



**HAL**  
open science

## **Taxonomy 2.0: computer-aided identification tools to assist Antarctic biologists in the field and in the laboratory**

Thomas Saucède, Marc Eléaume, Quentin Jossart, Camille Moreau, Rachel Downey, Narissa Bax, Chester Sands, Borja Mercado, Cyril Gallut, Régine Vignes-Lebbe

► **To cite this version:**

Thomas Saucède, Marc Eléaume, Quentin Jossart, Camille Moreau, Rachel Downey, et al.. Taxonomy 2.0: computer-aided identification tools to assist Antarctic biologists in the field and in the laboratory. *Antarctic Science*, 2021, 33 (1), pp.39-51. 10.1017/S0954102020000462 . hal-03090197

**HAL Id: hal-03090197**

**<https://hal.science/hal-03090197>**

Submitted on 7 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Taxonomy 2.0: computer-aided identification tools to assist Antarctic biologists in the field and in the lab

Journal:	<i>Antarctic Science</i>
Manuscript ID	Draft
Manuscript Type:	Synthesis Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Saucède, Thomas; University of Burgundy, Biogeosciences Eléaume, Marc; Muséum National d'Histoire Naturelle Vignes-Lebbe, Régine; Muséum National d'Histoire Naturelle Jossart, Quentin; ULB Université Libre de Bruxelles Moreau, Camille; ULB Université Libre de Bruxelles Downey, Rachel; Australian National University Bax, Narissa; University of Tasmania Sands, Chester; British Antarctic Survey Mercado, Borja; University of Valencia Gallut, Cyril; Muséum National d'Histoire Naturelle
Keywords:	Data base, field guide, knowledge base, software, taxonomic key, Xper3
Abstract:	Species inventories are essential to the implementation of conservation policies to mitigate biodiversity loss, maintain ecosystem services, and their value to the society. This is particularly topical with respect to climate change and direct anthropogenic effects on Antarctic biodiversity, the identification of the most at risk taxa and geographic areas becoming a priority. Identification tools are often neglected and considered helpful only for taxonomists. However, the development of new online information technologies and computer-aided identification tools provide an opportunity to enlarge them to a wider audience, also considering the emerging generation of scientists who apply an integrative approach to taxonomy. This paper aims to clarify essential concepts and provide convenient and accessible tools, tips and suggested systems to use and develop Knowledge Bases (KB). The software Xper3 was selected as an example of user-friendly KB management system to give a general overview of existing tools and functionalities through two applications, the "Antarctic Echinoids" and "Odontasteridae Southern Ocean (Asteroids)" KB. We highlight the advantages provided by KB over more classical tools, and future potential uses are given including the production of field guides to aid in the compilation of species inventories for biodiversity conservation purposes.

SCHOLARONE™  
Manuscripts

**Taxonomy 2.0: computer-aided identification tools to assist Antarctic biologists in  
the field and in the lab**

THOMAS SAUCÈDE<sup>1</sup>, MARC ELÉAUME<sup>2</sup>, RÉGINE VIGNES-LEBBE<sup>2</sup>, QUENTIN  
JOSSART<sup>3</sup>, CAMILLE MOREAU<sup>1,3</sup>, RACHEL DOWNEY<sup>4</sup>, NARISSA BAX<sup>5</sup>,  
CHESTER SANDS<sup>6</sup>, BORJA MERCADO<sup>7</sup> AND CYRIL GALLUT<sup>2</sup>

<sup>1</sup> *Biogéosciences, UMR 6282, CNRS, EPHE, Université Bourgogne Franche-Comté, 6  
boulevard Gabriel, 21000 Dijon, France*

<sup>2</sup> *Institut de Systématique, Evolution, Biodiversité (ISYEB UMR 7205), Muséum national  
d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 57 rue  
Cuvier, 75005 Paris, France*

<sup>3</sup> *Marine Biology Lab, CP160/15 Université Libre de Bruxelles (ULB) 50, Av. F.D.  
Roosevelt, B-1050 Brussels, Belgium*

<sup>4</sup> *Fenner School of Environment & Society, Australian National University, Linnaeus  
Way, Acton 2601, Canberra, ACT, Australia*

<sup>5</sup> *Institute for Marine and Antarctic Studies, University of Tasmania & Centre for Marine  
Socioecology, Hobart, TAS, Australia*

<sup>6</sup> *British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, UK*

<sup>7</sup> *Biodiversity and Cnidarians Evolution Group, Cavanilles Institute of Biodiversity and  
Evolutionary Biology (ICBiBE), University of Valencia, 22085, 46071 Valencia, Spain*

*thomas.saucede@u-bourgogne.fr*

**Abstract:**

Species inventories are essential to the implementation of conservation policies to mitigate biodiversity loss, maintain ecosystem services, and their value to the society. This is particularly topical with respect to climate change and direct anthropogenic effects on Antarctic biodiversity, the identification of the most at risk taxa and geographic areas becoming a priority. Identification tools are often neglected and considered helpful only for taxonomists. However, the development of new online information technologies and computer-aided identification tools provide an opportunity to enlarge them to a wider audience, also considering the emerging generation of scientists who apply an integrative approach to taxonomy. This paper aims to clarify essential concepts and provide convenient and accessible tools, tips and suggested systems to use and develop Knowledge Bases (KB). The software Xper3 was selected as an example of user-friendly KB management system to give a general overview of existing tools and functionalities through two applications, the "Antarctic Echinoids" and "Odontasteridae Southern Ocean (Asteroids)" KB. We highlight the advantages provided by KB over more classical tools, and future potential uses are given including the production of field guides to aid in the compilation of species inventories for biodiversity conservation purposes.

**Key words (6 max):** data base, field guides, knowledge base, software, taxonomic key,

Xper3

## Introduction

Species inventories are essential to the implementation of conservation policies to mitigate biodiversity loss, maintain ecosystem services, and their value to the society (Balmford & Gaston 1999, Gaston 2005, May 2011): the prevention of species declines or invasions (Chapman 2005); and forming the basis of management planning regimes (e.g., species recovery plans). This growing importance of conservation issues has increased the need for accurate identification of endangered, migratory, and invasive species (Balmford & Gaston 1999).

Considering the fast pace of climate and oceanographic changes already affecting (Turner *et al.* 2014) or predicted (Allan *et al.* 2013, Gutt *et al.* 2015) to affect the Southern Ocean and its marine life, the identification of the most at risk taxa and geographic areas has become a priority (Griffiths 2010). Taxon inventories form an essential basis for studying the multiple abiotic and biotic stressors on species and ecosystems (Pendlebury & Barnes-Keoghan 2007, Kargel *et al.* 2014, Molinos *et al.* 2015, Schram *et al.* 2016, Byrne *et al.* 2016). In recognition of the need for an improved understanding of species and community responses to environmental change, many key questions in the Scientific Committee for Antarctic Research (SCAR) ‘Antarctic life on the precipice’, cover topics relating to the inclusion of biodiversity research for the 2020 Science Horizon Scan (Kennicutt *et al.* 2015). Coordinating biodiversity collections, and increasing access to taxonomic and genetic identification resources, are key strategic

1  
2  
3 aims.

4  
5 Over the last decade, many efforts have been made to improve our knowledge of  
6 Antarctic marine life, including the use of new sampling tools for data collection (i.e.,  
7 sampling gears, remote sensing and autonomous vehicles) and the development of  
8 analytical techniques (e.g., -omics approaches and isotope analyses) (Gutt *et al.* 2017;  
9 Kaiser *et al.* 2013). Following the synergistic momentum of the Census of Antarctic  
10 Marine Life (CAML) and of the International Polar Year (IPY), the inventory of  
11 Antarctic marine species has been significantly enlarged (Schiaparelli *et al.* 2013),  
12 highlighting both the uniqueness and high level of Antarctic biota diversity (Clarke &  
13 Johnston 2003, Brandt *et al.* 2007, Cary *et al.* 2010, Kennicutt *et al.* 2015, De Broyer *et*  
14 *al.* 2014). Newly collected data are now compiled in information networks and initiatives  
15 such as the Barcode of Life Data System (BOLD at <http://www.boldsystems.org/>), the  
16 SCAR-Marine Biodiversity Information Network (SCAR-MarBIN at  
17 <https://www.scarmarbin.be/>, Griffiths *et al.* 2011), the Register of Antarctic Marine  
18 Species (RAMS, <http://www.marinespecies.org/rams/>, Jossart *et al.* 2015; De Broyer *et*  
19 *al.* 2020) and Biodiversity.aq (Van de Putte *et al.* 2020). They were also disseminated  
20 through the publication of the Biogeographic Atlas of the Southern Ocean (De Broyer *et*  
21 *al.* 2014), and a number of other open access databases (Danis *et al.* 2013, Gutt *et al.*  
22 2013, Van de Putte *et al.* 2020). Altogether these efforts have greatly advanced the  
23 documentation of Antarctic marine life (Gutt *et al.* 2011, De Broyer *et al.* 2011, 2014).  
24 However, many biodiversity collections in museums and scientific institutes remain  
25 unsorted, unidentified and thus far, unsearchable across these platforms, often due a  
26 deficit of funding towards taxonomic research costs.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 Naming species and using taxonomic and biodiversity information is not only a concern  
6 of taxonomists. It is of importance for many scientists and amateurs in life sciences, as it  
7 underpins any biodiversity study (Chapman 2005). It is central to countless numbers of  
8 publications and various fields such as phylogenetics, wildlife ecology, biogeography,  
9 conservation, natural resource management, biomedical research, and education  
10 (Chapman 2005, Costello *et al.* 2015). These various user-groups require the  
11 development of flexible, searchable, easily accessible and usable DataBases (DB)  
12 following the FAIR (Findable, Accessible, Interoperable, and Re-usable) guiding  
13 principles for scientific data management (Wilkinson *et al.* 2016). These DB also need to  
14 be easily amendable as systematics and taxonomy are regularly being revised and  
15 updated, at least for a number of invertebrate taxa (Griffiths 2010).  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

31 In the scientific literature, species descriptions can be incomplete, patchy and  
32 heterogenous in their methodology, and old-fashioned drawings or images are sometimes  
33 confusing and misleading. The interpretation of this literature and the observation of  
34 collection specimens by experts, are therefore needed to build structured and comparable  
35 taxonomic descriptions. By using one or more DB, a knowledge base (KB), which is an  
36 application that stores and manages the expert's knowledge for describing, classifying  
37 and identifying organisms, is utilised. In KB, the expert's knowledge is stored as formal,  
38 structured and complex information sources used by computer systems. Over the last 20  
39 years, the development of new information technologies has been an opportunity to  
40 develop such tools and generate identification guides and species checklists that can be  
41 used in distribution information systems (Chavan & Penev 2011).  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6 Decision criteria (*i.e.* diagnostic characters) are required in order to attribute a taxon  
7  
8 name to a specimen collected in the field or accessed in academic collections. Decision  
9  
10 criteria can be based on a wide variety of information sources and types, including  
11  
12 morphological and ecological (*e.g.*, sound, behaviour, ...) features when identifying  
13  
14 specimens in the field. Other features, like molecular data are of importance for post-field  
15  
16 identifications (Teletchea 2010). It is important to clarify that formal taxonomic  
17  
18 identification differs greatly from laymans' identification, even when the layman is an  
19  
20 experienced field researcher. Unfortunately, it is often difficult to distinguish between the  
21  
22 two in publications unless clear taxonomic methodology, involving relevant specialist  
23  
24 literature citations, are used. Identifications based on field guides are a useful starting  
25  
26 point but they should not be used in ecological analyses, unless verified by a taxonomic  
27  
28 expert.  
29  
30  
31  
32

