

ICES DIG REPORT 2015

ICES SCIENCE COMMITTEE

ICES CM 2015/SCICOM:02

REF. SCICOM, ACOM

Report of the Data and Information Group (DIG)

18-20 May 2015

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. 2015. Report of the Data and Information Group (DIG), 18-20 May 2015, Copenhagen, Denmark. ICES CM 2015/SCICOM:02. 47 pp. <https://doi.org/10.17895/ices.pub.8333>

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.

© 2015 International Council for the Exploration of the Sea

Contents

Executive summary	3
1 Opening of the meeting.....	5
2 Adoption of the agenda	5
3 Progress on ICES Data Plan	6
3.1 ASC theme session proposals	6
3.2 Other topics	6
4 ICES Data Centre	6
4.1 InterCatch and Regional Database (RDB) Fishframe	6
4.2 DATRAS.....	7
4.3 Quality control database.....	8
4.4 Spatial facility	8
4.5 Data archaeology	8
4.6 Other topics	9
5 Digital citation.....	9
5.1 Recent developments	9
5.1.1 Marine Scotland	9
5.1.2 Other organisations	10
5.2 The 8 principles on Data Citation.....	10
5.2.1 Agreement	10
5.2.2 Incorporation of citation guidelines in ICES Data Policy.....	10
6 Data guidelines	12
6.1 Data guidelines submitted to IODE/JCOMM/ICES clearing house (ocean data practices)	12
6.2 Data guidelines not yet submitted to IODE/JCOMM/ICES clearinghouse	12
7 Other topics.....	12
7.1 Training ‘Making the most of ICES data’	12
7.2 Data availability for and data transfer between ICES groups	13
7.2.1 “Data availability for ICES groups”	14
7.2.2 “Data transfer between ICES groups”	15
7.2.3 First step forward.....	16
8 Follow-up of actions and recommendations	17
8.1 2013 recommendation to DIG	17
8.2 2013 and 2014 recommendations from DIG.....	17

8.3 DIG Actions	19
Annex 1: List of participants.....	22
Annex 2: DIG Agenda 2015	25
Annex 3: DIG Draft Terms of Reference for the next meeting.....	27
Annex 4: Recommendations and Actions	28
Annex 5: Data Plan tables	30
Annex 6: SCICOM response on endorsement of 8 principles on Data Citation	39
Annex 7: Training ‘Making the most of ICES Data’	42
Annex 8: Template for inventory of data(products) used by ICES Expert Groups	44
Annex 9: Data availability in ICES groups.....	46

Executive summary

The Data and Information Group (DIG) met in Copenhagen, 18-20 May 2015. 15 people representing 9 different countries, a representative from OSPAR, Head of ICES Data Centre, and ca. 10 members of the ICES Data Centre joined the meeting.

During the 2015 plenary meeting, the group reviewed the progress related to the data deliverables of ICES Strategic Plan and furthermore met up with ICES Data Centre, and discussed digital data citation, progress on ICES Data Guidelines, and data availability between and within ICES Groups.

Data availability in ICES groups

As concluded by a number of groups and recently highlighted at the Bureau meeting in February 2015, there is an issue in the ICES strategic plan implementation in the systematic understanding of what data sources are being used, by whom, what is the quality of these data, how access is provided to these data, and when, and where the gaps in provision of data and data products are. This undermines the advice process, and is likely a cause of inefficiencies and duplication of effort.

To (1) have an overview of the datasets/-products used and/or created by all ICES Expert Groups, and (2) gain insight in the data flows between the groups, DIG proposed an approach to ACOM and SCICOM chairs.

Eight pre-selected ICES Expert Groups will be asked to fill in meta data of the datasets/-products they use and/or create and/or manage in an online catalogue, which will be publicly available and searchable. After the eight groups have provided the information, the information and the filling process will be evaluated by DIG, and other groups will be asked to add to the catalogue. The catalogue will also be pre-filled with information about existing ICES managed datasets and data products (stock assessment graphs, survey indices, ICES database regional datasets, etc.).

ICES dataset collections and portals

Tools and facilities that have been developed by the ICES Data Centre were presented and discussed:

- Commercial catch sampling: It is important to align Regional Database (RDB) Fish-frame and InterCatch as data submitters now have to submit the same data twice, on different aggregation levels. DIG supports the decision made by ICES Council to fund an extra person for 1.5 years to work on the RDB.
- Biological trawl survey: The most recent developments of the Database on Trawl Surveys ([DATRAS](#)) were presented. In October 2014, the Workshop on Integrating DATRAS Products ([WKIDP](#)) took place. This group defined new output products from the DATRAS system, and encouraged ICES Data Centre to make tools developed outside ICES available via the DATRAS webpage.
- Quality Control: The database contains all checks that are carried out on ICES databases, to create insight in quality flagging. A DIG subgroup will, intersessionally, think along with ICES Data Centre in 2015 to make quality information on databases available to end-users.
- Data archaeology: For the EMODnet Biology project, Danish benthic data 1910-1963 and trawl data for eel, including stomach data and other fish species caught were digitised as a pilot to investigate the time needed to digitise historic data.

ICES can only carry out data archaeology at ICES when project funding is available. DIG pointed out that historic data are very relevant for baseline studies for MSFD.

- Marine litter: The litter (seafloor, microplastics) reporting format is ready and tested. Two input formats were developed: one using the [DATRAS](#) id key, and an option for [environmental reporting format ERF3.2](#). Data will be presented through the data portal as one marine litter data type.
- Hydrochemical: The ICES Data Centre is facilitating making the HELCOM Eutrophication Assessment operational through the [EUTRO-OPER](#) project.

ICES Data Guidelines

ICES exposes its Data Guidelines on the ICES website (<http://tinyurl.com/md2hhgb>) and, since mid-2014, also on the repository of IODE/JCOMM/ICES (<http://www.oceandatapRACTICES.net/>). However, a short survey in 2014 learned that the existence of the ICES Data Guidelines is greatly unknown. DIG asked ICES and IODE to provide usage statistics of the Data Guidelines from both the ICES website as well as from the OceandatapRACTICES Repository. Based on the outcome of the statistics, DIG will decide on whether and how to review and update the existing Data Guidelines.

Digital Data Citation

Citation of data can give proper credit to data providers who have made data available to the scientific community. Operational examples of Digital Data Citation were discussed.

ICES Data Plan

On most topics scheduled for 2015, progress had been made. On some topics, ICES Data Centre could not make progress due to budget restraints.

1 Opening of the meeting

The Data and Information Group (DIG) met in Copenhagen, 18-20 May 2015. 15 people representing 9 different countries, a representative from OSPAR, Head of ICES Data Centre, and ca. 10 members of the ICES Data Centre joined the meeting.

The participants' list is in Annex 1.



DIG participants in 2015.

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 (WKIDP, 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 2015. 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.

Main topics discussed during the meeting were:

- i. Progress on ICES Data Plan (Chapter 3, Annex 5, related to ToR a)
- ii. Information exchange with ICES Data Centre (Chapter 4, related to ToR b)

- iii. Digital citation, mainly focusing on data, within ICES (Chapter 5, Annex 6, related to ToR e)
- iv. Availability, versioning and visibility of ICES Data Guidelines (Chapter 6, related to ToR d)
- v. Develop training 'Making the most of ICES Data' (Chapter 7.1, Annex 7)
- vi. Data flow between ICES groups (Chapter 7.2, Annex 8)

The agenda of the meeting is in Annex 2.

3 Progress on ICES Data Plan

The ICES Data Plan was reviewed and the status of the different topics was identified. All items having a deadline or milestone in 2015 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 ASC theme session proposals

The DIG ASC theme session proposal was rejected. Over the last years, it has become clear that only combined data sessions with other (scientific) groups were accepted. As Data is one of the ICES Strategic pillars it is important that this comes forward at the ICES ASC. DIG finds it very important to have Data as a topic present at the ICES ASC, whether it is as a theme session or in a different format. This point will be brought forward during Bureau meeting in June and will be discussed with the ASC organisers.

