SIOFA | APSOI
Southern Indian Ocean Fisheries Agreement
Accord relatif aux Pêches dans le Sud de l'Océan Indien

10th Annual Meeting of the Scientific Committee (SC10)

Concarneau, France, 17-26 March 2025

SC-10-INFO-11

Data Curation, Fisheries and Ecosystem-based Management: The Case Study of the Pecheker Database

Delegation of France Overseas Territories

Document type	Working paper ✔ Information paper ☐
Distribution	Public ✔ Restricted [1] ☐ Closed session document [2] ☐

Abstract

The scientific monitoring of the Southern Ocean French fishing industry is based on the use the Pecheker database. Pecheker is dedicated to the digital curation of the data collected on field by scientific observers and which analysis allows the scientists of the Muséum national d'Histoire naturelle institution to provide guidelines and advice for the regulation of the fishing activity, the protection of the fish stocks and the protection of the marine ecosystems. The template of Pecheker has been developed to make the database adapted to the ecosystem-based management concept. Considering the global context of biodiversity erosion, this modern approach of management aims to take account of the environmental background of the fisheries to ensure their sustainable development. Completeness and high quality of the raw data is a key element for an ecosystem-based management database such as Pecheker. Here, we present the development of this database as a case study of fisheries data curation to be shared with the readers. Full code to deploy a database based on the Pecheker template is provided in supplementary materials. Considering the success factors we could identify, we propose a discussion about how the community could build a global fisheries information system based on a network of small databases including interoperability standards.
https://www.ijdc.net/index.php/ijdc/article/view/674/630

[1] Restricted documents may contain confidential information. Please do not distribute restricted documents in any form without the explicit permission of the SIOFA Secretariat and the data owner(s)/provider(s).
[2] Documents available only to members invited to closed sessions.

IJDC | *General Article*

Data Curation, Fisheries and Ecosystem-based Management: The Case Study of the Pecheker Database

Alexis Martin
Muséum National d'Histoire Naturelle

Charlotte Chazeau
Muséum National d'Histoire Naturelle

Nicolas Gasco
Muséum National d'Histoire Naturelle

Guy Duhamel
Muséum National d'Histoire Naturelle

Patrice Pruvost
Muséum National d'Histoire Naturelle

Abstract

The scientific monitoring of the Southern Ocean French fishing industry is based on the use the Pecheker database. Pecheker is dedicated to the digital curation of the data collected on field by scientific observers and which analysis allows the scientists of the Muséum national d'Histoire naturelle institution to provide guidelines and advice for the regulation of the fishing activity, the protection of the fish stocks and the protection of the marine ecosystems. The template of Pecheker has been developed to make the database adapted to the ecosystem-based management concept. Considering the global context of biodiversity erosion, this modern approach of management aims to take account of the environmental background of the fisheries to ensure their sustainable development. Completeness and high quality of the raw data is a key element for an ecosystem-based management database such as Pecheker. Here, we present the development of this database as a case study of fisheries data curation to be shared with the readers. Full code to deploy a database based on the Pecheker template is provided in supplementary materials. Considering the success factors we could identify, we propose a discussion about how the community could build a global fisheries information system based on a network of small databases including interoperability standards.

Received 22 August 2019 ~ *Revision received* 12 March 2021 ~ *Accepted* 7 April 2021

Correspondence should be addressed to Alexis Martin, UMR BOREA/Muséum national d'Histoire naturelle, 43 rue Cuvier, 75005 Paris, France. Email: alexis.martin@mnhn.fr

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Introduction

The French fishing industry reaches the European Union's third rank in terms of volume and value (FranceAgriMer, 2019; Plazaola and Rignols, 2019). An important part of this value, almost 10%, is produced by fisheries located in the Southern Ocean and the south of the Indian Ocean (Lefebvre Saint-Felix and Maghin, 2019). The scientific monitoring of this activity, started in 1978, aims to ensure the sustainable exploitation of the fish stocks and to limit the impact on the ecosystems by the provision of guidelines to the public managing structures involved in the regulation of the fishing industry. To produce these guidelines, analysis are based on the data collected on board commercial vessels or during scientific cruises. Data collection is ensured by fishery observers and scientists according to survey protocols provided by the Muséum national d'Histoire naturelle[1] (MNHN). Datasets are uploaded and maintained into the "Pecheker" database. Pecheker includes various sets of raw data about the fishing activity and the biodiversity of the impacted ecosystems. Pecheker is developed and maintained by a team of scientists of the MNHN involved in biocomputing, with the support of IT engineers.

Most of the French and world's fishery monitoring programs are based on three main survey paradigms:

- Collection of declarative data of fish catches and fleets fishing effort, provided by the fishers, the fishing companies or the fish markets (FAO, 1999). Fish catches are usually recorded as weights of caught species or weights of fish products (FAO, 1999). Fishing effort reflects the vessels activity level and can be recorded by many variables such as time of fishing, number of fishing gears deployed, number of days of fishing or fishing spatial and geographical coverage (FAO, 1999). Such declarative data aims to be integrated into centralised national or regional databases for a global monitoring of the exploitation level of the various species (FAO, 2020; EUMOFA, 2020). This paradigm is supposed to be based on a full coverage of the activity and to be low cost for the monitoring programs. But its main limitation is the quality of the raw data, which can be impacted by inaccuracy (Sampson, 2011), non-declaration (Belhabib et al., 2014; Pramod et al. 2014), mis-declaration or mis-labelling (Helyar et al. 2014; Bénard-Capelle et al., 2015) of the catches. This can result in underestimation of the exploitation level of species and a distorted overview of catches (Pauly and Zeller, 2016), specially when low quality raw data must be aggregated into groups including various species (Freire et al., 2020). Because of their rarity, this may affect sensitive species in particular or cryptic species difficult to identify without taxonomic expertise (Iglésias et al., 2010; Filonzi et al., 2010).

- On-board partial sampling of the fish catches by scientific observers (Davies and Reynolds, 2002; Taconet et al., 2002; NMFS, 2020). Such data aims to be extrapolated to obtain a general evaluation of the species exploitation level (Gauduchon et al., 2020; Patterson et al., 2019), by crossing the partial sampling of catches with the full recording of the fishing effort of the fleets, by highlighting trends (Baum and Blanchard, 2009) or by predictive modelling (Mannocci et al., 2020). The data from this type of monitoring are usually

1 Muséum national d'Histoire naturelle: https://www.mnhn.fr

uploaded into databases each dedicated to a well identified fishery and/or integrated in national databases (AFSC and ARO, 2019). This paradigm is based on more accurate primary data and allows the collection of data not only on target-species but also on the by-catch species, which is an improvement regarding ecological impact assessment (MPI, 2019). But its main limitations are the complexity of the extrapolation process, the sensitivity to sub-sampling, and bias or uncertainties due the low statistical power when sampling data for small fisheries or rare species (Cheal and Emslie, 2020; Silburn et al., 2020; Suuronen and Gilman, 2020; Fernandes et al., 2021).

- Sampling of the sales in the fish markets by scientific observers (Holden and Raitt, 1974; Gulland and Rosenberg, 1992). Such approach allows the collection of length frequency data and corrections on the declarative data, which can be crossed with the sampling of the sales to provide corrected statistics on the level of exploitation of the target-species[2]. The main limitations of this paradigm are the lack of data for the non-commercial by-catch species and the lack of the accurate geographical information for the catches of target-species. This strongly restricts the possibility to assess the impact level of the fisheries on the ecosystems (Mackinson et al., 2018).

The French southern fisheries monitoring program led by the MNHN corresponds to a fourth paradigm, the ecosystem-based management approach. In response to growing expectations from society regarding the impact of fishing on marine ecosystems, this paradigm aims (1) to surpass the limits of the classical monitoring approaches by improving the data quality and (2) to consider the environmental context of the fisheries, to ensure their sustainable development by taking into account economical, social and ecological issues (Staples, 2009). On field, this implies the collection of full data of catches for target species and by-catch species by on-board scientific observers, with full time series of data, full recording of the fishing effort and spatial information with continuous geographical coverage (Kupschus et al., 2016). Furthermore, supplementary biodiversity observations are recorded according to protocols designed to perform ecological impact assessments (Clua et al., 2005). The completeness of the data, in time, space and taxa, constitutes a major asset to improve the stock assessment and the statistical modelling of the fish population dynamics (Duhamel et al., 2019; Hill et al., 2017). In this way, advice given to the managers for sustainable exploitation of stocks and contributions to the international fish stock assessment working groups are enhanced (Okuda and Massiot-Granier, 2019). Moreover, by crossing the living organisms observations and the fishing effort data, the monitoring program allows researchers to compute models and statistics based on abundance distributions (Chazeau et al., 2019). This is of particular interest to study the evolution of the species spatial distribution patterns and the evolution of their ecological niche over a long period (Guinet et al., 2015; Welsford et al., 2011).

The need for an integrated information system to store and manage all the various data of the program in a common format led to the development of the Pecheker database in 2007. The main objective consisted of the creation of a suitable system for the curation of the French southern fisheries datasets considering their specificity, completeness and complexity. Available tools for fisheries data curation, based on the three main survey paradigms presented above, could not be used for the Pecheker project due to the specific constraints of the Southern Ocean ecosystem-based

2 French OBSVENTES program: https://sih.ifremer.fr/Ressources/ObsVentes

monitoring (Leblond et al., 2008; Schmidt et al., 2020; Taconet et al., 2017). Available international data templates also appeared to be not suitable. Templates such as the "Fishframe" format (Degel and Jansen, 2006), templates provided by the The Food and Agriculture Organization (FAO) of the United Nations (Taconet et al., 2002) or by regional international regulation authorities[3] [4], allow the formatting and exchange of core data between fisheries. But their design corresponds to a compromise of what is considered to be part of the core data to be transmitted to the regional structures involved in international regulation. Such templates do not include the possibility to store and manage full datasets from scientific protocols exceeding the core data resulting from these international compromises.

Biodiversity databases is a related huge domain in the world of data curation. Many tools, templates, and integrated systems are available for the scientific community, with a very dynamic development and continuous improvements. Covering all the components of the biodiversity, including marine biology, this domain provides templates such as the "Darwin Core" format (Wieczorek et al., 2012), integrated databases such as FishBase[5] (Froese and Pauly, 2000), SeaLifeBase[6], the Global Biodiversity Information Facility[7] (GBIF) or the Ocean Biodiversity Information System[8] (OBIS) and dedicated databases such as the Barcode Of Life Data system (Ratnasingham and Hebert, 2006) for molecular data or the World Register of Marine Species[9] for taxonomic data. However, even though they are very powerful, none of these tools appeared to be completely suitable for fisheries data curation. Biodiversity databases and related tools are a product from the naturalist culture and are usually focused on the taxa presence information (Costello et al., 2018). Taxa presence information is there considered as the core data around which the information systems are built. Other information, such as observation protocols, are usually considered as supplementary data or meta-data for the core data. This results in a lower level of structuration and precision of the templates for the recording of the observation contexts and protocols which, in the fisheries world, are related to the calculation of the fishing effort, the main part of the core data.

