

SC/68D/GDR/01

Sub-committees/working group name: GDR

Databases and data holdings in the IWC: A review of the current status, issues and future plans

IWC Secretariat



**INTERNATIONAL
WHALING COMMISSION**

Papers submitted to the IWC are produced to advance discussions within that meeting; they may be preliminary or exploratory.

It is important that if you wish to cite this paper outside the context of an IWC meeting, you notify the author at least six weeks before it is cited to ensure that it has not been superseded or found to contain errors.

Databases and data holdings in the IWC: A review of the current status, issues and future plans

Dr. Iain J. Staniland, Mark Tandy, Imogen Webster, Moussa Ki Sow, Isidora Katara and Lydia O'Loughlin

IWC Secretariat, The Red House, 135 Station Road, Impington, Cambridge, CB24 9NP, UK

Introduction

Discussions at SC68B raised the issue of how data and databases were managed across the IWC. Several sub-committees discussed their desire to create new databases and requests were made for the Secretariat to take on the management of external databases. Whilst the SC research fund has been used to help fund the establishment of databases little or no resources have been set aside for ongoing maintenance and support.

The Global Databases and Repositories Group were unable to meet during SC68A or SC68B, but did meet in 2021 at SC68C. The need for a holistic approach to data collection and management was highlighted in particular around streamlining different submission processes and developing guidelines for data availability.

The Committee agreed the financial support of SC databases requires a robust, long-term approach and encouraged GDR to work with the Secretariat and SC Chair to propose a more appropriate funding model at SC68D. It also supported the submission of a funding proposal from the Secretariat to facilitate this work in the interim.

This paper aims to summarise the current status, highlight issues and stimulate discussions on establishing a more holistic approach to the capture, management and dissemination of data within the IWC.

Review of existing IWC databases

A summary of the IWC's existing databases and web applications can be seen in Table 1 and a brief overview of the major work areas are detailed below:

Southern Hemisphere Blue Whale Catalogue

This database and associated applications were migrated to the Secretariat managed servers in their current form (Python application). An external developer is currently upgrading and improving the application code to provide more options for photo recognition. The Secretariat is supporting the development by providing access to both the live and development servers.

Ship Strikes database

The ship strikes database was previously managed by a contract funded through the SC research fund. However at SC68B (HIM) it was felt that this had now reached a stage where a dedicated full time co-ordinator was required. Using funds from a variety of sources including the SC and CC a new data manager has been employed, initially on a one year contract starting from the 19th April. During SC68C funding was put forward to fund the person for 2022-2023.

The person has overseen the Ship Strikes database and begun advancing data use through developing metadata, data use documentation and exploring dashboards to visualize the data.

Database of Recommendations

The Secretariat has, after extensive work, now made the DoR available online [here](#). Historical records for Commission, SC and CC meetings are summarised in Table 2. This is still a work in progress but Resolutions and Recommendations from workshops will now be prioritised for back

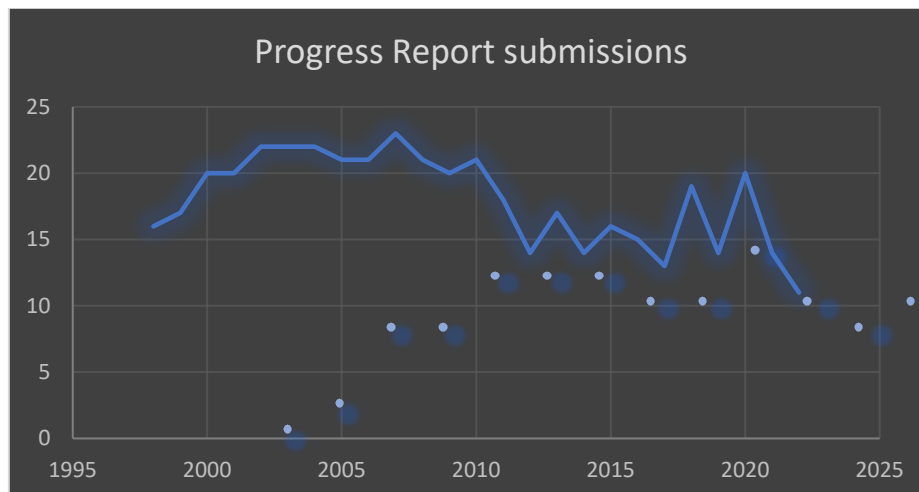
entry. Bulk upload via an Excel spreadsheet template has also been designed to assist data entry. It is hoped that, with use, the management of the DoR and the usefulness of the recommendations held within it will improve.

