

ICES DIG REPORT 2014

ICES SCIENCE COMMITTEE

ICES CM 2014/SCICOM:02

REF. SCICOM, ACOM

Report of the ICES Data and Information Group (DIG)

16-28 May 2014

Copenhagen, Denmark



ICES
CIEM

International Council for
the Exploration of the Sea

Conseil International pour
l'Exploration de la Mer

International Council for the Exploration of the Sea Conseil International pour l'Exploration de la Mer

H. C. Andersens Boulevard 44–46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2014. Report of the ICES Data and Information Group (DIG), 16-28 May 2014, Copenhagen, Denmark. ICES CM 2014/SCICOM:02. 39 pp.
<https://doi.org/10.17895/ices.pub.8328>

For permission to reproduce material from this publication, please apply to the General Secretary.

The document is a report of an Expert Group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

© 2014 International Council for the Exploration of the Sea

Contents

Executive summary	1
1 Opening of the meeting.....	3
2 Adoption of the agenda	3
3 Progress on ICES Data Plan	3
3.1 Regional facilitation.....	3
3.2 International Standards and Interoperability	4
3.3 Knowledge transfer and professional development.....	5
4 Information exchange with ICES Data Centre.....	5
4.1 Oceanographic data.....	5
4.1.1 Operational oceanographic products.....	5
4.1.2 ICES report on ocean climate (IROC).....	5
4.2 DATRAS (Database on Trawl Surveys).....	5
4.3 Standard graphs.....	5
4.4 Eggs and larvae database	6
4.5 INSPIRE services.....	6
4.6 Other topics	6
5 Digital data citation within ICES	7
5.1 Provide options for data citation within ICES.....	7
5.1.1 New datasets	7
5.1.2 Historic datasets.....	8
5.1.3 Updates and time-series.....	8
5.2 Joint Declaration of Data Citation Principles.....	8
5.3 Update of 'Introduction to Digital Citation'	9
6 Use of VMS data within and outside ICES.....	10
6.1 VMS Data policy	10
6.2 Proposed procedure for working with VMS/logbook data	10
6.3 Publishing data	10
7 Availability, versioning and visibility of ICES Data Guidelines.....	11
7.1 IODE/JCOMM/ICES clearinghouse.....	11
7.2 Next steps	11
7.3 Outreach.....	11
7.4 Follow-up.....	12
8 Other topics.....	12
8.1 ACOM representation.....	12

8.2	EMODnet	12
8.3	OSPAR request.....	12
8.4	Follow-up of actions and recommendations.....	14
8.4.1	Action list	14
8.4.2	Recommendations from DIG	16
8.4.3	Recommendations to DIG	17
8.5	Social media.....	18
9	Between May 2014 and May 2015	19
9.1	Meetings.....	19
9.2	Intersessional work.....	19
	Annex 1: List of participants.....	20
	Annex 2: Agenda.....	22
	Annex 3: Terms of reference for next meeting.....	24
	Annex 4: Recommendations and Action items	25
	Annex 5: Data Plan	29
	Annex 6: Data Guidelines.....	38

Executive summary

The Data and Information Group (DIG) met in Copenhagen, 26–28 May 2014. 16 people representing 9 different countries, representatives from OSPAR and HELCOM Secretariats, Head of ICES Data and Information, and ca. 9 members of the ICES Data Centre joined the meeting.

During the 2014 plenary meeting, the group reviewed the progress related to the ICES Data Plan and furthermore met up with ICES Data Centre, and discussed strategic issues related to VMS data, digital data citation, and the ICES Data Guidelines.

ICES Data Plan

On most topics scheduled for 2014 progress had been made.

ICES Data Centre

New tools and facilities that have been developed by the ICES Data Centre were presented and discussed.

The web application of the DATRAS database has been improved. One result is three application programme interfaces (APIs) that provide direct access to DATRAS data from other software i.e. within R programmes.

The standard stock assessment graphs that go to the ICES assessment reports can be selected for certain years, stocks and ecoregions and downloaded from the website. Only published and validated stock graphs can be downloaded. Stock coordinators can upload the data, customise the output, and publish the results.

ICES Data Centre received Estonian Gulf of Riga Fish Larvae dataset to add to the eggs and larvae database.

To answer to INSPIRE directive readiness, an online available catalogue about services and data products hosted by ICES has been developed using a simple open source ESRI Geoportal. This service will be integrated with the catalogue of datasets (<http://geo.ices.dk/geonetwork/srv/en/main.home>) when development is complete.

Under the science committee (SCICOM) a Call for operational oceanographic products and services was issued, that shall among other potential uses feed into the integrated assessment groups (e.g. WGIAB, WGINOSE).

Digital Data Citation

Journal citation is an accepted and well-established practice that gives due credit to scientific work done by scientists, and also signposts where others can find this information. In a similar way, citation of data can give proper credit to data providers who have made data available to the scientific community, while also providing a mechanism for tracing back scientific knowledge to the data that underpins it. The Introduction to Digital citation was presented to SCICOM in March 2014. This document is a living document and DIG discussed if updates were necessary.

The 8 principles of the Joint Declaration of Data Citation were reviewed in order to advise ICES on the adoption of the Declaration. The principles are grouped so as to facilitate understanding, rather than according to any perceived criteria of importance.

VMS data

ICES has produced a VMS data policy stating the conditions of use for experts, to be signed by the chair of groups using the data. DIG reviewed this document and recommends that additionally each expert working with VMS/logbook data signs it before getting access to the data and that all signatures are collated to the same document.

ICES Data Guidelines

A Workshop on ICES Data Guidelines was scheduled for 2014. The workshop was cancelled and re-placed by an intersessional DIG activity. In March 2014, a meeting took place between IODE and ICES to discuss ICES Data Guidelines in relation to the IODE/JCOMM/ICES clearing house. The ambition is to have all data and data management related documents, including Manuals from instrument manufacturers and software (versions), available through the clearinghouse. It will thus provide an additional access point to the ICES documents, next to the ICES publications library.

1 Opening of the meeting

The Data and Information Group (DIG) met in Copenhagen, 26-28 May 2014. 16 people representing 9 different countries, representatives from OSPAR and HELCOM Secretariats, Head of ICES Data and Information, and ca. 9 members of the ICES Data Centre joined the meeting.

The participant list is in Annex 1.

2 Adoption of the agenda

The Terms of Reference of the group were as follows:

- a) Review priorities on the Data Centre action list
- b) Provide guidance and feedback to the ICES Data Centre
- c) Advise on other data regulations and their impact on ICES Data Strategy, ICES Data Policy
- d) Review output from offspring groups (WKIDG, DUAP, LinkedIn Data and Information Forum) if relevant
- e) Promote new technologies and data management infrastructure development

Products (e.g. updated data management guidelines, reviews of ICES Data Strategy, ICES Data Policy, etc.) from the meeting as well as a written report to SCICOM will be delivered before 15 July 2014.

The group reports to SCICOM during the SCICOM mid-term meeting March 2014 as well as the SCICOM meeting at ICES ASC 2014. The group reports to ACOM by correspondence and via the ACOM representative.

Main topics discussed during the meeting were:

- i) Progress on ICES Data Plan (Chapter 3, Annex 5, related to ToR a)
- ii) Data service provision: Information exchange with ICES Data Centre (Chapter 4, related to ToR b)
- iii) Digital citation, mainly focusing on data, within ICES (Chapter 5, related to ToR e)
- iv) Use of VMS data within and outside ICES (Chapter 6, related to ToR c)
- v) Availability, versioning and visibility of ICES Data Guidelines (Chapter 7, Annex 6, related to ToR d)

The full agenda of the meeting is in Annex 2.

3 Progress on ICES Data Plan

All items having a deadline or milestone in 2014 were discussed by DIG and the status was added to the table. This chapter only contains some general discussions and the actions resulting from the discussions. The data plan tables including the status can be found in Annex 5.

3.1 Regional facilitation

The relation between the work carried out by ICES and Member State responsibilities was discussed and explained. The national representatives discussed the MSFD issues

with ICES. The member states have to make the underlying data available used in indicators and assessment in the MSFD, but as regional coordination is often limited, the European Commission has asked ICES to take a role in that process through the Common Implementation Strategy (CIS).