The choice of building a specific template for Pecheker soon appeared to be the best development strategy. This started in 2007 with the launch of the atlas of the French national fisheries Information System (Pruvost et al., 2011). Funding from the "Direction des Pêches Maritimes et de l'Aquaculture"[10] (DPMA), a state service from the French ministry involved in alimentation, agriculture and fishing, was provided to the MNHN to format, aggregate and transmit French southern fisheries data to the national atlas of fisheries. Part of this funding, limited in time, was used to develop the Pecheker template and integrate the historical data. Maintenance and development of the database was later funded by a permanent grant from the DPMA to the MNHN. A first functional version of Pecheker was delivered in 2008 and a major update was released in 2013.

3 CCAMLR data forms: https://www.ccamlr.org/en/data/forms
4 SIOFA summary and assessment of data submitted: http://www.apsoi.org/sites/default/files/documents/meetings/MoP-04-13%20Vessel%20Catch%2C%20Effort%20and%20Scientific%20Data%20-%20A%20summary%20of%20reports%20to%20the%20Secretariat.pdf
5 Fishbase: https://www.fishbase.se
6 SeaLifeBase: https://www.sealifebase.ca
7 Global Biodiversity Information Facility home page: https://www.gbif.org
8 Ocean Biodiversity Information System: https://obis.org/
9 World Register of Marine Species: http://www.marinespecies.org
10 Ministère de l'Agriculture et de l'Alimentation: https://agriculture.gouv.fr/administration-centrale

Here, we present the Pecheker project as a case study of fisheries data curation to be shared with the readers. For fishery monitoring teams already involved in similar projects or planning a similar development, we highlight the key features we have identified from our own experience. This is of special interest for teams involved in the implementation of the ecosystem-based management approach within their fishery monitoring programs. For scientists involved in data curation, we present an example of this field applied to the world of the fisheries, which is related to very important issues regarding humanity's food resources management in the context of global biodiversity erosion (Hiddink et al., 2008; Tromeur and Doyen, 2016). For our own team, we would be happy to obtain from our readers in the scientific community of data curation any feedback allowing us to improve our own work.

The "Fishing and scientific context" section includes a short general description of the background of the project. This allows a quick overview of Pecheker as a data resource and information system for readers involved in ecology or fisheries science and potential users of the database contents. For readers involved in fisheries data curation, this allows researchers to compare the context of the French southern fisheries to their own field, to assess how our respective needs, constraints and solutions can be compared and shared. For digital curation scientists, this section provides an example of a description of a fishery as a context for the creation of a database in terms of related scientific services and field applications. The "Historical data integration" section includes a general description of the compilation, integration and consolidation process performed on the data collected before the creation of Pecheker and uploaded into the new database. We highlight the key elements we discovered during this important part of the project as a feedback for readers starting a similar work. In the "Template development" section we present the various steps of the building of the data storage. We focus on the main points which appeared us to be considered first when planning such a development. The "Online access and archiving" section presents a quick overview of the strategy we adopted for the long-term curation of the sources and the distribution of the data to the user community. Finally, the "Beyond Pecheker, a network of fishery databases?" section includes a discussion on how we could develop a global fishery information system, with a point of view based on the success factors we identified from our own experience.

Full code to deploy a fishery monitoring database based on the Pecheker template is provided in supplementary materials. The code includes the SQL queries to reproduce all the tables of the model (Annex 1) and the Oracle © PL/SQL code for basic sequences and triggers (Annex 1). A synthetic documentation of table contents is provided as commentaries within the code. Vocabulary used to describe non-core data and stored in generic description tables can be provided on demand by the corresponding author, as well as supplementary documentation and advice to deploy the database.

Fishing and Scientific Context

The French fisheries of the Southern Ocean and the southern Indian Ocean cover different areas and territories, with various fishing activities and distinct legal frameworks. The whole area (Figure 1) is divided into three French Exclusive Economic Zones (EEZ) and two international areas ruled by two distinct international organisations. The EEZ fisheries are subject to the French laws and are managed by a local administration, the "Terres Australes et Antarctiques Françaises"[11] (TAAF). The international areas fisheries are subject to international agreements and the French operations are managed by the DPMA. Public managing structures provide the legal framework of the fishing regulation, such as fishing quotas, the list of authorized fishing periods, authorized techniques, gears and vessels, and the list of species and areas to be strictly or partially protected[12 13 14 15]. Furthermore, public managing structures collect taxes from the fishing activity and part of this funding is dedicated for the scientific monitoring and the recruitment of the scientific observers involved in the field work of the monitoring.

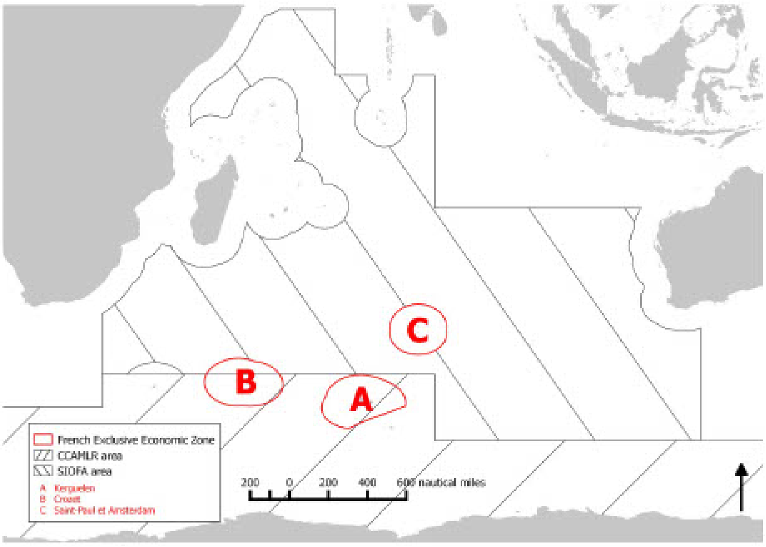


Figure 1. Geographical location of the French southern Exclusive Economic Zones of Kerguelen, Crozet and Saint-Paul et Amsterdam.

11 TAAF administration web page: https://taaf.fr

12 Legal decree of the Patagonian toothfish management plan and guidelines for regulation: https://taaf.fr/content/uploads/2019/11/JO-83_3%C3%A8me-trimestre-2019.pdf

13 Transposition of CCAMLR conservation measures into French law: https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000020331670?r=3Bd2Q7rVDp

14 Transposition of SIOFA fishing regulation measures into French law: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000039433567/

15 Legal decree for the creation of the TAAF national marine protected area: https://inpn.mnhn.fr/docs/espacesProteges/pprn/FR360016120170403.pdf

The Kerguelen EEZ (Surface: 567,732 km2) (Figure 1-A) constitutes the main area regarding the fishing activity, the amount of landings and the scientific research activity (Hureau, 2011). This subantartic EEZ is localized in the northern part of the Kerguelen Plateau. Modern commercial fishing started around the Kerguelen archipelago in the early 1970s when fishing grounds were discovered by the Soviet Union (Duhamel and Williams, 2011). Currently, no foreign fishing vessel operate in the area but seven French industrial automatic longliners, which are equipped with freezing capacities and transformation factories, target the Patagonian toothfish Dissostichus eleginoides Smitt, 1898 (Duhamel et al., 2011). This fleet is also involved in the exploitation of a smaller stock of Patagonian toothfish in the Crozet EEZ (Surface: 574 558 km²) (Figure 1-B) located up to the North, both in the subantarctic and subtropical regions (Pruvost et al., 2015a). The EEZ of Saint-Paul et Amsterdam (Surface: 509 015 km²) (Figure 1-C) is entirely located in the subtropical region. In this area, a single industrial fishing vessel uses traps, mainly, to target the rock lobster Jasus paulensis, Heller, 1862 (Pruvost et al., 2015b). The whole catch of these three fisheries are landed on the Île de la Réunion (French territory) to be exported to the Asian and North American markets.

The fleet involved in these three French EEZ fisheries also operates in the international seas. In the Indian Ocean, the commercial fishing activity is ruled by the Southern Indian Ocean Fisheries Agreement (SIOFA, 2006). In the Southern Ocean, activity is limited to prospecting and research fishing only, and ruled by the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR, 1980).

The scientific monitoring of these fisheries is realized by the MNHN which is a French State scientific institution. The monitoring, started in 1978 (Hureau, 2011), includes the collection of the fishing data and the provision of scientific guidelines for the public managing structures for the all three EEZ. The collection of data is based both on the survey of the commercial vessels activity and the organization of scientific campaigns (Duhamel, Williams, et al., 2011b). The guidelines include various components, such as fish stock assessment (which corresponds to exploited species biomasses estimation obtained with models), "Total Allowable Catch" evaluation by area and species (which corresponds to the maximum fishing limit allowing the fish population's ability to restore itself) and recommendations in the use of fishing gears or season closure (Duhamel et al., 2011). Advice and analysis are also provided on demand to the public managers regarding the biodiversity conservation issues in the fishing zones and the marine protected areas (Martin et al., 2019). In the CCAMLR and SIOFA areas, the monitoring includes also the scientific design of the official research plans and the reporting and analysis of the results during the dedicated international scientific workshops (Péron et al., 2018).

On board commercial vessels, the fishery observers, employed by the TAAF administration in the EEZ with addition of international ones in the CCAMLR area, collect scientific data under the MNHN supervision (Gasco, 2011). Data are archived on hard copies as well as on electronic logbooks which are sent to the MNHN at the end of each fishing trip. In the MNHN, the data is checked then uploaded and managed in the Pecheker database. The collection includes comprehensive data of catches for all target and bycatch species, fish and lobster measurements, description of vessels activity, description of gears and mitigation devices deployment, and observations of birds, mammals and benthic invertebrate species. Furthermore, the data to be uploaded into Pecheker also include the meta-data of the samples of organisms and photos which are collected on field. Adding to those commercial trips, dedicated scientific surveys (independent of the fishery activities) with a larger team of scientists and complementary protocols also provide data which are uploaded into the Pecheker database (Duhamel et al., 2011; Duhamel et al., 2019; Martin et al., 2019).

In 2018, Pecheker allowed the storage of data for 152,984 georeferenced stations over the whole areas of the French southern fisheries. Considering continuous areas of cells independently of the quantity of fishing operations, the main area appears to be the Kerguelen EEZ (Figure 7-A), followed by the western region of the Crozet EEZ (Figure 7-B). The Ob and Lena banks (Figure 7-3), the Del Cano ridge (Figure 2-1), the Elan bank (Figure 2-2) and the EEZ of Saint-Paul et Amsterdam (Figure 2-3) constitute secondary areas in terms of covered surface. Data is also managed into Pecheker for various isolated spots of prospecting fishing in the Indian Ocean and the Southern Ocean. These spots include two areas located on the Antarctica eastern continental shelf (Figure 2-4,5), which correspond to the most southern fishing stations recorded in the database.