Conservation Committee

At its meeting in 2018, the CC endorsed a proposal from New Zealand to replace the current system of voluntary reporting with an integrated database. The Secretariat requested further input from the CC during its 2020 virtual meeting including on the purpose of reporting and database fields. Based on these discussions the Secretariat has developed a prototype database in PHP which will continue to be refined ready for IWC68. An intersessional group was set up by the CC to assist in these efforts.

National Progress Reports

There has been a great deal of effort to improve the ease of data submission in recent years. Secretariat staff directly assisted countries in their submissions this year, have made additional outreach to African nations and provided training, resulting in the submission of 11 National Progress Reports, a decrease from the 14 received in 2021. Data was received from Mauritania for the first time as a result of the above efforts.¹



This reduction in submissions is alarming and highlights a need to review the whole National Progress reporting process to make the database a more effective tool for member Countries and the SC.

Australia has developed scripts in R (REF) that facilitate the automated visualization of annual NPR data to aid their own internal reporting. They are now developing these scripts further to summarise and identify trends using all their submitted NPR data. The applicability of these scripts to NPR data from other countries will be assessed. When complete, these scripts will be able to automatically generate multi-year tables and plots of national data and summaries of all the data within the NPR database.

Global entanglement database

Most countries/regions already have a national or regional database and it is felt that there was currently no need for the IWC to develop its own global database and the decision to do so has been

¹ Submissions for 2022 are Australia, Brazil, Croatia, Colombia, Germany, Italy, Mauritania, Mexico, Netherlands, New Zealand & USA

deferred. However, in lieu of an IWC database, all teams, networks and Countries are still encouraged to use the consensus data fields developed by the IWC entanglement expert panel, over several GWERN workshops.

Strandings database

The scientific committee asked the Secretariat to liaise with the strandings initiative in regard to the existing database development plans and agreed to consider this at SC68D. Scoping work has begun by the Ship Strikes and Strandings Data Manager. The Secretariat has met with ASCOBANS, ACCOBAMS, HELCOM, and the CARI'MAM Strandings Network to assess the current status of data and evaluate the potential for a combined database. Due to the wide variation on data format, data needs, and data sharing restrictions, additional scoping is planned before a proposal will be made.

Catch Database

The catch database was updated this year including historical data from post war whaling in Japan. Previous discussions have been held in regard to making this publicly downloadable through a webform but to date data are only available by request from the Secretariat. The database requires modification in order to make it public and the Statistics team have identified a plan to achieve this.

Other Database Discussions

Infectious disease & pollution

The focus session on infectious diseases within the sub-committee on Environmental Concerns concluded that to fill data gaps on pathogen prevalence and disease in under-reported areas, there is a need to enhance sampling and diagnostic capabilities in those regions. They suggested a broad surveillance programme and database should be developed. The sub-committee also places value in a contaminant mapping tool which is linked via the IWC website but hosted at St-Andrews University a plan for the future of this tool is needed. The recent Pollution 2025 workshop has recommended the E group seek SC research funds to support this tool.

Marine Debris

Following an IWC workshop on Marine Debris the SC recommended intersessional efforts should look at data collection from post mortem examinations related to Marine Debris and how these data might be managed in the context of other IWC databases i.e., Strandings.

IWC POWER/SOWER

There are currently >200,000 images from POWER cruises between 2010-2022 and >60,000 from SOWER cruises between 1988 and 2010 held in an Adobe Lightroom database. Pertinent information from natural marking / sighting / biopsy datasheets are coded into the image metadata, but not all information from the datasheets is included and not in every image has these data. However sightings data and biopsy data from these programmes are held in an outdated paradox database, paper records or on Excel spreadsheets. Money for any potential work could be made available from the cruise budget.

Abundances database

The abundances dataset holds abundance values and related uncertainty, endorsed by ASI. The values are in some cases related to time series of abundance estimates. ASI has recommended that these time series will also be available in the web version of the abundances dataset. The datasets need to be reorganised into a database.

Future progress

The Ad Hoc Working Group on Databases and Related Issues has been requested to review the status of existing IWC databases and to liaise closely with those sub-committee members who are developing ideas for databases. These actions are critical in order to consider new database specifications in light of ongoing work in the Secretariat and the potential for interoperability with existing databases. In order to facilitate this work the following measures could be considered

- It would be useful to define what constitutes an IWC database and what rules and procedures are associated with this. For example how do we deal with external databases funded by the IWC? Do data availability expectations change when an external database is moved to become IWC database?
 - The IWC should develop a clear policy in regards to data access for its databases – particularly for internal and external databases that seek SC and CC funds.
- A workshop/meeting of this committee to scope all current database and data holdings within the IWC scoping areas of potential integration and consolidation including data submission. The Terms of reference for this would include identifying need, intended users, overlap with external data holdings, the benefit to the work of the IWC now and over the next 10 years.
 - For example review the opportunity to fully integrate the CC and NPR reporting processes.
- Review the proforma and guidance regarding the development of new databases

Annex i.

