planned as a hands-on workshop in spring 2021, but had to be postponed due to the prolonged COVID19 pandemic. Finally, a virtual, whole day workshop was organized on the 9th of December 2021. A number of automated imaging instruments have been tested and verified against traditional methods (microscopy, particle counting, flow cytometry) by SYKE and FVB-IGB. The partners have had 6 meetings where these technologies have been discussed and compared. In the online workshop these developments as well as status updates from external experts on plankton imaging and/or artificial intelligence applications were presented and discussed.

If the pandemic situation allows, we aim to organize an additional hands-on workshop in spring 2022, where the imaging instruments, their software and image analysis techniques can be tested, compared and discussed.

2. The workshop

SYKE (Finnish Environment Institute) and IGB (Leibniz Institute of Freshwater Ecology and Inland Fisheries) organized an online workshop on Plankton Imaging and Artificial Intelligence.

The objective of the workshop was to present and discuss novel technologies for imaging plankton organisms from bacteria to zooplankton, using commercial instruments including CytoSense, Imaging Flow Cytobot, FlowCam, Holo imaging, ZooScan, MDPI, dElfi and the self-built low cost PlanktoScope, as well as arising Artificial Intelligence approaches for image processing and analysis.

In addition to the presentations (see list of the presenters below), the following topics were discussed:

- what the new technologies have enabled in routine analyses of particles and plankton, compared to the classical methods, and how reliable are the results
- how easy or difficult are the identification processes of taxa or traits, through-put rates and ease of analysis, etc.
- data management and FAIR principles in the image data
- artificial intelligence and machine learning in the image recognition

Invited speakers to the workshop:

- Kaisa Kraft, SYKE: Imaging Flow CytoBot IFCB and image recognition
- Tuomas Eerola, LUT University: Similarity learning for plankton recognition
- Stella Berger & Wojciech Uszko, FVB-IGB: Different types of FlowCams and AMNISFlowStream
- Hans Henrik Jakobsen, Århus University: Phytoplankton imaging
- Lumi Haraguchi, SYKE: FlowCam and CytoSense
- Felipe Artigas, ULCO-CNRS: CytoSense and FlowCAM imaging

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

- Klas Ove Möller, Hemholz-Zentrum Hereon: Plankton and particle imaging
- Bronwyn Lira Dyson, FVB-IGB: The low-cost PlanktoScope
- Rainer Kiko, Laboratoire d'Océanographie de Villefranche-sur-mer: The Underwater Vision Profiler, MorphoCluster and EcoTaxa
- Maiju Lehtiniemi, SYKE: Status of LISST Holo2 and ZooScan
- Tim Walles & Jens Nejstgaard, FVB-IGB: In situ MDPI profiler & dElfi laboratory scanner

The workshop had 66 attendees (Figure 1, please see Appendix 1), who had preregistered. The majority were AQUACOSM-plus partners, who had variable experiences of image analysis methods (listed in Appendix 2).