33 Diagnostic identification keys are the primary form of taxonomic identification  
34  
35 tool and outline a set of characters applied in a hierarchical system for species  
36  
37 determination. However, non-discrete and non-binary approaches also exist, like  
38  
39 automated identification tools. This computer-aided identification facilitates biodiversity  
40  
41 expertise. Identification keys can be produced manually, but algorithms and programs  
42  
43 can compute keys, and have existed since the 1970s. For example, the DELTA format  
44  
45 (DEscription Language for TAXonomy) or SDD (Structured Descriptive Data) are two  
46  
47 standards used for data exchange by Biodiversity Information Standards (TDWG).  
48  
49 Several programs were based on these formats including software producing  
50  
51 identification keys like Intkey, Lucid, Xper2 or Xper3 (Dallwitz 1980, 2000).  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4  
5 A workshop was organised in July 2017 in Leuven (Belgium) alongside the XIIth SCAR  
6 Biology Symposium. It was open to all scientists, either occasional or confirmed  
7 taxonomists, who have once needed to identify organisms or describe species from  
8 specimens collected in the field or accessed in collections. The objectives of the  
9 workshop were to present useful tools and software that can assist biologists in a wide  
10 range of applications to build and/or use KB. These included species identification, the  
11 formalization of diagnostic morphological characters, and the production of DB,  
12 catalogues, field guides, and monographs.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

24 The main outcomes of the workshop are synthesized in the present paper. The aim  
25 is not to make a comprehensive survey of all existing tools and possibilities, but to clarify  
26 essential concepts and provide Antarctic biologists with convenient and accessible tools,  
27 tips and suggested systems for them to follow up with the development of their own KB.  
28 The software Xper3 (Vignes-Lebbe *et al.* 2016) was chosen as an example of user-  
29 friendly knowledge base management system to illustrate these concepts and tools. A  
30 general overview of this computer-aided identification software is given along with  
31 existing tools and functionalities. Some applications and case studies are presented for  
32 the production and use of KB and future prospects are given including the production of  
33 field guides.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

## 49 **Computer-aided identification and decision criteria**

### 50 *Identification keys*

51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Recent advances in bioinformatics have led to the development of a high number of  
4 integrated electronic resources comprising online versions of data- and knowledge bases  
5 with interactive identification systems and digital imaging (Dallwitz 2000, Chapman  
6 2005, Nimis & Vignes-Lebbe 2010). The key advantages of this approach, over  
7 conventional systems, include no restriction in the order in which characters can be used,  
8 character deletion and changing, error tolerance and localising, uncertainty expressing,  
9 among many other functionalities (Dallwitz 2000). Reviewing all these programs  
10 available online is out of the scope of the present paper and comprehensive comparisons  
11 of main interactive identification programs are already available (see Dallwitz 2000 and  
12 Thiele & Dallwitz 2000 for instance, [https://www.delta-  
13 intkey.com/www/comparison.htm#Programs\\_compared](https://www.delta-intkey.com/www/comparison.htm#Programs_compared)).  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27

28 Taxon identification keys are generally of two kinds (Hagedorn *et al.*, 2010): (1)  
29 Conventional, single access keys (polytomous or dichotomous) have been classically  
30 used in taxonomic books, field guides, treatises, and monographs with one single possible  
31 access and decision criteria that must be followed. Such keys are not very flexible and  
32 cannot be updated easily; (2) Multiple access keys use any character within a set of  
33 possible descriptors and can be easily illustrated, updated and revised. However, such  
34 keys require the development of software-aided tools.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

46 Among many different DB projects, FishBase  
47 (<http://www.fishbase.org/keys/allkeys.php>) is an example of a global biodiversity  
48 information system that proposes different finfish identification tools on the web (Froese  
49 *et al.* 2018). Keys can be selected depending on the geographic area (following the Food  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 and Agriculture Organization of the United Nations (FAO) zonation), and taxon (order or  
4 family) of interest. Identification keys are also available to identify species using  
5 morphometric measurements. Finally, identification by picture is also possible for  
6 selected countries, families, and ecosystems. There is an important variety of  
7 identification keys available in FishBase but most of them are single access polytomous  
8 systems. Once selected, they only allow one single decision direction at a time, which  
9 must be followed.

19 Multiple access identification keys, using software-aided tools, can be found on  
20 the web for identification of various taxa. Many of them are devoted to the terrestrial  
21 realm. For instance, Navikey ([www.navikey.net](http://www.navikey.net), Neubacher & Rambold 2005) is  
22 implemented as a Java embedded application for accessing descriptive data in a DELTA  
23 format with 27 DB available for identification of taxa in botany, mycology, and zoology.  
24 The NorthWest African grasshoppers and locusts website  
25 (<http://acrinwafrica.mnhn.fr/SiteAcricri/Xper.html>) proposes a set of embedded multi-  
26 access identification keys using the software Xper3. It allows the selection of any  
27 character within a set of morphological descriptors for identification of Orthoptera  
28 (Louveaux et al. 2018). Species identification is achieved in several steps by gradual  
29 elimination of species names that do not meet the selection criteria at each step of the  
30 process. A first morphological descriptor and associated character state are selected with  
31 no particular order and the process is repeated until the name of a species is obtained  
32 according to the selection criteria. The e-Monograph of the Caricaceae website  
33 (<http://herbaria.plants.ox.ac.uk/bol/caricaceae/Keys>) (Carvalho *et al.* 2014) or the  
34 interactive key on Hymenoptera (<http://pteromalus.identificationkey.fr/mkey.html> )

1  
2  
3 (Klimmek & Baur 2018) are other examples of multi-access keys using the Xper3  
4 software.  
5  
6

7  
8 Multiple access identification keys of marine species can be found online at  
9  
10 <http://www.marinespecies.org/aphia.php?p=idkeys> as taxonomic facilities associated to  
11 the World Register of Marine Species (WoRMS editorial board 2020). These keys can be  
12 generated for any taxonomic group, at any chosen level, and allow the user to start  
13 identifying an organism by selecting any obvious character, then set the remaining  
14 characters relevant to the identification, thereby narrowing down to the final  
15 identification. The Mascarene's corals website is another example of marine species  
16 identification key providing computer-based assistance for the description, classification  
17 and identification of scleractinian corals from the SW Indian Ocean. Using a collection of  
18 ca. 3,000 specimens belonging to 185 species and 58 genera, two software options were  
19 proposed with different functionalities: Xper2 (Ung *et al.* 2010, see below) and IKBS  
20 (Iterative Knowledge Base System, Conruyt *et al.* 1998), distributed via a CD and a  
21 booklet ([http://coraux.univ-reunion.fr/IMG/pdf/Plaquette-Coraux\\_des\\_Mascareignes-  
22 mai08.pdf](http://coraux.univ-reunion.fr/IMG/pdf/Plaquette-Coraux_des_Mascareignes-mai08.pdf)).  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

#### 42 *Automated identification* 43 44 45 46

47 Recent advances in biodiversity identification tools have also led to the development of  
48 automated identification. Such tools allow experts to compile huge sets of data that leads  
49 to significant improvements in our understanding of biodiversity. Automated  
50 identification tools are based on the computerized recognition of clusters, ordination  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 patterns or artificial neural networks (Kumar *et al.* 2012; Hu *et al.* 2018; Joly *et al.* 2018).  
4

5         Pl@ntNet (<https://plantnet.org/en/>) is one of the most used systems proposing the  
6 automatic identification of plants through images. It is a citizen science project available  
7 as a smartphone application or as a web version. “By comparing visual patterns  
8 transmitted by users via photos of plant organs”, the Pl@ntNet system seeks to identify  
9 species. The images are analyzed and compared to an image bank produced  
10 collaboratively and enriched daily. The system then offers a possible list of species with  
11 illustrations (Joly *et al.* 2014). iNaturalist (<https://www.inaturalist.org>) also offers a web  
12 platform for citizen science project and collaborative identification by the participants.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

26 Automated identification tools were also developed for marine organisms based on  
27 different descriptor types. Some are based on morphological descriptors, such as a tool  
28 implemented for shark identification based on their fin shape (Hughes & Burghardt  
29 2016). However, automated identification tools are not restricted to morphological  
30 descriptors. The development of algorithms for plankton recognition, and classification  
31 on images using deep machine learning methods such as deep convolutional neural  
32 networks (Kuang 2015, Bueno *et al.* 2017); plankton scanning methods were also  
33 developed that rely on molecular screening tools using -omics approaches: proteomics,  
34 metabolomics, genomics, and metagenomics. The methods that have been developed for  
35 plankton scanning were accompanied by important scientific outputs and received a lot of  
36 media coverage. Several research programs were undertaken on the zooplankton  
37 automated identification (Gasparini *et al.* 2007, Gorsky *et al.* 2010) and the genomic  
38 monitoring of the plankton during expeditions of the RV Tara in 2009 and 2013 (Pesant  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 *et al.* 2015) is one of the most famous examples of the performance of such tools.  
4  
5  
6

## 7 **Knowledge Bases (KB)**

### 8 *The need of building KB*

9  
10  
11  
12  
13  
14 Taxonomic expertise is a rare commodity that is usually not well valued and passed on  
15 from generation to generation. KB allows for the storage, representation and preservation  
16 of the knowledge of recognized specialists in a specific field of interest as formal,  
17 structured and complex information sources. It allows for the management, manipulation,  
18 comparison, processing and transmission of knowledge. Based on a descriptive model to  
19 standardize the terminology and a DB to store the data such as comparable taxonomic  
20 descriptions, KB can use machine learning techniques and generate additional  
21 knowledge, such as diagnoses (by computing best discriminant descriptors) and decision  
22 trees for identification.  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34