It is felt that the recent restructuring within the Secretariat coupled with the increasing needs of the users provides an opportunity to reassess the data management arrangements and practices within the IWC.

Table 1: Summary of IWC databases and data holdings

Database/Application	Status	Work Required	SC Priority	Progress
Bibliographic reference database (EndNote)	Live	Import commission papers and circulars?	Medium	None
Blue Whale Song Library	Stalled	Develop database (£4000 available) – Assigned to tribalsystems.uk (not done). Recommendation added to the CRM https://crm.iwc.int/data/blue-whale-songs by IWC	High	None
Catch Summary Database	Live	Recommendation to be added to crm.iwc.int - Example crm.iwc.int/individualcatchdatabase	High	None
Cetacean Diseases of Concern Intranet	Stalled	Finalise website	Medium	None
Compendium of Whale Watching Regulations	Retired			
Conservation Reporting Database	Under Development	Finalise specifications, Develop database and Review at CC 2022. Recommendation: Add to IWC CRM https://crm.iwc.int/data/conservation-database		Started
Database of Recommendations	Live	Improved - https://recommendations.iwc.int		Done
Discovery Marking Data	Live			
Entanglement Response	Postponed	Develop database		
Individual Catch Database	Live	Document idiosyncrasies exist within data. Creation of a database of records for which there are no individual data and which conforms to all available summary data on the area, sex and month of these catches, to enable easy creation of catch series. Requires full documentation. Recommendation to be added to crm.iwc.int - Example crm.iwc.int/individualcatchdatabase and GitHub for versioning and documentation	High	None
InforMEA project database	Live	Ongoing development of Drupal application and database		Ongoing
IWC biopsy sampling database	Under Development	Updates only (in progress)	Medium	

IWC photographic cruise database and archive	Live	Updates only (in progress)	Medium	
Journal	Live	JCRM Issues 20-23 and Special Issues now available online. Backdating previous Issues required	Medium	Done
Meeting Documents	Live	Moved to SharePoint		Done
Meeting Registration	Live	For SC68D Microsoft tools were used for registration and automatic addition of users to our office 365 SC groups		Done
National Progress Reports	Live	Bulk import from CSV to be tested. Integration of standard FAO areas	Medium	Ongoing
New integrated sightings, photo-ID, database	Stalled	Updates only (funding available)	Medium	
Research Requests	Replaced	Replaced by simple form https://forms.office.com/r/wEmk6Py8U5		Done
SH Blue Whale Catalogue	Live	Migrated to IWC server		
Ship Strikes	Live	Bulk import tool (CSV) - Not done Allow data to be submitted via API directly from country databases	Medium	None
Small Cetaceans Catches (Bycatch & Direct)	Retired			
WNP gray whale catalogue	Under Consideration	Migration to new system	tbc	
WW Handbook	Live	Update to CraftCMS v3 and migrate database to PostgreSQL	High	None

Table 2: Status of the Database of Recommendations

Meeting	Committee	Year
Climate change workshop	CC/SC (ready for upload)	2021
Pollution workshop	SC (ready for upload)	2021
intersessional CC meeting	CC	2020
	SC	2021
	SC	2020
	SC	2019
	SC	2018
IWC67	WKM&WI	2018
	ASW	2018
	Infractions	2018
	CC	2018
	F&A	2018
	Resolutions	2018
	SC	2017
IWC66	WKM&WI	2016
	CC	2016
	Resolutions	2016
	SC (partial)	2016
IWC65	Resolutions	2014
IWC64	Resolutions	2012
IWC63	Resolutions	2011
IWC61	Resolutions	2009
IWC59	Resolutions	2007
IWC58	Resolutions	2006
IWC57	Resolutions	2005
IWC56	Resolutions	2004
IWC55	Resolutions	2003
IWC54	Resolutions	2002
IWC53	Resolutions	2001
IWC52	Resolutions	2000
1990's Resolutions	Ready for upload	1990's
Marine Debris workshop	SC	2019
Indian Ocean BMI Workshop		2019
IMMAs and Ship Strikes	SC	2019
4th workshop on large whale entanglement issues	WKM	2018
SWG on E Acoustic Masking and Whale population dynamics	SC	2016
Non-Hunting Related Aspects of Cetacean Welfare	WKM	2016
Workshop to Develop Practical Guidance for the Handling of Cetacean Stranding Events	WKM	2016
Mitigation and Management of the threats Posed by Marine Debris to Cetaceans	CC	2014
Workshop on Small Cetaceans and Climate Change	SC	2010

Annex i

Database Request Proforma

This proforma 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?

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		
--	--	--

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.

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.