The Big Data session at ASC 2014 was a moderate success. DIG should be prepared that there is sufficient expertise in and support for the session once it is accepted.

3.2 Other topics

An extra column has been added to the Data Plan tables (Annex 5) describing the progress till May 2015. Where relevant, a reference to a section in this report is provided.

4 ICES Data Centre

The ICES Data Centre presented last year's major developments and asked for feedback.

4.1 InterCatch and Regional Database (RDB) Fishframe

The main focus of the RDB on the short term is to support regional coordination meetings (RCMs). It is important to align RDB and InterCatch as data submitters now have to submit the same data twice, on different aggregation levels.

The EU funding for RDB only supports maintenance and hosting, and not development. Development should so far go via projects, but ICES' intention is that EU also will fund the RDB development. Currently a 1.5-year time-limited development period is funded from ICES budget. It is currently not clear how many stock coordinators are using the RDB. This should become clearer once metadata on the datasets used is being provided by the various expert groups. In the most optimal case, re-aggregation (e.g. in case of area and temporal level) of data from InterCatch could be done based on RDB.

Feedback about the use of InterCatch experienced by the expert working groups WKBALTCOD (Benchmark Workshop on Baltic Cod Stocks) and WGBFAS (Baltic

Fisheries Assessment Working Group) was presented (see also the relevant group reports).

- To conduct benchmark assessments for the two Baltic cod stocks within WKBALTCOD, a data call was issued. Instead of the normally used age based data, length based data should be uploaded to InterCatch. Since many countries imported the same catches again, but with different métiers (catch data was imported for the assessment as e.g. 'Trawl' and then again for WGBALTCOD as 'Active'), many double data sets were included in InterCatch. The functionality that data imports for the whole year would all be included was not communicated clearly enough to data providers with the data call. Correction of these errors as well as some additional database issues increased the workload of the stock coordinators using InterCatch. The final achievement of length-based catch data was not only delayed by several weeks but the results still showed large discrepancies. WKBALTCOD recommends that either InterCatch is enhanced and/or that investments are done in the further development of RDB. Simple, simultaneous storage and use of age and length data for a given year as well as holding and merging of more than one data set (stratum) per year, together with the possibility to select the needed strata when extracting data, should be possible.
- Based on the experiences when using InterCatch to prepare and aggregate stock data for the assessment of cod, flatfish and small pelagics in the Baltic, WGBFAS recommends that DIG considers a transition to the FishFrame/RDB platform with a clear timeline for implementation. This is based on the opinion that a database with access to raw data is preferred, so that data handling for benchmarks and exploration of data would be enhanced. In addition, several functions that are requested by stock coordinators are available in RDB, already.

DIG discussed the feedback and recommendations from the working groups and supported the decision made by ICES Council to fund an extra person for 1.5 years to work on the RDB.



ICES Data Centre presenting achievements to DIG.



4.2 DATRAS

The most recent developments of the Database on Trawl Surveys ([DATRAS](#)) were presented:

- In October 2014, the Workshop on Integrating DATRAS Products ([WKIDP](#)) took place. This group defined new output products from the DATRAS system. Products related to IBTSWG surveys and WGBIFS surveys are online available. The products from the Beam trawl survey data are almost ready and will be put online this summer.
- MSFD product for MSFD large fish indicator (LFI) is under review and will be available when agreed upon.
- During the Working group on demersal fish in North Sea and Skagerrak (WGNSSK) in 2015 it appeared that by the major submission agreed by IBTS Working Group changes occurred in the indices of some species. This was caused by an incomplete data submission by one of the countries. It is important to think about how to register changes in the database. Currently QA/QC version control is data submitter's responsibility, but there is a role for ICES too.
- There is a joint ICES/IMARES (Netherlands) project to automatize data submission to DATRAS. It will be made as generic as possible so other countries can easily join if this project is successful.
- Partial upload for DATRAS data will be ready this year.
- The current DATRAS web services were tested by VLIZ and no problems were found
- Changes suggested by WKDATR (2013) were successfully implemented in DATRAS in 2014.

4.3 Quality control database

The (internal ICES) database contains all checks that are carried out on ICES databases. The checks are however not easy to understand for non-experts. The reason for setting up the database was to get insight in quality flagging, so it should be investigated which information should be shared with the end-users. A DIG subgroup will, inter-sessionally, think along with ICES Data Centre in 2015 to make quality information on databases available to end-users.

4.4 Spatial facility

There is a new (Arc GIS based) spatial facility under development. Metadata are easily accessible via this facility and features are being developed to create user-friendly output. DIG discussed the pros and cons for a facility based on ArcGIS, as opposed to an open-source approach. A DIG subgroup is installed to support ICES Data Centre in the development of new features, and test the facility.

4.5 Data archaeology

For the EMODNET Biology project, Danish benthic data 1910–1963 and trawl data for eel, including stomach data and other fish species caught were digitised. It was a pilot to investigate the time needed to digitise historic data. The output of the ICES Working Group on the History of Fish and Fisheries (WGHIST) was used to locate historic data. ICES can only carry out data archaeology at ICES when project funding is available. DIG pointed out that historic data are very relevant for baseline studies for MSFD. In June an EMODNET workshop will be organised how to best retrieve information from written lists. OCR (optical character recognition) does not work properly for most paper entries. To be able to get better insight in historic data, DIG recommends that

WGHIST makes the list of historic datasets publicly available. It was mentioned in the discussion that crowdsourcing might be a way to digitise historic data, although this should be thoroughly investigated before implemented.

4.6 Other topics

- The litter (seafloor, microplastics) reporting format is ready and tested. Two input formats were developed: one using the [DATRAS](#) id key, and an option for [environmental reporting format ERF3.2](#). Data will be presented through the data portal as one litter data type.
- The stock assessment graphs are ready and being used by stock assessment groups and in the advice
- [Fish eggs and larvae database](#): major steps were made by the Workshop on the ICES Egg and Larval Database (WKIELD) in April 2015. Multiple datasets were made available to the database. It was decided that fecundity and atresia data will be incorporated in a separate ICES database. The addition of MIKey data (a small net attached to the regular MIK net used for herring larvae sampling during Q1 IBTS) was discussed and agreed upon.
- HELCOM Eutrophication Assessment Tool (HEAT): The ICES Data Centre is facilitating making the HELCOM Eutrophication Assessment operational through the [EUTRO-OPER](#) project. The ICES Data Centre will be hosting the assessment database where the individual indicators will be calculated base on a factsheet per indicator. The implementation will be finalized by October 2015.
- ICES is partnering in the H2020 project AtlantOS and has a task to create an acoustic (treated data) database. The acoustic database is under development in collaboration with WGFASST and WGIPS, and overseen by SSGIEOM

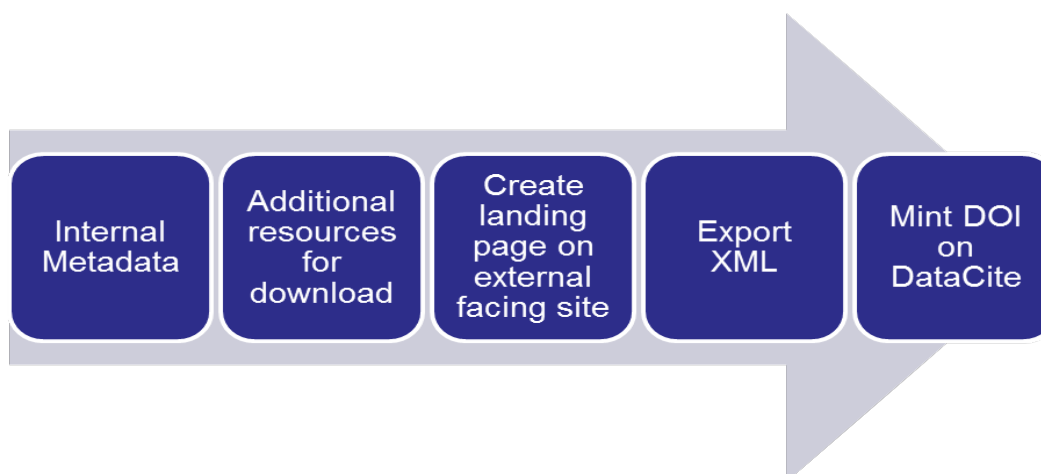
5 Digital citation

5.1 Recent developments

ICES investigated the possibility to mint Permanent Identifiers (PIDs) to datasets. A contract was drafted with the Danish PID provider (DTU), but it was not signed to budget cuts. This year no budget yet. Technically it is possible to mint PIDs to datasets at ICES.