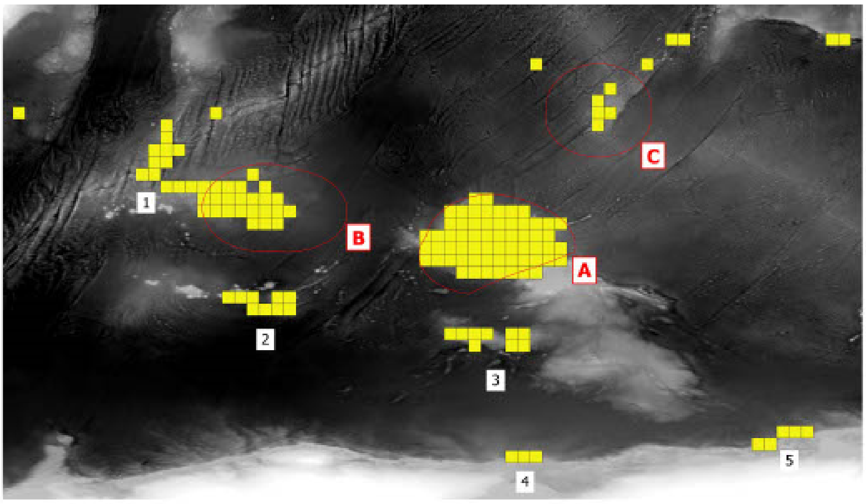


Figure 2. Geographical distribution of the Pecheker data; stations location are aggregated into a matrix of presence data according to grid of cells of 1°; A: Kerguelen, B: Crozet, C: Saint-Paul et Amsterdam, 1: Del Cano ridge, 2: Elan bank, 3: Ob et Lena Bank, 4 and 5: Antarctica shelf.

For the subantarctic and Antarctic regions, including the Kerguelen EEZ, the Crozet EEZ and the various fishing spots of the international seas of the CCAMLR, 121 905 operations are distributed from the 0-100 to the 2800-2900 meters depth bathome. For the Saint-Paul et Amsterdam EEZ, 31,079 fishing operations are distributed from the 0-100 to the 2400-2500 meters depth bathome. For the non-commercial by-catch taxa (demersal fishes and benthic invertebrates) and the observed species (marine mammals and birds), the availability of data related to the fishing stations is heterogeneous in time but similar regarding the spatial coverage. Pecheker appears to constitute an efficient information system for various components of the Southern Ocean biodiversity, with a continuous increase of the amount of taxa which are recorded and an important geographical coverage for deep-sea ecosystems, usually considered to be "data-poor" areas.

Historical Data Integration

On board computerization with a continuous quality checking process started in 2000 (Gasco, 2011). Data collected after that year was available for the Pecheker project into a series of Microsoft Excel© files, including a common integrated format and a previous consolidation procedure. For data collected between 2000 and 2007 (number of fishing operations: 23,606; number of vessels: 11), the integration process was limited to the development of an exchange file format and a tool to transcode and upload the datasets from the exchange files into Pecheker. This tool, based on SQL queries to be executed with the Oracle SQL Developer© client or a dedicated set of Microsoft© Visual Basic forms, was later adopted as the ordinary process to upload new datasets into Pecheker.

Data collected from 1978 to 2000 (number of fishing operations: 62,988; number of vessels: 67) was available in various physical supports (original paper support, floppy disks, hard drives, CD-ROMS) and formats (field notebooks, reports and synthesis, data tables in various formats, spreadsheets, files from obsolete software) with partial computerization only and no prior integrated quality checking procedure (ICES, 2018; Kosmala et al., 2016; McLure et al., 2014; Vandepitte et al., 2015). The integration of this historical data constituted a tricky technical challenge. Most of the time used for the first step of the Pecheker development was used for this part of the project. The integration process included various steps: computerization, obsolete files recovery, data formatting and standardization, completeness checking, quality checking, corrections and upload into Pecheker. The heterogeneity of this raw data prevented the possibility to define a simple work plan based on common tasks, tools and approaches: the scheduling of the various steps of the integration process depended on each dataset. The computer work turned out to be long to process and complex to organize.

The main difficulty was related to the absence of a prior curation policy for the sources: the various sources were not conserved in a unique storage room and any inventory including an up-to-date list of the whole sources was available, nor any meta-data to describe the contents (Day, 2005; Higgins, 2011; Peer and Green, 2012; Specht et al., 2018). The completeness checking of the sources, a key element for a fishery database dedicated to the ecosystem-based monitoring approach (Busch et al., 2003; Van Iseghem et al., 2011), appeared to be the more difficult task to perform. If quality level of a dataset can be estimated using mathematical tools, completeness level remains more difficult to estimate because it is related to a large set of factors (Vandepitte et al., 2015). To obtain a strong evaluation of the completeness level of a given dataset, very different informations concerning the context of the dataset need to be crossed, such as other similar datasets, or time tables, schedules, reports and scientific publications. This effort was made gradually with the other tasks part of the integration process. At this end of the development, this appeared to be a time consuming strategy. For our project, database construction preceded the implementation of the curation policy. In light of our experience, to facilitate the completeness checking and improve the efficiency of the work plan, we strongly recommend to start a similar project by setting up a curation policy for the sources, including an efficient storage, an inventory and the recording of metadata, prior to the construction of the database with full integration of the raw data.

Quality checking and corrections were processed on raw data with SQL queries. Various logical analysis and statistical tests were pursued to highlight errors and outliers. These tests were applied to the various numerical distributions included in the dataset, such as the spatial position data, the time series, the recording of catches and the recording of biometry. Corrections also included the standardization of the vocabulary used for the naming and the description of the data, and the standardization of the units for the recorded variables. Human expertise from the senior scientists of the team, involved in the production of this historical data, was also used intensively. The presence and accessibility to these colleagues, able to provide comments and analysis on the data quality and the acquisition context, was a key element for the success of the historical data integration, regarding both the quality of the treatment and the time spent on this part of the project. This constitutes a second recommendation we have drawn from our experience: in order to integrate historical data, computer work cannot replace human experience of the data and it is strongly recommended to start such a project when senior scientists are still present and accessible.

Furthermore, the original reports and field notebooks have been consulted to be compared to the computerized data. Computerization, checking and correction procedures were applied to all the historical data to be uploaded in Pecheker:

- fishing operations, including time, latitude, longitude and depth
- gears deployment, including full technical description,
- fishing campaigns dates and schedules,
- vessels technical description,
- catches of target and by-catch species,
- recording of the fish products,
- fish and crustaceans biometry,
- specimen sampling.

At the beginning of the project, various socio-economic datasets and legal data have been also included, for instance crew composition, selling prices of the fish products, detail of conservation measures and legal framework for fishing regulation. According to the evolution of the scope of the Pecheker project, this data has not been later updated nor collected.

Template Development

Pecheker is based on an Oracle 11G© relational database. Custom R (R Core Team, 2021) scripts and SQL views are used to access the database. Various sets of SQL queries allow to transcode and export the data to feed other fisheries Information Systems, such as the CCAMLR or SIOFA database[16] (CCAMLR, 2018a; CCAMLR, 2018b).

The former version of the template was based on a set of tables which structure was close to the format of the raw data as collected on field. In the first version, each table was dedicated to a thematic dataset (for instance: biometry table, stations geolocalisation table, fishing schedules table, etc). Each column contained the recording of a single information (for instance: total or standard length of an organism, latitude of a station, date of a vessel's departure from the harbour, etc). This template resulted in a simple structure with tables including a large amount of columns easy to read and query. But this template, because of its rigidity, soon appeared to be not suitable to record unstructured data nor to follow the rapid evolution of protocols due to the evolution of

16 SIOFA database: http://www.apsoi.org/node/461

fishing techniques. Furthermore, new needs in terms of stock assessment (Constable and Welsford, 2011) and statistical modelling (Mormede et al., 2011) resulted in a huge increase of data acquisition in terms of precision, diversity and quantity. The frequency of the changes also increased at the same time.

For instance, the rapid and recent diversification of the fish products resulted in a more complex recording process of the fish catches in the vessels factories, increasing the amount of variables to be stored in the database (Figure 4). The number of taxa recorded in Pecheker for caught organisms, corresponding to the demersal fishes and the benthic invertebrates, also increased continuously during the program (Figure 4). The minimum is four taxa recorded at the beginning of the scientific monitoring, corresponding to the target species of the trawl fisheries at this moment. The maximum value is 128 taxa sampled for a unique year, including target species, by-catch fishes and marine invertebrates impacted by fishing gears.

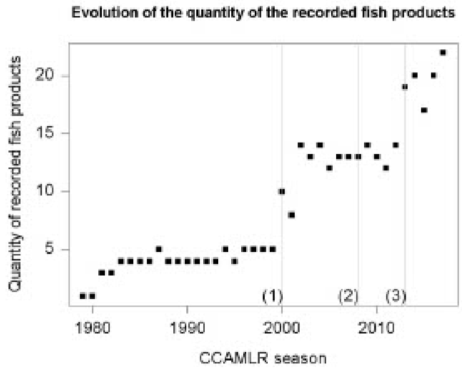


Figure 3. Evolution of the number of recorded fish products (for instance: fish fillet, fish cheeks) from 1978 to 2017 (min = 1, max = 22); (1) on-board computerization, (2) Pecheker first version, (3) Pecheker second version.

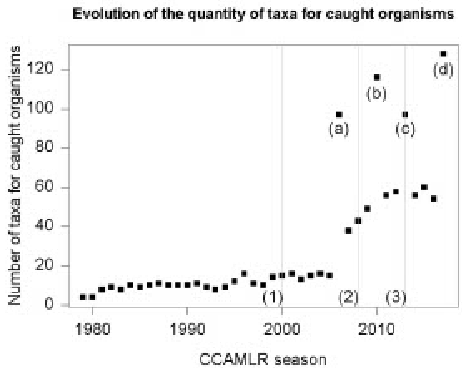


Figure 4. Evolution of the quantity of recorded taxa (fishes and benthos) data from 1978 to 2017; (1) on-board computerization, (2) Pecheker first version, (3) Pecheker second version; (a-b-c-d); years of scientific surveys: a, b, c and d.

Various sets of data from Pecheker show a similar profile with a strong and continuous increase of recordings over time. For the description of the fishing operations and gears deployment, the seven core variables recorded since the beginning of the scientific observer program are the ID number of the fishing operation, the date, the latitude and the longitude (in degrees and minutes) and the gear type. The recording of the full technical description of the fishing operations reaches today 124 information for each operation (for instance: number of baited hooks, size of the hooks, composition of the baits, sky brightness during the fishing gears retrieving, etc). For the mitigation techniques (deployment of devices used to reduce the impact of fishing activity on birds and marine mammals) and biodiversity monitoring data, the minimum of information is zero, from 1978 to 2002, before this data was collected. The recording of this data started in 2003 and reaches today 101 information for each operation, with related biodiversity observations for 32 species of mammals and birds.