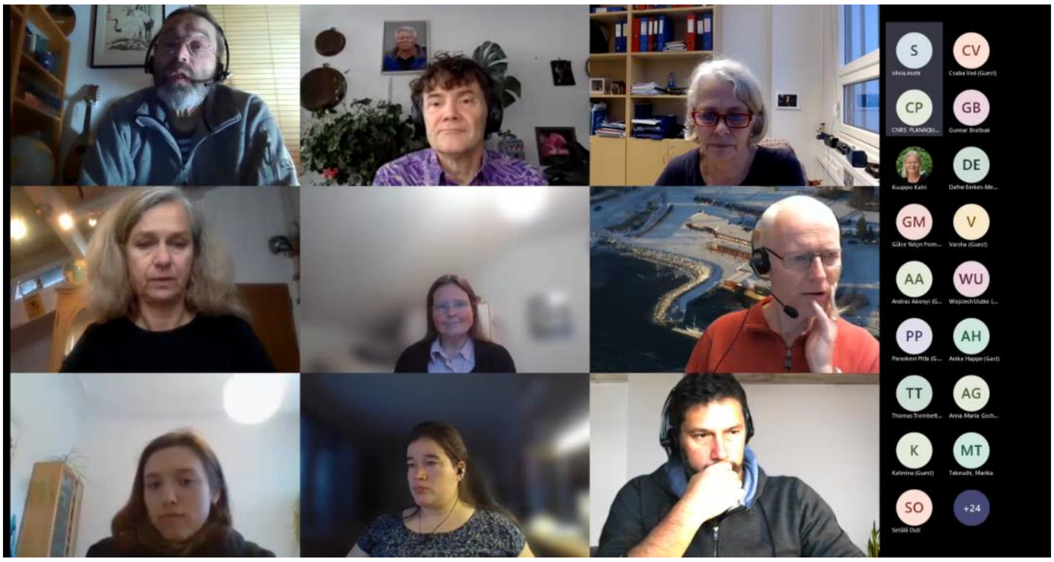

Figure 1: Participants of the online workshop on automated imaging and artificial intelligence, 9 December 2021

3. Preparations before the Workshop: Testing and development of image analysis instrumentation and AI-based recognition techniques

Preparations for the workshop included specific testing of the automated imaging instruments (SYKE, FVB-IGB), and their verification against traditional methods (microscopy, particle counting, flow cytometry, Figure 2). The methods included:

- FlowCams for phytoplankton and microzooplankton
- IFCB for phytoplankton
- CytoSense for phytoplankton and microzooplankton
- MDPI for zooplankton *in situ*
- Self-built dElfi zooplankton laboratory scanner for zooplankton
- PlanktoScope for phytoplankton and beads
- LISST Holo2 for zooplankton
- ZooScan for zooplankton