- Regional operational products for Marine Strategy Framework Directive (MSFD) and Data Collection Framework (DCF)/Multi-annual programme (DC-MAP)
 - (a) MSFD workflow: Collaboration between ICES Data Centre and Regional Sea Conventions/other organisations with respect to MSFD (WISE-Marine production process). This assumes a good flow of data/data harvesting into the data centre, and this can imply more resources in certain data types where data are not readily provided. Documentation on OSPAR methodology for D8 (Hazardous substances) is still missing. Chris Moulton and Neil Holdsworth will speed up the first draft of the documentation on the methodology.
 - New processes/products from existing data Advisory and Science with respect to MSFD: As marine litter is never monitored in a dedicated monitoring programme, it is difficult to coordinate the dataflow. Groups coordinating offshore surveys on which litter is being collected regard it as a by-product for which they do not have the full responsibility after the survey. DIG encourages the ICES initiative to set-up a dedicated ICES working group on marine litter. Although the working group should cover all scientific aspects of marine litter, it is likely that in terms of coordination of collection of data/methods the group's task should be restricted to offshore marine litter. The WGZE 2014 report might be of interest with respect to micro plastics.
- End-to-end workflow for scientific advice production:
 - Tool for drafting, approving and publishing Advice (RA-CMS), linking to data outputs from Expert groups (connecting the scientific reports to advice production): the system was delayed due to a mismatch between the database specifications and user interface and the real-life situation. DIG considers it very important to prevent these situations as time, money and effort is spent to develop a system. DIG recommends that ACOM and SCICOM clearly communicate to all their expert and working groups that requests for new data (storage) facilities need the best description possible and the request should be sent to ICES Data Centre using a request form as available via <http://ices.dk/marine-data/guidelines-and-policy/Pages/Requesting-data-from-ICES.aspx>
 - Mobilising aquaculture specific data: Ingeborg de Boois to check with chair(s) of WGAQUA if they see a need for an aquaculture database. If or when ICES Data Centre starts to develop an aquaculture database, experiences from Norway and Scotland might help in the development process.

3.2 International Standards and Interoperability

- Encouraging the broader use of ICES datasets by implementing IODE quality flagging schema
 - DIG commented on the reasoning behind the use of the IODE quality scheme. DIG recommends that the ICES Data Centre finds the scheme

fitting best to its needs, and make sure the chosen scheme can be mapped to other existing data quality schemes.

3.3 Knowledge transfer and professional development

- Training and reference guides for scientists and data managers: Ingeborg de Boois, Neil Holdsworth and Jens Rasmussen will explore the options for and the focus of a training course 'Making the most of ICES Data'.

4 Information exchange with ICES Data Centre

4.1 Oceanographic data

4.1.1 Operational oceanographic products

Under the science committee (SCICOM) a Call for operational oceanographic products and services was issued, that shall among other potential uses feed into the integrated assessment groups (e.g. WGIAB, WGINOSE). This call is essential, as there are existing products/expertise that could be lever-aged, and as the Data Centre cannot do it fully on its own, but help and input from the ICES network is needed. Existing data sets and products like those from the Working group on Oceanic Hydrography (WGOH) and Zooplankton expert groups are considered in this call. However, ICES wants to move away from a situation, where the presence of such data sets is basically known, but it is not known how these data sets are operated/quality controlled and the regularity of updates. Therefore this formal Call was issued to move towards a more secure footing for the provision of these data services to the ICES assessment groups.

4.1.2 ICES report on ocean climate (IROC)

The digital IROC is available on the web. The products can be browsed online.

4.2 DATRAS (Database on Trawl Surveys)

The web application of the DATRAS database has been improved. One result is three application programme interfaces (web service APIs) that provide direct access to DATRAS data from other software i.e. within R programmes. When the web services are complete, ICES Data Centre will communicate this via LinkedIn, ICES news and expert group meetings and working groups.

A question was raised whether providing the same data under DATRAS as well as via EMODnet-Biology will end in duplicates, but EMODnet -Biology is skimming the surface of datasets (species location and abundance information) and for more detailed data there will be a point-back to DATRAS. However, it was discussed that there are Norwegian trawl surveys that do not form part of the DATRAS system, and it should be encouraged that these data are made available to DATRAS.

DATRAS data products that apply species codes had an extra column implemented for the valid species WoRMS codes. The uploaded codes (including the invalid ones) will be kept in original data sets, but introduction of the actual valid ones will improve the species-based data analysis.

4.3 Standard graphs

A new, upgraded Stock Assessment Graphs database went live earlier this year. Data from stock assessment can now be uploaded directly, by the help of an excel template that converts the data in xml; these data are used to produce the standard stock graphs. Specific settings can be changed by the working group on a stock level.

The standard stock assessment graphs that go to the ICES assessment reports can be selected for certain years, stocks and ecoregions and downloaded from the website. Graphs and data table can be downloaded only for validated and published stocks.

DIG suggested investigating the possibility to provide Persistent Identifiers (PIDs) for these products (action Neil Holdsworth). Each species-year-ecoregion combination could be assigned a PID with relative ease. Some metadata would need to be generated, but it is believed that the majority of information required for e.g. the DataCite schema could be derived from existing data/information within the standards graph system. This would be an ideal area for ICES to start developing and deploying identifiers/PID, and would make the individual stock assessments easier to cite in reports as well.

4.4 Eggs and larvae database

ICES Data Centre received Estonian Gulf of Riga Fish Larvae Dataset to add to the eggs and larvae database, and furthermore, the German Baltic Sea Fisheries institute is working on Baltic egg data to add to the database. There were two actions related to this topic on the DIG action list, i.e. upload by WGMEGS and IBTS-MIK of historical and most recent data. Both data sets will be uploaded soon, the Data Centre is in contact with the relevant experts. Missing MIK-Data is worked on, the responsible experts reported problems with data quality and therefore decided to re-format the data before sending them to ICES. WGMEGS promised data set in April 2014, but experts are currently reshaping the relevant database.

WGEGGS2 sent a recommendation to DIG that is forwarded to ICES Data Centre, related to the adoption of MIKey net data in the eggs and larvae database. The Data Centre needs a clear description of the data (metadata) to be able to estimate the work related to implementation. WGESSG2 will be asked to send in a data request form related to the MIKey data.

4.5 INSPIRE services

As an action from DIG 2013, to answer to INSPIRE directive readiness, an online available catalogue about services and data products hosted by ICES has been developed using a simple open source ESRI Geoportal. This geoportal seems to be more stable and clear than Geonetwork. It also integrates better with other tools than Geonetwork. Data managers were asked to describe the services related to datasets in a form of INSPIRE metadata. The tool is still under development and currently no external link to the website is available. This service will be integrated with the catalogue of datasets (<http://geo.ices.dk/geonetwork/srv/en/main.home>) when development is complete.

4.6 Other topics

- a) ICES has been accredited with Associated Data Unit (ADU) status by IODE in February 2014.
- b) A very big and important feasibility study "Scientific data storage and transmission under the DCF" has been drafted by the EU Commission, and under internal consultation.
- c) The methods and processes to ensure the HELCOM Eutrophication (EU-TRO OPER) dataflow were presented. The intention is that all data have the same quality.
- d) The first draft of an interactive map on the web was presented that allow an easy finding of Popular stock advice sheets. This was asked by the Communications department for going online for the "Science in the City" event.

- e) The publication of ETC marine GIS reference layers is under way, including MSFD regions and sub regions. Open issues are missing definitions for a complete baseline of EU member states. For practical reasons, in those cases the areas are drawn from the coastline, instead. ICES Data Centre plans to align MSFD regions and ICES Ecoregions. EEA will publish the finalised MSFD regions.
- f) A new database for preliminary catch statistics was re-designed and published earlier this year. The official catch statistics are available to ICES community more than a year after catches have been made. However, the stock assessments need to be done earlier. Since 1970 ICES collects the preliminary catch data (rec-12). The new database facilitates verification and direct upload of the xml data files, and download of data per species.
- g) For advertising and highlighting the different services from the Data Centre to the world outside ICES a leaflet is being drafted, showing available data collections, etc. Gaynor Evans and Simon Claus volunteered to give feedback on the draft. As the leaflet should be ready for the “Science in the City” Event in mid-June, deadline for comments is 6 June 2014.

5 Digital data citation within ICES

Journal citation is an accepted and well-established practice that gives due credit to scientific work done by scientists, and also signposts where others can find this information. In a similar way, citation of data can give proper credit to data providers who have made data available to the scientific community, while also providing a mechanism for tracing back scientific knowledge to the data that underpins it. The Introduction to Digital citation was presented to SCICOM in March 2014. This document is a living document and DIG discussed if updates were necessary.