5.1.1 Marine Scotland

At Marine Scotland data citation has been made possible. UK and Scottish government has in general been moving into more open data and open data publication and through participating in ICES discussions on mechanisms for data citation has enabled us to implement persistent identifiers, in this case DOIs. Marine Scotland launched the marine data publication portal in November 2014 and is now gradually building up published datasets and reports from Marine Scotland. The data publication portal at Marine Scotland marinedata.scotland.gov.uk is built using a workflow that comprises by both internal metadata and a public facing data portal which provides the landing pages for the individual DOIs:



5.1.2 Other organisations

BODC (UK) producing DOIs for some time (see https://www.bodc.ac.uk/data/published_data_library/)

NIOZ (NL) developed a repository for publication which can also hold underlying datasets. Data part is not well known, but technically in place.

VLIZ (BE) created guidelines and steps on data citation.

5.2 The 8 principles on Data Citation

5.2.1 Agreement

The 8 principles of the [Joint Declaration of Data Citation](#) were reviewed by DIG in 2014 and in 2015 SCICOM discussed and took note of the 8 principles.

DIG has taken note of SCICOM's comments on the 8 principles and the DIG commentary – and asks SCICOM to disregard the previous DIG commentary and let the 8 principles stand alone. A DIG subgroup will work intersessionally on the interpretation and implementation of the 8 principles within the ICES Data Policy. The next paragraph contains a first inventory of overlaps between the current ICES Data Policy and the 8 principles of data citation.

5.2.2 Incorporation of citation guidelines in ICES Data Policy

The 8 principles should be added as aspirational guidelines to section 7 of the Data Policy. That is:

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

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

The principles are to a greater or lesser extent already covered in the existing ICES Data Policy. The following comments apply to this coverage.

Principle 1:

- Not explicitly stated but discussed in Section 7, Citation.
- 3 (d) also states data must be cited.

Principle 2:

- ICES will store DOIs that are supplied to them – at what level the DOIs are created at will then not be in their control.
- The provision of PIDs/DOIs should be best practice not a requirement.
- ICES need to implement a DOI system before agreeing on details of how it will work.
- DOIs at the highest level will probably be the most practical but they will prove more difficult for people to track down specific data when it is only a small fraction of a total data set.

Principle 3:

- Not explicitly mentioned but making data citable enables this.

Principle 4:

- The data policy on citation does not mention using Persistent IDs.
- We should add information on PIDs or DOIs to the citation section as an aspiration but not a requirement.
- Where DOIs exist these should be used but otherwise the existing citation guidelines should be used.
- Specific examples of citing using the DOIs should be given – specifically how different levels in the hierarchy should be cited.

Principle 5:

- Making PIDs/DOIs a “best practice” for data citations would facilitate this.

Principle 6:

- Again, the references to PIDs/DOIs would enable this.

Principle 7:

- Data Policy Section 4(a) assumes that data will be available by default.

- PIDs/DOIs should be available even for data that is not currently openly available.

Principle 8:

- Covered in 4(d) of the Data Policy

6 Data guidelines

6.1 Data guidelines submitted to IODE/JCOMM/ICES clearing house (ocean data practices)

ICES exposes its Data Guidelines on the ICES website (<http://tinyurl.com/md2hhgb>) and, since mid-2014, also on the repository of IODE/JCOMM/ICES (<http://www.oceandatapactices.net/>). However, a short survey in 2014 learned that the existence of the ICES Data Guidelines is greatly unknown.

Given that the Data Guidelines seem to be rather unknown by the wider public, DIG decided to ask ICES and IODE to provide usage statistics of the Data Guidelines from both the ICES website as well as from the Oceandatapactices Repository. Based on the outcome of the statistics, DIG will decide on whether and how to review and update the existing Data Guidelines. DIG also decided to investigate the need for additional new guidelines.

DIG discussed ways to increase the awareness of the existence of the Data Guidelines, through e.g. the use of social media, ICES training courses, other relevant ICES communication channels, etc. A news item was published at the ICES website: <http://ices.dk/news-and-events/Blogs/Inotherwords/default.aspx>.

DIG also reviewed the draft Oceandatapactices Repository Policy Document v0.5 written by Pauline Simpson and colleagues at IODE. The subgroup discussed, and produced a series of questions concerning the contents of this Policy Document. A reaction to that document has been sent to IODE directly. The response of IODE will be discussed intersessionally by a DIG subgroup.

6.2 Data guidelines not yet submitted to IODE/JCOMM/ICES clearinghouse

Currently, 13 ICES data guidelines have been uploaded to the clearing house. Two other ICES publication types (Survey protocols SISP and TIMES) numbering 64 in total are still planned to be submitted to the clearing house. However, to meet the metadata standard required by oceandatapactices, ICES would need to carry out some additional work i.e. create abstracts for the 54 TIMES publications. There is currently no resource to do this, so the action cannot be completed.

7 Other topics

7.1 Training 'Making the most of ICES data'

DIG discussed the options for a training to make data stored in ICES databases easier accessible for a wider audience. A first draft for the training (Annex 7) was created and should be put into practice in 2016 for at least a selection of databases.

The training is supposed to be modular, reflecting the numerous databases available via ICES. It would be beneficial to host a survey to gauge interest in the different topics proposed in Annex 7 before finalising the courses to be prepared for the first year.

7.2 Data availability for and data transfer between ICES groups

In the ICES community, many groups use data for various purposes. Some groups use data in databases hosted by ICES, other groups collate data from different sources, or only use data products.

There is a need to create an overview of the data and data products used by the ICES groups for a number of reasons. For example, duplication of effort should be prevented whenever possible, documentation on the data(products) used is needed to create insight in the quality, data(product) needs should match the data products.

For all datasets managed at ICES, and for most of the data products stored at and/or created by ICES, [metadata](#) records are available describing the data in the database, but not the use and/or the user(s). Some effort has been undertaken by the SCICOM/ACOM Steering Group on integrated ecosystem observation and monitoring ([SSGIOEM](#)) to describe the relations between ICES survey planning groups and ICES stock assessment groups. This does not give information on data used by other ICES groups. DIG discussed the data availability in ICES groups and Data (product) transfer between ICES groups using a stepwise approach: (1) two mind maps were created (section 7.2.1, 7.2.2) and (2) a potential first step on the way forward was proposed (section 7.2.3).