The rapid obsolescence of the first template due to this strong increase of data led us to an interpretation which had important consequences in the way we consider our need to think the technical evolutions of our information system. We understood that the development of Pecheker started at the end of a former period when data acquisition programs were strongly based on well delimited templates, which design necessary preceded and constrained the field work (FAO, 1999). The changeover to a new time in the world of databases occurred at the end of the 2000s with the new era of "Big Data" (Candela et al., 2015; Davenport et al., 2012; Marx, 2013; Waller and Fawcett, 2013). In the world of fisheries science, this new era is characterized by a reversal of the hierarchy between the data and the database, as we observed it at the beginning of the Pecheker project: nowadays, the design of a fishery database must not be only performed upstream the organisation of the data flow, but needs also, in same time, to be considered downstream the chain; the main goal is to obtain a tool not only able to generate the data flow, but a tool also able to absorb the data flow, however it is shaped, diverse and abundant.

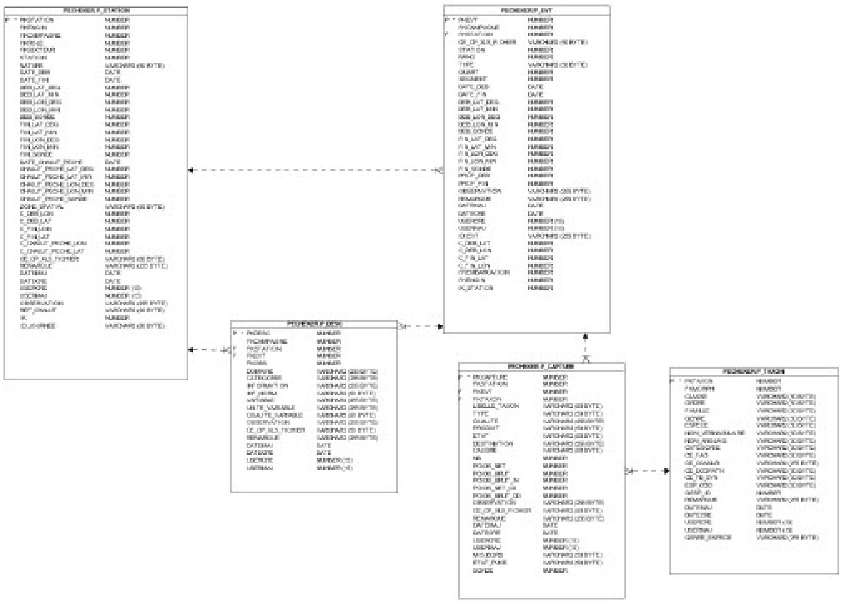


Figure 5. Extract of the relational model of Pecheker used to record the data of fishing operations (table "pecheker.p_station"), field events (table "pecheker.p_evt") and fish catches (table "pecheker.p_capture" and table "pecheker.p_taxon"); the generic description table "pecheker.p_desc", represented here linked by foreign keys to the tables "pecheker.p_station" and "pecheker.p_evt", can be used to store all informations exceeding the core data; the table includes 6 generic columns to describe and structure information: "pecheker.p_desc.domaine" (domain of information, e.g. "fishing gears"), "pecheker.p_desc.categorie" (category of information, e.g. "hooks"), "pecheker.p_desc.information" (name of information, e.g. "hooks size"), "pecheker.p_desc.variable" (value of the variable, e.g. "20/0"), "pecheker.p_desc.unite_variable" (unit of the variable, e.g. "Mustad hook size numbering system") and "pecheker.p_desc.qualite_variable" (a quality index for each recording, e.g. "exact value for all the longline"); catches are recorded to be linked to a hauling event and a taxon.

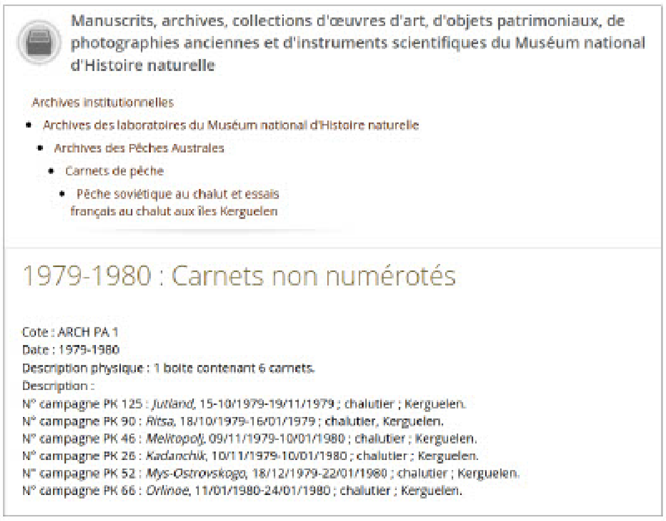
To take account of these constraints, a new relational model of Pecheker, which is currently in use, was delivered in 2013. The core data remains stored within thematic tables and structured in columns each dedicated to a single type of information. But the new template also includes various flexible structures allowing researchers to describe, to organize and to name the information exceeding the core data with a set of generic columns in non-thematic tables. For instance in the table "pecheker.p_desc" (Figure 5), each information contained in a row is named in the column "information" for which a variable, a unit and a quality index can be associated ; all the information are grouped by "categories", and categories are grouped by "domains". This is suitable to store unstructured data, opportunistic observations, data from non permanent acquisition protocols or data from permanent protocols but implying a huge amount of information. This is of particular interest to store into compact tables the data of the catches and fishery products, the gears deployments full description, the biodiversity observations and the description of the various events which can randomly occur during the fishing operations (for instance: multiple fishing gears retrieving due to the break of a longline).

Such table contents are less human-readable and queries structure is more complex. To extract a single information, a query has to read the contents of the six generic columns and to take account of their hierarchical nested organisation, when a single column is needed to extract a core variable from a thematic table. Furthermore, this implies a more important effort to maintain a controlled vocabulary in the database, due to the increase of terms and due to their standardization which can be more easily corrupted by users. But with no need to modify the template and no need to add new columns or new tables when a new field protocol is introduced, this new model appears to be more suitable to prevent upcoming evolutions in data acquisition. Moreover, this template also facilitates interoperability, which is an important issue to allow the integration of Pecheker data within regional databases and to facilitate the reuse of data for global analysis. Indeed, full description of information is included not only in external documentation but also within the structure of the data and it is based on controlled vocabulary: thus, queries used to format the data to fit the exchange templates can be more easily maintained by database administrators and metadata can be more easily produced (Daraio et al., 2016; Ide and Pustejovsky, 2010). Furthermore, this improvement of interoperability also contributes to make the database independent from the electronic logbooks allowing both tools to be developed separately, which is a key element for their sustainability.

The new template includes also the possibility to record metadata and a structured description of the field protocols, allowing us to link the raw data to the scientific context of their production. Through the metadata tables, it's possible to group the Pecheker contents not only according to data structure but also according to historical, scientific or technical consistency, which facilitates the data mining (for instance: "Soviet Union trawling activity around the Kerguelen archipelago between 1978 and 1989", "Birds counting during sunset in the Crozet EEZ between 2000 and 2010", "Patagonian toothfish fillet production in Kerguelen EEZ in 2017") (Guerrero et al., 2017; Losiewicz et al., 2000). Moreover, the tables dedicated to store the protocols description allow to organise and structure all the instructions provided to the scientific observers, including the possibility to record all the changes over time in the application of the protocols. This is of major importance to understand the raw data uploaded into Pecheker, for evaluation of the data quality, for the statistical analysis and the tuning of the models based on the raw data. For instance, the conversion factor calculation protocols to extrapolate processed weights of fish products recorded in the factories into full weights of caught animals in the ecosystems had many changes over time. These protocols are dependent on both the evolution of the legal framework, the evolution of fish transformation techniques and the field constraints of the observers. For instance, years ago, when computers memory was strongly limited, values were rounded when computerized. Furthermore, according to the period, the conversion factor calculation has been calculated by scientific observers for each vessel on every fishing campaign or officially set by regulatory authorities for all the fleet every fishing season. Knowing such details of these evolutions in time is very important to assess the raw data quality, for the standardization of the datasets extracted from Pecheker to be used as inputs for the fish biomass estimation models (McCluskey and Lewison, 2008; Rosenberg et al., 2005). Variations in the preliminary treatment of these inputs may induce variations in the final biomass assessment results (Colvin et al., 2012; Romagnoni et al., 2012), which is particularly sensitive regarding both economic and biodiversity conservation issues for the various stakeholders: fishing companies, environmental NGO, public managing structures, customers and citizens.

Online Access and Archiving

The database is hosted in a remote server. The server is managed by the IT Department of the MNHN, which provides a high level of service regarding IT engineering, security strategy and conservation. The multi-site backup process is based on mirrored servers allowing a continuous saving system with full-time availability of the database, and a tape storage service with a weekly frequency. Furthermore, hard copies of the full data and original field notebooks are conserved in the archive department of the MNHN library. This classic curation system is the key element for the long-time preservation of the data (Noonan and Chute, 2014). For data from 1978 to 2017, this represents a documentary collection of 192 archive boxes. The annual increase represents approximately 16 boxes. The archive boxes are managed by conservation professionals, stored in secured rooms with limited access and controlled atmosphere. The list and contents description of this documentary collection can be accessed online through the "Calames" website dedicated to the management of the MNHN archives (Figure 6). The online metadata includes the physical identification of the conservation boxes and a short description of their contents: references of the field notebooks, identification number of the campaigns for interoperability with Pecheker, name of the vessel, fishing dates, fishing technique and fishing area (Figure 7).


Manuscrits, archives, collections d'œuvres d'art, d'objets patrimoniaux, de photographies anciennes et d'instruments scientifiques du Muséum national d'Histoire naturelle

Archives institutionnelles

- Archives des laboratoires du Muséum national d'Histoire naturelle
  - Archives des Pêches Australes
    - Carnets de pêche
      - Pêche soviétique au chalut et essais français au chalut aux îles Kerguelen

1979-1980 : Carnets non numérotés

Cote : ARCH PA 1
Date : 1979-1980
Description physique : 1 boite contenant 6 carnets.
Description :
N° campagne PK 125 : *Jutland*, 15-10/1979-19/11/1979 ; chalutier ; Kerguelen.
N° campagne PK 90 : *Ritsa*, 18/10/1979-16/01/1979 ; chalutier, Kerguelen.
N° campagne PK 46 : *Melitopolj*, 09/11/1979-10/01/1980 ; chalutier ; Kerguelen.
N° campagne PK 26 : *Kadanchik*, 10/11/1979-10/01/1980 ; chalutier ; Kerguelen.
N° campagne PK 52 : *Mys-Ostrovskogo*, 18/12/1979-22/01/1980 ; chalutier ; Kerguelen.
N° campagne PK 66 : *Orlinoe*, 11/01/1980-24/01/1980 ; chalutier ; Kerguelen.

