

Annex R

Ad hoc Working Group on IWC Global Data Repositories and National Reports (GDR)

Members: Double, Miller, B. (co-Convenors), Aisha, Allison, Baker, Bell, Brownell, Cipriano, De Freitas, Donovan, Enmynkau, Fortuna, Fruet, Hielscher, Hubbell, Hughes, Iñiguez, Kelkar, Leaper, Lovell, Lundquist, Mallette, Moronuki, Olson, Palka, Panigada, Redfern, Reeves, Ritter, Rosenbaum, Rowles, Slugina, Zerbin, Zharikov.

1. INTRODUCTORY ITEMS

1.1 Convenors opening remarks

Double introduced the terms of reference for the group, which were to conduct an assessment of the utility and support requirements of all IWC databases relevant to the work of the Scientific Committee. Specifically the group will:

- (1) collate summary information on all IWC databases relevant to the Scientific Committee (including data availability considerations);
- (2) summarise data use by the Scientific Committee for each database;
- (3) provide recommendations to improve integration, content and workflows;
- (4) review technical progress on existing databases;
- (5) consider needs and specifications for potential new databases, including developing simple technical guidelines on new proposals; and
- (6) produce a budget and work plan for the implementation and development of existing and new databases.

1.2 Election of Chair and appointment of rapporteurs

Double was elected Chair and Brendan Miller as co-Chair. Leaper acted as rapporteur.

1.3 Adoption of Agenda

The adopted agenda is given as Appendix 1.

2. PROGRESS WITH EXISTING IWC DATABASES

2.1 Review of existing IWC databases

The Secretariat currently holds or is developing 18 databases (Table 1). In addition the Secretariat holds three web applications which include databases (Table 2). The Working Group reviewed the list of databases in Table 1 and focussed on the technical and financial support required and the priority to complete that work (Table 3).

2.2 National Progress reports

There was considerable discussion about National Progress reports. The number of countries reporting has dropped from around 20 in 2000 to around 15 in recent years with only 12 in 2017. The drop in reporting coincided with the change from paper submissions to fully electronic. The current system does not identify a Scientific Committee document as a progress report (i.e. there is no SC/67a/ProgRep XXXX). The Working Group **agreed** that returning to having an individually identifiable SC report from each country would be very valuable. This would help coordinators in the country pass back information to contributors and would allow the Committee to easily see each country's report. It was agreed to develop a system to generate a PDF file from the data submitted for

each country individually. The PDF report would also include additional information such as the names of national and regional coordinators for each country (the coordinators would be viewed as the 'authors' of the report which it was felt might improve reporting rates). In addition the system could also aggregate data on specific issues. For example tables of bycatch and ship strikes that are currently considered each year by the Committee (within the HIM Working Group). It was noted that the Commission Bycatch Mitigation Initiative coordinator might also assist with promoting submission of information in National Progress reports.

It was **agreed** that large whale direct catch data no longer needed to be submitted through National Progress reports because these data are required by the Schedule and entered into the catch database. However it was also **agreed** that these data needed to be easily accessible to the Committee and will require tables of catches for the year to be available as an SC report and also appended to each National Progress SC report.

A number of other changes to National Progress reports were agreed. These are listed in Table 4. The intention is to reduce the workload of data entry while still retaining all the data that are used by the Committee.

3. POTENTIAL FUTURE IWC DATABASES (OBJECTIVES AND WORK PLAN INCL. COSTS)

3.1 Development of simple technical guidelines for new database proposals

The Working Group developed a pro-forma for new database requests and major alterations to existing databases. The pro-forma will be completed by the proponents and reviewed by the relevant sub-committee or working group, together with technical input from the Secretariat, similar to the procedures for funding proposals. The pro-forma is given in Appendix 2. This included the following technical guidelines:

- all new scientific database requests must be submitted to the Scientific Committee for discussion, approval, prioritisation and funding. They must be submitted using a database proforma **after** being discussed by the relevant sub-committee;
- all agreed development work will be overseen by the IWC Secretariat and the Committee where specific input is needed;
- where possible, all databases must use open source software;
- programming languages, database engines and other technologies used must be discussed and agreed with IWC Secretariat to minimise development, infrastructure and maintenance costs;
- all completed source code and database schemas must be provided to and held by the IWC Secretariat;
- all databases will have a 5-year review cycle to ensure code and databases are kept up-to-date and secure;
- where appropriate, new databases should be developed in a way that will allow expansion and interaction from other databases and applications. This is to be discussed with the IWC Secretariat;
- where user accounts are required, they should be authenticated by a central IWC authentication server to minimise login credentials for users of multiple databases.

Table 1
Databases/repositories held by the IWC Secretariat.

Database	Status	Data owner/manager	Database technologies	Application technologies	Display technologies	Storage location	Technical lead	Technical support required	Financial support required
Progress Reports (~6,780 records)	Live	Scientific Committee/ Greg Donovan	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	-	Server infrastructure and basic maintenance: £1,200/year
Ship Strikes (>1,000 reports, >190 cases)	Live	Ship Strike Data Review Group	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	Creation of bulk uploader	Server infrastructure and basic maintenance: £1,200/year; bulk uploader development costs: £1,200/year
Research Requests	Live	Greg Donovan	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	-	Server infrastructure and basic maintenance: £1,200/year
IWC photographic cruise database and archive (over 110,000 images)	Live	Greg Donovan	Lightroom	-	-	IWC storage	Greg Donovan	Jess Taylor (contract)	Funding available now but may need after 2018
IWC biopsy sampling database	Under development	Greg Donovan	At present, Access	-	-	IWC storage	Greg Donovan	Jess Taylor (contract)	Funding available now but may need after 2018
SH blue whale catalogue	Awaiting deployment	-	PostgreSQL	Python	HTML, JavaScript	IWC web servers (awaiting deployment)	-	Server setup	Server infrastructure and basic maintenance: £1,800/year
WNP gray whale catalogue	Under consideration	Greg Donovan	TBD	TBD	TBD	TBD	Greg Donovan	Will be needed	Will be funded by a voluntary contribution
Document Web Archive	Live	-	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	-	Server infrastructure and basic maintenance: £1,200/year
Bibliographic reference database (about 49,000 references)	Live	Greg Donovan	EndNote	-	-	IWC storage	Greg Donovan	-	Minimal
Individual Catch Database (~2.3 million records)	Live	Cherry Allison	Text/Excel	-	-	IWC storage	Cherry Allison	-	-
Catch Summary Database (~3 million catches)	Live	Cherry Allison	Excel	-	-	IWC storage	Cherry Allison	-	-
Discovery Marking Data	Live	Cherry Allison	Text/Excel	-	-	IWC storage	Cherry Allison	-	-
Sightings Data (IWC-DESS)	Live	Cherry Allison	Paradox	-	-	IWC storage	Cherry Allison	-	-
New integrated sightings, photo- identification	Under development, SC Steering Group	Greg Donovan	TBD	TBD	TBD	TBD	TBD	Yes	Yes but money for development available in Cruise budget
Small Cetaceans Catches (Bycatch and Direct)	Retired	Cherry Allison	MS Word	-	-	IWC storage	Cherry Allison	-	-
Compendium of Whale Watching Regulations	Outdated	-	Text	-	-	-	-	New data manager	-
IWC Database of Recommenda- tions	In planning	Sarah Smith/ Sarah Ferriss	-	-	-	IWC web servers	Brendan Miller	-	-
Entanglement Response	In planning	David Mattila	-	-	-	IWC web servers	Brendan Miller	-	-

Table 2
Web applications which include databases held by IWC Secretariat.