During the plenary discussion of PID's and the document DIG produced in 2013, Neil Holdsworth noted that ICES was not in the business of minting PIDs. Building on that discussion, a subgroup discussed the use of PIDs, specifically for data (products), within ICES. There was recognition that ICES is not the owner of data in the ICES repository and therefore PIDs linked to particular data set in the repository should not have ICES as the owner.

5.1 Provide options for data citation within ICES.

There are multiple ways how to cite the originator for a particular data set. Some examples are listed in the textbox. It was recognized that some data will be coming into the ICES repository with PIDs and these should be part of the data stored in ICES. As an example, in 'Datacite' a contact or provider is used in their PIDs, but not the owner. It was also pointed out that keeping ownership or the data originator/creator as part of the metadata was important.

5.1.1 New datasets

DIG recommends that for new data sets, ICES Data Centre anticipates that data submitted be accompanied with appropriate Persistent Identifiers (PIDs), e.g. Digital Object Identifier (DOI). If not available, then ICES Data Centre should request one from the group submitting the data, but this should not stall the submission of data to the ICES Data Centre.

At SeaDataNet Common Data Index (CDI) are associated with SeaDataNet data for each station activity. The CDIs are a locally generated entity and are up to the individual institution. When adding PIDs to a dataset, it is important to keep the granularity at a moderate size and not to make it too fine for PIDs. Horizon 2020 asks for a standard for issuing PIDs to be generated by the projects by 2020.

5.1.2 Historic datasets

With respect to all data in the ICES system, DIG agrees that it is currently not realistic or necessary for ICES Data Centre to provide PIDs for all of the data currently in the repository.

There is an evolving need to provide a PID for data requested by users to be able to backtrack datasets to the data source. DIG recommends that ICES Data Centre investigates the possibilities to create a mechanism so a person can get a PID from ICES if the data is actually be used in a publication. For this mechanism, guidelines have to be made available to enable this process to be understood by the data user.

5.1.3 Updates and time-series

An issue about versioning of datasets after publication with persistent identifiers came up. As many data types are updated through data addition or quality improvement, keeping a fixed copy of these data can be a challenge. BODC have been working on a model for handling this challenge, this should be further discussed in the subgroup.

Regarding PIDs for time-series, BODC has a way to assign PIDs to increasing lengthened time series. This works for BODC as long as original data has not been changed. ICES does get time-series data, but usually the data sets are received yearly for the same station. So a PID could be created for each year's data set by the data submitter or ICES could have a PID for the time series and just add the data without changing the PID as long as the data are frozen. It was suggested that, to avoid the granularity problem, the same PID could be associated with every time step in the data set.

5.2 Joint Declaration of Data Citation Principles

A "Synthesis Group" of experts and organizations organized in mid-2013 has developed a common set of high-level principles on data citation that they seek endorsement for.

The 8 principles of the Joint Declaration of Data Citation were reviewed in order to advise ICES on the adoption of the Declaration. The principles are grouped so as to facilitate understanding, rather than according to any perceived criteria of importance.

- 1) Importance: Data should be considered legitimate, citable products of research. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications.

DIG agrees with this principle.

- 2) Credit and Attribution: Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data.

DIG interpreted this as everyone who has hand in creating the data gets credit for the data productions and think that this is appropriate.

- 3) Evidence: In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited.

DIG agrees about this statement.

- 4) Unique Identification: A data citation should include a persistent method for identification that is machine actionable, globally unique, and widely used by a community.

DIG agrees with this principle.

- 5) Access: Data citations should facilitate access to the data themselves and to such associated metadata, documentation, code, and other materials, as are necessary for both humans and machines to make informed use of the referenced data.

This statement does not say how the access is available. Is it free and open, or are there password protections or other impediments. The subgroup was opposed to these impediments except in the case where the data and information are sensitive or there are proprietary restrictions on them.

- 6) Persistence: Unique identifiers, and metadata describing the data, and its disposition, should persist -- even beyond the lifespan of the data they describe.

It was not clear to DIG on what 'lifespan of the data' implies. DIG agrees that identifiers should persist.

- 7) Specificity and Verifiability: Data citations should facilitate identification of, access to, and verification of the specific data that support a claim. Citations or citation metadata should include information about provenance and fixity sufficient to facilitate verifying that the specific time slice, version and/or granular portion of data retrieved subsequently is the same as was originally cited.

DIG agrees that this verification step is needed.

- 8) Interoperability and flexibility: Data citation methods should be sufficiently flexible to accommodate the variant practices among communities, but should not differ so much that they compromise interoperability of data citation practices across communities.

DIG agrees that this step is necessary.

- 9) DIG recommends that PUBCOM, SCICOM and ACOM read the document that lists the principles and after careful thought recommend that ICES endorse these principles (organisational endorsement).

The members of the DIG group should encourage their organisations to endorse the principles (organisational endorsement). The members of the DIG group are encouraged to endorse the principles (individual endorsement).

5.3 Update of 'Introduction to Digital Citation'

Based on the paragraphs about PID use in the ICES Data Centre, an update on the DIG/PUBCOM document "Introduction to Digital Citation – Towards an ICES policy on Data Citation DIG/PUBCOM Version 1.0 20 November 2013" will be created by DIG members (Peter Wiebe, Helge Sagen, Simon Claus, Gaynor Evans, Gisbert Breitbach) by correspondence before the 2014 ASC meeting.

6 Use of VMS data within and outside ICES

6.1 VMS Data policy

In 2013 and 2014, ICES has made two data calls for aggregated VMS/logbook data. One is a joint ICES-HELCOM data call, and the other is done by ICES for answering a request for advice made by OSPAR. In addition, through a written agreement with NEAFC, ICES receives VMS/logbook data from the NEAFC control area on a twice yearly basis. The data calls were successful, and ICES now stores the data on a secure server. The data will be made available to experts for the purposes of facilitating scientific advice and for answering specific terms of reference.

ICES has produced a VMS data policy stating the conditions of use for experts, to be signed by the chair of groups using the data. DIG reviewed this document and recommends that additionally each expert working with VMS/logbook data signs it before getting access to the data and that all signatures are collated to the same document.

6.2 Proposed procedure for working with VMS/logbook data

After completion of the requested VMS/logbook based product, intermediate VMS/logbook data used for the task are to be deleted from the expert's personal storage. The end results can be kept on the SharePoint, but it is suggested that intermediate data and results are stored in the ICES secure server in case of needed corrections to the product. To avoid having experts copying the data to their computers in the first place, it is recommended that VMS data can be directly accessed on a secured server, e.g. a VPN connection and a virtual PC are set up, so that the data can remain at the ICES secure server.

6.3 Publishing data

As the data are sensitive, care should be taken to what products are finally published and whether the data and results could be forwarded to clients requesting them. When handling requests care should be taken to follow the conditions specified in the data call, which should be clear and specific.

The terms of use are different in the two ICES data calls for VMS/logbook data. The "ICES/HELCOM Data call for VMS for fishing activities in the Baltic in support of ICES/HELCOM advice on the spatial distribution and impact of fisheries" outlines that the data is to be used by both ICES and HELCOM. The OSPAR data call for "VMS/logbook data for fishing activities in the OSPAR areas I-V in support of ICES advice on the spatial and temporal bottom fishing intensity as requested by OSPAR and standing requests from NEAFC and the EC" clearly states that the data will be used by ICES. The request is to provide OSPAR-wide mapping of the spatial and temporal intensity of fishing activities with mobile bottom contacting gears, in the form of a report which might be supplemented with the derived data product as a shapefile.

Both report and shapefile are to be considered as a product derived from the submitted data and can therefore be made available, as long as individual vessels can't be identified.

7 Availability, versioning and visibility of ICES Data Guidelines

During DIG 2013, a proposal for a dedicated workshop on Data Guidelines was written. The workshop was cancelled and replaced by an intersessional activity. In March 2014, a meeting took place between IODE and ICES to discuss ICES Data Guidelines in relation to the IODE/JCOMM/ICES clearing house (hereafter called 'clearinghouse'). Specific Action items and follow-up are in Annex 6.