Figure 6. Online metadata of the archive box contents.[17]

[17] See: http://www.calames.abes.fr/pub/mnhn.aspx#details?id=FileId-2172

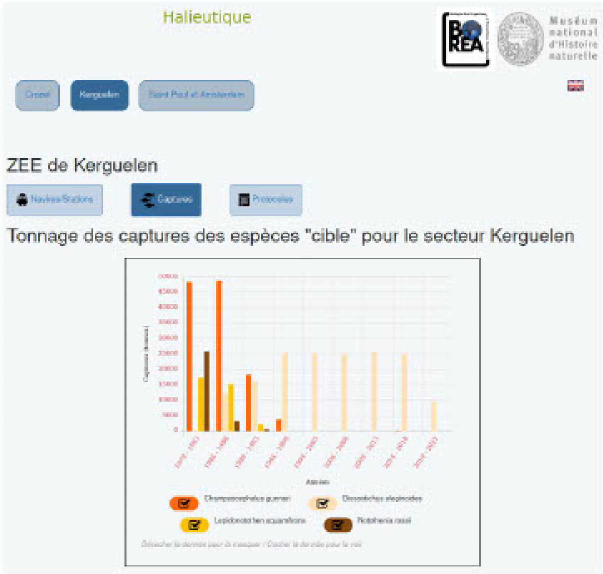


Figure 7. The "Halieutique" prototype web site including aggregated statistics dynamically computed from Pecheker.[18]

As a shared reporting system, a prototype of web site for now called "Halieutique" and which development is still in progress allows to access to a set of aggregated statistics dynamically computed from Pecheker with SQL queries (Figure 6). This series of statistics include various indicators about the three French EEZ southern fisheries. For each EEZ, indicators are presented into a series of dedicated graphics with comments allowing their interpretation. Time series of data are provided for fishing operations, catches and by-catches volumes, and biological sampling protocols. A short-term objective is to deploy a full online atlas about French southern fisheries based on the "Halieutique" project. Due to the commercial nature of the raw data, the confidentiality constraints about vessels activity and the protection of the fish stocks against illegal fishing, the raw data of Pecheker is not available online.

Data is transmitted on demand to the French management teams officially in charge of the regulation of the fisheries and the survey of the marine protected areas. Such data is transmitted aggregated into indicators or in a raw format allowing these users to perform their own analysis. To contribute to international regulation, data reporting of the French fishing activity is also provided to the SIOFA and to the CCAMLR according to the templates provided by their secretariats (CCAMLR, 2020a; SIOFA, 2019). Raw data can be also accessed on demand for research projects by scientific users through the signature of an agreement limited in time and setting out the conditions to use the data. The scientific users engage themselves not to keep any copy of the data after the end of the agreement. Access requests are sent to the MNHN with a description of the research project. When the access request is accepted the data is provided into an aggregated format which allows to protect the confidentiality of the vessels commercial activity.

18 Halieutique: http://halieutique.mnhn.fr

Beyond Pecheker: A Network of Fishery Databases?

Our experience of the development of Pecheker allows us to highlight a series of success factors. They partly overlap those which have been identified for the development of various projects by several authors (Bayraktarov et al., 2019; Costello et al., 2013; Costello et al., 2018; Vandepitte et al., 2018; Wilkinson et al., 2016). But we could also identify a series of success factors which contradict the conclusions of teams involved in such leading projects, suggesting specificities of the fisheries world as a scope of application for data curation.

Permanent funding, a permanent team and hosting of the database within a state institution appeared, unsurprisingly, to be a set of driving factors for the success of the project. Furthermore, the uniqueness of the database concerning spatial extent and scope, the availability of the whole data in a centralized repository and the high quality of the raw data with continuous data curation, gradually imposed Pecheker as the reference for French southern fishery data. The use of raw files from field, directly extracted from electronic logbooks by the managing structures and the scientific users, progressively disappeared to let Pecheker becoming the unique source of information for the user community. All these factors have been also identified as key elements by the above-mentioned authors involved in the development of various leading biodiversity databases.

However, other important factors should be considered. First of all, contrary to the principle of "bigger is better" in terms of database size, which has been identified as an important criteria for the success of various leading biodiversity information systems (Costello et al., 2013), Pecheker is based on the principle that "smaller is more sustainable". The spatial extent is restricted to a limited geographical area defined by political boundaries and not by large ecological regions. The contents of the database are restricted to a single monitoring program, including a limited set of field protocols and controlled data flows. Thus, if completeness of data reaches a high level considering the scope of the project, Pecheker remains a relatively small size information system. Furthermore, the community around Pecheker is also restricted, considering both the contributors and the users, and the database remains no open-access. Contributors are limited to the scientific field observers – a community of approximately 20 contributors every year, but with important turn over – and the eight members of the MNHN team involved in the scientific design of the protocols and the curation of the database. The user community is constituted by the TAAF and DPMA teams, the secretariats and working groups of SIOFA and CCAMLR, and some research teams involved in fishery science (Duhamel et al., 2011; Péron et al., 2016; Tixier et al., 2016) and/or Southern Ocean marine species biology: fishes (Straube et al., 2011; Tomaszkiewicz et al., 2011), birds (Michael et al., 2019), marine mammals (Gasco et al., 2019; Tixier et al., 2019) and invertebrates (Améziane et al., 2011; Chérel et al., 2011; Martin et al., 2019).

These various limitations allow us to provide a data curation service which reaches a high effectiveness level, with a high level of accuracy, adequacy and completeness of the database, justifying the existence of Pecheker for the various stakeholders and the funding from the French state. Furthermore, the limited community and the unique source of funding allow important flexibility for the engineers involved in the implementation of the technical evolutions which have to be done to maintain the information system efficient and up-to-date. The reassessment of the first template of Pecheker after only four years is an example of this flexibility, which would have been impossible if we have had to deal with a large collaborative community of users, various funding structures and external development.

Any global information system including raw data is available for fisheries, as the GBIF system for biodiversity or WORMS for taxonomy. Due to the specific constraints we discovered and faced when we developed Pecheker, we think that such a project should be based on a set of principles and steps specifically designed for the world of fisheries. Rather than a huge database including a large amount of contents, we think that a global information system for fisheries should be based on a decentralized network of small databases, each dedicated to a fishery or a monitoring program, but including interoperability standards. First, such a network should allow to keep the raw data confidential and not open-access, which remains a primary constraint to protect fish species and ecosystems from illegal fishing, legal overexploitation and marine fishing grounds destruction. Secondary, such a network structure should allow an efficient data curation activity, based on teams each specialist of a fishery, working close to the field and able to define the local needs and constraints to optimise the tools and the monitoring. International regulatory authorities, such as the CCAMLR or the RFMOs, could build the nodes of the network where data should be available aggregated by squares and in the form of quantitative and spatial indicators (Watson, 2017). These nodes could be the level around which larger user communities could be organized (Borgman, 2012; Cragin et al., 2010; Kratz and Strasser, 2014; Wallis et al., 2013; Zimmerman, 2007).

Interoperability should be based on three main steps. The exchange of the storage templates between the teams involved in fishery data curation, with associated meta-data describing the template itself and the fishing context, could be the first step. This should allow to build a global community of fishery database curators, to share experience and solutions, and make the raw data storage templates converge as much as possible. A second step should be the development of a common reference database, including an international standardization of the vocabulary used in the fishery databases, taking account of the diversity of the human languages and words (Hamm, 2012; Hughes et al., 2008; Shamsfard and Barforoush, 2004). This should include a full thesaurus and classification of the terms and units used to describe the fishing techniques, gears, tools, vessels, fish products and species. Experience from taxonomic or life-traits databases, for instance WORMS or FishBase, could be a precious aid, given various common issues have been there solved, such as the nested classification of terms, the management of synonymy and the management of a large community of contributors (Costello et al., 2013; Vandepitte et al., 2018).

Moreover, an international database of protocols should be planned. Such database should include a standardized and detailed description of the protocols used within the various fisheries to collect data. Contents of the conservation measures provided by national or international regulatory authorities could be used for the launch of the project, given they already include a large set of standardized and detailed description of the protocols used to report mandatory data for fishing regulation[19] [20] (CCAMLR, 2020b). This could improve interoperability by allowing to facilitate aggregations of datasets provided by distinct monitoring programs. Furthermore, such project would constitute an important criteria for the long-lasting usability of the fishery database contents (Chao, 2015; Tessarolo et al., 2017; Zimmerman, 2007; Zimmerman, 2008). Indeed, the long-term availability of information about the protocols used to produce the raw data, which implies a centralized information system and not local storage on

19 CCAMLR Conservation Measure 10-05 (2018): https://www.ccamlr.org/en/measure-10-05-2018
20 CCAMLR Conservation Measure 22-06 (2019): https://www.ccamlr.org/en/measure-22-06-2019

disparate documents, will be decisive to maintain the long-term possibility to treat fishery catches data in terms of abundances and trends, allowing long-term predictions, ecological systems variations analysis and future historical comparisons.

Conclusion

On board computerization of collected data constituted an important technical improvement for the French Southern Ocean fisheries monitoring program. A few years later, the Pecheker project constituted another important step for the program, with the integration of the data into a complete dedicated information system, allowing the setting of a digital curation policy and a strong improvement of the information availability. Over the next few years, upcoming evolutions in the nature of the collected data will constrain the fishery database curators to face new technological challenges. Succeeding to the current era of collection of recordings to be stored into data tables, new field protocols based on new technologies will induce supplementary constraints for the design of the fishery information systems. For instance, the availability and accessibility of Next-Generation Sequencing technologies will allow the production of massive molecular data. Using techniques such as Close Kin Mark Recapture and environmental DNA, molecular approaches will allow to develop new methods to assess the fish populations size or to perform ecological surveys on the impacted ecosystems. This will result in the production of new massive data flows with the need of managing specific file formats to be linked to the databases that are currently in service. Moreover, the increase of the use of marine video monitoring systems may lead to develop new solutions also dedicated to manage massive video data, based on specific file formats. Regarding long term conservation issues, this will constitute a strong challenge for curators, given such massive digital informations cannot be saved on hard copies, as data based on recording of variables within tables can be.

Acknowledgements

We would like to acknowledge the Direction des Pêches Maritimes et de l'Aquaculture du Ministère de l'Agriculture et de l'Alimentation and the Muséum national d'Histoire naturelle for support and funding. We also would like to acknowledge all the fishery observers from the Terres Australes et Antarctiques Françaises who provide raw data to be integrated into Pecheker.

We dedicate this article to Jean-François Murail, deceased in 2021, senior Research Engineer of the Muséum national d'Histoire naturelle and computer science pioneer, who co-created and developed the IT infrastructure of our institution.

