JNCC Report
No. 686

The BIG PICTURE II: Benthic Imagery Workshop Report

Golding, N., Dinwoodie, K., Golob, E., O'Connor, J., Taylor, J., van Rein, H. & Webb, K.

July 2021

© JNCC, Peterborough 2021

ISSN 0963 8091

For further information please contact:
Joint Nature Conservation Committee
Monkstone House
City Road
Peterborough PE1 1JY
https://jncc.gov.uk/

This report should be cited as:
Golding, N., Dinwoodie, K., Golob, E., O'Connor, J., Taylor, J., van Rein, H. & Webb, K. 2021. The BIG PICTURE II: Benthic Imagery Workshop Report. *JNCC Report No. 686*, JNCC, Peterborough, ISSN 0963-8091.

JNCC EQA Statement:
This report is compliant with JNCC's Evidence Quality Assurance Policy
https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Summary

On 2 – 4 March 2021, the Joint Nature Conservation Committee (JNCC) hosted the BIG PICTURE II workshop. This online event represented two years of benthic imagery innovation and collaboration within the Big Picture Group (BPG), marked a milestone in the implementation of the Benthic Imagery Action Plan (BIAP), and an opportunity to share progress of the group. Since its inception, membership of the BPG has expanded to include 121 individuals, split across 42 government, academic and private sector organisations.

In addition to sharing progress, the objectives of the workshop were to identify and explore future funding routes for the BPG, identify potential global outreach opportunities and discuss future challenges to the delivery of the BIAP.

The workshop was successful in achieving these objectives. Some outstanding products have already been delivered and were showcased during the event, such as the Epibiota Identification Protocol, a tool to improve the consistency of taxonomic identification from marine benthic imagery. Advances in imagery annotation software development and opportunities through the continued development of artificial intelligence and autonomous systems were common themes throughout the event. The importance of communication and coordination across the Project Working Groups (PWGs), tasked with delivering the BIAP, and role of the Action Plan Coordination Committee (APCC) in providing strategic oversight were also agreed.

Recommendations coming out of this workshop will be taken on board by each PWG and the APCC. They will be incorporated within the next iteration of the BIAP, and following this period of review, the BPG, its PWGs and members will work towards delivering this action plan over the coming year.

Copies of the workshop resources, including the workshop programme, videos of the sessions, and PDF copies of the presentations are available from The BIG PICTURE II webpage.

Contents

1 Background

In 2019, the Joint Nature Conservation Committee (JNCC) brought together 29 organisations from across the marine monitoring and assessment community in the UK and beyond, to discuss approaches to the collection, analysis and dissemination of benthic imagery. The workshop, held on 19 – 22 March 2019 and called the 'BIG PICTURE'[1], enabled experts to grapple with a broad range of issues, seeking collaborative solutions and opportunities, while considering how to incorporate new technologies, such as computer vision and machine-learning, into existing workflows.

The two main outcomes of the BIG PICTURE workshop were the creation of a Benthic Imagery Action Plan[2] for the UK and the formation of the Big Picture Group. The Benthic Imagery Action Plan collates and focuses all recommendations from the BIG PICTURE workshop into 87 tasks, organised across seven coherent workflows. The Action Plan is endorsed by UK's Healthy & Biologically Diverse Seas Evidence Group (HBDSEG), and progress towards it is reported to the North East Atlantic Marine Biological Analytical Quality Control Scheme (NMBAQC[3]). The Big Picture Group is a cross-organisation, interdisciplinary benthic imagery working group tasked with implementing the Action Plan. Work within this Group is carried out by cross-organisation Project Working Groups (PWG), each focused on a different theme. These are currently:

- Identification approaches for benthic imagery
- Benthic imagery data flows, archives and catalogues
- Imagery annotation software
- Artificial intelligence approaches for benthic imagery
- Benthic imagery workflow guidance
- Benthic imagery analysis training scheme
- Enumeration approaches for benthic imagery taxa
- Quality Assurance Framework for benthic imagery

The main point of contact for each of these PWGs periodically reports progress and flags up issues to the newly formed Action Plan Coordination Committee. This Committee provides an internal governance framework for the Big Picture Group and a means for sharing progress across the whole Group.

[1] The BIG PICTURE workshop 2019: http://www.nmbaqcs.org/media/1785/the-big-picture-workshop-2019.pdf
[2] Benthic Imagery Action Plan: http://www.nmbaqcs.org/scheme-components/epibiota/benthic-imagery-action-plan/
[3] http://www.nmbaqcs.org/

2 Purpose and scope of the workshop

The BIG PICTURE II workshop, held on the 2 – 4 March 2021, built on the successes of the first workshop in 2019, and represented two years of benthic imagery innovation and collaboration. It was also a major milestone in efforts to progress the implementation of the Benthic Imagery Action Plan. Over the past two years, membership of the Big Picture Group, which emerged as an entity following the 2019 workshop, has expanded to include 121 individuals split across 42 organisations, each with an interest in benthic imagery.

The BIG PICTURE II workshop engaged and brought together 74 stakeholders, over three days, with participation spanning eight countries including Belgium, Netherlands, Canada and the United Arab Emirates (Figure 1).

The workshop objectives were:

1. to share progress towards achieving the Benthic Imagery Action Plan with the Big Picture Group;
2. to identify effective funding routes for the Project Working Groups;
3. to identify global outreach opportunities to enrich and support the Project Working Groups;
4. to identify future challenges to the delivery of the Benthic Imagery Action Plan and propose solutions.

The first objective was delivered through a series of update presentations made by relevant experts and leads of the relevant Project Working Groups, whilst the latter three objectives were achieved through open discussion.

The BIG PICTURE II workshop was organised and facilitated by the JNCC. This report – produced by Aquarius Survey & Mapping, in partnership with JNCC – summarises the presentations, discussions and recommendations produced by all participants during the course of the workshop. Further details are available by viewing videos of the workshop sessions on the BIG PICTURE II workshop webpage.

Figure 1: Some of the participants of the BIG PICTURE II workshop, hosted online.

3 Session 1: The journey from the 2019 BIG PICTURE Workshop to the UK's Benthic Imagery Action Plan, the Big Picture Group (BPG) and Project Working Groups (PWGs)

This session was facilitated by Kirsten Dinwoodie (JNCC).

A video recording of the session is available on 'The Big Picture II Webpage', along with links to PDF copies of the presentations.

3.1 Development of the UK's Benthic Imagery Action Plan

Henk van Rein (JNCC) presented a summary of how the UK's Benthic Imagery Action Plan (BIAP) was developed, following the BIG PICTURE workshop in 2019. A total of 87 tasks, organised into seven themes, were considered and prioritised. The scope of the BIAP was focussed on delivering outputs for the marine biodiversity conservation community over the next five years. However, the resulting tools and outputs will be useful for a much wider range of stakeholders. The BIAP took approximately one year to establish and sign-off.

Workshop participants highlighted that while it was important to stress the biodiversity monitoring context, hence the group sitting under the NMBAQC and HBDSEG governance framework, these existing tasks and potential future tasks could also have value to the wider scientific community.

3.2 Formation of the Big Picture Group (BPG)

Henk van Rein (JNCC) presented an overview of the formation of the Big Picture Group (BPG). The original BIG PICTURE workshop in 2019 was oversubscribed with potential attendees exceeding the workshop capacity. However, it was possible to integrate all interested parties/organisations through the creation of the BPG; the current membership stands at 121 individuals split across 42 organisations (see Figure 2). Government bodies form approximately one third of the group's membership (the blue shaded segments in Figure 2). Research and academia also form approximately one third of the group's membership (green shaded segments in Figure 2).

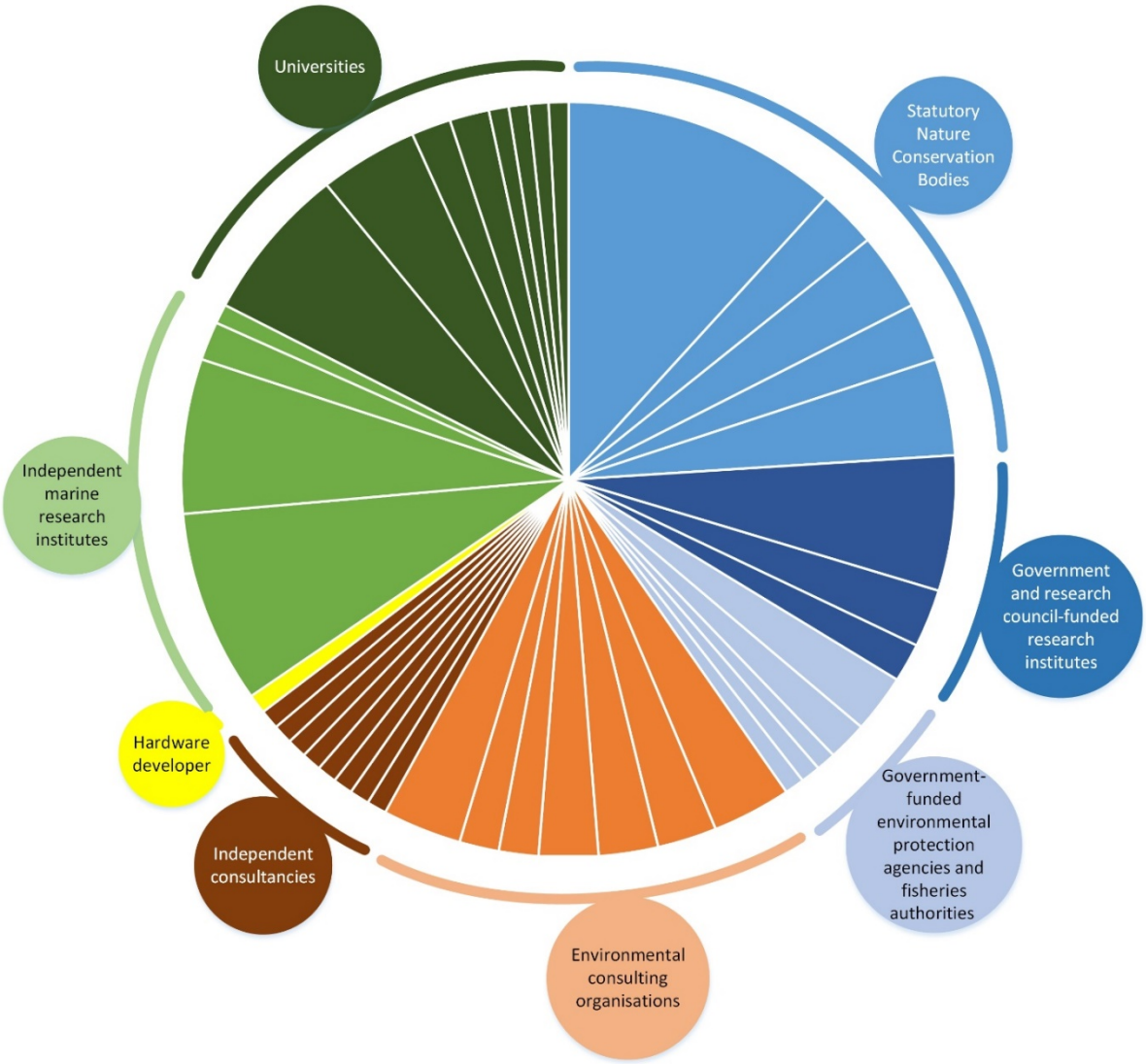

Figure 2: A breakdown of the Big Picture Group membership, by organisation type.