7.1 IODE/JCOMM/ICES clearinghouse

The ambition is to have all data and data management related documents, including Manuals from instrument manufacturers and software (versions), available through the clearinghouse. It will thus provide an additional access point to the ICES documents, next to the ICES publications library. The target audience of the clearinghouse consists of data and information management professionals. Documents available through the clearinghouse come from trusted sources and should be of a high quality, therefore ranking of the documents is not necessary. Searching the documents should be possible based on (a) a category, (b) keywords and (c) free text search:

- a) The categories are based on the different steps in a typical data flow. Suggestion for categorization within the IODE/JCOMM/ICES clearinghouse:
 - Collection
 - Processing
 - Assembly
 - Dissemination/reporting
 - Policies/plans
- b) The keyword list will be based on the current list of subjects available in the clearinghouse with additional keywords provided by ICES.

7.2 Next steps

The ICES Data Centre will provide the data guidelines before 15 June to the clearinghouse and TIMES and survey protocols when ready. The data guidelines should be updated subsequently:

- The information about GETADE should be rewritten as this group no longer exists.
- The part of the data guidelines regarding provision of data to ICES. NODC's that have expertise in the topics for the different data guidelines should be identified and asked to update these to state of the art methods for data collection.

When updated versions of the documents are available they should be provided to the clearinghouse.

Abstracts for the ICES Data Guidelines were provided by ICES Data Centre and reviewed by DIG during the meeting.

7.3 Outreach

Outreach on the existence will be done through the ICES LinkedIn site and the ICES website. DIG also suggests that IODE (Taco de Bruin) reaches out to all NODCs and informs them about the existence of this clearinghouse and its content and asks them to inform their network of data managers.

7.4 Follow-up

As only survey protocols in the SISP format will be made available through the new clearing house, it is recommended that the creation of survey protocols in the correct format as well as the review process for survey protocols gets high on the priority list of survey expert groups as well as the overarching Steering Group (SSGESST).

8 Other topics

8.1 ACOM representation

The ACOM representation for DIG moved from Christopher Zimmermann to Christian von Dorrien, alternate member for Germany in ACOM. The group thanks Christopher Zimmermann for the long-term participation and appreciates his contributions done throughout the years.

8.2 EMODnet

The current activities of the Emodnet – Biology project was presented. It is a marine data initiative funded by DG Mare to support the blue economy in Europe. Filling the data bases and information slots can only be done together with the data providers and users, so their contributions are asked for.

8.3 OSPAR request

The request to DIG was to provide *Advice on possible OSPAR way forward with the handling of data and/or statistics, under OSPAR monitoring programmes and for the benefit of OSPAR assessment procedures, from monitoring devices generating large amounts of data.*

ICES is requested to advice OSPAR (for the benefit of HASEC [Hazardous Substances Committee] and the ICG EUT [Intersessional Correspondence Group on Eutrophication]) on (1) options and (2) a possible OSPAR way forward with the handling of data and/or statistics, from monitoring devices generating large amounts of data, taking account of the increasing reliance by Contracting Parties on such devices, especially in eutrophication monitoring (Eutrophication Monitoring Programme, part of CEMP [Co-ordinated Environmental Monitoring Programme]) and Ocean Acidification (part of pre-CEMP, i.e. voluntary measurements). Primary focus is on devices used under OSPAR monitoring programmes and for the benefit of OSPAR assessment procedures.

The revised OSPAR Common Procedure (COMP) for eutrophication assessment makes reference to the use by Contracting Parties of data generated by 'novel' monitoring devices such as satellite data, automated moorings. These data may be considered to form part of the Eutrophication Monitoring Programme, although it is recognized that there are e.g. issues related to QA and confidence rating to be addressed in PIDng so (see COMP Annex 8 which states "*Such data are often characterised by a very high measurement frequency in space and time but potentially low data accuracy*").

While it may not be appropriate that the raw monitoring data from such devices are reported by Contracting Parties to the ICES database for OSPAR purposes, ICES is requested to advise whether:

- 1) There would be benefits and possibilities for Contracting Parties to report summary statistics and key metadata on such datasets to ICES for OSPAR purposes;

DIG comments:

- When parameters are measured on a subregional scale, it is recommended to include these compiled data products in the final assessment.
 - Where there are already well established repositories (e.g. Argo float, Satellite imagery data, automated mooring) and QA process it is suggested to make a data product from the data which can be included in the assessment.
 - When there is not existing data repositories and QA process, it is recommended to report the underlying data to ICES.
- 2) Whether there are any international mechanisms already in place to discover such datasets and metadata that can inform about the relative confidence Contracting Parties can have in using such data from international data sources;

DIG comments:

- WGOOFE, EU RSC Project (data flows from member states to Regional Seas Conventions), EMODnet (Physics) and MyOcean can all provide data or mechanisms to discover data that could be utilised in the assessment. This would require involvement from the groups mentioned along with HELCOM EUTRO-OPER who face a similar situation in the Baltic Region.
- 3) How OSPAR could address best the traceability of the use of such data under the OSPAR Common Procedure, taking account of the intention of several Contracting Parties to make such datasets part of the data used in the 3rd COMP application.

DIG comments:

- The use of PIDs might allow traceability of data sources, where they are available. Creation of PIDs for those without existing PIDs could be investigated. Further elaboration on PID availability and application could be provided by DIG.

DIG concludes that to accurately reply to this request, input from all of the groups mentioned in the comments above is required. The combined knowledge could be brought together at a (physical or virtual) workshop to provide more complete guidance on:

- Availability and suitability of existing datasets,
- Expert knowledge on combination of differing data types for assessment,
- Translating the traceability of data into a procedural methodology

8.4 Follow-up of actions and recommendations

8.4.1 Action list

ACTION	ADDRESSED TO	COMPLETE BEFORE	STATUS
1. Send the draft Data Plan to Bureau and SCICOM (for info)	Ingeborg de Boois, Neil Holdsworth	15 June 2013	Completed, June 2013
2. Revise the Data Plan (WebEx)	Ingeborg de Boois, Neil Holdsworth, Marcin Wichorowski, Lena Szymanek, Peter Wiebe	Bureau meeting July, and dep. on feedback maybe before 1 September	Completed, August 2013 The Data Plan is now a chapter in the ICES Strategy 2014-2018, which has been published. The text of the Data Plan is available on Sharepoint under Reports.
3. Revise ASC 2014 proposal in collaboration with PUBCOM and send the proposal to SCICOM	Jens Rasmussen, Liam Caffrey, Marcin Wichorowski, Helge Sagen, Neil Holdsworth	1 July 2013	Completed, July 2013
4. Send digital citation document to PUBCOM	Ingeborg de Boois, Neil Holdsworth, Helge Sagen	1 July 2013	Delayed-communicated to PUBCOM and SCICOM June 2013, Version 1 to PUBCOM Sept 2013
5. Arrange WebEx with PUBCOM on digital citation document	Helge Sagen, Ingeborg de Boois, Neil Holdsworth, Taco de Bruin	1 September 2013	Delayed, completed March 2014 presented to SCICOM
6. Provide the Data Centre action list (sharepoint)	Neil Holdsworth	1 March 2014	Completed, March 2014
7. Finish proposal workshop on ICES Data Guidelines and send to Ingeborg	Sjur Ringheim Lid, Taco de Bruin, Hjalte Parner, Peter Wiebe	1 July 2013	Completed, August 2013
8. Send the workshop proposal to SCICOM	Ingeborg de Boois	1 August 2013	Completed, August 2013
9. Add INSPIRE descriptions for services to ICES data sets	ICES Data Centre (Periklis), Jens Rasmussen	Start before September 2013	On agenda DIG 2014, see chapter 4.5
10. Ask WGMEGS and IBTSWG-MIK to add data to eggs and larvae database	Ingeborg de Boois, Anna Osypchuk	1 July 2013	Completed, May 2013, see chapter 4.4
11. Start discussion on LinkedIn how to improve DATRAS (products, webfacility, quality, etc.)	Neil Holdsworth	15 October 2013	Completed, December 2013 Working on interface, feedback=webservices! Tomorrow
12. Put comments from action 11 in a document for discussion by ICES Data Centre and DIG	Neil Holdsworth, Ingeborg de Boois	1 March 2014	On agenda DIG 2014