References

Améziane, N., Eléaume, M., Hemery, L.G., Monniot, F., Hemery, A., Hautecoeur, M., &, Dettaï, A. (2011). Biodiversity of the benthos off Kerguelen Islands: Overview and perspectives. In Duhamel G., Welsford D. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 157-67). doi:10.26028/cybium/2011-35SP-016

Alaska Fisheries Science Center & Alaska Regional Office. (2019). North Pacific Observer program 2018 Annual Report. AFSC Processed Rep. 2019-04. Retrieved from https://repository.library.noaa.gov/view/noaa/20199

Baum, J. K., & Blanchard, W. (2010). Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fisheries Research, 102(3), 229-239. doi:10.1016/j.fishres.2009.11.006

Bayraktarov, E., Ehmke, G., O'Connor, J., Burns, E. L., Nguyen, H. A., McRae, L., Possingham, H. P., & Lindenmayer, D. B. (2019). Do big unstructured biodiversity data mean more knowledge? *Frontiers in Ecology and Evolution, 6.* doi:10.3389/fevo.2018.00239

Belhabib, D., Koutob, V., Sall, A., Lam, V. W. Y., & Pauly, D. (2014). Fisheries catch misreporting and its implications : The case of Senegal. Fisheries Research, 151, 1-11. doi:10.1016/j.fishres.2013.12.006

Bénard-Capelle, J., Guillonneau, V., Nouvian, C., Fournier, N., Le Loët, K., & Dettai, A. (2015). Fish mislabelling in France: Substitution rates and retail types. *PeerJ, 2, e714.* doi:10.7717/peerj.714

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology, 63*(6), 1059-1078. doi:10.1002/asi.22634

Busch, W.-D. N., Brown, B. L., & Mayer, G. F. (Eds.). (n.d.). Strategic Guidance for Implementing an Ecosystem-based Approach to Fisheries Management. Prepared for the Marine Fisheries Advisory Committee by the Ecosystem Approach Task Force, NOA. Retrieved from https://ocean.floridamarine.org/efh_coral/pdfs/MafacReportEcoMgmt.pdf

Candela, L., Castelli, D., Manghi, P., & Tani, A. (2015). Data journals: A survey. *Journal of the Association for Information Science and Technology, 66*(9), 1747-1762. doi:10.1002/asi.23358

CCAMLR. (1980). The Convention for the Conservation of Antarctic Marine Living Resources. Retrieved from https://www.ccamlr.org/en/organisation/camlr-convention-text

CCAMLR. (2018a). Fishery Report 2018: Dissostichus eleginoides Crozet Island French EEZ (Subarea 58.6). CCAMLR Fishery Reports. Retrieved from https://www.ccamlr.org/fr/node/99089

CCAMLR. (2018b). Fishery Report 2018: Dissostichus eleginoides Kerguelen Islands French EEZ (Division 58.5.1). CCAMLR Fishery Reports. Retrieved from https://www.ccamlr.org/fr/node/103788

CCAMLR. (2020a). CCAMLR Catch and Effort Data Submission Instructions. Retrieved from https://www.ccamlr.org/ru/system/files/Catch%20and%20effort%20data%20submission%20final%20formatted_1.pdf

CCAMLR. (2020b). Schedule of Conservation Measures in Force 2020/21. Retrieved from https://www.ccamlr.org/en/system/files/e-schedule2020-21_0.pdf

Chao, T. (2015). Mapping methods metadata for research data. *International Journal of Digital Curation, 10*(1), 82-94. doi:10.2218/ijdc.v10i1.347

Chazeau C., Iglésias S., Péron C., Gasco N., Martin M., & Duhamel G. (2019). Shark by-catch observed in the bottom longline fishery off the Kerguelen Islands in 2006–2016, with a focus on the traveller lantern shark (Etmopterus viator). In Welsford D., Dell J., Duhamel G. (Eds), *The Kerguelen Plateau: marine ecosystem and fisheries, proceedings of the Second Symposium.* Australian Antarctic Division, Kingston (pp. 311-327).

Cheal, A. J., & Emslie, M. J. (2020). Counts of coral reef fishes by an experienced observer are not biased by the number of target species. *Journal of Fish Biology, 97*(4), 1063-1071. doi:10.1111/jfb.14466

Chérel, Y., Gasco, N., & Duhamel, G. (2011). Top predators and stable isotopes document the cephalopod fauna and its trophic relationships in Kerguelen waters. In Duhamel G., Welsford D. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 99-108). doi:10.26028/cybium/2011-35SP-016

Clua, E., Beliaeff, B., Chauvet, C., David, G., Ferraris, J., Kronen, M., Kulbicki, M., Labrosse, P., Letourneur, Y., Pelletier, D., Thébaud, O., & Léopold, M. (2005). Towards multidisciplinary indicator dashboards for coral reef fisheries management. *Aquatic Living Resources, 18*(3), 199-213. doi:10.1051/alr:2005026

Colvin, M. E., Pierce, C.L., & Stewart, T.W., (2012). Semidiscrete biomass dynamic modeling: An improved approach for assessing fish stock responses to pulsed harvest events. *Canadian Journal of Fisheries and Aquatic Sciences, 69*, 1710– 1721. Retrieved from https://pdfs.semanticscholar.org/1204/2d97a540b5d62a972119bf6d26df41689017.pdf

Constable A. J., Welsford D. (2011). Developing a precautionary ecosystem approach to managing fisheries and other marine activities at Heard Island and McDonald Islands in the Indian Sector of the Southern Ocean. In Duhamel G., Welsford D. (Eds), The Kerguelen Plateau: Marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris, (pp. 233-255).

Costello, M. J., Bouchet, P., Boxshall, G., Fauchald, K., Gordon, D., Hoeksema, B. W., Poore, G. C. B., van Soest, R. W. M., Stöhr, S., Walter, T. C., Vanhoorne, B., Decock, W., & Appeltans, W. (2013). Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and Related Databases. *PLoS ONE, 8*(1), e51629. doi:10.1371/journal.pone.0051629

Costello, M. J., Horton, T., & Kroh, A. (2018). Sustainable biodiversity databasing: International, collaborative, dynamic, centralised. *Trends in Ecology & Evolution, 33*(11), 803-805. doi:10.1016/j.tree.2018.08.006

Costello, M. J., Michener, W. K., Gahegan, M., Zhang, Z.-Q., & Bourne, P. E. (2013). Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology & Evolution, 28*(8), 454-461. doi:10.1016/j.tree.2013.05.002

Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368*(1926), 4023-4038. doi:10.1098/rsta.2010.0165

Daraio, C., Lenzerini, M., Leporelli, C., Naggar, P., Bonaccorsi, A., & Bartolucci, A. (2016). The advantages of an ontology-based data management approach: Openness, interoperability and data quality. *Scientometrics, 108*(1), 441-455. doi:10.1007/s11192-016-1913-6

Davies, S.L., & Reynolds, J.E. (Eds.) (2002). Guidelines for developing an at-sea fishery observer programme. FAO Fisheries Technical Paper 414. Retrieved from http://www.fao.org/3/y4390e/y4390e.pdf

Davenport, T. H., Barth, P., & Bean, R. (2012). How 'big data' is different. *MIT Sloan Management Review 54*(1). Retrieved from https://sloanreview.mit.edu/article/how-big-data-is-different/

Day, M. (2005). "Metadata", DCC Digital Curation Manual. Retrieved from https://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/metadata/metadata.pdf

Degel, H., & Jansen, T. (2006). FishFrame Fisheries and stock assessment data framework. ICES CM 2006/M:02. Retrieved from https://www.ices.dk/sites/pub/CM%20Doccuments/2006/M/M0206.pdf

Duhamel G., Péron C., Sinègre R., Chazeau C., Gasco N., Hautecœur M., Martin A., Durand I., & Causse R. (2019). Important readjustments in the biomass and distribution of groundfish species in the northern part of the Kerguelen Plateau and Skiff Bank. In Welsford D., Dell J., Duhamel G. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries, proceedings of the Second Symposium. Australian Antarctic Division, Kingston (pp. 135-184).

Duhamel, G., Pruvost, P., Bertignac, M., Gasco, N., & Hautecoeur, M. (2011). Major fishery events in Kerguelen Islands: Notothenia rossii, Champsocephalus gunnari, Dissostichus eleginoides – Current distribution and status of stocks. In Duhamel G., Welsford D. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 275-286).

Duhamel G., & Williams, R. (2011). History of whaling, sealing, fishery and aquaculture trials in the area of the Kerguelen Plateau. In Duhamel G., Welsford D. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 15-28).

Kratz, J., & Strasser, C. (2014). Data publication consensus and controversies. *F1000Research 2014*, 3:9. doi:10.12688/f1000research.3979.3

EUMOFA. (2000). The EU Fish Market. European Union, Maritime affairs and fisheries. Retrieved from https://www.eumofa.eu/documents/20178/415635/EN_The+EU+fish+market_2020.pdf/fe6285bb-5446-ac1a-e213-6fd6f64d0d85?t=1604671147068

FAO. (1999). Guidelines for the routine collection of capture fishery data. Prepared at the FAO/DANIDA Expert Consultation. Bangkok, Thailand, 18-30 May 1998. FAO Fisheries Technical Paper 382. Retrieved from http://www.fao.org/3/x2465e/x2465e.pdf

FAO. (2020). The State of World Fisheries and Aquaculture 2020. doi:10.4060/ca9229en

Fernandes, A. C., Oroszlányová, M., Silva, C., Azevedo, M., & Coelho, R. (2021). Investigating the representativeness of onboard sampling trips and estimation of discards based on clustering. *Fisheries Research, 234*. doi:10.1016/j.fishres.2020.105778

Filonzi, L., Chiesa, S., Vaghi, M., & Nonnis Marzano, F. (2010). Molecular barcoding reveals mislabelling of commercial fish products in Italy. *Food Research International, 43*(5), 1383-1388. doi:10.1016/j.foodres.2010.04.016

FranceAgriMer. (2019). The fisheries and aquaculture sector in France. Retrieved from https://www.franceagrimer.fr/fam/content/download/61053/document/a4-CC%20p%C3%AAche%202019%20eng-CORRIGE.pdf?version=8

Taconet, M., Bensch, A., Balestri, E., Gentile, A. & Prado, J. (2002). Fisheries Inventory: Method and Guidelines. Retrieved from http://www.fao.org/tempref/FI/DOCUMENT/FIGIS_FIRMS/Method_Guidelines/FisheriesInventoryGuidelines.pdf

Freire, K. M. F., Belhabib, D., Espedido, J. C., Hood, L., Kleisner, K. M., Lam, V. W. L., Machado, M. L., Mendonça, J. T., Meeuwig, J. J., Moro, P. S., Motta, F. S., Palomares, M.-L. D., Smith, N., Teh, L., Zeller, D., Zylich, K., & Pauly, D. (2020). Estimating global catches of marine recreational fisheries. *Frontiers in Marine Science, 7*, 12. doi:10.3389/fmars.2020.00012

Froese, R. & Pauly, D. (Eds.) (2000). FishBase 2000: Concepts, design and data sources. ICLARM, Los Baños, Laguna, Philippines. Retrieved from http://pubs.iclarm.net/resource_centre/WF_311.pdf