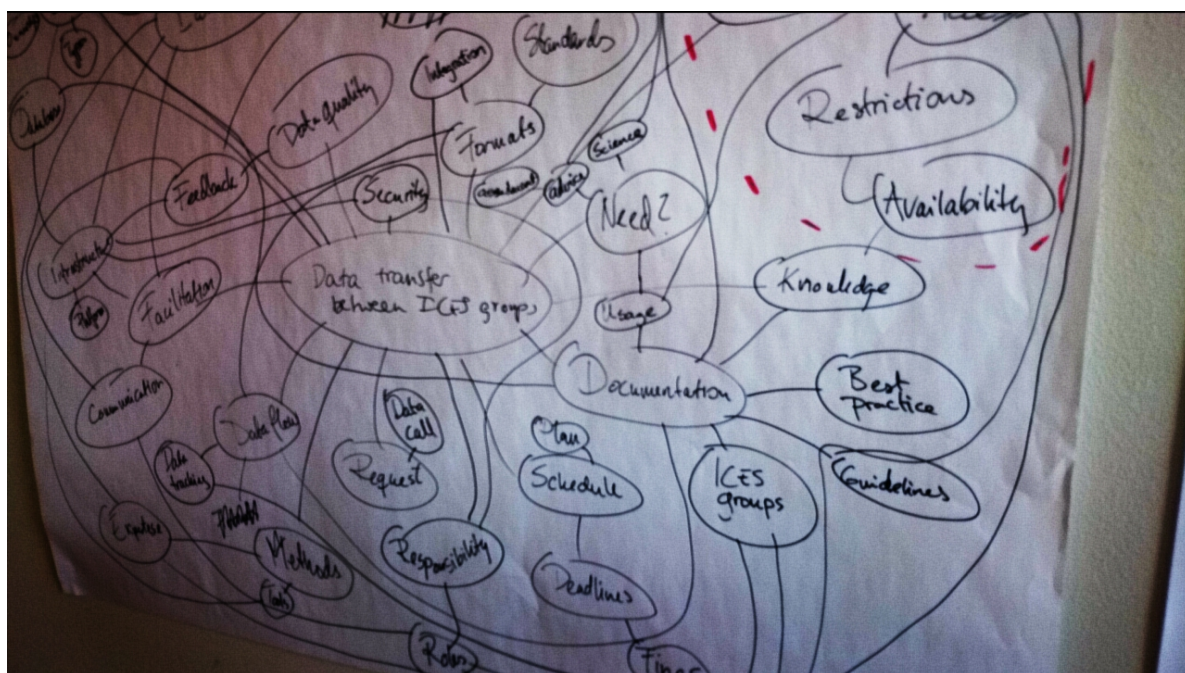
Mindmaps in progress.



3. Once we know what's available then communication and training can be used to raise awareness of it within working groups.
4. The ICES Data Policy should be a vehicle to identify and promote best practice.
5. The Operational Oceanographic Products and Services (OOPS) ongoing initiative is proving a successful way to identify and fill existing data gaps.

It was agreed that buy-in from the wider ICES community, including PGDATA and SSGIEOM, would be necessary for any successful solution.

7.2.2 "Data transfer between ICES groups"



A subgroup under DIG worked on a mind-map on "Data transfer between ICES groups". Many interrelated issues were mapped. Summarizing, this results in the following issues:

1. Facilitation of communication between ICES groups depending on each other's data:
 - o Clear descriptions of the data (product) need for ICES groups. What is the purpose (advice, science, stock assessment), what should the product look like, which algorithms have to be used, etc.
 - o In some cases a formal Data request/data call can help to streamline the data(product) flow
 - o Clear definition of the roles and responsibilities of the ICES groups with respect to data
 - o Insight in data(product) flow, data tracking
 - o Infrastructure, platform(s) for data(product) exchange
2. Information and standardization:
 - o Formats and standards for data and data products
 - o Integration
 - o Data issues: different data types, data sources, data limitations, data citation, databases, , ownership

- o Documentation of data: best practices, data guidelines, data usage to
- o Data quality
- 3. Insight in and knowledge about available data
 - o Metadata including description of methods, tools
 - o Data security
 - o Data policy: availability of data, restrictions, access, legislation and rules.
 - o Need to have a schedule/plan and deadlines. Fines were also mentioned.
 - o Expertise on data, both from the ICES Data Centre and in the ICES groups

7.2.3 First step forward

The first step forward in this process is to get insight in the data and data products used by the various ICES Expert Groups. DIG has developed a template to collect the necessary information (Annex 8). ICES Data Centre will create an online facility and a database to enter the information into. The database should in the end be searchable on dataset, data product and ICES Expert Group. Eight ICES Expert Groups were identified to start a pilot with. Those groups (a) represent a wide range of expert fields and data use and (b) are groups in which a DIG member is involved. The groups identified are: WGSFD, WGBYC, WGOH, WGZE, WGINOSE, WGBEAM, WGDEC and a benchmark group. The selection of a specific benchmark group is to be discussed with the Benchmark Steering Group.

The first information should be added by the end of 2015 so DIG 2016 can evaluate the results, the use of the online facility, and decide on further steps.

SCICOM, ACOM and SSGIOEM chairs will be informed about this proposal, and be asked for support. The full proposal is in Annex 9.

8 Follow-up of actions and recommendations

Below is the follow-up of recommendations to DIG (8.1), from DIG (8.2), and DIG action items (8.3).

8.1 2013 recommendation to DIG

One open 2013 recommendation to DIG by WKESST could be found in the ICES recommendations database:

‘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).’

Both DIG and Seadatanet see it as an advantage to use the CSRs for other purposes than only reporting on the cruise.

1. First of all, there must be a CSR.
2. And if so, sufficient geographical information should be in the CSR
3. If a CSR is available and sufficient geographical information is supplied in the CSR, the SeaDatNet geonetwork tools are capable to provide such a mapping tool. The mapping of the geographic information in then CSR to the level needed for WKESST might cause some problems. The existing CSR has just Marsden Squares as minimal geographical information (e.g. would just give "North Sea", not "German Bight").

There are three possibilities to solve the mapping issue:

- a. CSRs must provide bounding boxes of the survey area, this is already implemented in CSR online
- b. CSRs must include specific area names such as German Bight
- c. SSGIOEM provides a list of needed regional seas names, which would be included in the CSRs online tool in collaboration with SeaVOX. The ICES Eco Regions should also be added to that list.

Suggestion c appears as the most practical way. After that, BSH would provide a WMS based on the CSRs, which could be used by ICES and SSGIOEM. In general, we like to cooperate with SSGIOEM giving CSRs an additional function and use.

8.2 2013 and 2014 recommendations from DIG

NR	RECOMMENDATION	ADDRESSED TO	STATUS
2014-250	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. 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.	ACOM, SCICOM	No action
2014-251	Find a quality flagging scheme fitting best to the needs	ICES Data Centre	In progress
2014-252	To make the inclusion of the MIKey Net data in the database feasible (this recommendation was originally sent from WGECCS2 to DIG, 2013)	ICES Data Centre	In progress, see section 4.6 of this report
2014-253	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).	ICES Data Centre	In progress, see section 5.1 of this report

	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)		
2014-254	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.	ICES Data Centre	In progress, see section 5.1 of this report
2014-255	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	In progress, see Chapter 5 of this report
2014-256	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).	Nils Olav Handegard (SSGESST chair)	In progress
2014-257	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.	ICES Data Centre	Complete
2014-258	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.	ICES Data Centre	Only possible when major investments are being done. There is a secure server with the data, but people work on their own computers as all tools are available at their computers.
2013-163	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 “(…) to inform ICES Data Centre (this can be sent out as an online questionnaire to EG’s to fill in) 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) 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) i. IF YES: in which database and under which coding (List will be supplied by ICES Data centre)”	SCICOM and ACOM	No action, new recommendation proposed see section 7.2.3 of this report

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.