ACTION	ADDRESSED TO	COMPLETE BEFORE	STATUS
13. Prepare discussion on how to proceed with VMS products at DIG 2014	Josefine Egekvist	1 April 2014	On agenda DIG 2014
14. Discuss possibilities for combining Mining fisheries data ASC2014 session proposal with Big Data session	Ingeborg de Boois	ASC 2013	Completed The 2014 ASC data session is a combination of several theme session proposals and will be on Big data. As a result, DIG members found it difficult to contribute to this session and there may not attend.
15. Approach WGACEGG to contribute to eggs and larvae database	Neil Holdsworth	1 November 2013	In progress. ICES Data Centre (Carlos) will attend the 2014 meeting. Also passed through to SSGESST, in relation to acoustic data.
16 Keep a shortlist of topics that might be put on the ICES webpage on behalf of DIG	Ingeborg de Boois	1 December 2013	No list. Difficult to keep track during the year when one person is responsible. Add action item to provide input to Ingeborg
17. Pick dates and venue for WKIDG	Lesley Rickards, Taco de Bruin, Sjur Ringheim Lid, Hjalte Parner	26 September 2013	Completed
18. Final review of DOI document	Helge Sagen, Christopher Zimmermann, Liam Caffrey, Neil Holdsworth, Peter Wiebe	15 October 2013	Completed, November 2013

8.4.2 Recommendations from DIG

RECOMMENDATION	ADRESSED TO	STATUS
<p>1. Add a term of reference in 2015 to all ICES working groups to add metadata to the ICES metadata database, to create an overview of all datasets available within ICES. DIG suggests the following term of reference</p> <p>DIG suggests the following term of reference</p> <p>“(…) to inform ICES Data Centre (this can be sent out as an online questionnaire to EG’s to fill in)</p> <ol style="list-style-type: none"> 1. If the expert group is using datasets in its work (YES/NO) 2. If the expert group collates and manages the dataset(s) (YES/NO) <ol style="list-style-type: none"> i. IF YES: metadata records should be supplied to ICES Data Centre for the various datasets/databases to be included in ICES Metadata portal (these may already exist in other data portal systems, so references to these should be provided) 3. If the dataset(s) are stored in a database managed by ICES Data Centre (YES/NO) <ol style="list-style-type: none"> i. IF YES: in which database and under which coding (List will be supplied by ICES Data centre)” <p>Question(s) to elicit the details of the database, what kind of data management design and protocols are in place, etc. might be added. DIG could reflect the standards that it intends to have applied to ICES data onto source data systems.</p>	SCICOM and ACOM	<p>Nils Olav Handegard:</p> <p>There were some reluctance in SCICOM/ACOM to add at ToR to all EGs, but the questionnaire will be brought up at the next webex meeting for the SSG. Ideally we find one EG that is willing to do a first cut on this, and then we can use that as a template for the others. We can present it at the business meeting at the ASC, and if the other EGs feels that this is an ok idea, we can follow up with a at the SCICOM meeting after the ASC.</p>

8.4.3 Recommendations to DIG

YEAR	FROM	RECOMMENDATION	TO	STATUS
2013	WGBEAM	5. As the differences in coding (TSN vs. WoRMS, accepted codes vs. unaccepted codes) might not only influence DATRAS but also other databases hosted by ICES, it is recommended that ICES Data Centre and DIG define the most suitable way for ICES Data Centre, data-submitters and data-users to cope with the frequent updates of WoRMS.	ICES Data Centre; DIG	Complete, resolved for DATRAS, but not yet for all other databases stored at ICES
2013	WKESST	There is an urgent need for a database for the basic species interpreted data of the hydroacoustic surveys, e.g. NASC values (See Section 4.2 in the 2013 WKESST report).	DIG; ICES Data Centre	Complete. This recommendation is tied together with Horizon2020.
2013	WKESST	There is a need for a permanent reference for versions of survey reports in the form of DOIs or any other persistent identifiers.	PUBCOM; ICES Data Centre; DIG	Data guidelines discussion
2013	WKESST	There is a need for an interactive overview (online mapping capability) of the survey effort deployed in each regional sea (See Section 4.7 in the 2013 WKESST report).	DIG; ICES Data Centre	In progress. Better product description by WKESST chair required. Investigate if CSRs might be sufficient
2013	WGEGGS2	1. To make the inclusion of the MIKey Net data in the database feasible	DIG	Recommendation forwarded to ICES Data Centre, see chapter 4.4
2014	WGMEGS	WGMEGS recommends that the fecundity and atresia historic time series be stored in a central database at ICES. WGMEGS provided a formal request to accessions@ices.dk	DIG, ICES Data Centre	In progress by ICES Data Centre
2014	WGMEGS	WGMEGS recommends that the images for the fecundity and atresia estimation are stored in a central database (approximately 200 gb per survey).	DIG, WGMEGS participants	In progress. See section 8.4.3.1
2014	OSPAR	Advice on possible OSPAR way forward with the handling of data and/or statistics, under OSPAR monitoring programmes and for the benefit of OSPAR assessment procedures, from monitoring devices generating large amounts of data		On agenda DIG 2014

8.4.3.1 Storage of images fecundity and atresia

With respect to the recommendation that ICES Data Centre stores the pictures of the atresia and the fecundity sampling, DIG asked for clarification on a number of topics (questions in italics, answers by WGMEGS underneath):

1. Which scientific question will be answered by the availability of the pictures?

The pictures give us the realized fecundity estimation needed to calculate the fish stock spawning biomass of mackerel from egg production. For the estimation of realized fecundity (potential fecundity – atresia) we collected formaldehyde fixed samples of the ovaries. a) For potential fecundity we collect small pipette samples which are photographed completely and then analysed. All institutes have problems with long time storage of formaldehyde samples. b) For atresia estimation a part of the ovary is preserved and cut for histological slides. Of the histological slide a part (we need a certain area to be analysed) of the sample is photographed and analysed. This is not the whole histological section, thus for quality assurance we need to store the pictures.

2. Where are the pictures currently stored?

Pictures are currently stored at the individual institutes on external hard drives.

3. Is there a need for reference material, or pictures for exchanges?

Every third year we have a WKFATHOM workshop before the triennial survey. We use (part of) these pictures as reference and practice material for the workshops and exchange the pictures for quality assurance.

4. Is there any budget available with this request?

Sadly no budget available

Based on the answers DIG concluded that the request is mainly to only store the images. It was discussed if images are 'data' or 'samples'. The group concluded that in this case, the pictures could be seen as samples, and so, it the institutes should feel responsible for proper sample storage. ICES will only store (as a failsafe backup) these samples if there is a serious risk to lose the samples, e.g. when an institute will no longer exist. Additionally, it was pointed out that it is important to add to survey metadata that pictures are available at institutes.

For picture exchange, it was concluded that WebGR (hosted by AZTI) is available. As this programme is developed within an EU project that has since ended, it is not maintained and may need some development/updating. Only when it is a fully operational and proper system, ICES might consider to host that facility but first a project is needed to update the programme.

8.5 Social media

Jens Rasmussen provides a short overview of the usage statistics of the DIG LinkedIn group. Currently, the DIG LinkedIn group has 224 subscribers.

PUBCOM requests a tweet by the end of each meeting day or by the end of each meeting. DIG tried to do so, but it is difficult to tweet 'on command'. The group however provided four tweets at the end of the meeting, which will be published by ICES in the weeks after the meeting.

DIG created summaries on VMS data, data citation and data guidelines for LinkedIn ICES Data Group, LinkedIn ICES and/or ICES Newsletter. The information will be published by ICES soon after the meeting.

9 Between May 2014 and May 2015

9.1 Meetings

During ICES ASC in A Coruña a progress meeting for DIG will take place. Date and time will be decided based on the availability of DIG members at ICES ASC 2014.

The next plenary DIG meeting will take place 18-20 May (Mo. 13.00-Wed. 18.00) at ICES, Copenhagen.

9.2 Intersessional work

The following DIG members are available for ad-hoc requests: Lena Szymanzek, Gaynor Evans, Jens Rasmussen, Malin Werner.

The action list contains a number of actions that will take place before May 2015 and will be followed up. If necessary, WebEx meetings will be arranged by ICES Secretariat.