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081

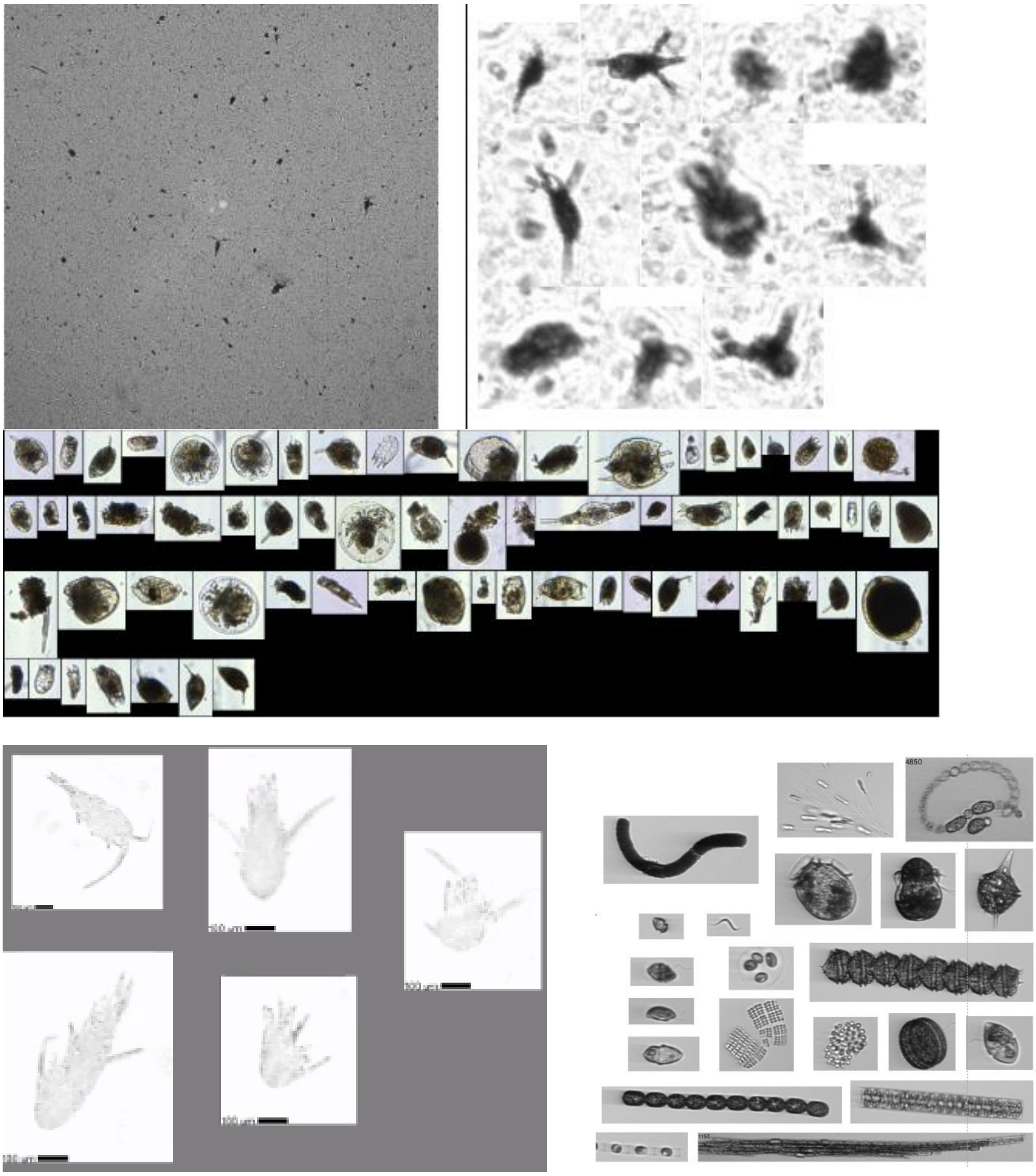

Figure 2*: Example pictures of live plankton: in situ MDPI of copepods and cladocerans (top, photos Tim Walles), sorted rotifers and copepods in the FlowCam (middle photo Maiju Lehtiniemi) and LISST Holo 3D pictures of copepods (below left photo Anne-Mari Lehto) and IFCB image (below right photo Kaisa Kraft).*

The images derived with these instruments have been material for development of image recognition using Artificial Intelligence methods:

- Device and species agnostic image recognition using machine learning
- Compact convolutional neural networks (CNN), and comparison to the classical Random Forest method

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

4. Results of the workshop

1.1 Instruments

The presentations and discussions in the virtual workshop showed that several imaging techniques are already routinely used especially for phytoplankton and particle counting both *in situ* and in the laboratory (e.g. FlowCam and IFCB). *In situ* zooplankton imaging is trickier, as zooplankters are fast moving and sparser in the environment. To overcome this problem the Mini Deep-focus Plankton Imager (MDPI) with a focus depth of 10 cm and a pixel resolution of 20 µm have routinely been used in the LakeLab at FVB-IGB. For analysis of fixed zooplankton samples in the Laboratory the use of the self-built digital Elfi (dElfi) based on a commercial camera and the commercial ZooScan was briefly discussed.

There are clear advantages in most of the instruments for particle imaging compared to the traditional examination of plankton using microscopes:

- High throughput and faster processing of samples in 'real time'
- Semi-automated to automated characterization
- The observed organisms are alive and thus have their characteristic details
- Detection of problematic particles
- Analysis of several parameters per particle
- Production of digital images allowing detailed analysis later

In addition, the autonomous instruments allow the analysis plankton of fine spatio-temporal scales, which is essential for understanding the fast-growing organisms in constantly fluctuating environment such as water.

The development of instruments and methods is continuous, leading to improved resolution and precision, as well as faster processing of samples. A general issue is the software and licenses provided by the instrument manufacturers. They are often expensive and do not have all features needed for research. For example, in order to extract images for detailed analyses from some instruments one needs to apply 'external' scripts (Python, Matlab).