8.3 DIG Actions

DIG 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	N/A
2. Keep a shortlist of topics that might be put on the ICES webpage on behalf of DIG	Ingeborg de Boois	All year	N/A
3. to speed up the first draft of the documentation on the methodology OSPAR Hazardous substances	Chris Moulton, Neil Holdsworth	1 July 2014	Complete (18 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	In progress, see
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	Complete; no final leaflet. Communications budget was cut and brochure was put on a hold. Idea to have little cards by dataset. Draft will be sent to DIG
6. Investigate the possibility to provide PIDs for standard stock assessment graphs	Neil Holdsworth	1 April 2015	On hold
7. Make a list of keywords that will be used on the data guidelines, TIMES and survey protocols.	ICES Data Centre	15 June 2014	Complete
8. Announce to IODE that a keyword list is going to be provided by ICES	Taco de Bruin	15 June 2014	Complete
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	Complete
10. Add ICES as a source for documents in the IODE clearinghouse	Taco de Bruin	15 June 2014	Complete
11. Provide the data type guidelines to IODE clearinghouse	ICES Data Centre	15 June 2014	Complete
12. Add TIMES and survey protocols to	ICES Data	When ready	In progress

the IODE clearinghouse	Centre		
13. Update parts regarding “getting data to ICES” for all data guidelines	ICES Data Centre	1 October 2014	Complete
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	Complete
17. Provide updated Data Guidelines to clearing house	ICES Data Centre	When ready	Complete and ongoing
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	Complete
19. Suggest conveners and presenters for ASC 2014 Big Data Session to Jens	All DIG members	6 June 2014	Complete
20. Provide suggestions for ASC 2015 theme sessions to Peter	All DIG members	15 June	Complete
21. Write working document on how to Data Guidelines group	Taco, Sjur, Lesley, Marcin, Ruth, Hjalte/Neil	1 April 2015	Complete
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	Complete
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	Complete
24. Encourage their organisations to endorse the principles (organisational endorsement) (see section 5.2)	All DIG participants	1 December 2014	Partly done, see section 5.1 of this report
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	Complete
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).’ (based on answer on action 23)	Friedrich Nast, Hjalte Parner	1 October 2014	Complete
27. Ask WGESG2 to send in a data request form related to the MIKey data.	Ingeborg de Boois	1 July 2014	Complete Eggs and larvae

28. Ask WGAQUA if a database is needed	Ingeborg de Boois	1 July 2014	In progress (mail sent, WGAQUA will discuss, no response yet) No action needed yet, WGAQUA sees no need for a database.
--	-------------------	-------------	--

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
Örjan Bäck	SMHI Sven Källfelts gata 15 42671 Västra Frölunda Sweden	+46 31-751 8956	orjan.back@smhi.se
Taco de Bruin	NIOZ P.O. Box 1790AB Den Burg Netherlands	+31-222-369479	Taco.de.Bruin@nioz.nl
Simon Claus	VLIZ Wandelaarkaai 7 8400 Oostende Belgium	+32 59-340159	Simon.claus@vliz.be
David Currie	Marine Institute Rinville Oranmore Co. Galway Ireland	+353 (0) 91 38 7403	David.currie@marine.ie
Christian von Dorrien	Thuenen-Institute of Baltic Sea Fisheries Alter Hafen Sued 2 18069 Rostock Germany		christian.dorrien@ti.bund.de
Josefine Egekvist	DTU Aqua Charlottenlund Slot Jægersborg Allé 1 2920 Charlottenlund Denmark	+45 35883438	jsv@aqua.dtu.dk
Neil Holdsworth	ICES H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark	+45 3338 6718	neilh@ices.dk
Hans Mose Jensen	ICES H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark	+45 33386736	hans.jensen@ices.dk
Henrik Kjems-Nielsen	ICES H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark	+45 33386734	henrikkn@ices.dk

NAME	ADDRESS	PHONE/FAX	EMAIL
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
Chris Moulton	OSPAR		Chris.Moulton@ospar.org
Friedrich Nast	Bundesamt fuer Seeschifffahrt und Hydrografie Bernhard-Nocht-Str. 78 20359 Hamburg Germany	+49 40 3190 3420	friedrich.nast@bsh.de
Hjalte Parner	ICES H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark	+45 33386737	hjalte@ices.dk
Carlos Pinto	ICES H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark	Tel. +45 33386713	carlos@ices.dk
Jens Rasmussen	Marine Scotland Science Marine Laboratory 375 Victoria Road Aberdeen AB11 9DB United Kingdom	+44 (0) 1224 295 440	Jens.rasmussen@scotland.gov.uk
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
Sjur Ringheim Lid	Havforskningsinstituttet Postboks 1870 Nordnes 5817 Bergen Norway	+47-99556943	Sjur.Ringheim.Lid@imr.no
Helge Sagen	Institute of Marine Research PoBox 1870, Nordnes 5817 Bergen Norway	+47 95215046	Helge.Sagen@imr.no
Marilynn Sorensen	ICES H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark	+45 33386720	marilynn@ices.dk

NAME	ADDRESS	PHONE/FAX	EMAIL
Lena Szymanek	National Marine Fisheries Research Institute, Kollataja 1 81-332 Gdynia Poland	+48 58 7356135	lena.szymanek@mir.gdynia.pl
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

Annex 2: DIG Agenda 2015

Monday 18 May

13.00 – 13.30	Logistics, round the table, etc.
13.30 – 14.30	Recommendations and actions 2014 (see Annex 1)
14.30 – 15.00	Data guidelines: progress [actions 7-18, 21 and Data Plan –Annex 2] (Taco, Hjalte)
15.00 – 15.20	Digital citation: how to proceed? Discussion [actions 6/24/25 and Data Plan]; feedback SCICOM
15.20	Safety, logistics etc. (Vivian)
15.30 – 16.00	Tea
16.00 – 16.30	Feedback from workshops, working groups, etc. and potential actions for DIG: • WKIDP [Data Plan] Jens/Ingeborg • WGSFD? Josefine • WKBALTCOD/WGBFAS Christian
16.30 – 17.00	Outline Training ‘How to make more of ICES Data?’ (Jens/Ingeborg/Neil)
17.00 – 17.30	Progress Data Plan if not covered in other agenda items, focus on highlighted topics in the table below
17.30	Identification of subgroup topics and subgroup participants

Tuesday 19 May

9.00 – 12.00	(including a coffee break) Update from and feedback to ICES Data Centre (ICES Data Centre) Including: DATRAS developments (based on WKIDP outcomes, and participation in WGBIFS, IBTSWG, WGBEAM Litter from trawl surveys: state of the art [Data Plan] QC database [Data Plan] New datasets [Data Plan] Reflection on Data Centre Work Plan MSFD Workflow [Data Plan], maybe include actions by ICES Data Centre resulting from EU projects BALSAM, IRIS-SES, JMP NS/CS? ?Data archaeology [Data Plan]
12.00 – 12.30	Plenary: summary of morning session, highlights, follow-up, etc.

12.30 – 13.30	Lunch
13.30 – 14.00	Optimising Data transfer between ICES groups (see Annex 3)
14.00 – 16.45	Subgroups (may be changed as a result of progress on Monday and outcomes Tuesday morning session, including tea break): • Data guidelines • Digital citation • Training ‘How to make more of ICES Data?’ • Optimising Data transfer between ICES groups
16.45 – 17.00	Theme sessions on conferences [Data Plan]: Ideas for 2016? ASC, other conferences?
17.00 – 17.30	Plenary: Wrap-up from and feedback to subgroups

Wednesday 20 May

09.00 – 12.00	Subgroups (may be changed as a result of progress on Tuesday, including coffee break): • Data availability in ICES EGs • Optimising Data transfer between ICES groups
12.00 – 13.00	Lunch
13.15 – 15.00	Preparing presentations of DIG work:
13.15	Plenary: agree on topics for outreach • Make questionnaire for ASC 2015 to investigate which databases are most interesting Ruth, Sjur • Guidelines; ‘In other words’ Taco • Data citation; explanation on 8 principles, Marine Scotland DOIs Örjan, Jens • Meta database; needs more time, but some text for ACOM and SCICOM and SSGIOEM needed Ingeborg • Mindmap ideas for report: Josefine, David • Finalising report sections
13.30	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 2015
15.00	Plenary: review subgroup texts and decide on final versions for social media and ICES Inside out Planning DIG meeting at ASC 2015 (who will be there?) and setting dates for DIG 2016. No one there planned. Next year’s meeting: 23–25 May 2016 (Monday 13.00-18.00, Tuesday 9-18, Wed 9-18)

Annex 3: DIG Draft Terms of Reference for the next meeting

The **Data and Information Group (DIG)**, chaired by Ingeborg de Boois, Netherlands, will meet in Copenhagen, Denmark, Monday 23 May (13:00)–Wednesday 25 May (18:00) 2016 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 (LinkedIn Data and Information Forum) if relevant
- e) Promote new technologies and data management infrastructure development (e.g. IODE/JCOMM/ICES Clearing house, data citation, training)

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 XX July 2016.