Database	Status	Data owner/manager	Database technologies	Application technologies	Display technologies	Storage location	Technical lead	Technical support required	Financial support required
Cetacean Diseases of Concern Intranet	In development	Claire Simeone	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	Content completion	Server infrastructure and basic maintenance: £1,200/year
WW Handbook	In development	Sarah Ferriss	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	Content completion, database development, application development	Server infrastructure and basic maintenance: £1,200/year
JCRM Submission Site	Live	Jessica Peers	MySQL	PHP	HTML, JavaScript	IWC web servers	Brendan Miller	-	Server infrastructure and basic maintenance: £1,200/year

Table 3
Database usage by the Scientific Committee, any database development and data entry required, and the priority to complete that work.

Database	Status	Use by Scientific Committee	Work required	Priority to complete
Progress Reports	Live	Time series data on bycatch and entanglements and other anthropogenic impacts on large and small cetaceans	Complete changes agreed at SC/67a (see Table 4)	High ; Approx. 1.5 weeks developer time
Ship Strikes	Live	Time series data of ship strikes on large whales	Migration script or 'brute force' data entry of 100+ records from other repository	In progress by Secretariat
Research Requests	Live	Portal to request data or samples held by the IWC	None required	N/A
IWC photographic cruise database and archive	Live	Keyworded data archive linked to cruise records, e.g. photo-ID, biopsy, scarring, health status etc.	Updates only	In progress by Secretariat
IWC biopsy sampling database	Under development	Facilitate stock structure analyses	Server setup and deployment – SC budget allocated	High
SH blue whale catalogue	Awaiting deployment	MR abundance estimation for population assessments, population structure		
WNP gray whale catalogue	Under consideration	MR abundance estimation for population assessments, population structure	Possible migration to new system	Process to be specified
Document Web Archive	Live	For everything	Updates only	In progress by Secretariat
Bibliographic reference database	Live	SC reports, communication	Updates only	In progress by Secretariat
Individual Catch Database	Live	Population assessments, catch limits, distribution and movement	Online portal. Idiosyncrasies within data. Creation of a database accounting for differences between total catches and those with individual data. Requires full documentation.	High
Catch Summary Database	Live	Population assessments, catch limits, distribution, movement	Online portal.	High
Discovery Marking Data	Live	Populations structure and movement, catch allocation	None required	N/A
Sightings Data (IWC-DESS)	Retired (see below)	Population assessments, abundance estimation	Data will be integrated into the new sightings, photo-ID, database	N/A
New integrated sightings, photo-ID, biopsy	Under development, SC Steering Group	Population assessments, abundance estimation	Updates only once developed	In progress by Secretariat, funding available
Small Cetaceans Catches (bycatch and direct)	Retired	Not used currently	None	N/A
Compendium of Whale Watching Regulations	Outdated	Global comparison of whale-watching regulations, assessment of best practice	None – not a database	N/A
IWC Database of Recommendations	In planning	Communication with Commission, assessment of progress and response	Develop database	Proposal will be presented at SC/67b
Entanglement Response	In planning	Develop best practice, information sharing and capacity development	Develop database; funding already available	High – final proposal will be presented at SC/67b
Cetacean Diseases of Concern	Intranet in development		Finalise website	Part of document archive; complete at SC/67b
WW Handbook	In development		Finalise website, develop database	Part of document archive; in progress by Secretariat
JCRM Submission Site	Live	JCRM Journal management system (submission to publication)	Customise some features	Medium ; current system functional

Table 4

Recommended changes to the current National Progress Reports schema.

All sections

Create facility to create pdf reports for each nation, and for each section

Submit pdf reports by nation as papers to Scientific Committee including national and regional coordinators as authors

Add 'Choose not to enter data/No data' option for all sections for national coordinator

Provide examples of data entry and the appropriate ways to aggregate data as alternative to submitting individual records

Add hyperlink to web page explaining 'RMP Small Area'

Option to use drag and drop pointer to indicate 'local area' - provide guidance on how to complete

Cetacean databases and archives

New section. Facility to pre-populate from past records from 2019 on. Proposed fields:

- Database/archive name
- Description of database/archive: content, temporal & geographic coverage, data access – provide guidance on how to complete
- Data/archive type: sightings, photo-id, tissue samples, genetic, telemetry, strandings, other
- Number of records - text field to allow descriptors
- Database/archive manager/curator
- Contact details

Systematic surveys

New section. Proposed fields:

- Survey name
- Description of survey: purpose geographic area, methods used, survey dates. Provide guidance on how to complete. Include associated databases in 'Cetacean databases' section
- Contact details

Sightings

Remove section from National Progress Reports - see new 'Systematic surveys' section above

Natural marking

Remove section from National Progress Reports - see new 'Cetacean database' section above

Telemetry and artificial marking

Remove section from National Progress Reports - see new 'Database' section above

Tissue and biological samples

Remove section from National Progress Reports - see new 'Database' section above

Direct catches of large whales

Remove section from National Progress Reports – these data are collected through a separate process – a summary of catches will be submitted at each Scientific Committee meeting and a summary table appended to each National Progress Report

Non-direct anthropogenic mortality of large whales

Change title to 'Vessel strike of large whales'

Include prompt to ship strike database once basic data entered

Change field name to 'Submitted to IWC or national Ship Strike Database?'

Replace sex specific fields to No. individuals seriously injured, injured, unknown

Fishery bycatch of large whales

Include guidance to include entanglement information in this section

Direct catches of small cetaceans

No changes

Non-direct anthropogenic mortality of small cetaceans

Change title to 'Vessel strike of small cetaceans'

Replace sex specific fields to 'No. individuals seriously injured, injured, unknown'

Fishery bycatch of small cetaceans

No changes

Strandings (and floating carcasses)

Remove field: Total stranding events

Add comment field to describe the stranding event(s) – add guidance on how to complete

Amend field title to 'Total number of individuals stranded'

There are a number of databases which receive funding from IWC but are not hosted by the Secretariat (e.g. Pollution 2020). The new pro-forma is intended to adequately describe the form and function of these external databases and specify data availability arrangements with the Scientific Committee. This information will assist the assessment of any associated funding proposals.

3.2 Global database for disentanglement activities

The Committee has recommended the development of a database for the IWC's Global Whale Entanglement Response Network (GWERN). This was discussed as an example of a well-advanced proposal for a new database that could be used as a test of the new pro-forma. Mattila agreed to fill out the pro-forma using the specification for the GWERN database.

It was noted that there had been consideration of a global bycatch database. It had been agreed in 2016 that this was a very ambitious project that could be developed in a modular fashion starting with components (such as the GWERN database) for which there was already a clearly defined need

and objectives. The GWERN database would be constructed in a modular fashion beginning with the data currently recommended for collection in GWERN's consensus field data form. If, as anticipated this initial database is successful, then the database could be expanded in the future to include other modules and sources of data. Hence the initial structure needs to be carefully designed to allow for future expansion.

3.3 Development of a database of IWC recommendations

Another database proposal is in development by the Scientific and Conservation Committees jointly. This will be for a web-accessible database of Scientific Committee recommendations. Rendell noted that the proposal for this was still in draft form but he would complete the pro-forma for the 2018 meeting. This would provide another opportunity to review the pro-forma and refine as needed.

4. ADOPTION OF REPORT

The report was adopted on 17 May at 15:00.