35 Multi-access keys using computers as proposed by Xper3 (Vignes *et al.* 2016),  
36 Navikey (Neubacher & Rambold 2005), Lucid [<https://www.lucidcentral.org>] and other  
37 programs (Dallwitz 2000) require an algorithm to automatically analyze and compare  
38 taxonomic descriptions. New descriptions produced by taxon specialists, result in new  
39 observations and measurements, that can then be easily added to previous descriptions. In  
40 practical terms, a KB can be stored in database management systems (Xper3 used  
41 MySQL) or XML files (data exchange standards like SDD) or text files (Delta files) (see:  
42 [https://gfbio.biowikifarm.net/wiki/Data\\_exchange\\_standards\\_protocols\\_and\\_formats\\_relevant\\_for\\_the\\_collection\\_data\\_domain\\_within\\_the\\_GFBio\\_network](https://gfbio.biowikifarm.net/wiki/Data_exchange_standards_protocols_and_formats_relevant_for_the_collection_data_domain_within_the_GFBio_network)).  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 Building and using KB can provide many facilities to address taxonomic issues compared  
6 to more classical tools. (i) For instance, identification books need to be occasionally  
7 updated because taxonomic descriptions change with taxonomy, new collections, and our  
8 knowledge of species biogeography. Revising, editing, producing, and publishing new  
9 book versions is time- and energy-consuming while updates can be regularly and readily  
10 done in KB. (ii) Some identification books may not be “user friendly”, leading to species  
11 misidentifications. Using KB, complicated identification keys can be substantially and  
12 regularly improved based on users’ feedback. Furthermore a software like Xper3 can  
13 explain identification results and display the discordant character states of eliminated  
14 taxa. (iii) The spreading of taxonomic information in the formal scientific literature can  
15 be patchy and data organized either geographically or taxonomically, which is a serious  
16 limitation to species identification, especially as clear identification keys are still needed  
17 for most taxa, and especially those from the deep sea and polar regions. Scattered  
18 taxonomic data can be easily merged and combined in KB. (iv) Lastly, there are  
19 paradigm shifts in the way in which we undertake taxonomic research. This is related to  
20 recent advances in new technologies and DNA sequencing techniques that can make us  
21 question traditional, morphology-based taxonomy, and look for better ways in which to  
22 identify specimens. This process is part of this general shift in taxonomy to a data  
23 intensive, expert-led, logically constructed, unambiguous taxonomic ID process,  
24 meaning that non-experts can be involved in the global categorizing of the living world.  
25 Near global access to computers, the internet, and collaborative scientific platforms,  
26 enables taxonomists and data-scientists to easily and regularly update KB.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### *Building a KB*

Building a KB includes several steps that are synthesized as a flow chart below (Fig. 1).

(1) The metadata. The first step consists in defining the metadata that is, the content, limits and usages, which means the taxonomic coverage, geographic area of interest, phenological states concerned by the key, and the public and prerequisite of the knowledge.

(2) Items. The objects to be described must be listed as items. Most often these items are taxa, but sometimes they can correspond to groups (e.g. for the crowdsourcing project on pollinisators, Spipoll key includes morphological groups when it is too difficult to identify at the species level, see <http://spipoll.snv.jussieu.fr/mkey/mkey-spipoll.html>).

(3) The descriptive model. Character description must be formalized along with the corresponding terminology (descriptors/characters) using a glossary to describe the taxa. This means that the characters, the character states in case of categorical characters, and the unit for numerical characters must be defined. Producing descriptive models and rules are a crucial point in KB development, in order to obtain comparable and intelligible descriptions and easy keys to use.

(4) Character dependency and weights. Consistency rules must be edited, which means the knowledge explanation when a character is relevant. Characters can be weighted to take into account observation difficulties, or to influence the key algorithm by suggesting an order in the identification process. In this way, the KB is not limited to a



1  
2  
3 simple matrix, and translate more of the taxonomic expertise.  
4

5 (5). Descriptions. This step corresponds to the items description linking taxa to  
6 descriptors, including polymorphism. A single non-hierarchical identification system can  
7 be generated for all species of a genus or family using a restricted number of characters,  
8 most of which are declined into many character states. However, generating an  
9 identification key for a higher taxonomic level (i.e. order, class) is more ambitious  
10 considering the incompatibility (specificity) of some characters. In such a case, a  
11 hierarchical structure of identification keys is possible by using internal URL links to  
12 connect identification keys from different taxonomic levels.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

23 (6). Additional resources. Additional resources such as comments or images can  
24 be provided. Both overview pictures of specimens and close-ups of some characters can  
25 be added to help the user in the identification process or to illustrate the species with  
26 living specimens exhibiting vivid colours taken in the field.  
27  
28  
29  
30  
31  
32

33  
34  
35 **Figure 1. Processing flowchart to build a Knowledge Base**  
36  
37

38  
39  
40 *Xper3 software as an application tool*  
41  
42

43  
44 Xper3 is a free and portable KB software that provides an easy-to-use interface. It allows  
45 us to check the consistency of existing descriptions, compare descriptions or sets of  
46 descriptions, create and use either single or multi-access keys, and publish identification  
47 keys and taxonomic forms. There have been three different versions of Xper (Forget *et al.*  
48 1986, Ung *et al.* 2010, Vignes *et al.* 2016). Two of them were made available at  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 www.infosyslab.fr. Xper2 was developed in years 2003-2007 and is a one-user local  
4 system. It is available for Windows, Mac Os and Linux environments, with a user  
5 interface in 5 languages (French, English, German, Spanish and Chinese). Xper3 is the  
6 last version available since 2014. It offers a collaborative, multi-users, interface in French  
7 and English without any installation step. The system is directly opened in an internet  
8 navigator, and the KB is stored in a remote MySQL database. An exchange format in  
9 XML (SDD = Structured Descriptive Data, Hagedorn *et al.* 2005) ensures data exchange  
10 with the local version Xper2 (Ung *et al.* 2010). The main differences between the two  
11 versions are listed below (Table 1).  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

26 Currently, Xper3 requires an internet connection, however, when there is no connection  
27 available, which is often the case at sea and in remote field locations, the easiest solution  
28 is to export the KB in SDD format to use it with Xper2, the local version of Xper. As  
29 soon as the internet is reconnected, it is possible to come back to Xper 3, with an easy  
30 switch between the two versions. Another solution is to export the SDD file to use it with  
31 Biodiversicle, software which was developed as a free access key with android mobile  
32 phones. The LIS (define) team is planning other solutions for the future.  
33  
34  
35  
36  
37  
38  
39  
40  
41

42 Future developments are in progress that will modify the Xper architecture in  
43 order to keep a single core code and different user interfaces. This will allow the same  
44 functionalities in both local and remote access situations. Xper3 currently offers an  
45 approachable way to define terminology for a taxonomic context. A prototype to export  
46 descriptive models in ontological language, OWL, was developed in 2018 (Pellen *et al.*  
47 2018). At the French national level, a link between Xper3 and collections (e-Recolnat  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 project) is in progress, and it will offer the possibility to tag collection images with  
4 controlled vocabularies structured in Xper3 KB (see ANNOTATE on  
5 <https://www.recolnat.org/fr/>).  
6  
7  
8  
9  
10  
11

## 12 **Table 1: Comparison between Xper2 and Xper3**

13  
14  
15  
16

### 17 *Two case studies*

18  
19  
20  
21

22 The following two case studies illustrate an attempt to create a KB and address common  
23 practical issues encountered when identifying species in the field or in the laboratory.  
24 Firstly, for most taxa in the Southern Ocean, taxonomic and distribution data from the  
25 literature are often patchy, which stresses the need for identification keys. In addition,  
26 when identification books are finally released, they do not include our current knowledge  
27 on the systematics and biogeography, and this data needs to be updated in a better format.  
28 Lastly, experience shows that traditional books are not always “user friendly” and that  
29 species can be misidentified. The two following examples are new KB that were built  
30 with different objectives. The first one, the "Antarctic Echinoidea" KB was created to  
31 update book content and facilitate access to taxonomic and biogeographic data. The  
32 second one, the “Odontasteridae Southern Ocean (Asteroids)” KB, digitized taxonomic  
33 information previously published.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

51 *Transferring a book content to a Knowledge Base: an example among Antarctic*  
52 *Echinoids*  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 In 2005, David *et al.* (2005a) published the book entitled “The Antarctic Echinoidea” as a  
6 special volume of the book series “Synopsis of the Antarctic benthos”. The book is a  
7  
8 synthesis and partial revision of the systematics and distribution of the Antarctic echinoid  
9  
10 species south of 45 °S latitude. The book content was also complemented by a database  
11  
12 entitled “Antarctic Echinoids: an interactive database” (David *et al.* 2005b). The book  
13  
14 started with a general introduction to the Southern Ocean and to the phylogeny and  
15  
16 biogeography of Antarctic echinoids, then followed with practical information on the  
17  
18 morphology and anatomy of echinoids, and included a useful section devoted to  
19  
20 anatomical terminology for echinoids. The systematics of Antarctic echinoids constituted  
21  
22 the main part of the book. This section proposes a system of three embedded  
23  
24 identification keys for the identification of taxa at family, genus, and species level  
25  
26 respectively. The systematic description of genera and species includes a synonymy list, a  
27  
28 diagnosis, a morphological description for each species, and species distribution data  
29  
30 provided on a map. The book has proven useful for non-expert taxonomists and was cited  
31  
32 in several biological studies for the identification of Antarctic echinoids (Linse *et al.* 2008,  
33  
34 Gutt *et al.* 2011, Hardy *et al.* 2011, Moya *et al.* 2012, Lecointre *et al.* 2013, among  
35  
36 others).  
37  
38  
39  
40  
41  
42  
43

44 During the last decade, occurrence data on Antarctic marine species have been  
45  
46 significantly improved following the increased number of field campaigns undertaken in  
47  
48 the Southern Ocean, with many of them were achieved under the CAML and IPY  
49  
50 initiatives (Schiaparelli *et al.* 2013). However, our knowledge of Antarctic echinoid  
51  
52 distribution has been improved substantially since the Antarctic echinoid book was  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 released (Saucède *et al.* 2015b; Fabri-Ruiz *et al.* 2017). Currently, echinoid systematics  
4  
5 are being revised based on new integrated molecular and morphological research (Díaz *et*  
6  
7 *al.* 2011, Saucède *et al.* 2014; Saucède *et al.* 2015a; David *et al.* 2016). An example of  
8  
9 this update is seen in the echinoid genus *Sterechinus* (Díaz *et al.* 2011; Saucède *et al.*  
10  
11 2015a), with now only three entities distinguished, compared to the six nominal species  
12  
13 initially attributed to the genus (David *et al.* 2005a).  
14  
15

16  
17 The content of the Antarctic echinoid book was transferred to a Xper3 KB, which  
18  
19 included echinoid systematics and occurrence data, which has been updated following  
20  
21 recent findings and revisions. Seventy-seven species are described using a descriptive  
22  
23 model based on morphological descriptors that are commonly used for echinoid  
24  
25 systematics (David *et al.* 2005a). In addition, for each species, a synonymy list, pictures  
26  
27 and anatomical drawings, a distribution map and references are given. This Xper3 KB  
28  
29 can be accessed at <http://echinoidea-so.identificationkey.org>. Transferring the book  
30  
31 content to the Xper3 KB proved easy, simple and relatively immediate. Adding pictures,  
32  
33 specimen plates, and figures (occurrence maps) using url links was handy as soon as the  
34  
35 files are ready and available online. Finally, the Xper3 "View Description Matrix" tool  
36  
37 proved particularly useful to check the assigned character states and the consistency of  
38  
39 the KB.  
40  
41  
42  
43  
44  
45  
46