The group reports to SCICOM during the SCICOM midterm meeting March 2016 as well as the SCICOM meeting at ICES ASC 2015. The group reports to ACOM by correspondence and via the ACOM representative.

Supporting Information

Priority	The Data and Information Group provides ICES with solicited and unsolicited advice on all aspects of data management including technical, data policy and data strategy and user oriented guidance. This operational group flies the flag for ICES in setting standards for global databases. It also provides an important interface for oceanographic, environmental, and fisheries data management in ICES, and promotes good data management practice.
Scientific justification	a), b), c), d), e) are direct results of DIG's main priority: The Group provides ICES with solicited and unsolicited advice on all aspects of data management including technical, data policy and data strategy and user oriented guidance.
Resource requirements	The resource required to undertake additional activities in the framework of this group is negligible.
Participants	The Group is expected to be attended by some 20–30 members, with good international and topical coverage.
Secretariat facilities	Meeting facilities, organization and facilitation of WebEx meetings (frequency and participants depending on topics to be discussed..
Financial	No financial implications.
Linkages to advisory committees	ACOM
Linkages to other committees or groups	As Data is an important topic for most groups under SCICOM and ACOM, this group links to a large number of groups, although often indirect.
Linkages to other organizations	There are linkages with relevant international bodies and programmes like PICES, IOC/IODE, GOOS, SeaDatanet, IPY, etc., with emphasis on IOC and its Working Committee on International Oceanographic Data and Information Exchange (IODE).

Annex 4: Recommendations and Actions

Recommendations

RECOMMENDATION	ADDRESSED TO
1. To be able to get better insight in historic data, it is recommended that the list of historic datasets is made publicly available.	WGHIST

Actions

(in *Italics*: focal point)

ACTION	ADDRESSED TO	ACTION BEFORE	STATUS
1. Prepare draft new version ICES Data Policy: include exceptions on data permissions, incorporation of 8 principles	Ingeborg de Boois, Christian von Dorrien, Chris Moulton, Joni Kaitaranta, Simon Claus, Neil Holdsworth	1 April 2016	
2. Data guidelines: compile statistics, approach to incorporate missing guidelines in the IODE/JCOMM/ICES clearing house	Taco de Bruin, Hjalte Parner, Ruth Lagring, Sjur Ringheim Lid	1 December 2015	
3. Data guidelines: prepare response on IODE draft Oceandatapactices Repository Policy Document v0.5	Taco de Bruin, Hjalte Parner, Ruth Lagring, Sjur Ringheim Lid	15 June 2015	
4. Send response on IODE draft Oceandatapactices Repository Policy Document v0.5 to IODE	Ingeborg de Boois, Taco de Bruin, Neil Holdsworth	25 June 2015	
5. QC database: develop output for end-users	Periklis Panagiotidis, Malin Werner, Simon Claus	1 April 2016	
6. Spatial facility: further develop and evaluate tools&widgets; evaluate technical aspects	Periklis Panagiotidis, Lena Szymanek, Jens Rasmussen, Lesley Rickards	1 April 2016	
7. Training 'Making the most of ICES data': develop a survey to gauge interest in the different topics (survey aimed for ICES ASC 2015)	Jens Rasmussen, Ingeborg de Boois, Carlos Pinto, Sjur Ringheim Lid, Ruth Lagring	1 August 2015	
8. Highlight the importance of the topic 'Data' during ICES ASC	Neil Holdsworth (Bureau), Ingeborg de Boois (ASC organisers)	1 August 2015	

9. Pilot metadata: identify benchmark group to be involved together with Jörn Schmidt, Carmen Fernandez	Ingeborg de Boois	15 June 2015
10. Pilot metadata: discuss approach with ACOM, SCICOM and SSGIOEM chairs	Ingeborg de Boois, Neil Holdsworth	15 June 2015
11. Pilot metadata: develop online form based on Annex 8	ICES Data Centre (Neil Holdsworth)	1 August 2015
12. Use online form to fill in the fields for the identified Expert Groups, and provide feedback	Ingeborg de Boois, Neil Holdsworth, Josefine Egekvist, Carlos Pinto, Christian von Dorrien, Peter Wiebe, Lena Szymanek, Jens Rasmussen	1 December 2015

Annex 5: Data Plan tables

Regional Facilitation						Resource implication		
Headline action	Detail	Performance measure	Timing	Status DIG 2015	Status DIG 2014	Data Centre	DIG	Other
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.	05/2015: a. Progress on all workflows i.e. EUTRO-OPER, (see also chapter 4 of this report). b. complete	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.	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.		
	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 ICES data/WG's and their operational product linkage)	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	05/2015: (offshore) litter: see section 4.6 of this report	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			

Regional Facilitation							Resource implication		
Headline action	Detail	Performance measure	Timing	Status DIG 2015	Status DIG 2014	Data Centre	DIG	Other	
					as a complement to existing groups/RSC processes				
	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.	05/2015: Microplastics & acoustic data: see section 4.6 of this report. Indicator calculation: see section 4.2 of this report		Depending on the level of ambition regarding establishing new international datasets and systems, additional resources may be required			
	- 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	05/2015: no action	05/2014: no action				
	- 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,		workshop in 2014 to list product requirements	05/2015: WKIDP took place and was	05/2014: workshop is planned in October and will be chaired by Clara Ulrich	Workshop participation and follow-up			

Regional Facilitation							Resource implication		
Headline action	Detail	Performance measure	Timing	Status DIG 2015	Status DIG 2014	Data Centre	DIG	Other	
	WGNEW, DIG proposal -action DIG chair)			successful. Re- port available via ICES website					
End-to-end workflow for scientific advice pro- duction	- RA-CMS linking to data outputs from Expert groups (connecting the scientific reports to advice pro- duction).	Successful implementation of inter- faces to a) scientific output from EG reports b) scientific output from assessment models	starting 2014 (depends on timing RA-CMS develop- ment).	05/2015: Standard graphs: see sec- tion 4.6 of this report	05/2014: Pro- cess delayed. Currently con- centrating on stock input and expanding standard graphs to other stocks. System renamed CARA.	Volume of activ- ity on RA- CMS would re- quire ad- ditional technical resource			
	- RA-CMS linking to data outputs from RDB-Fishframe	See (b) above	2015	05/2015: no ac- tion		Depend- ent on progress in devel- opment (and funding) of RDB- Fish- Frame		De- pend- ent on pro- gress in devel- opment (and fund- ing) of RDB- Fish- Frame	
Mobilising aquaculture specific data	- Aquaculture databases: exact de- scription to be decided. Related to WGAQUA.	Products and/or regional data man- agement established (where man- date is given)	starting from 2014.	05/2015: no ac- tion needed (agreed upon by WGAQUA as the group does not see the need for an aq- uaculture data- base)	05/2014: no ac- tion	Depend- ing on the level of ambi- tion re- garding new da- tasetes and sys- tems, ad- ditional	De- pend- ing on the level of ambi- tion re- garding new da- tasetes		

Regional Facilitation				Status DIG 2015 Status DIG 2014		Resource implication		
Headline action	Detail	Performance measure	Timing			Data Centre	DIG	Other
						resources may be required	and systems, additional resources may be required	
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	05/2015: Slow progress, some test files exchanged. The structure of the data committees is not clear. Meeting in October relate to the polar data forum; Helge Sagen and Taco de Bruin will attend	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	A higher level of technical support/guidance could be anticipated		

International Standards and interoperability				Status 2015	DIG	Status 2014	DIG	Resource implication		
Headline action	Detail	Performance measure	Timing					Data Centre	DIG	Other
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 by other expert groups, processes and member country activities	- All ICES datasets, including those that exist only within an expert group, are adequately described and the 'discovery' information are available through the ICES online portals	- Request to EG's to be filled 2015	05/2015: Technical complete; Jens Rasmussen helped validating the Data Centre's work. Not published yet. Content: no information from EGs; see also section 7.2.3 of this report for new approach		05/2014: ICES Data Services have an online system (INSPIRE compatible).		Some additional guidance and tools will be needed		ICES expert groups will need to incorporate into their work
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	05/2015: is in work plan – work planned after DIG 2015 meeting. See also section 4.3 of this report		05/2014: no progress				