The Big Picture Group (BPG) established a collaboration platform, utilising Microsoft Teams. The BPG Terms of Reference were finalised and approved in November 2020, and are available on the BPG Teams site. The Teams site has proved instrumental in allowing the group to continue to progress different aspects of the work, despite the impact of the global COVID-19 pandemic.

Parallels were drawn to the way the Big Picture Group functions with other groups such as the International Council for the Exploration of the Sea (ICES) working groups, FathomNet[4] and others. Looking forward, there may be benefits to making links with these wider groups to share knowledge, ideas and to coordinate.

[4] https://www.mbari.org/fathomnet/

3.3 Project Working Groups (PWGs) and looking ahead

Henk van Rein (JNCC) presented an overview of the formation of Project Working Groups (PWGs). Tasks within the Benthic Imagery Action Plan (BIAP), discussed in Section 3.1 above, have been tackled using two strategies:

1. passive implementation – e.g. Publication of JNCC enumeration comparison method studies. Work that is already ongoing within the community can be fitted/assigned to particular tasks.

2. active implementation – PWGs were assembled to tackle prioritised BIAP tasks, across the eight themes. These PWGs were discussed and finalised at the Big Picture Group (BPG) November workshop, held at the end of November 2020. Initial progress is being shared through this BIG PICTURE II workshop; a milestone event for the BPG.

These PWGs are semi-autonomous and are overseen by the Action Plan Coordination Committee.

A standing invitation was made for any interested individuals/organisations who may not yet have registered to contribute to a particular PWG, to get in touch during the BIG PICTURE II workshop with the workshop organising team.

4 Session 2: Early successes of the UK's Benthic Imagery Action Plan

This session was facilitated by Joey O'Connor (JNCC).

A video recording of the session is available on 'The Big Picture II Webpage', along with links to PDF copies of the presentations.

4.1 Development of standard benthic imagery purposes

Henk van Rein (JNCC) presented an overview of the development of standard benthic imagery purposes. Appropriate use of shared data requires knowledge of why data were collected, where they were collected, how they were collected and what processes/analysis was performed on them; through the development of standards.

Benthic imagery purposes were discussed at the BIG PICTURE workshop in 2019, and tasks were identified and then integrated within the Benthic Imagery Action Plan (BIAP). A specific high priority task that emerged from this process was:

To identify main purposes and develop standards for benthic imagery purposes.

The Big Picture Group (BPG) were consulted on benthic imagery purposes, via a questionnaire. There were 28 mixed structured questions, sent to 100 individuals across 39 organisations. From this, responses from 38 individuals were received. This questionnaire was then followed up with a one-hour interview, of which 26 were carried out.

Six core purposes were identified through the consultation process:

1. Marine archaeology and heritage
2. Fisheries stock assessment
3. Identification, training and public outreach
4. Research and/or development of novel technologies
5. Full seabed characterisation – human activities
6. Full seabed characterisation – habitats

A summary of these six purposes went out for review to the BPG and the Marine Monitoring Group – and following revisions for various reasons, the focus became full seabed characterisation, split across feature verification, habitat mapping and habitat/species monitoring.

It was noted that many more 'purposes' are possible and can be developed by the BPG for other purposes if deemed necessary. All these 'purposes' can be inserted into the Quality Assurance Framework for benthic imagery, explored in Section 4.2 below.

Following the presentation, a number of questions were posed by the audience. One participant enquired as to the rate of responsiveness to the questionnaire process, which was approximately 40%.

Another participant asked whether there were any plans in place for the four imagery purposes that were not progressed (e.g. Marine archaeology and heritage, Fisheries stock assessment, Identification/training/public outreach and Research and/or development of novel technologies). Henk responded that the reason they were not progressed at this time was the requirement to progress the project onto the next level, using the Quality Assurance

Framework (QAF) to test and assess data against those purposes. These four purposes could be developed in the future, it is just a question of resource.

Joey enquired as to whether there were lessons learned through this focussed work in the three areas (full seabed characterisation, split across feature verification, habitat mapping and habitat/species monitoring) that could be fed back into the wider work, such as the purposes not originally progressed or in imagery collection more widely, and is there a mechanism in place to allow this to happen. Henk responded that the primary mechanism will be the QAF; there will be mechanisms to feed in / review / enrich this work with new ideas. Outside the QAF mechanism, participants with ideas to explore were encouraged to get in touch with the BPG.

4.2 Development of a Quality Assurance Framework

Jessica Taylor (JNCC) and Graeme Duncan (JNCC) presented an overview of the Quality Assurance Framework (QAF) development work. This work built on the benthic imagery purposes work within the Benthic Imagery Action Plan (BIAP), discussed above in Section 4.1, and in particular, these three purposes; feature verification, habitat mapping and habitat/species monitoring.

The QAF was developed to include a quality assurance scheme for benthic imagery, and a suite of tools to assist in the standardisation of imagery data, including standardised proformas for the recording of information about benthic imagery. A series of automated proforma data checks have been developed as an integral part of this QAF.

Graeme then provided a demonstration of the online QAF data checking tools, the development of which is being led by the Marine Biological Association[5], as part of their work on DASSH - the UK Archive for Marine Species and Habitats Data[6], under the Marine Environmental Data and Information Network (MEDIN)[7]. A beta test version of the tools are available at https://www.dassh.ac.uk/qaf/. These online tools including a data completeness check, a species ID comparison with image quality check (in relation to the Epifauna Identification Protocol (EIP), discussed in Section 4.3 below) and a dataset quality assurance comparison tool.

A question was posed by the audience on the intended use of the dataset comparison tool. Graeme responded that its primary role was to assist in dataset quality assurance, for example when 10% of a dataset was re-analysed by a second laboratory. A discussion followed as to whether this method included consideration of sample units, such as unit area of seabed, and if not, how this method may be expanded to consider different seabed units. Linkages between this work and the **Quantification Working Group** were made.

The discussion highlighted how these tools were primarily aimed at achieving compliance for the easier elements of the imagery standards under development; picking out the more repetitive errors or those that may be easily missed by human eyes reviewing a dataset comparison. However, they were not so suited to those ecological / biological errors. It also highlighted how various add-ons, such as combining individual images to consider different sample unit areas, could be developed.
The increased use of annotation tools was highlighted and how some of these tools may link with these, particularly as some annotation software already have some of these tools available, such as QA/QC and data comparison tools. Therefore, the online tools developed

[5] https://www.mba.ac.uk/
[6] https://www.dassh.ac.uk/
[7] https://www.medin.org.uk/

as part of the QAF may not necessarily be used in this case but are available if needed. For any outputs of data QA undertaken, MEDIN compliance should be considered. There could also be a benefit to engage with annotation software developers on some of these quality assurance tools in the QAF. It was noted that there is a diverse range of annotation software available, so it would not be appropriate to link to just a single platform.

Next steps in the QAF workflow are:

- Producing guidance documents on using the proformas and tools.
- To make the outputs of the QAF available as part of the NMBAQC[3] epibiota component on the NMBAQC website.

After March, feedback will be sought on the QAF. In future, if there is a standard that doesn't exist in the QAF, this can be considered. The online tools will move from a beta-testing phase and release to a live testing phase.

All QAF products will be live at the end of March 2021 on the NMBAQC website.

Feedback is welcome

A general show of hands (22 raised in the Teams meeting) highlighted significant interest in being involved with future developments within the QAF.

4.3 The Epifauna Identification Protocol

Jessica Taylor (JNCC) presented an overview of the Epifauna Identification Protocol (EIP) work that has been progressed over the past few months. The EIP aims to improve the consistency of taxonomic identification of benthic taxa/morphotaxa from underwater imagery, such as video and still imagery, through assigning a taxonomic level identification recommendation for the various quality levels. Part of this work included reassessing the image quality levels laid out in NMBAQC[3] guidance to check they were fit for purpose for the EIP. As a consequence, the NMBAQC image/video quality categories were revised to increase specificity and clarity. The revised quality categories are available in the EIP guidance[8].

The EIP itself was developed through a series of dedicated workshops and extensive discussion with experts in the field between December 2020 to February 2021 and takes the form of a spreadsheet. For a given species, the latest taxonomic information from the World Register of Marine Species (WoRMS)[9] database is provided on the left-hand side, followed by a variety of morphological classifications that are in current use, such as CATAMI (Althaus *et al*. 2015[10]). Guidance/recommendations are then provided for the lowest level that the organism can be identified to, for both still imagery and video imagery, at a range of image qualities. These quality categories follow those revised NMBAQC definitions, though *Very Poor* and *Zero* visibility classes were not considered during the EIP process. The final

[8] http://www.nmbaqcs.org/scheme-components/epibiota/epibiota-quality-assurance-framework-and-documents/
[9] http://www.marinespecies.org/index.php
[10] https://doi.org/10.1371/journal.pone.0141039

column called ‘Notes’ provides useful contextual information to assist with identification, such as colour, morphology or biogeographic distribution.