47 **Fig. 2. Recent advancement in our knowledge of echinoid distribution in the**  
48  
49 **Southern Ocean comparing the databases from David *et al.* 2005a (white dots: 2,000**  
50  
51 **records) and Fabri-Ruiz *et al.* 2017 (red dots: 7,100 records).**  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 *Building a Knowledge Base from various sources of information: an example of Antarctic*  
4  
5 *Asteroids*  
6  
7  
8  
9

10 As for many Antarctic taxa, there is no single reference that merges all taxonomic  
11 information in order to easily identify Antarctic asteroid specimens collected in the field  
12 by non-expert taxonomists. Clark and Downey “Starfishes of the Atlantic” (1992),  
13 represents the most complete attempt to date for the identification of asteroids, however,  
14 it is restricted to Atlantic waters (including its southernmost parts, which are adjacent to  
15 the Southern Ocean). To find a more comprehensive work focusing on Antarctic  
16 Asteroids, one must mine through many different sources, including field guides  
17 dedicated to particular areas (ex: Brueggeman 1998 in McMurdo; Schories & Kohlberg  
18 2016 in King George Island), legacy works of original descriptions (ex: Koehler 1906,  
19 Fisher 1940), and contemporary scientific literature (Janosik & Halanych 2010, Mah &  
20 Foltz 2014, Mah *et al.* 2014). In addition to this patchy bibliography, most  
21 aforementioned references do not include actual identification keys. Therefore, the  
22 objective of the Antarctic Asteroids KB was to produce a Xper3-based identification tool  
23 as a unique and useful reference for those who need to identify common species of sea  
24 stars from the Southern Ocean.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 In the present case, a prototype of an identification key was built for the family  
45 Odontasteridae. Odontasteridae represents one of the most commonly encountered  
46 families in the Southern Ocean. In particular, the species *Odontaster validus* is one of the  
47 most studied invertebrates around the entire Antarctic continent and plays a major role in  
48 shallow-water food webs (McClintock 1994). The family was also a good choice for a  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 prototype key, as diagnostic characters were recently reinvestigated and validated  
4 (Janosik & Halanych 2010, Janosik *et al.* 2011, Janosik & Halanych 2013). We focused  
5  
6 on 15 species (3 genera: *Acodontaster*, *Diplodontias*, *Odontaster*) and 13 descriptors (e.g.  
7  
8 general morphology, number and organization of skeletal plates, morphology of spines,  
9  
10 number of teeth, and type of pedicellariae). We also took macro photographs to illustrate  
11  
12 whole specimens and descriptors, allowing a realistic identification. This Xper3  
13  
14 identification key can be accessed at <http://odontasteridae-so.identificationkey.org>.  
15  
16  
17  
18

19 Transferring the experts' knowledge from recent scientific papers to an interactive  
20  
21 identification key and building the KB through digitization was straightforward and  
22  
23 easier to implement than initially assumed. Supplementary tools provided in Xper3, such  
24  
25 as matrices comparing the diagnostic characters among a set of selected taxa ("Compare  
26  
27 Items" tool) proved particularly convenient. Following the completion of this first key, a  
28  
29 medium-term objective is to build an enlarged identification key including the  $\approx$  120  
30  
31 extent sea star genera in the Southern Ocean (Moreau *et al.* 2015).  
32  
33  
34  
35  
36  
37

38 **Fig. 3 – Xper3 identification key of the Odontasteridae (Asteroidea). Descriptors can**  
39  
40 **be selected on the left panel while remaining species (considering the descriptors**  
41  
42 **already selected) are listed on the right panel. Several macro pictures were uploaded**  
43  
44 **to show the whole-specimen morphology or close-up on particular characters.**  
45  
46  
47  
48

## 49 Knowledge bases as a development framework

50  
51 *Data papers*  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Building an Xper3 KB, as well as any DB, is time-consuming for taxon and data experts,  
4 as terminology must be defined unambiguously, and all credible data sources must be  
5 merged and organized. Such a work can be valued by data paper publications. A data  
6 paper is a relatively new publication type designed to describe a dataset such as, how data  
7 were collected, processed, and verified, and make other researchers aware of these data  
8 (Pasquetto *et al.* 2017, Kim 2020). Data papers were initiated in the journal Biodiversity  
9 by a GBIF (Global Biodiversity Information Facility) consortium in discussion with the  
10 editor Pensoft, in order to add the value to the process of collecting, identifying, and  
11 building a dataset that can be freely shared to all researchers. GBIF now lists 30 journals  
12 for data publications. Other domains now publish data papers, such as physics, chemistry,  
13 environmental science, and climatology (e.g., Nature's scientific data journal). The core  
14 of a data paper is built on the metadata of a dataset, which precisely describes dataset  
15 coverage, methodological protocols, data validation, field and museum collection  
16 information, and the relevant usage etc. Examples are available on the GBIF France  
17 website ([www.gbif.fr](http://www.gbif.fr)), and include very recent examples (Simier *et al.* 2019; Dillen *et al.*  
18 2019). Data papers on Xper3 identification keys and KB appear also in Jouveau *et al.*  
19 (2018), Klimmek & Baur (2018) and Engel *et al.* (2016), among others.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

#### 44 *Field guides*

45  
46  
47  
48  
49 The production of taxonomic field guides is another application that can be greatly  
50 facilitated by using an established KB. In practice, field guides can prove highly popular  
51 and useful to a wide category of users in the field (Féral *et al.* 2019), emphasizing the  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 interest of these users to the KB itself. Field guides are used for naming morpho-species  
4 from images and samples taken in the field. They can provide an important baseline  
5 resource to assist biologists in monitoring species, communities, and environments. They  
6 represent a useful first resource for non-specialist taxonomists during collection  
7 expeditions. However, they should not be treated as a taxonomic reference or considered  
8 complete, and firm identifications should be avoided without strong evidence. Using field  
9 guides must be considered an initial step for identification of specimens collected in the  
10 field before a more precise identification is performed using complementary tools. For  
11 instance, identification of species complexes or assessing cryptic diversity requires the  
12 combined use of molecular tools and morphological approach (Sands *et al.* 2008, Tan *et*  
13 *al.* 2010, Martin-Ledo *et al.* 2012; Jossart *et al.* 2019; Moreau *et al.* 2019).

24  
25  
26  
27  
28  
29 The production of photographic field guides for non-experts is complementary to  
30 the production of KB. The production of field guides and their updates can turn out to be  
31 expensive, thereby irregular, incomplete for certain regions, and the taxonomy of some  
32 field guides may not be checked by experts for major biological groups. Usually, field  
33 guides do not include detailed taxonomic information like the specimen depository in  
34 academic collections, barcoding associated to specimen depositories, and identification  
35 keys are not always provided. Barriers to making marine imagery discoverable and  
36 accessible also constitute a major issue, and specific recommendations were recently  
37 refined (Przeslawski *et al.* 2019). Such issues or potential limitations could be partly  
38 addressed using KB. Field guides and updates can be easily generated from a KB like  
39 Xper3 and, conversely, images from field guides can be integrated into existing KB.  
40 Therefore, field guides and KB can build upon each other to become better informed and  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 illustrated guides.  
4

5           The tools provided by KB to produce field guides could be particularly helpful for  
6 Antarctic biologists as there is a real need for generating new Antarctic field guides, or  
7 updating existing ones. A relatively low number of field guides to the flora and fauna of  
8 Antarctica was made available, in comparison to temperate and tropical regions  
9 (Brueggeman 1998, Hibberd & Moore 2009, Rauschert & Arntz 2015, Schories &  
10 Kohlberg 2016, Neill *et al.* 2016). In addition, regarding the marine fauna and flora of the  
11 Southern Ocean, it appears that the biodiversity of the Antarctic zone has been more  
12 often treated (Brueggeman 1998, De Broyer *et al.* 2014, Rauschert & Arntz 2015,  
13 Schories & Kohlberg 2016) than species of the sub-Antarctic islands (see however  
14 Fischer & Hureau 1985, and Hibberd & Moore 2009), in nearshore habitats in particular  
15 (Féral *et al.* 2019). Available Antarctic guides mainly focus on epifaunal species (Barnes  
16 2007, Schories & Kohlberg 2016, Hibberd & Moore 2009) or include only some benthic  
17 fauna (Fischer & Hureau 1985). Coverage of infaunal communities is minimal or only  
18 available for larger visible infaunal taxa at the sediment surface (i.e. burrowing bivalves  
19 and echinoderms). Many organisms remain undescribed, with either little to no published  
20 literature, or no other available information.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

42           Initiated under the aegis of SCAR, the SCAR-MarBIN Antarctic Field Guides  
43 (AFG) initiative was launched to overcome such a need. This is a collaborative project  
44 aiming at establishing and supporting a system of interoperable databases to provide free  
45 access to data to identify Antarctic organisms (<http://afg.biodiversity.aq>). Field guides  
46 can be easily built upon, updated, and customized from the content of controlled data  
47 sources, SCAR MarBIN (free and open access to Antarctic Marine Biodiversity data),  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 ANTABIF (Antarctic Biodiversity data), RAMS (access to taxonomic lists of species  
4 occurring in the Southern Ocean), GBIF (Global Biodiversity Information Facility), and  
5  
6 Panoramio (photo-sharing community). The generated field guides can be browsed online  
7  
8 or downloaded into a PDF format, published or printed to be taken in the field. Some  
9  
10 examples are provided, but the taxonomic coverage remains limited to common, or  
11  
12 emblematic organisms (<http://afg.biodiversity.aq/guides>).  
13  
14  
15  
16  
17  
18  
19

## 20 **Conclusions and future prospects**

21  
22  
23

24 The objectives of the Leuven taxonomic workshop were to present useful computer-aided  
25 identification tools and software to assist Antarctic biologists from species identification  
26 to the production of KB and field guides. The software Xper3 was chosen as an example  
27 of KB management system and development framework for the different applications. In  
28 general, the Xper3 software would greatly benefit from a wider audience. There is  
29 however a need to increase its accessibility and interoperability with other taxonomical  
30 informatic tools and field guides, and encourage the use of such tools in the scientific  
31 community. Moreover, another welcome addition would be the inclusion of a taxonomic  
32 backbone from reference databases (e.g. WoRMS, Catalogue of Life) directly into the  
33 key in the form of dynamic links. Identification tools are often neglected and considered  
34 helpful only for taxonomists. KB could enlarge them to a wider audience, making a better  
35 link between taxonomists and non-taxonomists and encourage the latest to create a  
36 preliminary KB from their observations. Broadly, we highlight several advantages  
37 provided by KB over more classical tools such as, better portability, update frequency,  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 and visibility.  
4