Annex 1: List of participants

NAME	ADDRESS	PHONE/FAX	EMAIL
Ingeborg de Boois (Chair)	IMARES P.O. Box 68 1970 AB IJmuiden Netherlands	+31 317 487070	Ingeborg.deboois@wur.nl
Taco de Bruin	NIOZ P.O. Box 1790AB Den Burg Netherlands	+31-222-369479	Taco.de.Bruin@nioz.nl
Sjur Ringheim Lid	Havforskningsinstituttet Postboks 1870 Nordnes 5817 Bergen Norway	+47-99556943	Sjur.Ringheim.Lid@imr.no
Josefine Egekvist	DTU Aqua Charlottenlund Slot Jægersborg Allé 1 2920 Charlottenlund Denmark	+45 35883438	jsv@aqua.dtu.dk
Marcin Wichorowski	Institute of Oceanology, PAS Powstańców Warszawy 55 81-712 Sopot Poland	+48 58 7311705	wichor@iopan.gda.pl
Ruth Lagring	Belgian Marine Data Centre (RBINS-OD Nature) Gulledelle 100 B-1200 Brussels Belgium	+32 2 773 21 44	r.lagring@mumm.ac.be
Malin Werner	Swedish University of Agricultural Sciences Department of Aquatic Resources Institute of Marine Research Turistgatan 5 S- 453 30 Lysekil Sweden	Tel. +46 10-478 4057 Mob. +46 76-12 68 048	Malin.Werner@slu.se
Friedrich Nast	Bundesamt fuer Seeschifffahrt und Hydrografie Bernhard-Nocht-Str. 78 20359 Hamburg Germany	+49 40 3190 3420	friedrich.nast@bsh.de
Peter H. Wiebe	MS#33 Woods Hole Oceanographic Institution Woods Hole, MA, 02543 USA	+1 508-289-2313 +1 508-457-2169	pwiebe@whoi.edu

NAME	ADDRESS	PHONE/FAX	EMAIL
Christian von Dorrien (also representing ACOM)	Thuenen-Institute of Baltic Sea Fisheries Alter Hafen Sued 2 18069 Rostock Germany		christian.dorrien@ti.bund.de
Helge Sagen	Institute of Marine Research PoBox 1870, Nordnes 5817 Bergen Norway	+47 55 23 85 00 +47 55 23 84 47	Helge.Sagen@imr.no
Gisbert Breitbach	Helmholtz-Zentrum Geesthacht Max-Planck-Str. 1 21502 Geesthacht Germany	+49 4152 87 1563	Gisbert.Breitbach@hzg.de
Lesley Rickards	British Oceanographic Data Centre (BODC) Joseph Proudman Building 6 Brownlow Street Liverpool L3 5DA United Kingdom	+44 151 795 4897	ljr@bodc.ac.uk
Chris Moulton	OSPAR		Chris.Moulton@ospar.org
Joni Kaitaranta	HELCOM		joni.kaitaranta@helcom.fi
Simon Claus	VLIZ Wandelaarkaai 7 8400 Oostende Belgium	+32 59-340159	Simon.claus@vliz.be
Gaynor Evans	British Oceanographic Data Centre Joseph Proudman Building 6 Brownlow Street Liverpool L3 5DA	+44 151 795 4881	gaev@bodc.ac.uk
Neil Holdsworth	ICES		neilh@ices.dk
Jens Rasmussen	Marine Scotland Science Marine Laboratory 375 Victoria Road Aberdeen AB11 9DB United Kingdom	+44 1224 295 440	rasmussenj@marlab.ac.uk

Annex 2: Agenda

Agenda DIG2014

Monday 26 May

9.00-9.30 Logistics, round the table, etc.

9.30-10.15 Recommendations and actions 2013

10.30-11.00 Coffee

11.00-11.30 Theme sessions on conferences [Data Plan]:

- a. 2014 ASC theme session on Big Data: state of the art [actions 3/14] Jens
- b. Ideas for 2015? IMDIS conference or ASC
- c. IOC Barcelona Peter

11.30-12.30

How to continue with VMS products created by WGSFD? [action 13 and Data Plan]
Josefine

12.30-13.30 Lunch

13.30-14.00

Digital citation: how to proceed? Discussion [actions 4/5/18 and Data Plan]

14.00-14.30

Data guidelines: progress [actions 7/17 and Data Plan] Taco

14.30 Identification of subgroup topics and subgroup participants

14.40-17.00 Subgroups (may be changed as a result of morning session, including tea break):

- Data guidelines
- VMS products
- Digital citation

17.00-17.30

Wrap-up from and feedback to subgroups (+140 characters(=tweet))

Tuesday 27 May

9.00-12.00 (including a coffee break)

Update from and feedback to ICES Data Centre ICES Data Centre

Including:

- Operational products:
 - How to improve (the visibility of) DATRAS products? [action 11/12]
 - Delivery of operational oceanographic products to IEA process
 - Standard graphs
 - MSFD data workshops
 -
- More data for eggs and larvae [actions 10/15]
- INSPIRE descriptions to ICES datasets [action 9 and Data Plan]
- QC database [Data Plan]

- Reflection on Data Centre Work Plan

12.00-12.30

Plenary: summary of morning session, highlights, follow-up, etc.

12.30-13.30 Lunch

13.30-14.30 Progress Data Plan, focus on highlighted topics in the table below

14.30-17.00 Subgroups (prelim. topics), including tea break

- Proposal theme session(s)
- OSPAR request: Joni, Friedrich, Hjalte,
- Actions from data plan

17.00-17.30

Plenary: Wrap-up from and feedback to subgroups (+140 characters)

Wednesday 28 May

9.00-11.30 Preparing presentations of DIG work:

9.00 Plenary: agree on topics for outreach

9.30 Fine-tune in subgroups:

1. Social media: LinkedIn, facebook, twitter and ICES website
2. ICES newsletter (e.g. <http://www.ices.dk/news-and-events/news-archive/newsletters/Pages/Newsletter-March-2014.aspx>)
3. Report DIG to SCICOM at ASC 2014

10.30 Coffee

11.00 Plenary: review subgroup texts and decide on final versions for social media and ICES Inside out

11.30-13.00

Planning DIG meeting at ASC 2014 (who will be there?) and setting dates for DIG 2015

Finalise report

12.30 End

Annex 3: Terms of reference for next meeting

The **Data and Information Group (DIG)**, chaired by Ingeborg de Boois, Netherlands, will meet in Copenhagen, Denmark, 18 (13.00) -20 (18.00) May 2015 to:

- a) Review priorities on the Data Centre action list
- b) Provide guidance and feedback to the ICES Data Centre
- c) Advise on other data regulations and their impact on ICES Data Strategy, ICES Data Policy
- d) Review output from offspring groups (WKIDP, LinkedIn Data and Information Forum) if relevant
- e) Promote new technologies and data management infrastructure development
- f) Products (e.g. updated data management guidelines, reviews of ICES Data Strategy, ICES Data Policy, etc.) from the meeting as well as a written report to SCICOM will be delivered before 15 July 2015.
- g) The group reports to SCICOM during the SCICOM mid-term meeting March 2015 as well as the SCICOM meeting at ICES ASC 2015. The group reports to ACOM by correspondence and via the ACOM representative.

Annex 4: Recommendations and Action items

Recommendations

RECOMMENDATION	ADRESSED TO
1. Clearly communicate to all their expert and working groups that new data (storage) facilities or products the request form as available via http://ices.dk/marine-data/guidelines-and-policy/Pages/Requesting-data-from-ICES.aspx is used. <u>Only in this way can be assured that requests for new data storage or data products end up in the ICES Data Centre request overview.</u>	ACOM, SCICOM
2. Find a quality flagging scheme fitting best to the needs (from Data Plan)	ICES Data Centre
3. To make the inclusion of the MIKey Net data in the database feasible (this recommendation was originally sent from WGEGBS2 to DIG, 2013)	ICES Data Centre
4. DIG recommends that for new data sets, ICES Data Centre anticipates that data submitted be accompanied with appropriate Persistent Identifiers (PIDs), e.g. Digital Object Identifier (DOI). If not available, then ICES Data Centre should request one from the group submitting the data, but this should not stall the submission of data to the ICES Data Centre. (see section 5.1 of DIG 2014 report)	ICES Data Centre
5. Investigate the possibilities to create a mechanism so a person can get a PID from ICES if the data is actually be used in a publication. For this mechanism, guidelines have to be made available to enable this process to be understood by the data user. (see section 5.1 of DIG 2014 report)	ICES Data Centre
6. DIG recommends that PUBCOM, SCICOM and ACOM read the document that lists the principles and after careful thought recommend that ICES endorse these principles (organisational endorsement). (see section 5.2 of DIG 2014 report)	PUBCOM, ACOM, SCICOM
7. As only survey protocols in the SISP format will be made available through the new clearing house, it is recommended that the creation of survey protocols in the correct format as well as the review process for survey protocols gets high on the priority list of survey expert groups as well as the overarching Steering Group (SSGESST). (see section 7.4 of DIG 2014 report)	Nils Olav Handegaard (SSGESST chair)
8. DIG reviewed the VMS data policy and recommends that additionally to the chair of a group signing, each expert working with VMS/logbook data signs it before getting access to the data and that all signatures are collated to the same document. (see section 6.1 of DIG 2014 report)	ICES Data Centre
9. To avoid having experts copying the data to their computers in the first place, it is recommended that ICES Data Centre investigates the possibilities to work with VMS data directly on a secured server, e.g. by using a VPN connection and a virtual PC are set up, so that the data can remain at the ICES secure server. (see section 6.1 of DIG 2014 report)	ICES Data Centre