The following taxa groups were reviewed as part of this work:

- Porifera
- Bryozoa
- Cnidaria
- Mollusca
- Arthropoda
- Echinodermata
- Chordata
- Annelida
- Platyhelminthes (small number)

The EIP will be continually developed and enhanced, and feedback on its taxonomic recommendations/guidance is welcome. In addition, any further information which could be beneficial to others for inclusion in the ‘Notes’ field is also welcome.

The EIP is also available as a standalone tool, for use during ID checks whilst completing annotations. There is also a QAF tool and comparison tool, complete with link to the EIP, as discussed above in Section 4.2.

Future work on the EIP will focus on integration of deep-sea species within the species list, including a desire to link with the Standardised Marine Taxon Reference Image Database (SMarTaR-ID; Howell *et al*. 2019) work which is currently underway. There will also be the functionality to make recommendations for additions to the EIP. It was recognised that there are gaps in the current list of taxa groups covered, such as no algae.

During the various EIP workshops, it became apparent that the current CATAMI[9] morphological classification structure was not really appropriate for sponges or bryozoans found in UK waters; there are plans for a workshop in future which will look to expand this from a UK perspective. From a ‘Big Data’ perspective, there is the potential to link the EIP to existing or new image catalogues, such as those being considered by the Big Picture Group as well as wider, to provide identification aids.

Jess thanked all the Big Picture Group (BPG) members who had devoted a large amount of time attending the EIP workshops and contributing to this work. There was interest in assisting with the development of the EIP into the future – 23 hands were raised within the Teams meeting in response to this question.

There was a question posed around the quality of imagery versus the quality of the specimen, so for example, there may be a very good image of a specimen, but it still isn’t possible to see some of the diagnostic features of that species. Jess clarified that within the EIP, characteristic features needed for species level identification have been noted, which should aid identification.

Another participant explained that often, even though you have a generally poor-quality video, there are moments of clarity when the camera settles, and whether this can be taken into account with the EIP. Jess clarified that the EIP only provides a recommendation, and is not prescriptive, so those moments of ‘clarity’ can be considered. It was highlighted that segmenting video into appropriate sections of similar quality may be useful in this respect.

With respect to image/video quality, a question was raised regarding at which scale should this be considered; set of images, per image, per specimen, per sample unit. Some cross-Project Working Group (PWG) discussion around this issue would be beneficial, rather than this being tackled by individual PWGs in isolation. Consideration of bias is also important in how/when video is segmented, and with annotation for that matter. It was recommended that this wider challenge be discussed at the Action Plan Coordination Committee (APCC), before being discussed at the PWG level.

ACTION: APCC to discuss challenges around appropriate scale for consideration of image quality

A question was posed about having a catalogue of images to use alongside the EIP. Jess responded that there could be a repository of images linked to the EIP – you could see example images of species at a range of image quality levels, such as excellent, satisfactory or poor. It was also flagged that many annotation software packages already come with image catalogues to assist with identification.

This led to a wider discussion around annotation tools (note the dedicated session, reported in Section 6.3, on this subject). There are lots of differences in annotation tool software and their outputs, but they also have some overlapping functionality. BPG members were encouraged to engage and discuss challenges with annotation platform developers; they would welcome feedback and they may be receptive to make changes, especially when brought to their attention as a group, rather than as an individual. Engagement with developers on other aspects of annotation software, such as image catalogues, was also recommended. Joey echoed these comments that annotation developers have been very receptive to feedback provided.

There was a question around annotation software standards – there are clear benefits, allowing users to exchange information between different platforms (to be discussed in more detail in Section 6.3). Standards could have benefits to streamlining the data flow process, for example creating Application Programming Interfaces (APIs) which would allow rapid transfer of information.

There was a discussion around quality consideration, and the EIP in particular, which seems to be focussed on the biological elements – but what about the abiotic/physical element – for example trying to identify the proportion of sediment, gravel, cobble, boulder, etc. It was also acknowledged that substrate is very important in driving ecological community distribution and future work could explore this in more detail. There may also be relevance in considering substrate identification with respect to artificial intelligence/machine learning, discussed in Section 5.1.

5 Session 3: Benthic Imagery Action Plan progress in the digital realm

This session was facilitated by Jessica Taylor (JNCC).

A video recording of the session is available on 'The Big Picture II Webpage', along with links to PDF copies of the presentations.

5.1 Artificial intelligence

Kerry Howell (UoP) provided an update from the Artificial Intelligence (AI) Project Working Group on progress and challenges in the application of artificial intelligence to marine survey and monitoring.

Use of cameras has seen massive growth – both benthic **and** pelagic, although benthic was the focus of the presentation. New technology creates vast datasets, creating a challenge for data processing and storage. In addition, manual interpretation and analysis is expensive and time consuming, and after extended periods, humans can and do make mistakes, as well as observer bias.

Computer vision and AI provides a complementary approach to human input. It has been created for other applications; Facebook and facial recognition is a good example. However, the same principles can be applied to ecological imagery.

Open-source tools are making it easier to access computer vision and AI, but it is still not altogether "easy". Kerry ran through a number of examples and approaches being trialled through the Big Picture Group (BPG) including (but not limited) work led by Nils Piechaud (NOC), who has been applying these approaches to AutoSub data, utilising imagery from AutoSub deployments undertaken on DeepLinks[11] project (2016). He looked at performance linked to a variety of other factors, such as number of training images needed to correctly classify. Scottish Government commissioned work looking at the automated identification of fish and other aquatic life, in a variety of scenarios (Blowers *et al*. 2020). Case studies looked at smolts and sprat in the back of trawl nets, and tracking individuals. Other case studies counted seapens on the seafloor from a towed camera system. The report from this work provided a great summary report of where we are with this. Other case study examples were also discussed.

Kerry highlighted a number of challenges in using these AI tools in marine benthic imagery analysis. While better tools may need to be developed, the current emphasis should be on applying and learning from the tools we currently have, rather than developing new ones.

Before developing new AI tools for marine benthic imagery analysis, we should apply and learn from those we currently have.

We need to remember the fact that bad training material creates bad AI models – training material is KEY!

The vision of the AI approaches Project Working Group is to utilise data collected by fleets of autonomous vehicles surveying our oceans, collecting imagery, processing and training as they go.

[11] https://deeplinksproject.wordpress.com/

There was a discussion around the applicability of AI in low visibility. Sensor technology may evolve to allow better 'vision' (not necessarily in the visible spectrum), but obscured vision is still obscured vision, whether seen by humans or machines.

Consistency was discussed next. Consistent errors can be dealt with, inconsistent errors are more challenging. There also needs to be agreement in the level of accepted error for using AI, before its use is mainstreamed, which is effectively a policy decision around levels of risk, of which errors in methodology form one element to consider.

There is a need to better collaborate with the computer science community, rather than trying to upskill ourselves to become computer programmers. However, scientists still need to become familiar with using these methods and this can be improved by making them more accessible.

Using a common dataset to explore and compare how effective different AI methodologies/ approaches are, is a key aspiration. Projects such as FathomNet[12] (discussed in Section 5.4) provide one such example of developing this common dataset approach.

Finally, there were discussions around using AI to look at sedimentology, rather than just biological species. This work is being progressed by the geological community. One example is Dan Buscombe's work at the United States Geological Survey (USGS), pyDGS[13].

5.2 Data flows

Dan Lear (DASSH) provided an update on the essential infrastructure and processes required to enable AI and other approaches. In particular, Dan was seeking a sense check from the wider Big Picture Group (BPG) on the proposed data follow process (shown in Figure 3), and whether it reflected the needs of the various users. Note this flow is conceptual and intended to kick-start conversations about the development of future data management infrastructure for benthic imagery.

It is essential to have a comprehensive data management plan in place before data acquisition. At the data acquisition phase, there are challenges around data storage, standards, QC processes and maintaining the provenance (what's been done to the data at each stage). At the data processing stage, there are challenges around storage, tools/software and interoperability (across the toolsets being used). The final two steps are the storage of records and reporting. There are challenges around the re-aggregation of data which have been split from the images they were derived from, to allow re-interpretation.

The data flow process can be summed up by trying to make benthic imagery FAIR; Findable, Accessible, Interoperable, Reusable.

[12] https://www.mbari.org/fathomnet/
[13] https://github.com/dbuscombe-usgs/pyDGS

We need to make benthic imagery FAIR

Findable, Accessible, Interoperable, Reusable

Storage can be seen as the consistent challenge running through the data flow process.

Data flow and re-aggregation were then discussed. Dan provided an overview of the data landscape in the UK, and how this rests within wider European programmes such as OBIS[14]/EurOBIS[15] and EMODnet Biology[16].

As a use-case, Dan explained that as image and annotated occurrence data gets ingested into MEDIN, and as this data is published on to aggregators such as EMODnet/OBIS, the link between image and associated data can be lost. MEDIN have been piloting a small project called MEDIN Automated Image Management System (MAIMS), using a digital asset management system called ResourceSpace[17] with a publicly available API. As it's an open-source application, the metadata that underpins it has been modified, and has been aligned to the MEDIN Data Guidelines. The pilot has been a success and ensures that the image and annotated data can be re-aggregated.

More generally, the re-aggregation of split data can be achieved through ensuring high quality metadata, and consistent use of Persistent IDentifiers (PIDs), such as the Digital Object Identifier (DOI)[18], ORCID iD[19] or Research Organisation Registry (ROR)[20]. Using this combination will ensure that disaggregated data can be re-aggregated in a machine actionable way – removing the barriers to the re-aggregation of data. This can also unlock the potential for data to be citable in its own right, ensuring that academics get the credit they deserve from data, outside of the normal peer reviewed publication process.