Gasco, N. (2011). Contributions to marine science by fishery observers in the French EEZ of Kerguelen. In Duhamel G., Welsford D. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 93-98). Retrieved from http://sfi-cybium.fr/fr/contributions-marine-science-fishery-observers-french-eez-kerguelen

Gasco, N., Tixier, P., Delord, K., Guinet, C., & Duhamel, G. (2019). Scientific observation program around Kerguelen Island, a ship-of-opportunity for bird and mammal data collection. In Welsford D., Dell J., Duhamel G. (Eds), The Kerguelen Plateau: marine ecosystem and fisheries, proceedings of the Second Symposium. Australian Antarctic Division, Kingston (pp. 365-379). Retrieved from http://heardisland.antarctica.gov.au/__data/assets/pdf_file/0016/229201/45-Gasco-fullMS.pdf

Gauduchon, T., Cornou, A.-S., Quinio-Scavinner, M., Goascoz, N., Dubroca, L., Billet, N., Rostiaux, E., Miossec, D., & Boiron-Leroy, A. (2020). Captures et rejets des métiers de pêche français. Résultats des observations à bord des navires de pêche professionnelle en 2018. Obsmer, Ifremer. Retrieved from https://archimer.ifremer.fr/doc/00619/73122/72281.pdf

Guerrero, J. I., García, A., Personal, E., Luque, J., & León, C. (2017). Heterogeneous data source integration for smart grid ecosystems based on metadata mining. *Expert Systems with Applications, 79,* 254-268. doi:10.1016/j.eswa.2017.03.007

Guinet, C., Tixier, P., Gasco, N., & Duhamel, G. (2015). Long-term studies of Crozet Island killer whales are fundamental to understanding the economic and demographic consequences of their depredation behaviour on the Patagonian toothfish fishery. *ICES Journal of Marine Science 72*, 1587-1597. Retrieved from https://academic.oup.com/icesjms/article/72/5/1587/761039

Gulland, J. A. & Rosenberg, A. A. (1992). A review of length-based approaches to assessing fish stocks. FAO Fisheries Technical Paper, 323. Retrieved from http://www.fao.org/3/T0535E/T0535E00.htm

Hamm, D. E. (2012). Development and evaluation of a data dictionary to standardize salmonid habitat assessments in the Pacific Northwest. *Fisheries, 37*(1), 6-18. doi:10.1080/03632415.2012.639679

Helyar, S. J., Lloyd, H. ap D., de Bruyn, M., Leake, J., Bennett, N., & Carvalho, G. R. (2014). Fish product mislabelling: Failings of traceability in the production chain and implications for illegal, unreported and unregulated (IUU) fishing. *PLoS ONE, 9*(6), e98691. doi:10.1371/journal.pone.0098691

Holden, M.J. & Raitt, D.F.S. (1974). Manual of fisheries science Part 2 – Methods of Resource Investigation and their Application. Retrieved from http://www.fao.org/3/f0752e/f0752e00.htm

Hiddink, J. G., MacKenzie, B. R., Rijnsdorp, A., Dulvy, N. K., Nielsen, E. E., Bekkevold, D., Heino, M., Lorance, P., & Ojaveer, H. (2008). Importance of fish biodiversity for the management of fisheries and ecosystems. *Fisheries Research, 90*(1-3), 6-8. doi:10.1016/j.fishres.2007.11.025

Higgins, S. (2011). Digital curation: The emergence of a new discipline. *International Journal of Digital Curation, 6*(2), 78-88. doi:10.2218/ijdc.v6i2.191

Hill, N., Foster, S., Duhamel, G., Welsford, D., Koubbi, P., & Johnson, C. (2017). Model-based mapping of assemblages for ecology and conservation management: A case study of demersal fish on the Kerguelen Plateau. *Diversity and Distributions, 23*, 1216-1230. doi:10.1111/ddi.12613

Hughes, L. M., Bao, J., Hu, Z.-L., Honavar, V., & Reecy, J. M. (2008). Animal trait ontology: The importance and usefulness of a unified trait vocabulary for animal species. *Journal of Animal Science, 86*(6), 1485-1491. doi:10.2527/jas.2008-0930

Hureau, J. C. (2011). Marine Research on the Kerguelen Plateau: from early scientific expeditions to current surveys under the CCAMLR objectives. In Duhamel G., Welsford D. (Eds) The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 5-13). Retrieved from http://sfi-cybium.fr/fr/marine-research-kerguelen-plateau-early-scientific-expeditions-current-surveys-under-ccamlr

ICES. (2018). Report of the Working Group on Spatial Fisheries Data (WGSFD), 11–15 June 2018, Aberdeen, Scotland, UK. Retrieved from https://archimer.ifremer.fr/doc/00586/69789/67681.pdf

Ide, N., & Pustejovsky, J. (2010). What Does Interoperability Mean, Anyway? Toward an Operational Definition of Interoperability for Language Technology. Retrieved from https://www.cs.vassar.edu/~ide/papers/ICGL10.pdf

Iglésias, S. P., Toulhoat, L., & Sellos, D. Y. (2010). Taxonomic confusion and market mislabelling of threatened skates: Important consequences for their conservation status. *Aquatic Conservation: Marine and Freshwater Ecosystems, 20*(3), 319-333. doi:10.1002/aqc.1083

Plazaola, J. de, & Rignols, E. (2019). Tableaux de l'économie française. Institut national de la statistique et des études économiques. Retrieved from https://www.insee.fr/fr/statistiques/3696937

Kosmala, M., Wiggins, A., Swanson, A., & Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology and the Environment, 14*(10), 551-560. doi:10.1002/fee.1436

Kupschus, S., Schratzberger, M., & Righton, D. (2016). Practical implementation of ecosystem monitoring for the ecosystem approach to management. *Journal of Applied Ecology, 53*(4), 1236-1247. doi:10.1111/1365-2664.12648

Leblond, E., Daures, F., Berthou, P., & Dintheer, C. (2008). The Fisheries Information System of Ifremer: A multidisciplinary monitoring network and an integrated approach for the assessment of French fisheries, including small-scale fisheries. ICES 2008 Annual Science Conference, 22-26 september 2008, Halifax, Canada. Retrieved from https://archimer.ifremer.fr/doc/00059/17002/

Lefebvre Saint-Felix, C. & Maghin, P. (2019). Rapport sur la gestion de la pêche de légine dans les Terres australes et antarctiques françaises. Contrôle général économique et financier N°181058. Retrieved from https://taaf.fr/content/uploads/2019/11/Rapport-CGefi-l%C3%A9gine-envoy%C3%A9-le-11.02.19.2.pdf

Losiewicz, PB., Oard, D.W., & Kostoff, R. (2004). Textual data mining to support science and technology management. *Journal of Intelligent Information Systems.* doi:10.1023/A:1008777222412

Mannocci, L., Forget, F., Tolotti, M. T., Bach, P., Bez, N., Demarcq, H., Kaplan, D., Sabarros, P., Simier, M., Capello, M., & Dagorn, L. (2020). Predicting bycatch hotspots in tropical tuna purse seine fisheries at the basin scale. *Global Ecology and Conservation, 24*, e01393. doi:10.1016/j.gecco.2020.e01393

Martin, A., Trouslard, E., Hautecoeur, M., Blettery, J., Moreau, C., Saucède, T., Améziane, N., Duhamel, G. & Eléaume, M. (2019). Ecoregionalisation and conservation of benthic communities in the French exclusive economic zone of Kerguelen. In Welsford D., Dell J., Duhamel G. (Eds) The Kerguelen Plateau: marine ecosystem and fisheries, proceedings of the Second Symposium. Australian Antarctic Division, Kingston (pp. 279-303). Retrieved from http://heardisland.antarctica.gov.au/__data/assets/pdf_file/0009/229158/34-Martin-FullMS.pdf

Marx, V. (2013). The big challenges of big data. *Nature, 498*(7453), 255-260. doi:10.1038/498255a

McCluskey, S.M., & Lewison R.L. (2008). Quantifying fishing effort: A synthesis of current methods and their applications. *Fish and Fisheries 9*, 188-200. Retrieved from https://onlinelibrary.wiley.com/doi/10.1111/j.1467-2979.2008.00283.x

McLure, M., Level, A. V., Cranston, C. L., Oehlerts, B., & Culbertson, M. (2014). Data curation: A study of researcher practices and needs. *Portal: Libraries and the Academy, 14*(2), 139-164. doi:10.1353/pla.2014.0009

Michael, P. E., Wilcox, C., Barbraud, C., Delord, K., de Grissac, S., & Weimerskirch, H. (2019). Future resilience scaled by surveillance: High sensitivity of albatross to regional illegal, unreported and unregulated fishing effort. In Welsford D., Dell J., Duhamel G. (Eds) The Kerguelen Plateau: marine ecosystem and fisheries, proceedings of the Second Symposium. Australian Antarctic Division, Kingston (pp. 279-303). Retrieved from http://heardisland.antarctica.gov.au/__data/assets/pdf_file/0006/229164/40-Michael-FullMS.pdf

Ministry for Primary Industries. (2019). Aquatic Environment and Biodiversity Annual Review 2018. Compiled by the Fisheries Science Team, Ministry for Primary Industries, Wellington New Zealand. Retrieved from http://docs.niwa.co.nz/library/public/AEBAR-2018.pdf

Mormede, S., Dunn, A. & Pinkerton, M. (2011). Generic modelling tools for the assessment of marine populations. In Duhamel G., Welsford D. (Eds) The Kerguelen Plateau: Marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris, pp 257-262. Retrieved from http://sfi-cybium.fr/fr/generic-modelling-tools-assessment-marine-populations

Noonan, D. & Chute, T. (2014). Data curation and the university archives. *The American Archivist 77*(1), 201–240. doi:10.17723/aarc.77.1.m49r46526847g587

Patterson, J. C., Beech, V. D., Fitzhugh, G. R., Colson, P. M., Noble, B. T. & Willett, N. A. (2019). North Pacific Observer Program 2018 Annual Report. Retrieved from https://repository.library.noaa.gov/view/noaa/20199

Okuda, T. & Massiot-Granier, F. (2019). Revised CASAL model for D. eleginoides with updated biological parameters at Division 58.4.4b. CCAMLR Working Group Working Group on Statistics, Assessments and Modelling. Retrieved from https://www.ccamlr.org/en/wg-sam-2019/30

Pauly, D., & Zeller, D. (2016). Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nature Communications, 7*(1), 10244. doi:10.1038/ncomms10244

Peer, L., & Green, A. (2012). Building an open data repository for a specialized research community: Process, challenges and lessons. *International Journal of Digital Curation, 7*(1), 151-162. doi:10.2218/ijdc.v7i1.222

Mackinson, S., Martin-Gonzalez, G., Balestri, E., Coull, K., Clarke, E. & Marshall, C.T. (2018). Feasibility study into a scientific self-sampling programme for the pelagic sector. Fisheries Innovation Scotland project. Retrieved from https://scottishpelagic.co.uk/wp-content/uploads/2019/01/Pelagic-self-sampling_FIS020-report_FINAL.pdf

Péron, C., Welsford, D., Ziegler, P., Lamb, T., Gasco, N., Chazeau, C., Sinègre, R. & Duhamel, G. (2016). Modelling spatial distribution of Patagonian toothfish through life-stages and sex and its implications for the fishery on the Kerguelen Plateau. *Progress in Oceanography 141*, 81-95. doi:10.1016/j.pocean.2015.12.003

Péron C., Yates P., Maschette D., Chazeau C., Ziegler .P, Welsford D., Gasco N., Duhamel G. (2018). Report on fish by-catch on exploratory fishing in Divisions 58.4.1 and 58.4.2. CCAMLR Working Group on Fish Stock Assessment. Retrieved from https://www.ccamlr.org/en/wg-fsa-18/28

Pramod, G., Nakamura, K., Pitcher, T. J., & Delagran, L. (2014). Estimates of illegal and unreported fish in seafood imports to the USA. Marine Policy, 48, 102-113. doi:10.1016/j.marpol.2014.03.019

Pruvost, P., Duhamel, G., Gasco, N. & Palomares, M. L. D. (2015a). A short history of the fisheries of Crozet Islands. In: Palomares, M., Pauly D. (Eds), Marine Fisheries Catches of SubAntarctic Islands, 1950-2010 (pp. 31-38).