5 In recognition of the need for an improved understanding of species and community  
6 responses to environmental change, many key questions in the SCAR ‘Antarctic life on  
7 the precipice’ cover topics relating to the inclusion of biodiversity research for the 2020  
8 Science Horizon Scan (Kennicutt *et al.* 2015). Coordinating biodiversity collections and  
9 increasing access to taxonomic and genetic identification resources are key strategic  
10 aims. The risk of non-native species introductions into Antarctica is increasing (Barlett *et*  
11 *al.* 2020). Recent evidence suggests that physical barriers to biological dispersal,  
12 previously assumed to isolate Southern Ocean fauna, can be crossed, a phenomenon  
13 enhanced by changing climate conditions (Fraser *et al.* 2018). Biofouling organisms can  
14 survive sea ice scraping of vessel hulls (Hughes & Ashton 2017) and anti-fouling  
15 treatment (Lee & Chown 2007). Finally, as shipping visits to Antarctica coastal waters  
16 increase, particularly in the tourist industry (Verbitsky 2013, Kruczek *et al.* 2018), the  
17 risk of non-native introduction increases (McCarthy *et al.* 2019). Facilitating species  
18 inventory using KB can help monitor native species, the community state, and track non-  
19 native species. Field guides also provide an important baseline resource to assist in  
20 monitoring Antarctic marine environments for invasive species by non-expert observers  
21 and conservation managers.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

#### 49 **Acknowledgements**

50 We are grateful to the organizing committee of the SCAR XIIth Biology Symposium for  
51 the excellent facilities provided for the holding of the taxonomic workshop.  
52  
53  
54  
55  
56  
57  
58  
59  
60

### **Financial Support**

TS benefited from the support of the “Refugia and Ecosystem Tolerance in the Southern Ocean” project (RECTO; BR/154/A1/RECTO) funded by the “Belgian Science Policy Office” (BELSPO). Participation of NB was made possible thanks to support from David Barnes from the British Antarctic Survey and the Antarctic Seabed Carbon Capture Change project.

### **Author contributions**

TS, ME and CG conceived and designed this work. TS, QJ, and CM conceived and built the two exemplified KB. TS, RVL, QJ, and JM prepared the figures and the table. TS wrote the manuscript with contributions from all authors.

**REFERENCES**

Allan, E.L., Froneman, P.W., Durgadoo, J.V., McQuaid, C.D., Ansorge, I.J. & Richoux, N.B. 2013. Critical indirect effects of climate change on sub-Antarctic ecosystem functioning. *Ecology and Evolution*, **3**, 2994–3004, doi: 10.1002/ece3.678

Balmford, A. & Gaston, K.J. 1999. Why biodiversity surveys are good value. *Nature*, **398**, 204–205.

Barnes, D.K.A. 2007. Benthic communities in the Southern Ocean. In *Encyclopedia of the Antarctic*. New York: Routledge, 142–144.

Bartlett, J.C., Convey, P., Pertierra, L.R. & Hayward, S.A. 2020. An insect invasion of Antarctica: the past, present and future distribution of *Eretmoptera murphyi* (Diptera, Chironomidae) on Signy Island. *Insect Conservation and Diversity*, **13(1)**, 10.1111/icad.12389.

Brandt, A., Gooday, A.J., Brandao, S.N., Brix, S., Brokeland, W., Cedhagen, T. *et al.* 2007. First insights into the biodiversity and biogeography of the Southern Ocean deep sea. *Nature*, **447**, 307–311.

Brueggeman, P. 1998. Underwater Field Guide to Ross Island & McMurdo Sound, Antarctica. *The National Science Foundation's Office of Polar Programs sponsored*

1  
2  
3 *Norbert Wu.—Univ. California, San Diego.*  
4  
5  
6

7  
8 Byrne, M., Gall, M., Wolfe, K. & Agüera, A. 2016. From pole to pole: the potential  
9  
10 for the Arctic seastar *Asterias amurensis* to invade a warming Southern Ocean. *Global*  
11  
12 *Change Biology*, **22(12)**, 3874–3887, doi: 10.1111/gcb.13304.  
13  
14

15  
16  
17 Bueno, G., Deniz, O., Pedraza, A., Ruiz-Santaquiteria, J., Salido, J., Cristóbal, G. *et al.*  
18  
19 2017. Automated Diatom Classification (Part A): Handcrafted Feature Approaches.  
20  
21 *Applied Sciences*, **7**, 753, 10.3390/app7080753.  
22  
23

24  
25  
26 Carvalho, F.A., Filer, D., Renner, S.S. 2014. Taxonomy in the electronic age and an e-  
27  
28 monograph of the papaya family (Caricaceae) as an example. *Cladistics*, **31(3)**, 321–  
29  
30 329, doi:10.1111/cla.12095  
31  
32

33  
34  
35 Cary, S.C., McDonald, I.R., Barrett, J.E. & Cowan, D.A. 2010. On the rocks: the  
36  
37 microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology*, **8**, 129–138.  
38  
39

40  
41  
42 Chapman, A.D. 2005. *Uses of Primary Species-Occurrence Data, version 1.0*. Report for  
43  
44 the Global Biodiversity Information Facility, Copenhagen,  
45  
46 <http://www.gbif.org/document/80545>.  
47  
48

49  
50  
51 Chavan, V. & Penev, L. 2011. The data paper: a mechanism to incentivize data  
52  
53 publishing in biodiversity science. *BMC bioinformatics*, **12(15)**, S2.  
54  
55

1  
2  
3  
4  
5 Clark, A. M. & Downey, M. E. 1992. *Starfishes of the Atlantic*. London, New York:  
6  
7 Chapman & Hall.  
8  
9

10  
11  
12 Clarke, A. & Johnston, N.M. 2003. Antarctic marine benthic diversity. *Oceanography*  
13  
14 *and Marine Biology: an Annual Review*, **41**, 47–114.  
15  
16

17  
18  
19 Conruyt, N., Grosser, D., Faure, G., Pichon, M., Geynet, Y., Guillaume, M., *et al.* 1998.  
20  
21 IKBS : an iterative knowledge base system for managing knowledge in remote  
22  
23 systematics. *Application to corals of the Mascarene Archipelago*. Perpignan, France:  
24  
25 International Society for Reef Studies (ISRS) European Meeting.  
26  
27

28  
29  
30 Costello, M.J., Vanhoorne, B. & Appeltans, W. 2015. Conservation of biodiversity  
31  
32 through taxonomy, data publication, and collaborative infrastructures. *Conservation*  
33  
34 *Biology*, **29(4)**, 1094–1099, doi: 10.1111/cobi.12496.  
35  
36

37  
38  
39 Dallwitz, M.J. 1980. A general system for coding taxonomic descriptions. *Taxon*, **29**, 41–  
40  
41 46.  
42  
43

44  
45  
46 Dallwitz, M. J. 2000. A comparison of interactive identification programs. [delta-](http://delta-intkey.com)  
47  
48 [intkey.com](http://delta-intkey.com).  
49  
50

51  
52  
53 Danis, B., Van de Putte, A., Renaudier, S. & Griffiths, H. 2013. Connecting biodiversity  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 data during the IPY: the path towards e-polar science. *In* Verde, C., di Prisco, G., eds.  
4  
5 *Adaptation and evolution in marine environments 2*. Berlin: Springer, 21–32.  
6  
7

8  
9  
10 David, B., Choné, T., Mooi, R. & De Ridder, C. 2005a. *Antarctic Echinoidea. Synopses*  
11  
12 *of the Antarctic benthos*. Königstein: Koeltz Scientific Books.  
13  
14

15  
16  
17 David, B., Choné, T., Festeau, A., Mooi, R. & De Ridder, C. 2005b. Biodiversity of  
18  
19 Antarctic echinoids: a comprehensive and interactive database. *Scientia Marina*,  
20  
21 **69**, 201–203.  
22  
23

24  
25  
26 David, B., Saucède, T., Chenuil, A., Steimetz, A. & De Ridder, C. 2016. The taxonomic  
27  
28 challenge posed by the Antarctic echinoids *Abatus bidens* and *Abatus cavernosus*  
29  
30 (Schizasteridae, Echinoidea). *Polar Biology*, **39(5)**, 897–912, 10.1007/s00300-015-1842-  
31  
32 5.  
33  
34

35  
36  
37 De Broyer, C. & Danis, B. 2011. How many species in the Southern Ocean? Towards a  
38  
39 dynamic inventory of the Antarctic marine species. *Deep Sea Research Part II: Topical*  
40  
41 *Studies in Oceanography*, **58(1-2)**, 5–17.  
42  
43

44  
45  
46 De Broyer, C., Koubbi, P., Griffiths, H.J., Raymond, B., Udekem d’Acoz, C. d’, Van de  
47  
48 Putte, A.P. et al. 2014. *Biogeographic Atlas of the Southern Ocean*. Scientific Committee  
49  
50 on Antarctic Research, Cambridge, XII+498 pp.  
51  
52  
53

1  
2  
3 De Broyer, C., Clarke, A., Koubbi, P., Pakhomov, E., Scott, F., Vanden Berghe, E. &  
4  
5 Danis, B. 2020. Register of Antarctic Marine Species,  
6  
7 <http://www.marinespecies.org/rams>.  
8  
9

10  
11  
12 Díaz, A., Féral, J.-P., David, B., Saucède, T. & Poulin, E. 2011. Evolutionary pathways  
13  
14 among shallow and deep-sea echinoids of the genus *Sterechinus* in the Southern Ocean.  
15  
16 *Deep Sea Research II*, **58 (1-2)**, 205–211, 10.1016/j.dsr2.2010.10.012.  
17  
18

19  
20  
21 Dillen, M., Groom, Q., Chagnoux, S., Güntsch, A., Hardisty, A., Haston, E. *et al.* 2019. A  
22  
23 benchmark dataset of herbarium specimen images with label data. *Biodiversity Data*  
24  
25 *Journal*, **7**, 10.3897/BDJ.7.e31817.  
26  
27

28  
29  
30 Engel, J., Brousseau, L. & Baraloto, C. 2016. GuiaTreeKey, a multi-access electronic key  
31  
32 to identify tree genera in French Guiana. *PhytoKeys*, **68**, 10.3897/phytokeys.68.8707.  
33  
34

35  
36  
37 Fabri-Ruiz, S., Saucède, T., Danis, B. & David, B. 2017. Southern Ocean Echinoids  
38  
39 database—An updated version of Antarctic, Sub-Antarctic and cold temperate echinoid  
40  
41 database. *ZooKeys*, **697**, 1–20.  
42  
43