Following the presentation, there were discussions about how many copies of the imagery should be retained. Some users keep all copies from all steps, although there are challenges around this from a storage perspective. The answer depends on each use case. For example, if you have a well-defined workflow one could potentially keep the original dataset and a record of the workflow steps undertaken.

Storage was recognised as being one of the key challenges. Reference to Moore's Law was made, drawing parallels between benthic imagery storage requirements and the observation that the number of transistors in a dense integrated circuit (IC) doubles about every two years[21]. Note that Moore's law is an observation and projection of a historical trend, rather than a law of physics; it is an empirical relationship linked to gains from experience in production.

Challenges around storage differ depending on whether centralised versus distributed storage is used. If we get standards and interoperability sorted, then a single, centralised storage solution is not required.

---

[14] https://obis.org/
[15] https://www.eurobis.org/
[16] https://www.emodnet-biology.eu/
[17] https://www.resourcespace.com/
[18] https://www.doi.org/overview/DOI_article_ELIS3.pdf
[19] https://orcid.org/signin
[20] https://ror.org/about/
[21] https://newsroom.intel.com/wp-content/uploads/sites/11/2018/05/moores-law-electronics.pdf

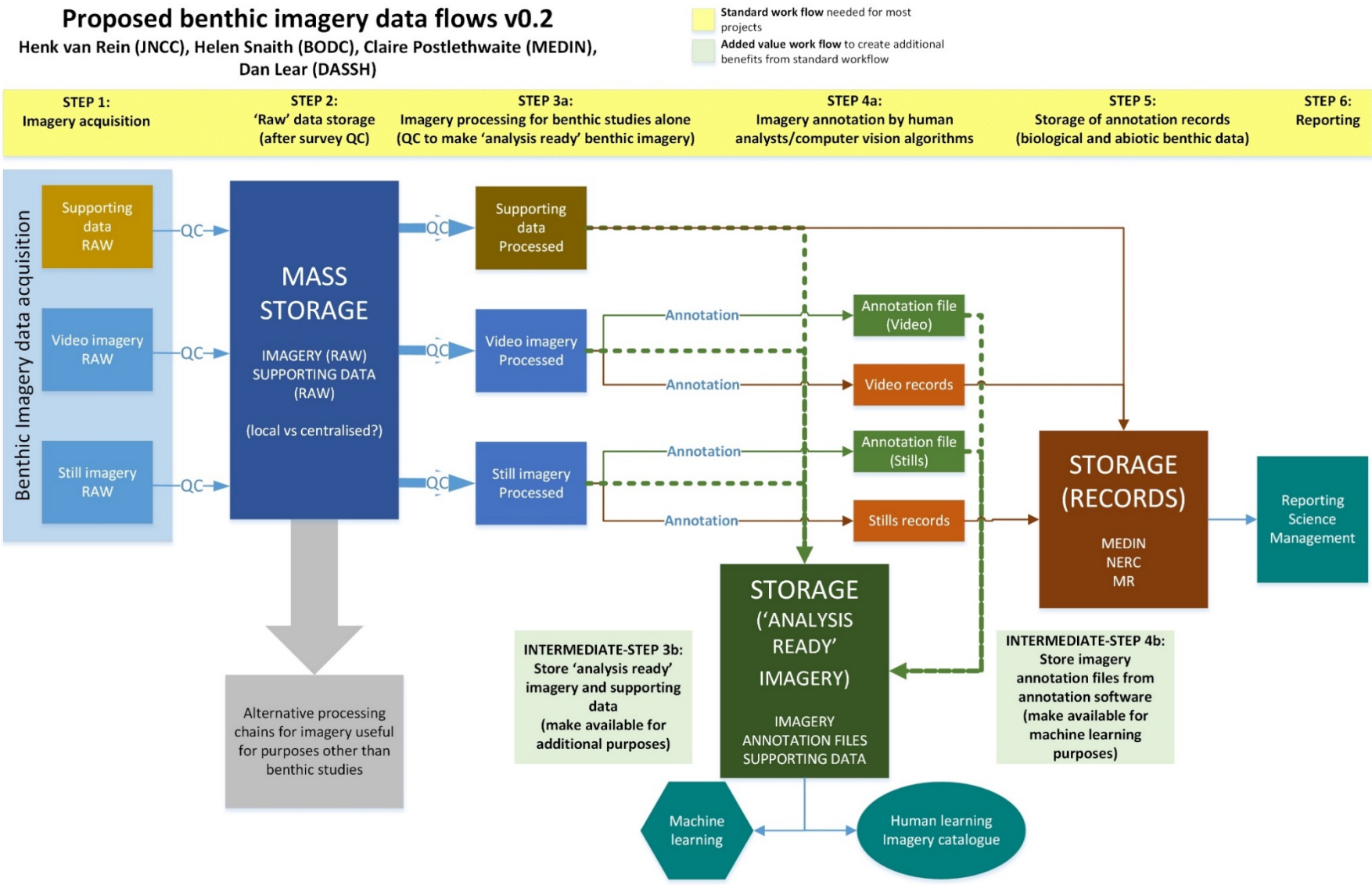

Figure 3: Proposed data follow process for marine benthic imagery

Discussions around storage continued, exploring the reasons why we like to store everything in case we may need it in the future. An example was given of analyst contractors being requested to store original imagery, but it not being requested by clients. However, there is the potential for future re-use to be considered. For example, we may be able to use imagery collected today in ten years' time to achieve things we haven't even thought of, let alone worked out how to achieve.

There was a question around more ethical considerations of data storage, such as energy use and climate change. While it was acknowledged that many cloud storage providers will strive to reduce energy usage to keep costs down, this is an area that needs further consideration into the future. There was also a comment that Kate Hendry (University of Bristol) has been running the Net-Zero Oceanographic Capability[22] series of workshop, looking at examples of good practice.

ACTION: Data Flow PWG to circulate proposed benthic imagery data flow chart

5.3 Imagery Catalogues

Graeme Duncan (JNCC) provided an update on the challenges associated with image catalogues, with a view of AI/Machine Learning as an end customer.

One of the challenges associated with image catalogues is storage. The benefits and challenges of a centralised versus distributed model were explored. Variations include utilising a distributed model but with addition of a broker providing the link to the various data suppliers; a model widely used in online shopping.

Graeme went on to discuss annotation – which is a specific thing at a specific place/area (at a specific time). So, an annotation should have a label and context. There are currently no known standard ways of annotating marine biodiversity images to the level required for AI/machine learning. There are annotation standards out there in existence, such as the W3C annotation model[23]. This would provide a useful start to build-on, but the current model is too generic and technical to use as is. Should the annotation be part of the image or stored as its own entity? The latter is the approach taken by W3C.

Image catalogues can also include reference catalogues. The images within these reference catalogues may also create excellent training datasets for AI/Machine Learning approaches.

Any storage solution will need to work closely and patiently with current software as much as possible, to guide it to ask for the correct information, and understanding the information it receives. This requires interoperability (knowing how to interact) and standards (knowing what you will get as a result), the former is well handled within the general imagery world, with the International Image Interoperability Framework[24]. The latter, standards, are important to describe what to search for, and what you'll get back (and there are still gaps here). There is the potential for the Big Picture Group (BPG) to run a pilot study of a small centralised system, linking a reference catalogue with annotations catalogue to test the concept.

[22] https://www.ukri.org/news/nerc-launches-scoping-project-for-net-zero-oceanographic-capability/
[23] https://www.w3.org/TR/annotation-model/
[24] https://iiif.io/

Looking ahead, we need to:

- fill in the missing gaps in the network, where these centralised image catalogues do not exist
- adapt or develop standards for marine imagery; we can apply imagery metadata as a use-case and develop other standards for the annotation of marine images.
- work towards a pilot framework for image catalogues, as alluded to above.

The image catalogue working group aims to provide the link between the data producers and customers (with machine learning as a 'key' customer).

There was a subsequent discussion about image catalogues and associated challenges, such as their management, conflicting with value. For example, there is low value from an academic perspective to making a catalogue available online, whereas it has high value being online for other users (in an open access context). Some academics have chosen to publish them in journals such as Biodiversity Data Journal, which places them in a restricted online space, of limited use, but provides impact factor to the academic institution. Other organisations such as MBARI have done the opposite, with an open access catalogue, and have funded a staff member dedicated to managing their catalogue. This is still a barrier – many academics have catalogues which are not available externally. One option to increase the sharing of image catalogues, at a reduced cost, may be to provide imagery to DASSH, to be made available online on your behalf. It would be important to ensure appropriate credit/provenance of information is provided for the authors, etc.

When considering submission of training image datasets for AI use, typically AI requires images with the same conditions (lighting, same perspectives, etc.) whereas image catalogues ideally show a variety of views. This may be where metadata can come in useful, allowing AI to specify which types of images may be compatible for their use.

There was a final note that metadata can mean different things to different people, for example computer scientists call annotations metadata, not where the image was taken.

Workshop participants watched a short video created by MBARI introducing the FathomNet[25] project, exploring our ocean using artificial intelligence. This was a prelude to a dedicated live evening session organised with Kakani Katija about the FathomNet project, discussed below in Section 5.4.

5.4 FathomNet – Exploring our ocean using artificial intelligence

An evening session about FathomNet[26], hosted by Kakani Katija, afforded workshop participants an opportunity to discuss this project with one of its principal investigators. A video recording of the FathomNet evening session is available on 'The Big Picture II Webpage'.

The grand vision for FathomNet is to provide a repository for underwater image training data that can be used for the development of computer vision and AI algorithms for automated detection and classification applications. In addition to the training data, FathomNet will also provide a machine learning model repository via GitHub, to enable users of the training data to contribute back to the community and creating an ethos of sharing equally. FathomNet is built as a distributed database using FAIR data principles, aggregating annotations and accessing imagery via public URLs hosted either by the contributor's institution, annotation

[25] https://www.youtube.com/watch?v=PljG1xGn9BE
[26] https://arxiv.org/abs/2007.00114

tool, or other database. Along with a well-defined REST API[27,28,29], the website[30] provides a portal through which contributors and users can interact with and explore the available data. The website currently requires a username/password, but this will not be a requirement once it is formally released.