Knowledge transfer and professional development				Status 2015	DIG	Status 2014	Resource implication		
Headline action	Detail	Performance measure	Timing				Data Centre	DIG	Other
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	05/2015: Proposal 2015 ASC was not accepted by SCICOM. There is a need for 'Data' as a topic at ASC, but may be in a different format than a theme session.		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	- ICES training courses: 'Making the most of ICES Data', modular, webinars?. - Online materials and guidance: WKIDG in 2014	(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	- Training: end 2017 - Workshop to produce reference guide in 2014 (WKIDG, proposed)	05/2015: DIG worked on a proposal for training development. See also section 7.1 of this report		05/2014: In progress.		Leading workshop	

Data stewardship and data management					Status DIG 2015	Status DIG 2014	Resource implication		
Headline action	Detail	Performance measure	Timing				Data Centre	DIG	Other
Data archaeology; identifying and making available datasets that are relevant to the marine community	- (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	(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	(a) Start 2014. (b) follow-on from 'IN-SPIRE readiness' activity under heading 3		05/2015: a. see section 4.5 of this report b. see section b. see section 4.5 and 7.2.3 of this report c. no action	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 DIG report 2014 chapter 5	Historic data recovery will require additional resources/funding and this may be possible in part through EMODnet biology		
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	05/2015: See section 5 of this report	05/2014: in progress. See chapter DIG report 2014 chapter 5				

Data stewardship and data management					Status DIG 2015	Status DIG 2014	Resource implication		
Headline action	Detail	Performance measure	Timing				Data Centre	DIG	Other
Maintaining the user rights, security and integrity of the data sources to ICES managed datasets	- 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	05/2015: No action needed, data policy update scheduled for 2016. See also section 5.2.2 of this report	05/2014: RDB-FishFrame data policy drafted but not agreed by all participating countries yet				

Annex 6: SCICOM response on endorsement of 8 principles on Data Citation

SCICOM endorsement of Data Citation Principles (requested by DIG)

ACTION:

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

SCICOM RESPONSE:

"SCICOM has read and taken note of the Data Citations Principles and support them as a general guideline".

Comments from SCICOM

Nils Olav Handegard I fully support the principles, but I am not shure that DIG's comment to point 2 is appropriate: "DIG interpreted this as everyone who has hand in creating the data gets credit for the data productions..." This probably needs clarification similar to the Vancouver declaration at some point, but I fully support the principle.

Jörn Schmidt I also fully support the list.

Graham Pierce I agree with most of the principles but am less sure about the comments made in relation to point 5. My perspective as an individual working in a university is that data collected by my research group cannot usually immediately be made publicly available - they are an investment, a resource for future publications, basically the foundation on which university careers are built., particularly for younger researchers whose future employment depends, quite literally, on their publications. I understand that where data collection is on a large-scale involving many researchers and is publicly funded a different approach is needed and indeed I am an end-user of such data sets so I very much appreciate their availability - but basically there is not a "one size fits all" solution for data access.

Daniel Duplisea I agree in principle but emphasise the point of Nils-Olav: DIG's interpretation of point 2 seems very encompassing. Does their interpretation suggest regular ship crew, researchers and Neil H. would all be treated equally for data citations for data held in ICES databases?

Henn Ojaveer As already indicated above, points 2 and 5 deserve some further attention. Especially point 5: this is a sensitive issue and 'data owner/supplier' should have a possibility to restrict the use/access to the data. This is perhaps not valid for monitoring data, but for data collected within dedicated research projects.

Nils Olav Handegard I am strongly in favour of an open data policy. For some reason the fisheries community is way behind the physical oceanographers where an open data policy have been the de facto standards for decades. As long as the source is acknowledged and cited (hence my previous comment), an open data policy only add value to the data (and to the data owner if properly cited). I would argue that ICES should follow this practice. A university may of course choose another policy, but I assume this discussion is about whether DIG/ICES should adopt this strategy or not.

Thomas Noji I agree that point 2 needs some specificity. EVERYONE involved in helping to collect data should not necessarily be referenced.

RE point 5, I support the principle, but concur with DIG's comments that certain restrictions to data accessibility may be needed in some cases, particularly if the data are proprietary.

Laura Uusitalo Hi. I agree with the principles, but some comments below:

I agree with Nils Olav's comments regarding open data policy and his and Thomas's comments about point 2 of the list. Monitoring data etc. should definitely be available for all right away. Regarding proprietary data such as that collected by PhD students, projects, etc., I think that aspect is covered by this document. I do think that if data is used as an argument in scientific work, it should be available for scrutiny - hence, the PhD student or postdoc could keep their data until they publish, but after the publication they should agree to let others evaluate the data if they request. Letting them use the data is another matter of course, but it's not really the scope of this document, right?

Pierre Petitgas What data are we talking about? DCF data, which are monitoring data have a policy attached to them which I am not sure is completely opened.

Pierre Petitgas What data are we talking about? DCF data, which are monitoring data have a policy attached to them which I am not sure is completely opened.

Pierre Petitgas What data are we talking about? DCF data, which are monitoring data have a policy attached to them which I am not sure is completely opened.

Mats Svensson I fully support the principles.

DCF data have their own policy, particularly principles to protect personalized data, and is not completely open.

Jan Jaap Poos I am not sure what we are endorsing here: the statements in the 8 principles, or the interpretations by DIG. Clearly most of the questions arise from the interpretations, which are "further reaching" than the original principles. Maybe somebody can clarify?

If we endorse point 5 as the interpretation of DIG, who decides what the cases are "where the data and information are sensitive or there are proprietary restrictions on them"?

Antonina dos Santos I am in the same situation of Jan Jaap Poos, not sure of what are we exactly endorsing here. I surely need a clarification.

On the other hand, I understand that we are discussing DIG/ICES data policy but this kind of policies happens to have always further implications on data policy everywhere.

My perspective is that all researchers, working in universities or research institutes, have careers that are based on publications and therefore a data policy with restrictions for an initial period is necessary. Universities are also public institutions in many places and, they also use public funds to collect data. Therefore, if a general data policy is implemented they should include all kinds of data.

Maria Begoña Santos Dear all, Data accesibility is a very sensitive topic, even for a research institute that gets part of its funding from public sources such as the DCF and others. We do have to be very careful about endorsing such statements that potentially have far reaching implications. I will agree as well with the need for clarification on what can be considered "sensitive information and proprietary restrictions"

Yvonne Walther I have noted your question marks which are quite relevant.

I have found more information on the Joint Declaration of Data Citations Principles here which i think gives some answers to your comments:

<https://www.force11.org/datacitation>

I suggest we provide DIG with the following statement " SCICOM has read and taken note of the Data Citations Principles and support them as a general guideline"

Thomas Noji I agree that a general statement like the one you propose, Yvonne, is acceptable.

Graham Pierce I also agree that a general statement along these lines is acceptable

Yvonne Walther Thank you for all your comments

I will forward the statement to DIG as suggested before.

"SCICOM has read and taken note of the Data Citations Principles and support them as a general guideline"

Annex 7: Training ‘Making the most of ICES Data’

What is it: Training! But provided in short, modular sessions to make the topics more digestible. We propose that each “module” is around 2-3 hours and held as informal drop-in session at the ASC, and are subsequently turned into online training modules. These sessions are short, focussed “taster” sessions to increase knowledge and familiarity with what data are available in ICES.

Audience: We believe that it would be beneficial to target new entrants to the ICES community. We often talk about the ICES family and the importance of getting involved during the ASC. This is an opportunity for people not familiar with some data topics to get an introduction from people working with the data and to get a hands-on experience as part of each short course.

Format: Each module is a 2-3 hour slot. The first hour is the presentation created by a member of the data centre in collaboration with someone from the ICES community who works with the data on a regular basis. After the 1 hour presentation/demo session, two exercises are presented to participants, and they can spend the following two hours working on the exercise and ask questions from the presenters. The two hour session is optional for attendants. If they find the course too basic or too complex, they are free to leave after the first hour.