44  
45  
46 Féral, J.-P., Poulin, E., De Ridder, C. & Saucède, T. 2019. A field guide to coastal  
47  
48 echinoderms of the Kerguelen Islands. In Oji, T., Motokawa, T., Fujita, T., Komatsu, M.,  
49  
50 Agatsuma, Y. & Kanazawa, K. eds. *Proceedings of the 16<sup>th</sup> International Echinoderm*  
51  
52 *Conference, 2018 May 28-June 1, Nagoya, Japan. Zoosymposia*, **15(1)**, 33–43,  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 doi: 10.11646/zoosymposia.15.1.6.  
4  
5

6  
7  
8 Fischer, W.K. 1940. Asteroidea. *Discovery Reports*, **20**, 69–306.  
9

10  
11  
12 Fischer, W.K. & Hureau, J.-C. 1985. *FAO species identification sheets for fishery*  
13  
14 *purposes Southern Ocean: Fishing Areas 48, 58 and 88 (CCAMLR Convention Area)*.  
15

16  
17 Prepared and published with the support of the Commission for  
18  
19 the Conservation of Antarctic Marine Living Resources, 1 & 2. Rome, FAO, 470 pp.  
20  
21

22  
23  
24 Forget, P.-M, Lebbe, J., Puig, H., Vignes, R., Hideux, F.L.S., Hibeux, M. 1986.  
25

26 Microcomputer aided identification: An application to trees from French Guiana.  
27

28 *Botanical Journal of the Linnean Society*, **93**, 205–223, 10.1111/j.1095-  
29

30  
31 8339.1986.tb01021.x.  
32

33  
34  
35 Froese, R. & Pauly, D. 2018. FishBase. World Wide Web electronic publication.  
36

37  
38 [www.fishbase.org](http://www.fishbase.org).  
39

40  
41  
42 Gasparini, S. & Antajan, E. 2007-2013. Plankton Identifier: a software for automatic  
43  
44 recognition of planktonic organisms.  
45

46  
47 [http://www.obs.vlfr.fr/~gaspari/Plankton\\_Identifier/index.php](http://www.obs.vlfr.fr/~gaspari/Plankton_Identifier/index.php).  
48  
49

50  
51 Gaston, K.J. 2005. Biodiversity and extinction: species and people. *Progress in Physical*  
52  
53 *Geography*, **29(2)**, 239–247.  
54  
55

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Gorsky, G., Ohman, M.D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.B. *et al.* 2010. Digital zooplankton image analysis using the ZooScan integrated system.

*Journal of Plankton Research March*, **32(3)**, 285–303.

Griffiths, H.J. 2010. Antarctic marine biodiversity – What do we know about the distribution of life in the Southern Ocean? *PLOS One*, **5**, 10.1371/journal.pone.0011683

Griffiths, H.J., Danis, B. & Clarke, A. 2011. Quantifying Antarctic marine biodiversity: The SCAR-MarBIN data portal. *Deep Sea Research Part II*, **58**, 18–29, 10.1016/j.dsr2.2010.10.008

Gutt, J., Barratt, I., Domack, E., d’Udekem d’Acoz, C., Dimmler, W., Grémare, A. *et al.* 2011. Biodiversity change after climate-induced ice-shelf collapse in the Antarctic. *Deep Sea Research II*, **58**, 74–83.

Gutt, J., Barnes, D.K.A., Lockhart, S. & van de Putte, A. 2013. Antarctic macrobenthic communities: A compilation of circumpolar information. *Nature Conservation*, **4**, 1–13, 10.3897/natureconservation.4.4499.

Gutt, J., Bertler, N., Bracegirdle, T.J., Buschmann, A., Comiso, J., Hosie, G. *et al.* 2015. The Southern Ocean ecosystem under multiple climate change stresses ? An integrated circumpolar assessment. *Global Change Biology*, **21(4)**, 1434-1453.

1  
2  
3  
4  
5 Gutt, J., Isla, E., Bertler, A.N., Bodeker, G.E., Bracegirdle, T., Cavanagh, R.D. *et al.*  
6  
7  
8 2017. Cross-disciplinarity in the advance of Antarctic ecosystem research. *Marine*  
9  
10 *Genomics*, **37**, 1–17.

11  
12  
13  
14 Hagedorn, G., Thiele, K., Morris, R. & Heidorn, P.B. 2005. Structured Descriptive Data  
15  
16 (SDD) w3c-xml-schema, Version 1.0. Biodiversity Information Standards  
17  
18 (TDWG) <http://www.tdwg.org/standards/116>  
19  
20

21  
22  
23 Hagedorn, G., Rambold, G. & Martellos, S. 2010. Types of identification keys. *In* Nimis,  
24  
25 P.L. & Vignes-Lebbe, R., *eds. Tools for Identifying Biodiversity: Progress and Problems.*  
26  
27 *Proceedings of the International Congress, Paris, September 20-22, 2010.* Trieste: EUT  
28  
29 Edizioni Università di Trieste, 59–64.  
30  
31

32  
33  
34 Hardy, C., David, B., Rigaud, T., De Ridder, C. & Saucède, T. 2011. Ectosymbiosis  
35  
36 associated with cidaroids (Echinodermata: Echinoidea) promotes benthic colonization of  
37  
38 the seafloor in the Larsen Embayments, Western Antarctica. *Deep Sea Research II*, **58**,  
39  
40 84-90.  
41  
42  
43

44  
45  
46 Hibberd, T. & Moore, K. 2009. *Field Identification Guide to Heard Island and*  
47  
48 *McDonald Islands Benthic Invertebrates, a guide for scientific observers on board*  
49  
50 *fishing vessels in that area.* The Department of Environment, Water, Heritage, and the  
51  
52 Arts. Australian Antarctic Division and the Fisheries Research and Development  
53  
54  
55

1  
2  
3 Corporation, Australia, 158 pp., ISBN 9781876934156.  
4  
5

6  
7  
8 Hu, Jing, Zhibo Chen, Meng Yang, Rongguo Zhang and Yaji Cui. 2018. A Multiscale  
9  
10 Fusion Convolutional Neural Network for Plant Leaf Recognition. *IEEE Signal*  
11  
12 *Processing Letters*, **25**, 853–857.  
13  
14

15  
16  
17 Hughes, B.& Burghardt, T. 2016. Automated Visual Fin Identification of Individual  
18  
19 Great White Sharks. *Computer Science - Computer Vision and Pattern Recognition*, 17 p,  
20  
21 2016arXiv160906323H  
22  
23

24  
25  
26 Hughes, K.A. & Ashton, G.V. 2017. Breaking the ice: the introduction of biofouling  
27  
28 organisms to Antarctica on vessel hulls. *Aquatic Conservation: Marine and Freshwater*  
29  
30 *Ecosystems*, **27(1)**, 158–164, 10.1002/aqc.2625.  
31  
32

33  
34  
35 Janosik, A.M. & Halanych, K.M. 2010. Unrecognized Antarctic biodiversity: a case  
36  
37 study of the genus *Odontaster* (Odontasteridae; Asteroidea). *Integrative and comparative*  
38  
39 *biology*, **50(6)**, 981–992.  
40  
41

42  
43  
44 Janosik, A.M., Mahon, A.R. & Halanych, K.M. 2011. Evolutionary history of Southern  
45  
46 Ocean *Odontaster* sea star species (Odontasteridae; Asteroidea). *Polar biology*, **34(4)**,  
47  
48 575–586.  
49  
50

51  
52  
53 Janosik, A.M. & Halanych, K.M. 2013. Seeing stars: a molecular and morphological  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 investigation into the evolutionary history of Odontasteridae (Asteroidea) with  
4 description of a new species from the Galapagos Islands. *Marine biology*, **160(4)**, 821–  
5  
6  
7  
8 841.  
9

10  
11  
12 Joly, A., Goëau, H., Bonnet, P., Bakić, V., Barbe, J., selmi, S. *et al.* 2014. Interactive  
13  
14 plant identification based on social image data. *Ecological Informatics*, **23**, 22–34.  
15  
16

17  
18  
19 Joly, A., Vrochidis, S., Karatzas, K., Karppinen, A. & Bonnet, P. 2018. *Multimedia tools*  
20  
21 *and applications for environmental & biodiversity informatics*. Cham: Springer  
22  
23 International Publishing, pp. 221.  
24  
25

26  
27  
28 Jossart, Q., Moreau, C., Agüera, A., De Broyer, C. & Danis, B. 2015. The Register of  
29  
30 Antarctic Marine Species (RAMS): a ten-year appraisal. *ZooKeys*, **524**, 137.  
31  
32

33  
34  
35 Jossart, Q., Sands, C.J. & Sewell, M.A. 2019. Dwarf brooder versus giant broadcaster:  
36  
37 combining genetic and reproductive data to unravel cryptic diversity in an Antarctic  
38  
39 brittle star. *Heredity*, **123(5)**, 622–633.  
40  
41

42  
43  
44 Jouveau, S., Delaunay, M., Vignes-Lebbe, R. & Nattier, R. 2018. A multi-access  
45  
46 identification key based on colour patterns in ladybirds (Coleoptera, Coccinellidae).  
47  
48 *ZooKeys*, **758**, 55–73, 10.3897/zookeys.758.22171.  
49  
50

51  
52  
53 Kaiser, S., Brandão, S.N., Brix, S., Barnes, D.K.A., Bowden, D.A., Ingels, J. *et al.* 2013.  
54  
55

1  
2  
3 Patterns, processes and vulnerability of Southern Ocean benthos: A decadal leap in  
4 knowledge and understanding. *Marine Biology*, **160**, 2295–2317, 10.1007/s00227-013-  
5 2232-6  
6  
7  
8  
9

10  
11  
12 Kargel, J.S., Bush, A.B.G., Cogley, J.G., Leonard, G.J., Raup, B.H., Smiraglia, C. M. *et*  
13 *al.* 2014. A world of changing glaciers: summary and climatic context. *In* Kargel, J.S.,  
14 Leonard, G.J., Bishop, M.P., Kääb, A. & Raup, B.H., *eds.* *Global land ice measurements*  
15 *from space*. Berlin Heidelberg: Springer, 781–840, 10.1007/978-3-540-79818-7\_33.  
16  
17  
18  
19  
20  
21

22  
23  
24 Kennicutt, M.C., Chown, S.L., Cassano, J.J., Liggett, D., Peck, L. S., Massom, R. *et al.*  
25  
26 2015. A roadmap for Antarctic and Southern Ocean science for the next two decades and  
27 beyond. *Antarctic Science*, **27**, 3–18, 10.1017/S0954102014000674  
28  
29  
30

31  
32  
33 Kim, J. 2020. An analysis of data paper templates and guidelines: types of contextual  
34 information described by data journals. *Science Editing*, **7(1)**, 16–23.  
35  
36  
37

38  
39  
40 Klimmek, F. & Baur, H. 2018. An interactive key to Central European species of  
41 the *Pteromalus albipennis* species group and other species of the genus (Hymenoptera:  
42 Chalcidoidea: Pteromalidae), with the description of a new species. *Biodiversity Data*  
43 *Journal*, **6**, e27722, 10.3897/BDJ.6.e27722.  
44  
45  
46  
47  
48  
49