The FathomNet data use policy has been defined to ensure that annotations and imagery are openly available to a wide array of users while also limiting the use of imagery to the mission of the database. Annotations (which includes the required fields of image URL, position of the localization or bounding box, the concept, and optional fields such as the depth, latitude and longitude of the observation, etc.) are licensed using a Creative Commons Attribution-No Derivatives 4.0 International license. The images are also licensed using a Creative Commons Attribution-Non-Commercial-No Derivatives 4.0 International license, and all of the images may be used for training and development of machine learning algorithms for commercial, academic, and government purposes. All other uses of the images require that potential users contact the original copyright holder, which is included in the Darwin CORE data provided during the submission of a data collection.

Current functionality of FathomNet

- Data can be filtered using concept name, location, taxonomic tree, date, depth, imaging type, contributor institution, verification status, and other fields on the website. Data are also accessed using FathomNet's REST API.
- Concepts can be searched using the text searchbar (known as the VARS Knowledgebase[31]) or expandable concept tree, where the concept tree is currently based on MBARI's Taxonomic Tree[32], and can include exact matches or all descendants.
- Locations can be selected using either the map functionality on the *Explore* page or the drop-down filter that uses defined locations from Marine Regions[33].
- FathomNet currently integrates WoRMS and MBARI's taxonomic trees via their respective APIs. MBARI's taxonomic tree, while used primarily for deep-sea species identification, also includes geological and substrate concepts, allowing the classification of seabed substrata within the same label tree. Additional taxonomic trees can be added to the database (e.g. CATAMI or SMarTaR-ID) provided that they have an API, which allows FathomNet to adapt flexibly to different classification protocols.
- Imaging type is a free-form field, that can be used by the research community to label their data based on their respective imaging community's standards (e.g. IFCB, ISIS).
- Verification status indicates what data have been quality controlled by a FathomNet user with appropriate permissions. Both image annotations and localisations are assessed during the verification process and verified status can be applied to a single bounding box, an image, or an entire collection.

Future functionality of FathomNet

- By establishing partnerships with widely used annotation tool developers (e.g. SQUIDLE+, BIIGLE 2.0, VARS, Tator), we hope to enable automated data pipelines

[27] http://fathomnet.org:8080/rapidoc
[28] http://fathomnet.org:8080/redoc
[29] http://fathomnet.org:8080/swagger-ui
[30] http://www.fathomnet.org
[31] https://www.mbari.org/products/research-software/video-annotation-and-reference-system-vars/knowledgebase/
[32] http://dsg.mbari.org/dsg/browsetree/concept/marine%20organism#marine%20organism
[33] https://www.marineregions.org

for users to submit their training data directly to FathomNet, as well as download FathomNet training data and ML models for their data analysis tasks. Note that these data pipelines are already in place for VARS and Tator.

- Further integration of AI into FathomNet to generate bounding box proposals for contributed image data that are not fully annotated (i.e. not every concept has been detected or classified).
- For those individuals or groups that lack the infrastructure to host their own data, FathomNet could locally host their data. However, under those circumstances, it would require the contributor relinquishing rights to the imagery, which is outlined in the FathomNet Data Use Policy.
- Create more flexible user permissions based on an individual's area(s) of expertise. Enable verifiers to select their areas of expertise and receive notifications anytime relevant data has been added to FathomNet and is available for verification.
- Achieve broader community engagement from the public to interact with, contribute, and verify data by leveraging partnerships with Aquariums and other citizen science organizations (e.g. National Geographic Society, iNaturalist).
- Host annual competitions (see Kaggle[34] and ImageNet[35]) or challenges where groups are charged with developing new algorithms to address novel tasks that are decided on by the FathomNet community.

[34] https://www.kaggle.com/competitions
[35] http://www.image-net.org/challenges/LSVRC/

6 Session 4: Continual progress of the Benthic Imagery Action Plan in day-to-day analysis

This session was facilitated by Stef Golob (JNCC).

A video recording of the session is available on 'The Big Picture II Webpage', along with links to PDF copies of the presentations.

6.1 Identification approaches

Joey O'Connor (JNCC) presented an overview of progress within the Benthic Imagery Identification Approaches Project Working Group (PWG). Joey discussed the identification of all living things visible within the benthic imagery collected. Identification includes the Linnaean approach, but also morphological classifications and using a combined approach, such as through utilising annotation software. The key identification requirements are accuracy and precision, and it is important to be able to make comparisons to mark changes in biological communities. A group interested in identification approaches met during the Big Picture Group workshop in November 2020 (20 attendees) and formed the Project Working Group (PWG) for this area. There was a subsequent PWG scoping meeting on 12th February 2021 to refine the remit of the group and discuss next steps (10 attendees with additional 2-3 interested in passive engagement).

After the presentation, a question/answer and discussion session followed. The group highlighted the importance of diversity and range of experience in this PWG, with attendance ranging from academics, consultants and Statutory Nature Conservation Bodies (SNCBs).

It was highlighted that there is still time to get involved in the PWGs such as this one – the group is scheduled to meet in early April 2021 to consolidate the PWG plan going forwards. Further development of the Epifauna Identification Protocol, discussed in Session 4.3, is a good example of a flagship product coming out from this PWG. There is significant scope to branch out, linking the EIP with other projects, other PWG outputs, and other catalogues/classification systems, etc.

6.2 Enumeration (Quantification) approaches

Henk van Rein (JNCC) and Ross Bullimore (Cefas) jointly presented an overview of developments in the Enumeration Approaches Project Working Group (PWG).

After the presentation, a question/answer and discussion session followed. It was agreed that the Enumeration PWG would be better renamed the Quantification PWG as not all things being considered are individuals.

A note of caution was flagged around drawing conclusions from the study commissioned by JNCC which considered the optimisation of benthic imagery analysis approaches (Moore *et al.* 2019). The aim was to explore the consequence of multiple observers using different image annotation methods on imagery from a sublittoral rocky reef community, to rank the methods, for a specific purpose. The study highlighted a series of challenges. There were some fundamental issues around not using appropriate sample units and not having sufficient replicates. It was recommended to revisit this study, for example perhaps re-analyse the underlying data or design some new aspect in relation to the study, in light of the current Quantification PWG priorities. It was also recognised that much was learnt during the completion of this report, and afterwards, so it is important to bear this in mind. One point

highlighted were the challenges around ensuring SACFOR scales are applied consistently at the data recording/analysis phase.

Pragmatically, it was suggested that perhaps a 'line in the sand' needs to be drawn regarding quantification methods and that old methods may not always be compatible with newer methods; there needs to be acceptance that data collected going forward may not be directly comparable with more historic data.

In relation to sample unit size, the appropriate quantification method may vary dependent on the community under investigation. For example, a sea-pen field may be quantified differently to an encrusting reef community, and differently from a deep-sea sponge garden.

The appropriate quantification method for benthic imagery may vary dependent on the community under investigation.

The ideal scenario would be the development of a decision tree for users to follow, to understand which quantification method would be most appropriate.

There was also a discussion around exploring the quantification of biomass data. For example, instead of using a point to count an object, what about drawing lines/polygons. With this information, someone may then be able to use this in future as a proxy for biomass.

6.3 Annotation software

Mark Burton (Natural Resources Wales) presented an update from the Image Annotation Software Project Working Group (PWG). Mark commented that being able to listen all the presentations and views exchanged over the past two days has been extremely valuable, as it can link with so many of the other PWGs. Nuno Gomes-Pereira *et al* (2016)[36] was highlighted as being a particularly useful paper exploring current and future trends for marine image annotation software.

Mark proposed that rather than trying to prescribe a particular software tool, the PWG could provide a list of desirable features that users should look for in a tool. There are clear links between this PWG and:

- the Benthic Imagery Identification Approaches PWG (particularly the EIP)
- the Benthic Imagery Data Flows PWG
- the Artificial Intelligence (AI) Approaches PWG
- the Quantification Approaches PWG

Picking up on an earlier discussion in Session 4.3 during the EIP, it was flagged that Cefas had developed a CATAMI based label tree after BIG PICTURE 2019. All users are encouraged to continue trialling this and provide feedback. The importance of tying these label trees to existing vocabularies, such as the ICES or the NERC vocabularies was also stressed, as this will assist down the line when these annotations eventually become data.

It was noted that there is a big opportunity here for the Big Picture Group (BPG) to be active, rather than passive, with respect to moulding annotation platforms. Annotation software developers are keen for this type of engagement, and it could deliver some very promising results into the future. There is also a second opportunity for the BPG to really push the adoption and use of annotation software tools in the wider community.

[36] https://www.sciencedirect.com/science/article/abs/pii/S0079661116301240

It was highlighted that back in the BIG PICTURE 2019, very few of the SNCBs were using annotation software, but now, several are using them, and we all better understand the requirement for standardisation across these new tools. There is a huge value in using these annotation tools, especially from a consistency perspective. There was also a call for those newly interested in participating in the Image Annotation Software PWG to get in touch with Mark via the BPG.

Discussion moved to the MAIA tool, within BIIGLE. There is some frustration that it doesn't just work as a 'black box', but it was noted that there isn't a 'silver bullet' with automated annotation at the moment. This is a huge area of research and requires significant amounts of training data. The MAIA tool is also only designed for object detection rather than identification. The only way to improve these tools is through a massive amount of testing and trialling these tools. Things do move quickly though in the AI world, and developments are being made all the time to the various annotation software packages out there.

Crowd counting, a technique used for seabirds, was discussed as potentially offering the ability to count these small clustered organisms, such as cup corals.

iNaturalist[37] was discussed; a global web-based tool for capturing citizen science. A new UK node is soon to be launched, a collaboration between the Marine Biological Association (MBA), the Biological Records Centre (BRC) and the National Biodiversity Network (NBN) Trust[38]. This UK node should provide better access to the back-end data and images.