Multiple measurements reveal more of the community dynamics of e.g., algal blooms. Images combined with Flow Cytometer profiles (Figure 3), in addition to integration of additional observations using *in vivo* sensors, e.g., multispectral fluorometers are useful for capturing short spatio-temporal scales of plankton.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

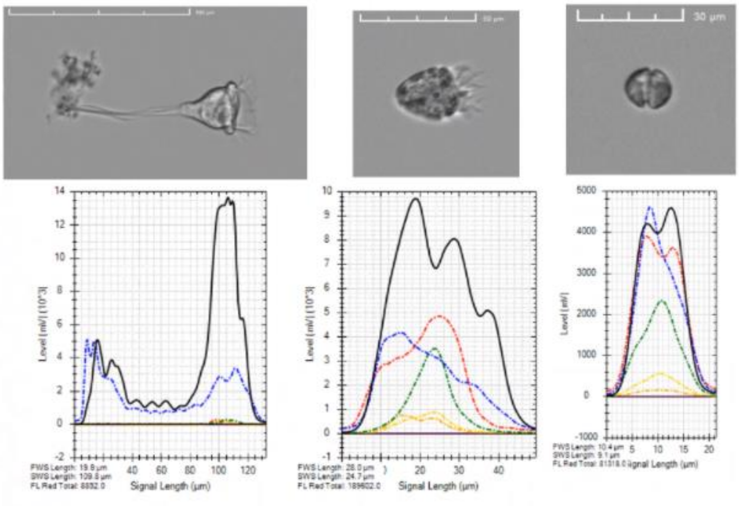

Figure 3*. Plankton images combined with signal profiles from the Flow Cytometer (photo Lumi Haraguchi).*

Whereas phytoplankton imaging is more or less routine at both SYKE and FVB-IGB, zooplankton samples are routinely analyzed at genus level or higher at FVB-IGB using the self-built dElfi, however there are issues which need consideration. Tests with ZooScan for zooplankton monitoring at SYKE, showed clear advantages such as the fast analysis of samples, easy routine scanning and measurement dimensions of individuals for biomass estimations. However, downsides include uncertainties in the observation of smaller zooplankton (small cladocerans, rotifers, copepod nauplii). The latter is especially a challenge in freshwater habitats. Here, too, the software can be cumbersome and difficult to use.

1.2 Data

Even if there may be challenges in imaging instruments themselves, the real issue is large amount of data they produce. For example, during a cyanobacterial observation campaign in the northern Baltic Sea (Kraft et al. 2020), an Imaging Flow Cytobot processed a sample every 20 min, 24/7, for several weeks. This poses challenges not only for data transfer but particularly its storage and analysis.

It is essential to include metadata to the image data from a given sample, and the immediate file output should then be archived. The archived data can then be uploaded into an image recognition application (e.g., EcoTaxa) where the work continues with the classification of the images. As it is not possible to handle millions of images by hand, the only option is to improve archiving of the raw data since the image recognition afterwards will be at some level automated.

Data sharing is a topic which needs consideration of the imaging 'community'.

The image data management can be approached either using centralized databases, in which there is personnel carrying out the analyses, or making freeware, which is then used by the clients at their own costs. In both, funding and payment structure are the problem.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

It is also important to keep in mind two parallel tracks in the development and innovation: new instruments and new methods for recognition. As imaging data have a clear advantage to fixed samples, in that they do not deteriorate below the original resolution over time, it is therefore crucial that the data remains available for later reanalysis. The digital data should be available applying the FAIR principles also for raw images. To enable this, good metadata is critical. Considering the very large data volume typically created by imaging approaches, it is perhaps more important than ever to have clear and rigid system to keep track of the data accumulation from the very beginning, and foresee large enough servers with good backup functions of also the raw data, to ensure that the data will not end up into a dead end, but be available when new algorithms are developed.

In addition to the ensuring proper metadata and large volume storage, another issue is currently sorting of the data. For example, the EU EMODNet does not accept machine-sorted data. As there is presently an explosive development towards imaging technologies and machine sorted data, these points need to be solved in collaboration, perhaps by suggesting clear tagging of pure machine sorted data, machine aided but human controlled data, as opposed to traditionally fully manually sorted data.

1.3 Image recognition

Image recognition and machine learning are usually dependent on training sets of images. It means that the start can be slow and laborious, as the results are verified using the traditional approaches and expert judgement. The machine learning algorithms convolution neural networks (CNN) and similarity learning approaches (Figure 4) are used.