50  
51 Koehler, R. 1906. Note préliminaire sur les Échinodermes recueillis par l'expédition,  
52 antarctique franYaise du Dr. Charcot (Échinides, Astéries et Ophiures). *Bulletin du*  
53  
54  
55  
56  
57  
58  
59



1  
2  
3 *Museum Paris*, **1905**, 464-470.  
4  
5  
6

7  
8 Kruczek, Z., Kruczek, M. & Szromek, A. 2018. Possibilities of using the tourism area life  
9  
10 cycle model to understand and provide sustainable solution for tourism development in  
11  
12 the Antarctic Region. *Sustainability*, **10(1)**, 89.  
13  
14  
15

16  
17 Kuang, Y. 2015. *Deep Neural Network for Deep Sea Plankton Classification. Technical*  
18  
19 *Report*. Stanford: Stanford University.  
20  
21  
22

23  
24 Neeraj, K., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C. *et al.*  
25  
26 2012. Leafsnap: A Computer Vision System for Automatic Plant Species  
27  
28 Identification. *ECCV*.  
29  
30  
31

32  
33 Lecointre, G., Améziane, N., Boisselier, M.-C., Bonillo, C., Busson, F., Causse R. *et al.*  
34  
35 2013. How operational is the species flock concept? The Antarctic shelf case. *PLoS ONE*,  
36  
37 **8(8)**, e68787, 10.1371/journal.pone.0068787.  
38  
39  
40

41  
42 Linse, K., Walker, L.J. & Barnes, D.K., 2008. Biodiversity of echinoids and their  
43  
44 epibionts around the Scottia Arc, Antarctica. *Antarctic Science*, **20**, 227–244.  
45  
46  
47

48  
49 Louveaux, A., Amédégnato, C., Poulain, S. & Desutter-Grandcolas, L. 2018. North West  
50  
51 Africa Grasshoppers and Locusts. <http://acrinwafrica.mnhn.fr/>.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Mah, C.L. & Foltz, D.W. 2014. New taxa and taxonomic revisions to the Poraniidae  
4 (Valvatacea; Asteroidea) with comments on feeding biology. *Zootaxa*, **3795(3)**, 327–372.  
5  
6  
7

8  
9  
10 Mah, C.L., Neill, K., Eléaume, M. & Foltz, D. 2014. New species and global revision of  
11 Hippasteria (Hippasterinae: Goniasteridae; Asteroidea; Echinodermata). *Zoological*  
12 *Journal of the Linnean Society*, **171(2)**, 422–456.  
13  
14  
15  
16

17  
18  
19 Martìn-Ledo, R., López-González, P., Sands, C.J. 2012. A new brooding species of brittle  
20 star (Echinodermata: Ophiuroidea) from Antarctic waters. *Polar Biology*, **36(1)**, 115–  
21  
22  
23  
24 126, 10.1007/s00300-012- 1242-z  
25  
26

27  
28 May, R.M. 2011. Why Worry about How Many Species and Their Loss? *PLoS Biology*,  
29  
30  
31 **9(8)**, e1001130, 10.1371/ journal.pbio.1001130.  
32  
33

34  
35  
36  
37  
38 McCarthy, A.H., Peck, L.S., Hughes, K.A. & Aldridge, D.C. 2019. Antarctica: The final  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
frontier for marine biological invasions. *Global change biology*, **25(7)**, 2221–2241.

56  
57  
58  
59  
60  
McClintock, J.B. 1994. Trophic biology of Antarctic shallow-water echinoderms. *Marine*  
*Ecology Progress Series*, **111(1)**, 191–202.

Molinos, J.G., Halpern, B.S., Schoeman, D.S., Brown, C.J., Kiessling, W., Moore, P.J. *et*  
*al.* 2015. Climate velocity and the future global redistribution of marine biodiversity.

1  
2  
3 *Nature Climate Change*, **6(1)**, 83–88, 10.1038/nclimate2769.  
4  
5

6  
7  
8 Moreau, C., Agüera, A., Jossart, Q. & Danis, B. 2015. Southern Ocean Asteroidea: a  
9  
10 proposed update for the Register of Antarctic Marine Species. *Biodiversity data journal*,  
11  
12 **3**, e7062, 10.3897/BDJ.3.e7062.  
13

14  
15  
16  
17 Moreau, C., Danis, B., Jossart, Q., Eléaume, M., Sands, C., Achaz, G. *et al.* 2019. Is  
18  
19 reproductive strategy a key factor in understanding the evolutionary history of Southern  
20  
21 Ocean Asteroidea (Echinodermata)? *Ecology and Evolution*, **9(15)**, 8465–8478.  
22

23  
24  
25  
26 Moya, F., Saucède, T. & Manjón-Cabeza, M.E. 2012. Environmental control on the  
27  
28 structure of echinoid assemblages in the Bellingshausen Sea (Antarctica). *Polar Biology*,  
29  
30 **35**, 1343–1357, 10.1007/s00300-012-1176-5.  
31

32  
33  
34  
35 Neill, K., Kelly, M. & Herr, B. 2016. *Amazing Antarctic asteroids. An interactive guide*  
36  
37 *to the starfish of the Ross Sea*. New-Zealand: NIWA, 49 pp., [https://niwa.co.nz/coasts-](https://niwa.co.nz/coasts-and-oceans/marine-identification-guides-and-fact-sheets/amazing-antarctic-asteroids)  
38  
39 [and-oceans/marine-identification-guides-and-fact-sheets/amazing-antarctic-asteroids](https://niwa.co.nz/coasts-and-oceans/marine-identification-guides-and-fact-sheets/amazing-antarctic-asteroids)  
40

41  
42  
43  
44 Neubacher, D. & Rambold, G. 2005. NaviKey – a Java applet and application for  
45  
46 accessing descriptive data coded in DELTA format, <http://www.navikey.net>.  
47

48  
49  
50  
51 Nimis, P.L. & Vignes-Lebbe, R. 2010. *Tools for identifying Biodiversity: Progress and*  
52  
53 *Problems*. Proceedings of the International Congress, Paris, September 20-22. 445 pp.,  
54  
55 <http://hdl.handle.net/10077/5849>.  
56

1  
2  
3  
4  
5 Pasquetto, I.V., Randles, B.M., Borgman, C.L. 2017. On the reuse of scientific data. *Data*  
6  
7  
8 *Scientific Journal*, **16**, 8, 10.5334/dsj-2017-008.  
9

10  
11  
12 Pellen, F., Bouquin, S., Mougnot, I. & Vignes-Lebbe, R. 2018. Building an OWL  
13  
14 ontology with Xper3. *Biodiversity Information Science and Standards*, **2**, e25614.  
15  
16 10.3897/biss.2.25614  
17  
18

19  
20  
21 Pendlebury, S.F. & Barnes-Keoghan, I.P. 2007. Climate and climate change in the sub-  
22  
23 Antarctic. *Papers and Proceedings of the Royal Society of Tasmania*, **141(1)**, 67–79,  
24  
25 0.26749/rstpp.141.1.67  
26  
27

28  
29  
30 Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G. *et al.*  
31  
32 2015. Open science resources for the discovery and analysis of Tara Oceans data.  
33  
34 *Sciences Data*, **2**, 10.1038/sdata.2015.23.  
35  
36  
37

38  
39  
40 Przeslawski, R., Barrett, N., Bax, N., Carroll, A., Foster, S., Heupel, M. *et al.* 2019. *Data*  
41  
42 *Discoverability and Accessibility: Report from July 2019 Workshop on Marine Imagery.*  
43  
44 Report to the National Environmental Science Program, Marine Biodiversity Hub.  
45  
46 Geoscience Australia, [https://www.nespmarine.edu.au/document/data-discoverability-](https://www.nespmarine.edu.au/document/data-discoverability-and-accessibility-report-july-2019-workshop-marine-imagery)  
47  
48 [and-accessibility-report-july-2019-workshop-marine-imagery.](https://www.nespmarine.edu.au/document/data-discoverability-and-accessibility-report-july-2019-workshop-marine-imagery)  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Rauschert M. & Arntz W.E. 2015. *Antarctic macrobenthos: a field guide of the*  
4 *invertebrates living at the Antarctic seafloor*. Wurster Nordseeküste: Arntz & Rauschert  
5  
6  
7  
8 Selbstverlag, 143 pp., ISBN 978-3-00-049890-9.  
9

10  
11  
12 Sands, C.J., McInnes, S.J., Marley, N.J., Goodall-Copestake, W.P., Convey, P. *et al.*  
13  
14  
15 2008. Phylum Tardigrada: an "individual" approach. *Cladistics*, **24**, 861–871.  
16  
17

18  
19 Saucède, T., Pierrat, B. & David, B. 2014. Chapter 5.26. Echinoids. *In* De Broyer, C.,  
20  
21 Koubbi, P., Griffiths, H.J., Raymond, B., Udekem d'Acoz d', C., Van de Putte, A.P. *et*  
22 *al.*, eds. *Biogeographic Atlas of the Southern Ocean*. Cambridge: Scientific Committee  
23  
24 on Antarctic Research, 213–220.  
25  
26  
27

28  
29  
30 Saucède, T., Díaz, A., Pierrat, B., Sellanes, J., David, B., Féral, J.-P. *et al.* 2015a. The  
31  
32 phylogenetic position and taxonomic status of *Sterechinus bernasconiae* Larrain, 1975  
33  
34 (Echinodermata, Echinoidea), an enigmatic Chilean sea urchin. *Polar Biology*, **38**, 1223–  
35  
36 1237.  
37  
38

39  
40  
41 Saucède, T., Griffiths, H., Moreau, C., Jackson, J., C. Sands, R. Downey *et al.* 2015b.  
42  
43 East Weddell Sea echinoids from the JR275 expedition. *Zookeys*, **504**,  
44  
45 10.3897/zookeys.504.8860.  
46  
47  
48

49  
50  
51 Schiaparelli, S., Danis, B., Wadley, V. & Stoddart, D.M. 2013. The Census of Antarctic  
52  
53 Marine Life: the first available baseline for Antarctic marine biodiversity. *In* *Adaptation*  
54  
55  
56

1  
2  
3 *and Evolution in Marine Environments, Volume 2.* Berlin, Heidelberg: Springer, 3–19.  
4  
5  
6

7 Schories, D. & Kohlberg, G. 2016. *Marine Wildlife King George Island Antarctica:*  
8 *Identification Guide.* Dirk Schories Publications, 348 pp., ISBN-13: 9783000508325.  
9  
10  
11  
12

13  
14 Schram, J.B., Schoenrock, K.M., McClintock, J.B., Amsler, C.D. & Angus, R.A. 2016.  
15 Testing Antarctic resilience: The effects of elevated seawater temperature and decreased  
16 pH on two gastropod species. *ICES Journal of Marine Science*, **73(3)**, 739–752,  
17 10.1093/icesjms/fsv233  
18  
19  
20  
21  
22  
23