Taking the example of cell frequency distribution and potential incompatibilities with annotation software, this highlights the importance of ensuring that all PWGs keep communicating with each other, to ensure that the developments being tackled in one group are compatible with other PWG requirements. This would seem to be an ideal action for the Action Plan Coordination Committee (APCC) to maintain a watching brief over.

ACTION: APCC to maintain a watching brief over PWG priorities and activities to ensure compatibility of outputs across all PWGs

The requirement for a UK focussed CATAMI classification was reiterated, with the potential for it to have its own PWG to drive delivery. This was highlighted as one of the outcomes from the EIP work. SMarTaR-ID will also look to incorporate CATAMI alongside taxonomic classifications, in particular for certain groups.

Finally, there was a discussion around the technical aspects and pros/cons of utilising some of the tools within annotation software.

[37] https://www.inaturalist.org/
[38] https://nbn.org.uk/news/introducing-inaturalist-for-the-uk/

7 Session 5: Training, standards and acquisition approaches for the future

This session was facilitated by Kirsten Dinwoodie (JNCC).

A video recording of the session is available on 'The Big Picture II Webpage', along with links to PDF copies of the presentations.

7.1 Deep sea training approaches

Jaime Davies (University of Plymouth) presented an update from the Benthic Imagery Analysis Training Project Working Group (PWG). In particular, she described the work that the University of Plymouth are doing, moving towards imagery-based taxa ID training, and the development of the SMarTar-ID project and Operational Taxonomic Units (OTUs). The PWG have not yet met but are planning to have their first scoping meeting later in March 2021.

It was noted that some of the tasks (originally classed as low priority) in the Benthic Imagery Action Plan (BIAP) are now being taken forward. This was recognised as happening primarily through circumstance rather than design, i.e., these tasks are happening anyway as part of other work.

While SMarTaR-ID is not fully functional as yet, when launched, it will be open access. However, the University of Plymouth image catalogue (Howell & Davies 2010) is available from the Deep Sea Conservation Research Unit (Deep Sea CRU) website and was last updated in 2016[39].

Jaime clarified that the AphiaID, stored in the species table within the SMarTar-ID framework, has a link to WoRMS to keep the nomenclature up to date.

There was a general comment that it was a good idea to work out that something was different to everything else and be able to recognise it again and again and only then do you worry about what to call it in a particular classification system.

A broader discussion ensued about whether it would be possible to integrate the SMarTar-ID framework with the EIP workflow, to become a reference training catalogue making recommendations on what things may be, dependent on image quality. This integration sounded possible, but would be dependent on funding, as the current SMarTar-ID work is being funded through an independent project. Once the pilot version of SMarTar-ID is delivered (due imminently), this could be shared internally within the BPG, where potential linkages could be identified.

7.2 Workflow guidance

Ross Bullimore (Cefas) presented an update from the Benthic Imagery Workflow Guidance Project Working Group (PWG). The purpose of this PWG is to consolidate guidance on the most suitable approaches to using benthic imagery for different purposes across the entire workflow of design, collection, annotation, analysis, interpretation, reporting, storage and dissemination. He outlined how many of the tasks of this PWG are aligned with many of Cefas's own work packages, due to the significant amount of image acquisition and analysis

[39] https://deepseacru.org/2016/12/16/deep-sea-species-image-catalogue/

they complete in-house. The PWG has not yet met but are also planning to have their first scoping meeting later in March/April 2021.

It was acknowledged that progress within this PWG may be sporadic, primarily down to a lack of specific funding, with many of the tasks being 'piggy-backed' on Cefas projects.

There was discussion within the chat window about a need for standards about when to keep/delete imagery. It was clear that different users have different standards of acceptable imagery, but also tied to the fact that image quality requirements vary with task/purpose.

Workshop participants highlighted concerns that some of the tools being developed, as described by Ross, were already developed in other institutions/organisations, and that it was important to avoid duplication of effort – some of these tools could be merged/shared and there are great opportunities here to collaborate/share.

There are clear challenges, in this PWG, but also in others, regarding funding – partly because tasks looking at workflow guidance isn't necessarily exciting or 'sexy' to fund, but it is necessary to get work completed efficiently, and would assist all PWGs. Other organisations have tackled this by bolting small, additional, 'unexciting' work packages onto other 'more exciting' work, to get the work funded. It basically comes down to a marketing strategy to get some of this work funded.

In this session, and others, it was apparent that many organisations have been developing training/user guides for BIIGLE and other annotation software packages. There are clear advantages to pooling this knowledge and developing a single comprehensive resource, rather than each organisation drafting their own guidance. By doing this, there will also be an opportunity to share lessons learned, knowledge and expertise which would benefit all members of the BPG.

ACTION: APCC to coordinate the pooling of knowledge on annotation tools across the Big Picture Group

7.3 Technology reviews

Henk van Rein (JNCC) discussed how these tasks have yet to come together in a formal Project Working Group (PWG) but highlighted the need for reviews of imagery acquisition methods and technology/equipment used. Many organisations are using different acquisition techniques/methods, and when this is the case, trying to standardise the outputs can be challenging.

As mentioned above, no PWG has yet been established yet, as it wasn't considered a priority at the November 2020 workshop. However, this will likely be of increasing priority due to interest and consensus developed at BIG PICTURE II. As part of this, interested organisations need to be identified along with leads. The requirement for this to be a fully-fledged PWG was echoed by many BPG members. Henk encouraged interested members to get in touch with the BPG, and this will help the PWG to get off the ground.

There are two sets of tasks identified:

- to ensure reviews are updated periodically; these include equipment reviews, data acquisition approaches and data processing reviews. Rather than prescriptive standards, a decision tree on the best system to use may be more appropriate.

- Develop quality control recommendations for survey.

There is a need for a more strategic approach for longer term benthic imagery acquisition to fit into future workflows.

There was a discussion around potential of SNCBs using the same technologies to collaborate on training and standards for that particular piece of kit. Collaboration is already happening sporadically across the various SNCBs, which is great to see.

It was also recognised that there is a lot of existing guidance already out there[40], although some, such as the Mapping European Seabed Habitats (MESH)[41] Recommended Operating Guidelines, may now be quite dated, and need updating.

Other work in relation to this work area is already underway. For example, Ulster University are involved with an Interreg project, along with SAMS and AFBI, called Marine Protected Areas Management and Monitoring (MarPAMM)[42], part of which will investigate non-invasive methods for monitoring the sub-tidal component of Natura 2000 sites, focusing on low cost and novel technologies.

Ross Bullimore suggested that the Benthic Imagery Workflow Guidance Project Working Group (PWG) could go some way to achieving some of these tasks, such as a shared resource of references or libraries. Having a central place where people can access this type of information could go a long way to addressing this need.

It was pointed out that there are still many 'wins' that can be got by changing the way existing equipment is used, rather than buying the latest and greatest. For example, by focussing on good lighting or ensuring your lasers are parallel.

While many organisations have also been collected environmental data whilst collecting seabed imagery, getting the most out of this has been challenging, partly because of the linkages between uncalibrated point data and wider models; some consideration on how these opportunities may be improved upon would be timely.

In the interim, while reviews are underway, having user stories from people explaining when things have gone wrong as well as right, with respect to seabed imagery equipment setups, could provide some quick wins, and help people avoid making the same mistakes.

A question was posted around whether changes in camera resolution over time should be an important consideration for monitoring? While they should be a consideration, as long as the resolution is suitable for the study design, then it shouldn't be an issue. For example, as long as you only look at changes that can be detected equally across all time points (and across both low resolution and high-resolution cameras).

[40] https://marine-sampling-field-manual.github.io/files/NESP-field-manuals-V2.pdf
[41] https://www.emodnet-seabedhabitats.eu/resources/mesh-archive/
[42] https://www.mpa-management.eu/

8 Session 6: The future of The BIG PICTURE

This session was facilitated by Henk van Rein (JNCC).

A video recording of the session is available on 'The Big Picture II Webpage', along with links to PDF copies of the presentations.

This session was focussed on summarising the past three days of the BIG PICTURE II workshop, and picking out the salient points, particularly with respect to future funding opportunities, global outreach opportunities and what the key next steps, and challenges, will be for the Big Picture Group.

8.1 Summary of funding opportunities to explore

There were extensive discussions over the three days regarding funding opportunities, as these were recognised as the key factor limiting progress in achieving the Benthic Imagery Action Plan.

In order for a fully strategic funding plan to be established, an understanding of the resource (financial and time) needed to achieve the Benthic Imagery Action Plan (BIAP) is required. Once any reviews/updates to the BIAP have been completed (discussed in Section 8.3), the Action Plan Coordination Committee (APCC) could task the Project Working Groups (PWG) to 'cost out' their work tasks, noting and acknowledging where tasks have been 'piggy-backed' onto external work priorities being led by partners. This information can then be fed back to the APCC to provide a true cost of delivering the BIAP. Only then can fully strategic decisions be made by the APCC, based on discussions around what type/level of funding is needed/achievable, and which areas could and should be prioritised for funding.

ACTION: APCC to task PWGs to 'cost out' their work tasks, providing the APCC with the information needed to make fully strategic decisions as to which areas could and should be prioritised for funding.

Timing in this plan is also critical; workshop participants discussed the fact that tasks will need to be achieved in a certain order to progress the BIAP as efficiently as possible and to avoid duplication of effort and having to revisit work. The latter may be inevitable recognising how the area of benthic marine imagery is evolving over time, but it would still be prudent to try and minimise this as far as possible.

It was discussed that some of the larger work tasks could be broken down into smaller, more achievable chunks, which could be worked on using more agile methods, perhaps across PWGs. This would mean that no one PWG is required to build a complete solution to a problem in one go, but that each group could build on the small steps made by other groups. This process may suit those areas that are more challenging to seek funding for.

While some aspects of the work undertaken by the Big Picture Group (BPG) may be relatively easier to seek funding for, it was acknowledged that other aspects, such as data infrastructure and workflow projects, are not as exciting or attractive to funders. For these, it may be necessary to piggy-back these tasks onto other bigger projects.