The similarity learning based on one instrument is currently being developed further (domain adaptation, domain generalisation). The objective of the similarity approaches is to avoid looking at every image individually. This means that the algorithm automatically finds similar images and groups them to named classes. Most images in the archive should be only checked when the software prediction finds out they are somehow deviant.

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

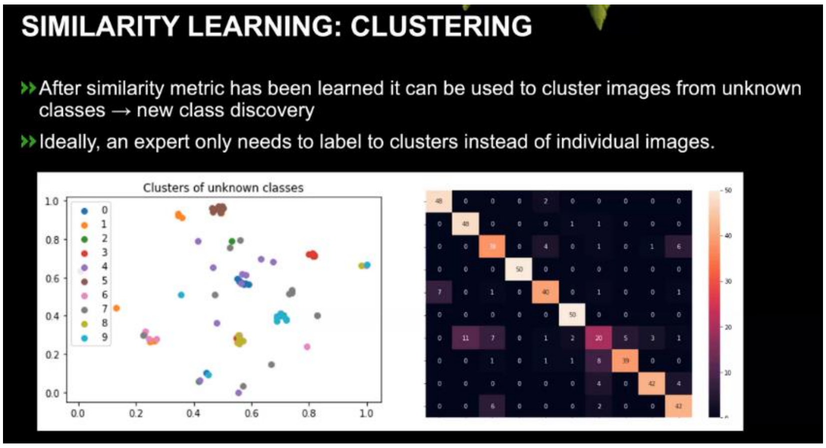

Figure 4*. Similarity learning principle (source: Tuomas Eerola).*

Ecotaxa offers many services for plankton imaging: a MorphoCluster with the following features: deep learning prediction, project management, daily backups, collaborative sorting, open source. Service to pelagic imaging professionals, sharing of instruments and data. Links: EcoTaxa, Zooprocess https://sites.google.com/view/piqv/home

5. Collaboration

JERICO-RI and FINMARI RI

- Development of IFCB, FlowCam and CytoSense imaging using natural plankton assemblages and operational, real-time observation of phytoplankton using the IFCB. at the Utö field Station (https://www.finmari-infrastructure.fi/field-stations/uto-fmi/)

Project FastVision plus LINK (funding from the Academy of Finland)

- The project uses images from the novel automated imaging instruments, to train image recognition software for species identification using efficient machine learning approaches that produce interoperable data across instruments and habitats

AQUA-KI (Funding BMU, Germany, at FVB-IGB)

- A pre-project to develop intelligent optical approaches for effective analysis of freshwater microorganisms

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

6. Dissemination Activities Related to this Deliverable

Presentation on the FINMARI Researcher Day 17.03.2021 by Kaisa Kraft: Imaging flow cytometry observations and image analysis of the Baltic Sea phytoplankton community.

7. References

Kaisa Kraft, Jukka Seppälä, Heidi Hällfors, Sanna Suikkanen, Pasi Ylöstalo, Sílvia Anglès, Sami Kielosto, Harri Kuosa, Lauri Laakso, Martti Honkanen, Sirpa Lehtinen, Johanna Oja, Timo Tamminen: First Application of IFCB High-Frequency Imaging-in-Flow Cytometry to Investigate Bloom-Forming Filamentous Cyanobacteria in the Baltic Sea. Front. Mar. Sci., 25 March 2021 https://doi.org/10.3389/fmars.2021.594144