24  
25  
26 Simier, M., Ecoutin, J. & Tito de Morais, L. 2019. The PPEAO experimental fishing  
27 dataset: Fish from West African estuaries, lagoons and reservoirs. *Biodiversity Data*  
28 *Journal*, **7**, 10.3897/BDJ.7.e31374.  
29  
30  
31  
32

33  
34  
35 Smith, V.R. 2002. Climate change in the sub-Antarctic: An illustration from Marion  
36 Island. *Climatic Change*, **52(3)**, 345–357, <https://doi.org/10.1023/A:1013718617277>  
37  
38  
39  
40

41  
42 Surender, M., Chandra Shekar, K., Ravikanth, K. & Saidulu, R. 2019. Automatic  
43 Identification of Bird Species from the Image Through the Approaches of Segmentation.  
44 *In Saini, H., Sayal, R., Govardhan, A. & Buyya, R., eds. Innovations in Computer*  
45 *Science and Engineering. Lecture Notes in Networks and Systems, vol 74*, Singapore:  
46 Springer.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Tan, D.S.H., Ang, Y., Lim, G.S., Ismail, M.R.B. & Meier, R. 2010. From ‘cryptic  
4 species’ to integrative taxonomy: an iterative process involving DNA sequences,  
5 morphology, and behaviour leads to the resurrection of *Sepsis pyrrhosoma*  
6 (Sepsidae: Diptera). *Zoologica Scripta*, **39(1)**, 51–61, doi:10.1111/j.1463-  
7 6409.2009.00408.x

8  
9  
10  
11  
12  
13  
14  
15  
16  
17 Teletchea, F. 2010. After 7 years and 1000 citations: comparative assessment of the DNA  
18 barcoding and the DNA taxonomy proposals for taxonomists and non-taxonomists.  
19 *Mitochondrial DNA*, **21(6)**, 206–226.

20  
21  
22  
23  
24  
25  
26 Thiele, K. & Dallwitz, M.J. 2000. A critique of Dallwitz’s ‘A comparison of interactive  
27 identification programs’. <http://delta-intkey.com>.

28  
29  
30  
31  
32  
33 Turner, J., Barrand, N., Bracegirdle, T., Convey, P., Hodgson, D.A., Jarvis, M. *et al.*  
34 2014. Antarctic climate change and the environment: an update. *Polar Record*, **50(3)**,  
35 237–259, <https://doi.org/10.1017/S0032247413000296>.

36  
37  
38  
39  
40  
41  
42 Ung, V., Dubus, G., Zaragüeta-Bagils, R. & Vignes-Lebbe, R. 2010. Xper<sup>2</sup>: introducing  
43 e-Taxonomy. *Bioinformatics*, **26(5)**, 10.1093/bioinformatics/btp715

44  
45  
46  
47  
48  
49 Van de Putte, A., Youdjou, N. & Danis, B. 2020. *The Antarctic Biodiversity Information*  
50 *Facility*. World Wide Web publication, <http://www.biodiversity.aq>.

1  
2  
3 Verbitsky, J. 2013. Antarctic tourism management and regulation: The need for change.  
4  
5 *Polar Record*, **49(3)**, 10.1017/S003224741200071X  
6  
7

8  
9  
10 Vignes-Lebbe, R., Chesselet, P. & Diep Thi, M.-H. 2016. Xper3: new tools for  
11  
12 collaborating, training and transmitting knowledge on botanical phenotypes. *In*  
13  
14 Rakotoarisoa, N.R., Blackmore, S. & Riera, B., eds. *Botanists of the twenty-first century:*  
15  
16 *roles, challenges and opportunities*. Paris: UNESCO, 228–239  
17  
18

19  
20  
21 Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A. *et*  
22  
23 *al.* 2016. The FAIR Guiding Principles for scientific data management and  
24  
25 stewardship. *Scientific Data*, **3**, 160018, 10.1038/sdata.2016.18.  
26  
27

28  
29  
30  
31 WoRMS Editorial Board. 2020. *World Register of Marine Species*.  
32  
33 <http://www.marinespecies.org> at VLIZ, doi:10.14284/170.  
34  
35



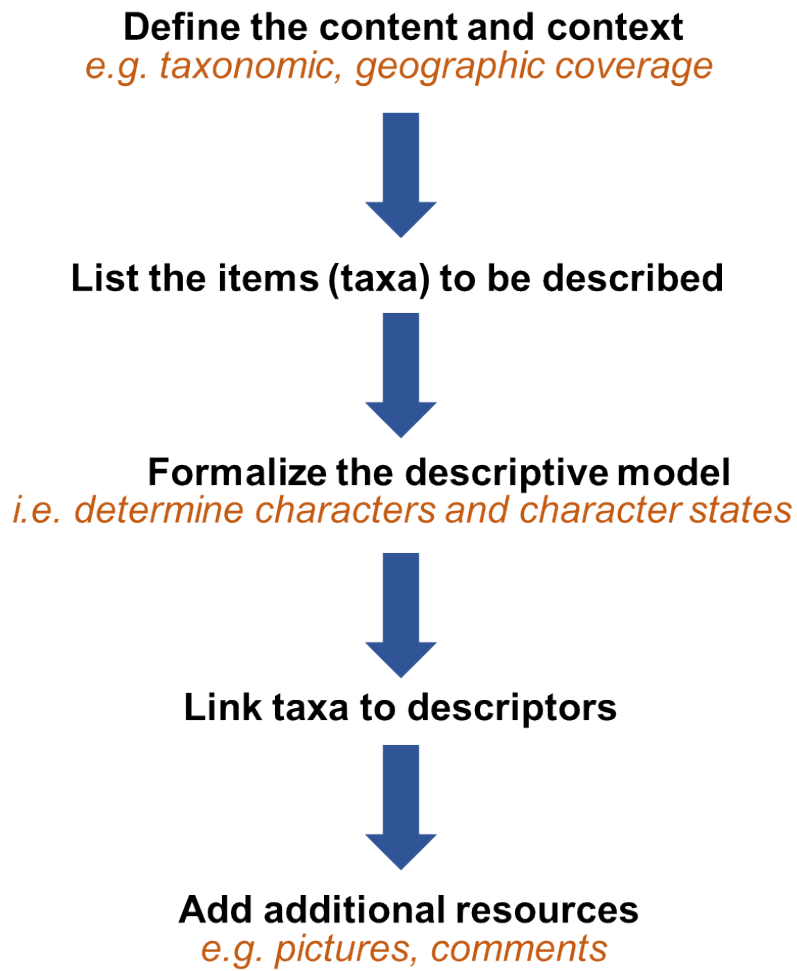
## Figure captions

Figure 1. Processing flowchart to build a Knowledge Base

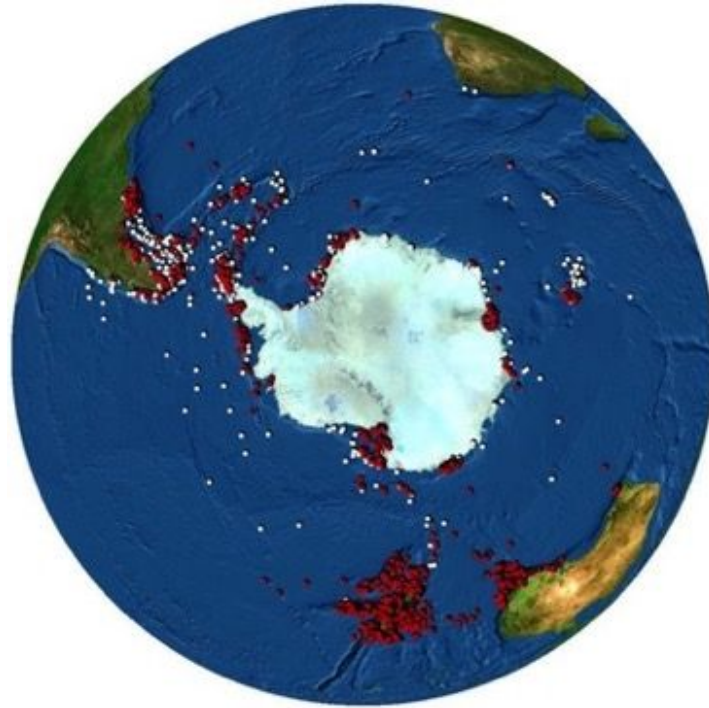
Figure 2. Recent advancement in our knowledge of echinoid distribution in the Southern Ocean comparing the databases from David *et al.* 2005a (white dots: 2,000 records) and Fabri-Ruiz *et al.* 2017 (red dots: 7,100 records).

Figure 3. Xper3 identification key of the Odontasteridae (Asteroidea). Descriptors can be selected on the left panel while remaining species (considering the descriptors already selected) are listed on the right panel. Several macro pictures were uploaded to show the whole-specimen morphology or close-up on particular characters.

Figure 1



**Figure 2**



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Figure 3

13 Descriptors History (0) Unselect Submit

Teeth of oral plate

Single tooth (12) ✓

Pair of teeth (3)

Crystal bodies on abactinal plate

15 Remaining taxa Among 15

- Acodontaster capitatus
- Acodontaster conspicuus
- Acodontaster elongatus
- Acodontaster hodgsoni
- Acodontaster marginatus
- Diplodontias dilatatus
- Diplodontias robustus
- Diplodontias singularis
- Odontaster aucklandensis
- Odontaster benhami
- Odontaster meridionalis
- Odontaster pearsei
- Odontaster penicillatus
- Odontaster roseus
- Odontaster validus

## Tables

Table 1: Comparison between Xper2 and Xper3

<b>Function</b>	<b>Xper2</b>	<b>Xper3</b>
Installation	Local. Windows, Mac OS, Linux	Not required. Remote access via internet navigator
User	Mono user	Multi users
Language interface	French, English, German, Spanish and Chinese	French, English
Items	List, hierarchy, group, and scope	List and hierarchy
Descriptors	Categorical, numerical	Categorical, numerical and calculated. Weight.
Illustrations	Local or remote files	Only remote files accessible via URL
Check base	Yes	Yes
Comparison of descriptions	Yes: two, several, or comparison of sets of descriptions	Yes: two, several, or comparison of sets of descriptions
Diagnose	Special features	Special features
Editor	Add, modify, Copy, rename, merge, delete, order	Add, modify, copy, rename, merge, delete, order
Detailed sheet	In editor and in identification	In identification only
Publishing	KB content and descriptors with their discriminator power, identification	KB content, identification
Statistics	Yes	Yes
Import formats	Xper, CSV, SDD	CSV, SDD
Export formats	Xper, CSV, SDD, Nexus, HTML, wiki	CSV, SDD, Nexus, HTML
Identification	Multiple (free) access key	Both single and multiple (free) access key