From the discussion sessions, a strategy which has had some success in the past for securing funding has been the consolidation of ideas into packages of work, tailored to be better aligned to meet particular funding calls. With those BPG work areas which are considered a high priority for funders like NERC and Government, such as Artificial

Intelligence (AI) and autonomy, considering the data infrastructure, standards and workflow requirements for these funding proposals, and costing them into proposals, will ensure that some resource is available if these projects are funded. Alternatively, those work areas that are challenging to fund, could be grouped together, to justify more funding, although generally this approach has had mixed success. Regardless of approach, the BPG agreed that how these proposals were branded/marketed was key to influencing their success.

It should be recognised that the strength of the BPG lies in its diverse membership. During the workshop discussions, it became apparent that individuals within the BPG had ideas to pursue funding opportunities. Members should be encouraged to explore these, ensuring that the PWG is kept updated with progress. PWG leads can then feed this information into the APCC; it will be essential that the APCC maintains oversight of these funding opportunities as they progress, in order to develop and maintain a strategic overview of progress.

Ideas for potential funding sources discussed included:

- NERC Marine Facilities Advisory Board – annual call for capital infrastructure/facilities/instrumentation projects.
- Centralised funds via central government.
- Explore calls for funding via UK Research & Innovation (UKRI)[43]
- NERC Digital Environment Call – while the initial call is over, there is the potential for this to be continued
- NERC Strategic Programme Fund[44] – takes ideas from scientific community for funding call topics.
- UKRI funding call for autonomous remote sensing technologies.
- UKRI / National Science Foundation[45] funding.
- Decade of Ocean Science[46] – part of this is about enabling new technologies, which ties in with some of the BIAP work areas.

The strength in an application to any of the above calls would be through the diverse range of interests that the BPG represents; it represents a coherent, united voice across the marine benthic imaging community, rather than independent organisations seeking funding alone. Funding would be relevant to all PWGs but may be more appropriate for an individual (or small group of) PWG to lead. Again, coordination through the APCC would be critical to ensure that funding calls are aligned with agreed BIAP targets.

8.2 Summary of global outreach opportunities

During the workshop, a range of global engagement opportunities were discussed.

[43] https://beta.ukri.org/opportunity/
[44] https://nerc.ukri.org/funding/available/programmes/
[45] https://www.nsf.gov/funding/
[46] https://en.unesco.org/ocean-decade/about

The BIG PICTURE II workshop brought together stakeholders and scientists from across eight countries including Belgium, Netherlands, Canada and the United Arab Emirates, highlighting the reach (outside the UK and Devolved Administrations) that the Big Picture Group (BPG) has developed since its inception.

A dedicated session (Section 5.4) with Kakani Katija from FathomNet (MBARI) afforded workshop participants an opportunity to understand more about this exciting project, further cementing the links between the BPG and this initiative, as well as providing individuals and organisations a chance to connect and discuss collaborations.

With Artificial Intelligence and Annotation software discussed as two areas of future development for the BPG, there may be opportunities to connect wider with relevant specialists by hosting a hackathon, or some form of competition, which could pose challenges to the computer science community as well as trying to answer some of the BPGs own challenges within its Project Working Groups.

Because of the BPGs diverse range of membership, there may be opportunities for these milestone events every one-two years, such as BIG PICTURE and BIG PICTURE II, to attract a wider audience as the PWGs progress their work, expanding outreach opportunities.

8.3 Summary of next steps and future challenges in the BIG PICTURE journey

Since its inception after BIG PICTURE 2019, the Big Picture Group (BPG) community has grown significantly, bringing together individuals from a range of backgrounds, but with a common goal of seeking collaborative solutions and opportunities, while considering how to incorporate new technologies, such as computer vision and machine-learning.

The Benthic Imagery Action Plan (BIAP) has been the central spine that holds the BPG together. It was noted that the BIAP was never considered to be static, and during the various discussions over the past three days, it was agreed that some parts of the BIAP may need to be reviewed and updated. The Project Working Groups (PWGs), after reflecting on the workshop discussions that took place, are best placed to tackle this over the next six months / one year. As discussed in Section 8.1 above, effective and strategic coordination through the Action Plan Coordination Committee (APCC) is essential.

ACTION: APCC to task PWGs with reviewing all work tasks from the BIAP under the remit of their group prior to progressing any further work.

The subject of benthic purposes, initially discussed in Section 4.1, was raised again during the summary stage; out of six core purposes considered, a handful were taken forward for considering within the Quality Assurance Framework (QAF). While it was useful to constrain thinking to this handful, the importance of not forgetting the others was stressed, and each Project Working Group (PWG) should be mindful of these purposes as they move forward with their plans.

With the wide range of PWGs in operation, some coordination in the future, via an active role, perhaps through the Action Plan Coordination Committee (APCC) is fundamental to ensuring that each PWG moves in the same direction (discussed in Section 8.1 above). This type of coordination/facilitation may require some resource, such as through government funding. To date this funding has been provided by Defra and has been essential in getting things to where they are now. Overall, there was a consensus view that some form of oversight will be crucial as the BPG moves forward into the future.

Communication, both internally with the BPG and externally to stakeholders, was a common theme across all PWGs. The need for PWGs to communicate effectively between themselves was reiterated regularly throughout the three days and echoed by the majority of participants. One participant summed it brilliantly when they said that “If everything talks, and everything connects, everything just works”. The concept of standards, services and systems was also discussed (Figure 4). While the BPG is not quite there yet, with respect to the ideal scenario laid out in Figure 4, it is well on the way and a solid foundation has been set. For example, since the BIG PICTURE 2019 workshop, there are now representatives from MEDIN and the BODC on board in addition to others.

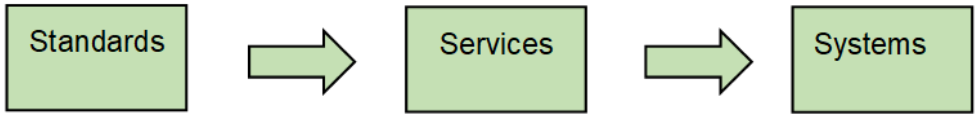

Figure 4: Standards lead to Services, which lead to Systems. You cannot build systems if you don’t have the standards or services to underpin them.

Looking forward, it was agreed that having a shared space to store / work / collaborate / chat / communicate will be critical for the BPG, moving forward. Microsoft Teams has been used to date by the BPG and generally this has worked well, although there are some known issues when working with people who don’t have fully fledged Teams licences.

With the Epifauna Identification Protocol (EIP) forming an integral part of the QAF, and being viewed as a flagship product, potential links were made between the EIP and other PWG outputs during the various discussions over the three days, particularly with respect to training and annotation software. The EIP was also linked to development of a UK focused CATAMI framework, and within the EIP itself, there were various ways this could be developed further, including considering additional taxonomic groups (such as algae) and expansion to include deep sea taxa. Linkages between SMarTaR-ID, once it comes online, and the EIP will be important to provide a tool which works both for shallow and deep-water habitats. In addition, having reference images linked to the EIP (showing a range of image qualities) would be an extremely valuable development for this tool.

Since the original BIG PICTURE workshop in 2019, there have been significant developments in the field of annotation software. A key goal from BIG PICTURE II was to encourage more members to get actively involved in using annotation software. Rather than dictate which software to use, a preferred approach was for the Annotation Software PWG to develop a decision tree to guide users, highlighting how and when they are appropriate to use. Engaging with annotation software developers was highlighted as an important way to influence the process; the Big Picture Group as a centralised entity is ideally situated (via the PWGs) to engage with these developers, rather than people/organisations engaging individually. Finally, an important element for the Annotation approaches PWG to address will be the storing/archiving of these annotations – this was discussed briefly in the workshop but will be taken up by the PWG when then they meet for the first time in April/May 2021.

An important point raised relates to the sheer amount of information out there with respect to marine benthic imagery. During the workshop, massive amounts of information have been shared, which is fantastic, and is effectively demonstrated by the number of links to additional information provided throughout this workshop report. However, there is also a danger of the BPG being overwhelmed by all this information sharing and a valuable point was made in the summary stages of restricting relevant information to the PWGs, who could then sift through and separate/summarise all the useful information.

While acknowledging that PWGs have an important role in assessing the best way of carrying out a task for one purpose/multiple purposes, the PWGs could also have a role in reviewing the work, which would provide an opportunity for workflows/procedures to be refined/improved for next time.

The development of training and standards forms a core part of the BIAP. The need for these, particularly in new areas such as image annotation were stressed. It was acknowledged by the group that a more coordinated approach would benefit all PWGs – for example a number of different organisations have been developing user guides for annotation software suites such as BIIGLE – consolidation and collaboration to create a master set of training material and guides would have massive benefit, whilst at the same time ensuring that we are not 'reinventing the wheel'.

The value of citizen science programmes, such as SeaSearch[47] and iNaturalist, was discussed. Existing citizen science training programmes such as SeaSearch are already engaging with DASSH/MEDIN through data flows, and it was highlighted that these existing training programmes could be expanded, creating bespoke training courses for the Epifauna Identification Protocol (EIP).

Finally, there was a general call for additional participation in the PWGs – some groups will be having their inaugural scoping meetings in the coming months, so now is a great time to get involved and help to drive forward developments in the field of marine benthic imagery innovation through collaboration. Interested individuals/groups were encouraged to reach out to the BPG, who will link them up to the relevant PWG lead.

[47] http://www.seasearch.org.uk/

9 Concluding remarks

Henk van Rein (JNCC), as workshop Chair, brought the proceedings to a close at the end of Session 6. Of all the challenges facing the Big Picture Group (BPG), securing adequate funding has been, and will continue to be a pressing issue, and has proved the primary limiting factor in progressing the Benthic Imagery Action Plan (BIAP). Up until this point all central coordination of the BPG has been funded by Defra. However, as time goes forward new funding routes must be explored to continue this fantastic work. Funding was an overarching theme across the three-day event, with new funding opportunities discussed at every opportunity, and a strong set of options are now available for the BPG to explore into the future.