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

8. Appendix

1.4 List of online workshop attendees 9 December 2021 (derived from Teams)

Andras Abonyi
Anika Happe
Anna-Maria Gschwandner
Annie Marie Cox
Ayoub El Ghadraoui
Bronwyn Lira Dyson
Camarena Gómez Teresa
Caroline Gorzerino
Cass Bromley
Christian Dilewski
Christian Preiler
Csaba Vad
Dafne Eerkes-Medrano
Dilvin Yıldız
Duanyai Pranchalee
Eric Edeline
Felipe Artigas
Francesca Vidussi
Georgia Ktistaki
Gülce Yalçın
Gunnar Bratbak
Haraguchi Lumi
Heidrun Feuchtmayr
Hương Huỳnh
I-Hao Chen
Iordanis Magiopoulos
Jean-Christophe Aymes
Jens Nejstgaard
Johan Wikner
Jorun Karin Egge
Jukka Seppälä
Justine Courboulès
Katerina Symiakaki
Katri Kuuppo
Kraft Kaisa
Kyle Mayers
Lisette Senerpont Domis
Maiju Lehtiniemi
Margarita Machairopoulou
Marika Takeuchi
Meryem Beklioğlu
Outi Setälä
Paraskevi Pitta
Patrick Galvan
Rainer Kiko
Robert Ptacnik
Sami Jaballah
Sanna Suikkanen
Sarah Fiorini
Silvia Mohr
Stella Berger
Siru Tasala
Thomas Trombetta
Tim Walles
Timo Tamminen
Tuomas Eerola
Vildan Acar
Virág Csiszár
Wojciech Uszko
Zoe Jones
Zsofia Horvath

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

1.5 Interests and experience of the attendees as indicated in the registration form

- FlowCam and Zoo Imaging
- we plan to buy equipment for plankton imaging for the lab in the long term
- plankton ecologist, experience in flow citometry
- I used to use light microscope and hemocytometer for counting and making image for microplankton
- FlowCam beginner.
- Flow cytobot
- FlowCam and Flow cytometry
- Imaging of prokaryotes
- ZooScan, Planktoscope, image recognition
- plankton image recognition
- ZooScan, Flowcam
- I'm interested in all, and I have no experience with these devices.
- Flow cams
- Phytoplankton microscopy
- Flow Cam, cyto sense
- Planktonic
- phytoplankton methods
- Some little experience with FlowCam
- Phytoplankton
- research assistant
- Light Microscopy, Electron Microscopy, Flow Cytometry
- FlowCam & Cytocense, experience on Microscopy and Flow Cytometry
- microplastics in aquatic environments
- We have traditionally only used samples and taxonomy to assess plankton communities. It would be interesting to find out if there is technology available that would be suitable for some of our studies.
- Imaging flow cytobot
- Flow cytometry and Flow Cam. I have been running the Flow cam for some time now, but am interested in a rapid method for classifying the images gained.
- Zooplankton Analyst
- general
- general interest in new methods; experience with phytoplankton imaging
- Phytoplankton Analysis
- FlowCam, experience on Microscope
- Video
- ZooScan
- ML, Underwater imaging
- I am interested in the use of image processing applied to mesocosm experiments. I am currently a PI of such project.
- Zooscan, FlowCam, CytoBuoy
- Freshwater algal taxonomist
- Automated algae and zooplankton recognition and measurement methods

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.

- Flow Cytometry, Microscopy
- I have no experience and skills in imaging, yet, I'm very interested in these methods that I want to work with
- Machine learning-based macroinvertebrate detection
- FlowCam application- beginner
- Ciliate community with FlowCam
- FlowCam, interest in plankton imaging
- FLOWCAM AND CYTOSENSE
- Flowcam and zooscan
- CytoSense
- Flowcam / a few months of use
- Mainly FlowCAM
- Phytoplankton image classification by machine learning
- Holo imaging with AI recognition, UVP, CPICS, PlanktoScope
- Flow cam and flow cytometry, experienced user of both
- Interested on several methods, practical experience level low
- General
- FlowCam
- Interest high, experience modest
- Interest in FlowCam, experience with classical (count) approach
- Zooplankton
- FlowCam, Amins Image Stream, phyto- and microplankton analyses plankton,
- Building and testing the PlanktoScope
- IFCB and image recognition
- AI
- Underwater imaging / very experienced
- Zooplankton imaging
- Zooplankton: Zooplankton in situ profiler, laboratory scanners

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 871081
Responsibility for the information and views set out in this report lies entirely with the authors.
The European Commission is not responsible for any use that may be made of the information it contains.