Action list

ACTION	ADDRESSED TO	COMPLETE BEFORE	STATUS
1. Communicate topics related to DIG and worth sharing with the wider audience either directly on LinkedIn page(s) or to Ingeborg	All DIG members	All year	
2. Keep a shortlist of topics that might be put on the ICES webpage on behalf of DIG	Ingeborg de Boois	All year	
3. to speed up the first draft of the documentation on the methodology OSPAR Hazardous substances	Chris Moulton, Neil Holdsworth	1 July 2014	
4. Explore the options for and the focus of a training course 'Making the most of ICES Data'	Ingeborg de Boois, Neil Holdsworth, Jens Rasmussen	1 April 2015	
5. Give feedback on leaflet advertising and highlighting the different services from the ICES Data Centre to the world outside ICES	Gaynor Evans, Simon Claus	6 June 2014	
6. Investigate the possibility to provide PIDs for standard stock assessment graphs	Neil Holdsworth	1 April 2015	
7. Make a list of keywords that will be used on the data guidelines, TIMES and survey protocols.	ICES Data Centre	15 June 2014	
8. Announce to IODE that a keyword list is going to be provided by ICES	Taco de Bruin	15 June 2014	
9. Suggest/request to IODE to have two or three broad collections in the clearing house that all documents can be grouped into.	Taco de Bruin	15 June 2014	
10. Add ICES as a source for documents in the IODE clearinghouse	Taco de Bruin	15 June 2014	
11. Provide the data type guidelines to IODE clearinghouse	ICES Data Centre	15 June 2014	
12. Add TIMES and survey protocols to the IODE clearinghouse	ICES Data Centre	When ready	
13. Update parts regarding "getting data to ICES" for all data guidelines	ICES Data Centre	1 October 2014	
14. Identify NODCs that have expertise in data collection for the different data guidelines	Taco de Bruin	15 July 2014	

15. Identified NODCs review the part of the data guidelines that ICES didn't review and update those	Taco de Bruin	15 November 2014
16. Update data guidelines, section 1.2 (GETADE)	Lesley Rickards	15 July 2014
17. Provide updated Data Guidelines to clearing house	ICES Data Centre	When ready
18. Inform potential users through LinkedIn and ICES website when data guidelines are available through the IODE clearinghouse.	ICES Data Centre via social media and webpage	When ready
19. Suggest conveners and presenters for ASC 2014 Big Data Session to Jens	All DIG members	6 June 2014
20. Provide suggestions for ASC 2015 theme sessions to Peter	All DIG members	15 June
21. Write working document on how to Data Guidelines group	Taco, Sjur, Lesley, Marcin, Ruth, Hjalte/Neil	1 April 2015
22. ASC 2015 theme session proposal 'Marine data (management) in support of Marine directives and marine ecosystem based management' first draft	Peter, Ingeborg, Christian, Simon	1 August 2014
23. Ask SSGESST chair for clear product description for product asked for in WKESST recommendation 'There is a need for an interactive overview (online mapping capability) of the survey effort deployed in each regional sea (See Section 4.7 in the 2013 WKESST report).'	Ingeborg de Boois	1 July 2014
24. Encourage their organisations to endorse the principles (organisational endorsement) (see section 5.2)	All DIG participants	1 December 2014
25. Create second version of Digital Citation Document (see section 5.3) for discussion DIG@ASC	Peter Wiebe, Helge Sagen, Simon Claus, Gaynor Evans, Gisbert Breitbach	1 September 2014
26. Investigate if CSRs might be sufficient for product asked for in WKESST recommendation 'There is a need for an interactive overview (online mapping capability) of the survey effort deployed in each regional sea (See Section 4.7 in the 2013 WKESST report).'	Friedrich Nast, Hjalte Parner	1 October 2014

27. Ask WGESG2 to send in a data request form related to the MIKey data.	Ingeborg de Boois	1 July 2014
28. Ask WGAQUA if a database is needed	Ingeborg de Boois	1 July 2014

Annex 5: Data Plan

Selection relevant paragraphs of the Data Plan

Goal 4: Promote the advancement of data and information services for science and advice needs

Promoting the advancement of data and information services for science and advice needs on both regional and subregional levels, such as providing operational products for marine spatial planning, the Data Collection Framework and for the Marine Strategy Framework Directive;

Provision of regional workflows for the coordination of data collection, collation and data product/indicator production on an agreed basis with assured delivery for i.e. MSFD reporting timelines. A collaboration between ICES Data Centre, Regional Sea Conventions and national data originators. Supporting the establishment of integrated ecosystem observation and monitoring systems that enable coordinated data collection in support of scientific and advisory needs, and which have strong links with ICES data centre and national data centres.

New processes/products from existing data to serve both MSFD and DCF needs for Advisory and Science services. Primarily, this may entail calculations for indicators in collaboration with the relevant working groups, but also the automatisisation of data acquisition i.e. data calls that provide information directly into the regional data assembly mechanisms with minimal use of human effort.

ICES is entering a critical phase in ensuring that the production of its advice and science is managed within a robust IT business model. This means the roll-out of the Report and Advisory Content Management System (RA-CMS) will be a major feature of the implementation plan through to 2018.

See **Regional Facilitation**

Gearing up for new/expanding areas of dataset collections, such as new datasets from integrated ecosystem monitoring, including marine litter, and anthropogenic noise in the marine environment;

New datasets and products are being/will be requested under the MSFD. To ensure that the advisory and science services are able to respond to this, DIS will be looking at data inventories, new data assemblies and working arrangements to ensure the provision of regional data products in areas such as underwater noise, microplastics and acoustics (fish).

In addition to reviewing and revising the ICES Data policy, Data and Information Services will also be collaborating with the Publications Committee to ensure ICES has a common strategy towards using and providing digital citation resources.

See **Regional Facilitation, Data stewardship and data management.**

Goal 5: Catalyse best practices in marine data management, and promote the ICES data nodes as a global resource

Ensuring the use of International standards/interoperability to enable the use and application of ICES datasets, products, and services to an expanded international user base, and to provide tools and knowledge to facilitate this use.

All ICES datasets and data services, including datasets and data products that exist only within an expert group, are adequately described and the 'discovery' information

are available through the ICES online portals. This will also allow ICES member countries to draw on these services for their own reporting needs i.e. under MSFD.

Initiate plans for training and reference guides for scientists and data managers.

This will help raise the profile of both ICES data and data services, but also of the importance to the ICES community of the value of good data management.