All agreed that it was also important not to lose sight of the fact that within the BPG, there are a lot of individuals working together to deliver the BIAP, and so much has already been achieved through *pro bono* work alone, in addition to funded work.

Developments in the world of annotation software (to improve comparability of data analysis) and Artificial Intelligence (AI) have been some of the enduring themes throughout much of the workshop, and it was recognised that training will be key to its uptake.

The BPG reflected on the various connections between individuals, organisations and work streams made back in 2019 during the work of the plan development group. It is important that this information is not lost or forgotten, and for those who have not engaged, there may be advantages to revisiting and encouraging these connections, particularly as they have already been acknowledged as important by the BPG.

Finally, the BIG PICTURE II workshop team thanked all the participants for sharing their time, insight, knowledge and ideas in a constructive and progressive way. The workshop met its objectives as a milestone event, sharing progress with the BIAP across the various PWGs, and discussing ideas for future funding and global outreach possibilities. There was also an important acknowledgement of the future challenges which the BPG may face, but recognition that a great start has been made in getting the BPG to where it is today.

10 References

Blowers, S., Evans, J. & McNally, K. 2020. Automated Identification of Fish and Other Aquatic Life in Underwater Video. Scottish Marine and Freshwater Science Vol 11 No 18. Marine Scotland Science, ISSN: 2043-7722. DOI: 10.7489/12333-1.

Howell, K.L. & Davies, J.S. 2010. Deep-sea species image catalogue. DeepSeaCRU, Marine Biology and Ecology Research Centre, University of Plymouth. On-line version 2, 2016

Howell, K.L., Davies, J.S., Allcock, A.L., Braga-Henriques, A., Buhl-Mortensen, P., Carreiro-Silva, M., Dominguez-Carrió, C., Durden, J.M., Foster, N.L., Game, C.A., Hitchin, B., Horton, T., Hosking, B., Jones, D.O.B., Mah, ~C., Marchais, C.L., Menot, L., Morato, T., Pearman, T.R.R., Piechaud, N., Ross, R.E., Ruhl, H.A., Saeedi, H., Stefanoudis, P.V., Taranto, G.H., Thompson, M.B., Taylor, J.R., Tyler, P., Vad, J., Victorero, L., Vieira, R.P., Woodall, L.C., Xavier, J.R. & Wagner, D. 2019. A framework for the development of a global standardised marine taxon reference image database (SMarTaR-ID) to support image-based analyses. PLoS ONE 14(12): e0218904. https://doi.org/10.1371/journal.pone.0218904

Moore, J., van Rein, H., Benson, A., Sotheran, I., Mercer, T. & Ferguson, M. (2019). Optimisation of Benthic Image Analysis Approaches, JNCC Report, No. 641, JNCC, Peterborough, ISSN 0963-8091.

Appendix 1: Workshop participants

A list of workshop attendees and affiliation.

Name	Affiliation
Gillian Allen	Gardline Limited
Magnus Axelsson	Natural England
Lin Baldock	Marine Ecological Solutions Ltd
Michael Ballard	Fugro GB Marine Limited
Alison Benson	Envision Mapping Ltd
Eloise Boblin	Envision Mapping Ltd
Erin Browne	University of Plymouth
Ross Bullimore	Cefas
Mark Burton	Natural Resources Wales
Amy Cartwright	University of Plymouth
Alfonso Nebra Costas	Instituto Español de Oceanografía
Emma Curtis	University of Southampton
Matthew Curtis	Cefas
Jaime Davies	University of Plymouth
Steven Dewey	Seastar Survey Ltd
Kirstin Dinwoodie	JNCC
Graeme Duncan	JNCC
Jennifer Durden	National Oceanography Centre
Rachael Eyley-Roberts	Gardline Limited
Nicola Foster	University of Plymouth
Clive Fox	Scottish Association of Marine Science
Mike Fraser	Natural England
Chloe Game	University of East Anglia
Andrew Gates	National Oceanography Centre
Neil Golding	Aquarius Survey & Mapping

Name	Affiliation
Estefania Golob	JNCC
Laura Hearnden	Gardline Limited
James Highfield	Natural England
Terry Holt	Independent marine consultant
Samantha Hormbrey	Eastern Inshore Fisheries and Conservation Authority
Tammy Horton	National Oceanography Centre, Southampton
Kerry Howell	University of Plymouth
Veerle Huvenne	National Oceanography Centre, Southampton
Robyn Jones	Ocean Ecology Limited
Danae Kapasakali	Royal Belgian Institute of Natural Sciences
Georgios Kazanidis	University of Edinburgh
James King	Natural Resources Wales
Jenny Kuru	Fugro
Rebecca Langton	Marine Scotland Science
Dan Lear	Marine Biological Association
Charles Lindenbaum	Natural Resources Wales
Tim Mackie	Department of Agriculture, Environment and Rural Affairs (Northern Ireland)
Fionnuala McBreen	JNCC
Ryan McGeady	Ulster University
Chris McGonigle	Ulster University
Niall Mcleod	Marine Scotland Science
Giacomo Montereale-Gavazzi	Royal Belgian Institute of Natural Sciences
Jon Moore	Coastal Assessment, Liaison & Monitoring Ltd
Christine Morrow	MERC Consultants
Barbara Neves	Department of Fisheries and Oceans, Canada
Tammy Noble-James	Cefas
Claude Nozeres	Department of Fisheries and Oceans, Canada

Name	Affiliation
Joey O'Connor	JNCC
Jen O'Dell	Seastar Survey Ltd
Tabitha Pearman	University of Southampton
Bernard Picton	National Museums (Northern Ireland), MERC Consultants and Seastar Survey Ltd
Nils Piechaud	National Oceanography Centre, Southampton
Madeleine Purver	University Centre of the Westfjords, Iceland
Katie Searle-Evans	Mott MacDonald
Shrini Shrinivaasu	SGS Gulf Limited
Lucy Shuff	Gardline Limited
Rona Sinclair	NatureScot
Helen Snaith	National Oceanography Centre
Paris Stefanoudis	University of Oxford
David Stirling	Marine Scotland Science
Alison Tamkin	APEM Ltd
Jessica Taylor	JNCC
Stephen Thompson	Eastern Inshore Fisheries and Conservation Authority
Udo van Dongen	Bureau Waardenburg, Netherlands
Henk van Rein	JNCC
Jessica Vevers	Swansea University
Oscar Ward	Marine Space
Karen Webb	JNCC
Adrian Weetman	Marine Scotland Science
Nadescha Zwerschke	JNCC

Appendix 2: Workshop agenda

Day one: Tuesday, 2 March

Time	Session
0930 – 1030	Welcome, practicalities, working style and programme Reviewing the context and background, aims and objectives of the workshop
1030 – 1100	Tea & coffee break
1100 – 1200	**Session 1: The journey from BIG PICTURE to BIAP to Big Picture Group to Project Working Groups (PWGs)** Session facilitator: **Kirsten Dinwoodie** Presentations (by Henk van Rein; Joint Nature Conservation Committee): • Development of a UK Benthic Imagery Action Plan • Formation of the Big Picture Group • Project Working Groups and looking ahead Questions on presentations, funding opportunities, global outreach and next steps
1200 – 1300	Lunch
1300 – 1500	**Session 2: Early successes of BIAP: Benthic imagery purposes, Quality Assurance Framework, Epifauna Identification Protocol** Session facilitator: **Joey O'Connor** Presentations: • Development of standard benthic imagery purposes – Henk van Rein (JNCC) • Development of Quality Assurance Framework – Jess Taylor (JNCC) and Graeme Duncan (JNCC) • The Epifauna Identification Protocol – Jess Taylor (JNCC) Questions on presentations, funding opportunities, global outreach and next steps
1500 – 1530	Tea & coffee break
1530 – 1600	Day one summaries and close

Day two: Wednesday, 3 March

Time	Session
930 – 1000	Arrival and plan for the day
1000 – 1200	**Session 3: BIAP progress in the digital realm: AI, Data flows and Catalogues** Session facilitator: **Jess Taylor** Presentations: • Artificial intelligence approaches for benthic imagery – Kerry Howell (University of Plymouth) • Benthic imagery data flows – Dan Lear (Marine Biological Association of the UK / Data Archive for Seabed Species and Habitats) • Benthic imagery catalogues – Graeme Duncan (JNCC) Questions on presentations, funding opportunities, global outreach and next steps
1200 – 1300	Lunch
1300 – 1500	**Session 4: Continual progress of BIAP in day-to-day analysis: Annotation, Identification and Enumeration** Session facilitator: **Steph Golob** Presentations: • Identification approaches used for benthic imagery data – Joey O'Connor (JNCC) • Enumeration approaches used for benthic imagery data – Jon Hawes (Centre for Environment, Fisheries and Aquaculture Science) • Imagery annotation software – Mark Burton (Natural Resources Wales) Questions on presentations, funding opportunities, global outreach and next steps
1500 – 1530	Tea & coffee break
1530 – 1600	Day two summaries and close

Day three: Thursday, 4 March

Time	Session
930 – 1000	Arrival, recap Plan for the final day
1000 – 1200	**Session 5: Training, standards and acquisition approaches for the future** Session facilitator: **Kirsten Dinwoodie** Presentations: • Training approaches for deep sea imagery – Jaime Davies (University of Plymouth) • Development of benthic imagery workflow guidance – Ross Bullimore (Cefas) and Mike Fraser (Natural England / Environment Agency) • Overview of benthic imagery acquisition approaches – Henk van Rein (JNCC) Questions on presentations, funding opportunities, global outreach and next steps
1200 – 1300	Lunch
1300 – 1500	**Session 6: The future of The BIG PICTURE** Session facilitator: **Henk van Rein** Summaries of all funding, global outreach and next steps from all five sessions presented by session facilitators (Kirsten Dinwoodie, Joey O'Connor, Jess Taylor and Steph Golob) Break out discussions of key topics focusing on future delivery of Benthic Imagery Action Plan
1500 – 1530	Tea & coffee break
1530 – 1600	Workshop close