Pruvost, P., Duhamel, G., Manach, F. L., Palomares, M. (2015b). La pêche aux îles Saint-Paul et Amsterdam. In: Palomares, MLD., Pauly D. (Eds), Marine Fisheries Catches of SubAntarctic Islands, 1950-2010 (pp. 39-48).

Pruvost, P., Martin, A., & Denys, G. (2011). SIMPA – A tool for fisheries management and ecosystem modeling. In Duhamel G., Welsford D. (Eds) The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 3-270). Retrieved from http://sfi-cybium.fr/fr/simpa-%E2%80%93-tool-fisheries-management-and-ecosystem-modeling

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Ratnasingham, S., & Hebert, P.D.N. (2007). Bold: The barcode of life data system. *Molecular Ecology Notes 7*(3), 355-364. doi:10.1111/j.1471-8286.2007.01678.x

Romagnoni, G., Mackinson, S., Hong, J. & Eikeset, A. M. (2015). The Ecospace model applied to the North Sea: Evaluating spatial predictions with fish biomass and fishing effort data. *Ecological Modelling 300*, 50-60. doi:10.1016/j.ecolmodel.2014.12.016

Rosenberg, A. A., Bolster, W. J., Alexander, K. E., Leavenworth W. B., Cooper, A. B. & McKenzie, M. G. (2005). The history of ocean resources: Modeling cod biomass using historical records. *Frontiers in Ecology and the Environment 3*(2), 78-84. doi:10.1890/1540-9295(2005)003[0078:THOORM]2.0.CO;2

Sampson, D. B. (2011). The accuracy of self-reported fisheries data: Oregon trawl logbook fishing locations and retained catches. *Fisheries Research, 112*(1-2), 59-76. doi:10.1016/j.fishres.2011.08.012

Schmidt, S., Maudire, G., Nys, C., Sudre, J., Harscoat, V., Dibarboure, G., & Huynh, F. (2020). Streamlining data and service centers for easier access to data and analytical services: The strategy of ODATIS as the gateway to French marine data. *Frontiers in Marine Science, 7*, 548126. doi:10.3389/fmars.2020.548126

Shamsfard, M., & Barforoush, A. A. (2004). Learning ontologies from natural language texts. *International Journal of Human-Computer Studies, 60*(1), 17-63. doi:10.1016/j.ijhcs.2003.08.001

Silburn, J., Johnson, D. D., Booth, D. J., & Taylor, M. D. (2020). The effect of subsampling when monitoring bycatch in a penaeid trawl fishery. *Fisheries Research, 224*, 105459. doi:10.1016/j.fishres.2019.105459

SIOFA. (2006). Final act of the Conference on the Southern Indian Ocean Fisheries. Retrieved from https://www.apsoi.org/sites/default/files/documents/SIOFA%20AGREEMENT_EN.pdf

SIOFA. (2019). Templates for data submission. Retrieved from http://apsoi.org/sites/default/files/documents/meetings/SC-04-09%20Templates%20for%20data%20submission.pdf

Specht, A., Bolton, M., Kingsford, B., Specht, R., & Belbin, L. (2018). A story of data won, data lost and data re-found: The realities of ecological data preservation. *Biodiversity Data Journal, 6*, e28073. doi:10.3897/BDJ.6.e28073

Staples, D. (2009). Ecosystem approach to fisheries and aquaculture: Implementing the FAO code of conduct for responsible fisheries. Food and Agriculture Organization of the United States, Regional Office for Asia and the Pacific. Retrieved from http://www.fao.org/3/i0964e/i0964e00.pdf

Straube, N., Duhamel, G., GaSco, N., & Kriwet, J. (2011). Description of a new deep-sea Lantern Shark Etmopterus viator sp. Nov. (Squaliformes : Etmopteridae) from the Southern Hemisphere. In Duhamel G., Welsford D. (Eds) The Kerguelen Plateau: Marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 137-150). Retrieved from http://sfi-cybium.fr/en/node/2934

Suuronen, P., & Gilman, E. (2020). Monitoring and managing fisheries discards: New technologies and approaches. *Marine Policy, 116*, 103554. doi:10.1016/j.marpol.2019.103554

Taconet, P., Chassot, E., Blondel, E., & Barde, J. (2017). Global datasets for tuna fisheries. IOTC-2017-WPDCS13-32. Retrieved from https://www.researchgate.net/publication/324218072_Global_datasets_for_tunas_fisheries

Tessarolo, G., Ladle, R., Rangel, T., & Hortal, J. (2017). Temporal degradation of data limits biodiversity research. *Ecology and Evolution, 7*(17), 6863-6870. doi:10.1002/ece3.3259

Tixier P., Gasco N., Duhamel, G. & Guinet, C. (2016). Depredation of Patagonian toothfish (Dissostichus eleginoides) by two sympatrically occurring killer whale (Orcinus orca) ecotypes: Insights on the behavior of the rarely observed type D killer whales. *Marine Mammal Science 32*, 983-1003. doi:10.1111/mms.12307

Tixier, P., Welsford, D., Lea, M.-A., Hindell, M. A., Guinet, C., Janc, A., Richard, G., Gasco, N., Duhamel, G., Arangio, R., Villanueva, C., Suberg, L., & Arnould, J. P. Y. (2019). Fisheries interaction data suggest variations in the distribution of sperm whales on the Kerguelen Plateau. In Welsford D., Dell J., Duhamel G. (Eds) The Kerguelen Plateau: Marine ecosystem and fisheries, proceedings of the Second Symposium. Australian Antarctic Division, Kingston (pp. 259-27). Retrieved from https://archimer.ifremer.fr/doc/00502/61411/

National Marine Fisheries Service. (2020). National Observer Program FY 2018 Annual Report. NOAA Tech. Memo. NMFS-F/SPO-206. Retrieved from https://spo.nmfs.noaa.gov/sites/default/files/TMSPO206.pdf

Tomaszkiewicz, M., Hautecoeur, M., Bonillo, C., Dettaï, A., Mazzei, F., Ghigliotti, L., Pisano, E., Couloux, A., Chanet, B., Lecointre, G., Detrich, H., Duhamel, G., & ozouF-cosTaz, C. (2011). Comparative cytogenetic studies of the Nototheniidae (Teleostei: Acanthomorpha) from the Indian (Kerguelen-Heard Plateau) and Atlantic (South Georgia, South Sandwich, Falkland/Malvinas, Bouvet Islands) sectors of the Southern Ocean. In Duhamel G., Welsford D. (Eds) The Kerguelen Plateau: marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 109-120). Retrieved from https://glecointre.mnhn.fr/docs/090_Tomaszkeiwicz%202011.pdf

Tromeur, E., & Doyen, L. (2016). Optimal biodiversity erosion in multispecies fisheries. *Cahiers du GREThA 2016*(20). Retrieved from http://cahiersdugretha.u-bordeaux.fr/2016/2016-20.pdf

Van Iseghem, S., Quillérou, E., Brigaudeau, C., Macher, C., Guyader, O., & Daurès, F. (2011). Ensuring representative economic data: Survey data-collection methods in France for implementing the Common Fisheries Policy. *ICES Journal of Marine Science, 68*(8), 1792-1799. doi:10.1093/icesjms/fsr112

Vandepitte, L., Bosch, S., Tyberghein, L., Waumans, F., Vanhoorne, B., Hernandez, F., De Clerck, O., & Mees, J. (2015). Fishing for data and sorting the catch: Assessing the data quality, completeness and fitness for use of data in marine biogeographic databases. *Database, 2015*. doi:10.1093/database/bau125

Vandepitte, L., Vanhoorne, B., Decock, W., Vranken, S., Lanssens, T., Dekeyzer, S., Verfaille, K., Horton, T., Kroh, A., Hernandez, F., & Mees, J. (2018). A decade of the World Register of Marine Species – General insights and experiences from the Data Management Team: Where are we, what have we learned and how can we continue? *PLoS ONE, 13*(4), e0194599. doi:10.1371/journal.pone.0194599

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics, 34*(2), 77-84. doi:10.1111/jbl.12010

Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS ONE, 8*(7), e67332. doi:10.1371/journal.pone.0067332

Watson, R. A. (2017). A database of global marine commercial, small-scale, illegal and unreported fisheries catch 1950–2014. *Scientific Data, 4*(1), 170039. doi:10.1038/sdata.2017.39

Welsford, D., Candy, S., Lamb, T., Nowara, G., Constable, A., & Williams, R. (2011). Habitat use by Patagonian toothfish (Dissostichus eleginoides Smitt 1898) on the Kerguelen Plateau around Heard Island and the McDonald Islands. In Duhamel G., Welsford D. (Eds) The Kerguelen Plateau: Marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris (pp. 125-136). Retrieved from http://sfi-cybium.fr/fr/habitat-use-patagonian-toothfish-dissostichus-eleginoides-smitt-1898-kerguelen-plateau-around-heard

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin Core: An evolving community-developed biodiversity data standard. *PLoS ONE, 7*(1), e29715. doi:10.1371/journal.pone.0029715

Wilkinson, M. D., Dumontier, et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*(1), 160018. doi:10.1038/sdata.2016.18

Zimmerman, A. (2007). Not by metadata alone: The use of diverse forms of knowledge to locate data for reuse. *International Journal on Digital Libraries, 7*(1-2), 5-16. doi:10.1007/s00799-007-0015-8

Zimmerman, A. S. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology, & Human Values, 33*(5), 631-652. doi:10.1177/0162243907306704