When: First appearance at the Riga ASC in 2016, perhaps with 2-4 modules to happen during the event. The topics of the training courses should be aligned with the formal ASC programme to avoid topic clashes (e.g. try to avoid running training in a topic parallel to a session on the same topic). Then turn the modules into online content/training shortly after the ASC (ideally, the content is ready 1 month after). It will be important to gauge reception of the format. If it proves popular, there’s scope for adding new modules in the following years. DIG would be well placed to review outcome and ensure new modules are prepared each year, gradually building up a comprehensive set of training modules for entrant level ICES community members over the coming years.

Gauging topics and interest: Via an online questionnaire shared on social media. Present a topic list + allow respondents to make their own suggestions. Then coordinate people to arrange/prepare the material in advance of the ASC.

Preparation: It is envisaged that preparing material is a collaborative tasks between a member form the ICES community with expertise, and a member of the Data Centre. Both people would need to be available at the ASC to run the course and assist with exercises. Since we are targeting entry level information about data, it should not be a hugely onerous task to prepare a 1 hour introduction to a particular data tasks.

After the ASC: Especially for the first time, it will be important to get feedback form attendants. This can be achieved simply with attendants adding their email to a list, and subsequently contacted after the event to fill in a short survey.

Turning the events into online training: After the ASC sessions, the material and exercises should be turned into an online course. This should be relatively simple to do by taking a screencast of the presentations and followed up with a quiz or result to compare your exercise outcomes with.

Requirements/Resources: The courses should be freely accessible. Attending the course at the ASC will be free/open to everyone registered for the ASC. The only resource requirement during the ASC would be a room/location with presentation screen

and internet connectivity. Subsequent publishing of online content could be done via the ICES website.

There is a question about resources for the experts to prepare and provide the training.

Promoting: The modules could be promoted through social media, both prior to the ASC, and afterwards when released online.

Guidelines for the presentations: Things that should be included in the 1 hour presentation

1. Workflow
2. Which parts of the community works with the data (creation/submission, use/extraction and reporting). Advice and Science.
3. What systems?
4. How to reach them?
5. Methods for extracting data (e.g. web interface, web service, etc.)
6. Formats, vocabularies, references
7. Documentation and/or data guidelines.
8. Citation
9. Reporting issues and making data requests.
10. Examples of use.

Proposed topics to begin with

Topic	Outcomes	ICES Data Centre	WG Person
Ecosystem data	- How to use the system - Extract data	Carlos	Jens?
DATRAS		Vaishav	Ingeborg
Oceanography		Hjalte/Else	Leslie/Oerjan
Combining Oceanography and Biology		Hjalte/Hans/Vaishav	Lena
Commercial Fisheries Data		Henrik	Christian
Eggs and Larvae		Carlos	Cindy van Damme?
DOMÉ		Marilyn/Hans	Rob Fryer?
Spatial Data Facility		Perikles	Joni, Chris
Reference data (Ecoregions, statistical squares, RECO)		Mike? Mehdi	
Stock assessment graphs			
Fish Stomach database			
Plankton		Joergen	
Using Web services		Carlos	Jens

Please note, no names are confirmed at this stage.

Annex 8: Template for inventory of data (products) used by ICES Expert Groups

	FIELD	TYPE	USAGE	
Geographical context	ICES Ecoregion	Fixed list	Required	
	Other geographical reference system	Fixed list	Optional	
	Geographic Reference value	Fixed list	Optional	
About Dataset	Dataset title	Free text	Required	
	Basic dataset description	Free text	Required	
	Link to existing metadata record for dataset	URL free text	Optional*	If this is used then potentially all fields can be left blank
	Link to online resource for dataset	URL freetext	Optional	
	Dataset custodian (Organisation)	Fixed list	Required	but also 'unknown' option
	Dataset custodian (Person contact)	Free text	Optional	
	Dataset policy	Fixed list	Required	ICES, other open, restricted, not known etc.
	Dataset policy reference (URL)	Free text	Optional	
	Availability	Fixed list	Optional	Online, digital but not online, on paper etc.
About Data Product	Basic product description	Free text	Required	
	Link to existing metadata record for product	URL free text	Optional*	If this is used then potentially all fields can be left blank
	Link to online resource for product	URL free text	Optional	
	Documentation	Free text	Optional	i.e. survey manual, data product methodology, guidelines
	Availability	Fixed list	Optional	Online, digital but not online, on paper etc.
Usage	Expert group (provider)	Fixed list	Optional	EG list picker
	Expert group (consumer)	Fixed list	Optional	EG list picker
	Dataset caveats and limitations	Free text	Optional	Issues with data, correct usage etc, policy applicability
	Use in ICES publication	Fixed list	Optional	URL of primary ICES publication
Temporal information	Time series range (start value)	Fixed list	Optional	

	FIELD	TYPE	USAGE	
	Time series range (end value)	Fixed list	Optional	
	Latest update/publication of dataset	Date time	Optional	
	Update frequency	Fixed list	Optional	Continuous, Monthly, Yearly
	Sampling frequency	Fixed list	Optional	Continuous, Monthly, Quarter, Yearly
About Record	Who filled this out	Fixed list	Required	EG list picker
	Date edited	auto-complete	Required	

Annex 9: Data availability in ICES groups

To: Chairs ACOM, SCICOM, SSGIOEM (Eskild Kirkegaard, Yvonne Walther, Nils Olav Handegaard)

From: DIG

Subject: Data availability in ICES groups

Problem identification

As concluded by a number of groups (i.a. SSGIOEM, BSG) and recently highlighted at the Bureau meeting in February 2015, there is an issue in the ICES strategic plan implementation in the systematic understanding of what data sources are being used, by whom, what is the quality of these data, how access is provided to these data, and when, and where the gaps in provision of data and data products are. This undermines the advice process, and is likely a cause of inefficiencies and duplication of effort.

Problem resolution

It is therefore important that ICES takes steps to gain an understanding of these issues. This can be initially addressed by having an overview of the data availability and usage in ICES working groups, and of datasets/dataset derived products used by ICES working groups from sources outside ICES.

For this, DIG recommended in 2013 that ACOM and SCICOM “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” (recommendation 2013-163). This term of reference was never put into practice, so DIG proposes a different approach to collate information on the data availability in the ICES Expert Groups.

The final objectives are (1) to have an overview of the datasets/-products used and/or created by all ICES Expert Groups, and (2) gain insight in the data flows between the groups. This is beneficial for ICES as a whole (insight in process, identification of key datasets and dependencies) as well as for the individual Expert Groups (improved documentation of datasets and -products for existing science and advice outputs, identification of datasets from other expert groups that could be used). Datasets as well as data products should be taken into account.

Approach

Eight pre-selected ICES Expert Groups will be asked to fill in meta data of the datasets/-products they use and/or create and/or manage in an online catalogue, which will be publicly available and searchable. After the 8 groups have provided the information, other groups will be asked to add to the catalogue. The catalogue will also be pre-filled with information about existing ICES managed datasets and data products (stock assessment graphs, survey indices, ICES database regional datasets etc).

The 8 groups are WGSFD, WGOH, WGBEAM, WGBYC, WGDEC, WGZE, WGINOSE and a benchmark group (to be identified). The groups represent a range of expertise fields and in all groups a DIG member is involved to guide the process.

DIG have developed a template to fill the catalogue. The required information is minimal: ICES Ecoregion, dataset title, Dataset description, Dataset custodian, Dataset policy (may be ‘unknown’), the Product description and the Expert Group that filled out the form. Additionally, more information on Geographical context, Data usage, Temporal information and information about the dataset/-product like the methodology used can be provided.

Steps to be taken

Before the system can be operational, the form has to be put into an online system, the system should be prepared to host the new information, the terminology in the template has to be aligned with ISO standards where possible, and with other ICES catalogues.

A DIG subgroup will support ICES Data Centre to make the system available as soon as possible.