Appendix 1

AGENDA

1. Introductory items
 - 1.1 Opening remarks
 - 1.2 Election of the Chair
 - 1.3 Appointment of Rapporteurs
 - 1.4 Adoption of agenda
 - 1.5 Documents available
2. Terms of Reference: This *ad hoc* working group will conduct an assessment of the utility and support requirements of all IWC databases relevant to the work of the Scientific Committee. Specifically it will:
 - collate summary information on all IWC databases relevant to the Scientific Committee (including data availability considerations);
 - summarise data use by the Scientific Committee for each database;
 - provide recommendations to improve integration, content and workflows;
 - review technical progress on existing databases;
 - consider needs and specifications for potential new databases, including developing simple technical guidelines on new proposals; and
 - produce a budget and workplan for the implementation and development of existing and new databases.
3. Progress with existing IWC databases
 - 3.1 Review of existing IWC databases including:
 - 3.1.1 Content and engagement
 - 3.1.2 Use by the Scientific Committee
 - 3.1.3 Management and workflows
 - 3.1.4 Potential for integration with other databases and workflows
 - 3.2 Recommendation for further development or integration of existing databases
 - 3.3 Indicative future support and funding requirements for existing IWC databases
4. Potential future IWC databases (objectives and work plan including costs)
 - 4.1 Development of simple technical guidelines for new database proposals
 - 4.2 Global database for disentanglement activities
 - 4.3 Global bycatch database

Appendix 2

DATABASE REQUEST PRO FROMA

This pro forma is to be used for new database requests, and major alterations to existing databases

Date	Request type	Related sub-committee
New database/alteration (<i>delete as appropriate</i>)		

1. Database title

Please provide the title of the database.

2. Brief overview of the database or alteration

Give a very brief overview on your proposal and its expected usage within the IWC community. Be succinct and clear as this may be used to summarise your request in a report.

3. Identified scope and usage within the IWC and its committees

Please explain what data the database will hold, and how the database will be used within the IWC and its committees; to what questions of importance to the IWC will this database contribute?

4. Proposed database schema or architecture

Please provide an overview of the proposed database schema or architecture. Where possible, please consider providing an entity relationship diagram.

5. Interaction with other databases

Will the database be required to interact with any other databases or applications?

6. How will the data be populated?

Please explain how the data will be populated. Please consider the following: will the data be entered by the public or a select group? Will the data be verified after entry? Will it be web or mobile accessible?

7. Other similar databases

Please list any other databases that capture similar data either within the IWC or externally.

8. Timetable for key planning activities

Activity to be undertaken	Key person(s)	Start (mm/yy)	Finish (mm/yy)

9. Proposed completion dates

Expected outputs	Completion data (mm/yy)

10. Associated people

Name	Affiliation	Role within database

11. Data ownership and sharing

Please state your expected data availability arrangements including data ownership and data sharing agreements.

12. Total costs

Please provide a breakdown of costs. These may require discussion within the IWC Secretariat.

Type	Description	Cost (GBP)
Planning costs (e.g. travel/subsistence)		
Development costs (e.g. salaries, contractors, software)		
Ongoing costs (e.g. servers, back-ups, maintenance)		
Equipment costs		
Expected ongoing data co-ordinator costs		
Expected ongoing and one-off data entry costs		
Other costs		
Total		

13. Funds and contributions

Please provide details of any existing funds or potential contributions

Type	Description	Cost (GBP)
Existing funds		
Potential contributions		
Total		

14. Have you read and agree to the IWC Database Guidelines? (please tick)

The guidelines can be found at the end of this pro forma. To ensure you have the latest version please contact the IWC Secretariat.

Yes	
No	

IWC Database Guidelines Version 2017.05

To standardise the creation of databases, repositories, catalogues and applications, the following guidelines should be adhered to.

- All new scientific database requests must be submitted to the scientific committee for discussion, approval, prioritisation and funding. This must be submitted using a database proforma **after** being discussed by the relevant sub-committee.
- All agreed development work will be overseen by the IWC Secretariat, and the Committee where specific input is needed.
- Where possible, all databases must use open source software.
- Programming languages, database engines and other technologies used must be discussed and agreed with IWC Secretariat to minimise development, infrastructure and maintenance costs.
- All completed source code and database schemas must be provided to and held by the IWC Secretariat.
- All databases will have a 5-year review cycle to ensure code and databases are kept up-to-date and secure.
- Where appropriate, new databases should be developed in a way that will allow expansion and interaction from other databases and applications. This is to be discussed with the IWC Secretariat.
- Where user accounts are required, they should be authenticated by a central IWC authentication server to minimise login credentials for users of multiple databases.