See **International Standards, Knowledge transfer**

Regional Facilitation				Status
Headline action	Detail	Performance measure	Timing	
Regional operational products for Marine Strategy Framework Directive (MSFD) and Data Collection Framework (DCF)/Multi-annual programme (DC-MAP)	(a) MSFD workflow: Collaboration between ICES Data Centre and Regional Sea Conventions/other organisations with respect to MSFD (WISE-Marine production process). This assumes a good flow of data/data harvesting into the data centre, and this can imply more resources in certain data types where data are not readily provided. (b) Leading to a joint MSFD data flow vision paper. Also depends on WISE-Marine. Link to secretariat plan.	(a) Workflow(s) operational and ready for uptake into WISE-Marine (b) Joint paper strategy accepted by stakeholders at EU level	- (a) OSPAR Hazardous substances: milestone 2014 - (a) HELCOM Eutrophication: milestone 2014 - (a) OSPAR Eutrophication (2015) - (b) MSFD Data vision paper: 2014.	05/2014: a. Progress on all workflows i.e. EUTRO-OPER, (see also chapter 4 of this report). Online tools are developed. Documentation on methodology is still not there. b. Started, drafted template and vision paper under development. Vision paper accepted by WGDike.
	New processes/products from existing data Advisory and Science with respect to MSFD: calculations for indicators. Needed: data selections, algorithms, calculation examples. Challenge: who is going to decide on the final calculations and data selections? Workshop on MSFD related DC-MAP indicators. Refer to table (MSFD table of	a) Uptake of ICES dataset products in EG's responsible for MSFD indicators b) Operational provision of datasets, including discovery and download services	Fish and litter Timeframe: 2014-2015 for development, and from 2016 onwards fine-tuning	05/2014: (Offshore) litter: In progress. Drafted extension to trawl survey format for marine litter, needs further iteration. ICES will try to establish a WG on Marine litter as a complement to existing groups/RSC processes

Regional Facilitation				Status
Headline action	Detail	Performance measure	Timing	
	ICES data/WG's and their operational product linkage)			
	New datasets and products Advisory and Science: MSFD - master data holdings; data storage, calculations for indicators. Noise, microplastics, acoustic fish data (WGFAST). Needed: data collection guidelines, data, responsible WGs for data, algorithms, calculation.	Products and/or regional data management established (where mandate is given)	2015 for setup, implementation from 2016 onwards.	
	- Data requirements with regard to multi-species assessments (input for assessments). Currently, multi-species assessments are applied in e.g. Baltic, but insufficient spatial data products are available. Baltic, other areas. (action plan to be created). Needed: clear data request (unless no data are available)	(a) Successful data call(s) (b) Provision of spatial data products	Baltic: 2014-2015	05/2014: no action

Regional Facilitation				Status
Headline action	Detail	Performance measure	Timing	
	- Data requirements for e.g. one species from all fish surveys (WGEF, WGNEW); search facility over all data, not only for raw data but also for products. (joint WGEF, WGNEW, DIG proposal -action DIG chair)		Workshop in October 2014 to list product requirements	05/2014: workshop is planned in October and will be chaired by Clara Ulrich
End-to-end workflow for scientific advice production	- RA-CMS linking to data outputs from Expert groups (connecting the scientific reports to advice production).	Successful implementation of interfaces to a) scientific output from EG reports b) scientific output from assessment models	Starting 2014 (depends on timing RA-CMS development).	05/2014: Process delayed. Currently concentrating on stock input and expanding standard graphs to other stocks. System re-named CARA.
	- RA-CMS linking to data outputs from RDB-Fish-frame	See (b) above	2015	

Regional Facilitation				Status
Headline action	Detail	Performance measure	Timing	
Mobilising aquaculture specific data	- Aquaculture databases: exact description to be decided. Related to WGAQUA.	Products and/or regional data management established (where mandate is given)	Starting from 2014.	05/2014: no action
Mobilising Arctic specific data	- In cooperation with AMAP, getting data from small arctic research institutes. Implementing data formatting tool.	Milestone: implementing the tool, first half 2014. Performance measure: receiving data	starting 2014	05/2014: In progress. Some testing and need further documentation of SIMON system Helge Sagen (DIG) nominated to Committee on Information and Data Service (CDIS) of SAON

International Standards and interoperability				Status
Headline action	Detail	Performance measure	Timing	
Ensuring INSPIRE readiness for ICES managed datasets/data services	- describe and make available all ICES/ICES expert group managed datasets, data products or services through ISO/INSPIRE standards to allow their discovery and reuse	- All ICES datasets, including those that exist only within an expert group, are adequately described and the	- Request to EG's to be filled 2015	05/2014: ICES Data Services have an online system (INSPIRE compatible). See also chapter 4.

	by other expert groups, processes and member country activities	'discovery' information are available through the ICES online portals		
Encouraging the broader use of ICES datasets by implementing IODE quality flagging schema	building on the quality control database that is in the process of being populated and then exposing this to online users in a digestible way to make the linkage between type of data, type(s) of QC performed and the QC flags applied to the data	- QC database online - QC flags included in data downloads	2014-2018	05/2014: no progress

Knowledge transfer and professional development

Headline action	Detail	Performance measure	Timing	Status
Input to key data symposia and science meetings	- Data theme sessions (ASC, IMDIS etc): annual theme session proposal ASC by DIG	(a) presentation and promotion of ICES work at key events (b) requests for new services/projects resulting from those activities	-IMDIS runs in 2015, 2017 - ASC annual cycle	05/2014: IMDIS will not take place in 2015 so a proposal for ICES ASC 2015 was prepared by DIG 2014

Training and reference guides for scientists and data managers	<ul style="list-style-type: none"> - ICES training courses: 'Making the most of ICES Data', modular, webinars?. - Online materials and guidance: WKIDG in 2014 	<ul style="list-style-type: none"> (a) metrics on usage of reference materials (b) requests for new services/projects resulting from reference materials/training (c) Increased awareness of data management/ICES services in new sectors 	<ul style="list-style-type: none"> - Training: end 2017 - Workshop to produce reference guide in 2014 (WKIDG, proposed) 	05/2014: In progress. For data guidelines see chapter 7.
---	--	--	---	--

Data stewardship and data management

Headline action	Detail	Performance measure	Timing	Status
Data archaeology; identifying and making available datasets that are relevant to the marine community	<ul style="list-style-type: none"> - (a) benthic historic data recovery. Plan ready, no timeframe. Connected to BEWG, DGMARE (DC-MAP related), perhaps EMODnet biology? - (b) Legacy data: data that are in other systems, but not available to the wider world. Linking to other data archives i.e. through metadata -(c) other historic data 	<ul style="list-style-type: none"> (a) inclusion of pilot project in EMODnet biology (b) Providing discovery services for archived information (through EG's) (c) Where resource, to run data recovery projects 	<ul style="list-style-type: none"> (a) Start 2014. (b) follow-on from 'IN-SPIRE readiness' activity under heading 3 	05/2014: a. benthic historic data recovery proposal was ready. After discussion not put there due to wrong focus. Work package is on hold. b. See chapter 4
Ensuring ICES data are citeable in the digital age, and therefore making the datasets easier to discover	Digital data citation and publication: ensuring ICES data are citeable in the digital age, and ensuring contributing data sources are duly credited, as well as guiding the ICES member countries on how to approach digital citation	Creating a strategy for digital citation of data resources, in agreement with PubCom	2014-2015	05/2014: in progress. See chapter 5.
Maintaining the user rights, security and integrity of the data sources to ICES managed datasets	<ul style="list-style-type: none"> - Data policy, facilitation of rights issues - Data security, and implications if data portfolio changes in nature (i.e. VMS, VME etc.) 		Annual basis, 2014-2018	05/2014: RDB-FishFrame data policy drafted but not agreed by all participating countries yet

Annex 6: Data Guidelines

Subgroup on data guidelines and connection to IODE clearinghouse.

Participants: Lesley, Marcin, Neil, Taco, Ruth, Jens, Sjur

1. Action items ICES IODE clearinghouse meeting in March 2014 (Taco, Peter Pierissens, Neil):
 - a. Neil/Katie ICES will need to confirm that through its Publication Committee (PUBCOM) that they are happy to allow duplicate ICES publications to be made available on other web portals

Follow-up:

PUBCOM has agreed that copies of the ICES documents can be made available through the IODE clearinghouse.

Requirements:

1. Originator of the document should be made clear
2. Usage statistics should be provided to the document originator
- b. Hjalte/DIG ICES data type guidelines need to be split out and published on ICES library

Follow-up: The data guidelines have been split out and published on the ICES library
- c. Neil, Peter: Check that ICES Publication metadata are compatible with IODE Clearinghouse format

Follow-up: Metadata schema at ICES have been expanded with keywords, abstract and URI to comply with the clearinghouse metadata schema
- d.
 - a: Neil Check XML export format from ICES system
 - b: Peter Check XML import format required for IODE system

Follow-up: Will be finished when content is in place.
- e. Peter: Inform Steering Group on OceanDocs about this joint initiative and request them to produce a (short) list of suitable categories or an ontology at their meeting in mid-March

Follow-up: Discussion around vocabulary for categories is taking place, nothing decided
- f. Neil: Send email to Roy Lowry and Adam Leadbetter asking their advice for a suitable ontology

Follow-up: Not done, dependent on action item e.