METHODS AND TECHNIQUES

Producing a plant diversity portal for South Africa

M. Marianne le Roux,[1,2] Paul Wilkin,[3] Kevin Balkwill,[4] J. Stephen Boatwright,[5] Benny Bytebier,[6] Denis Filer,[7] Cornelia Klak,[8] Ronell R. Klopper,[1,9] Marinda Koekemoer,[1] Laurence Livermore,[10] Roy Lubke,[11] Anthony R. Magee,[1,2] John C. Manning,[1] Alan Paton,[3] Tim Pearce,[3] Jasper Slingsby,[12] Ben-Erik van Wyk[2], Janine E. Victor[1] & Lize von Staden[1]

1 *Biosystematics Research and Biodiversity Collections Division, South African National Biodiversity Institute, Private Bag X101, Pretoria, 0001, South Africa*
2 *Department of Botany and Plant Biotechnology, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa*
3 *Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, United Kingdom*
4 *School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Private Bag 3, Johannesburg, 2050, South Africa*
5 *Department of Biodiversity and Conservation Biology, University of the Western Cape, Private Bag X17, Bellville, 7535, South Africa*
6 *Bews Herbarium, School of Life Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, 3209, South Africa*
7 *Department of Plant Sciences, University of Oxford, University College Oxford, OX1 4BH, United Kingdom*
8 *Bolus Herbarium, Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch, 7701, South Africa*
9 *Department of Plant and Soil Sciences, University of Pretoria, Pretoria, 0001, South Africa*
10 *Biodiversity Informatics, Natural History Museum, Cromwell Road, London, SW7 5BD, United Kingdom*
11 *Department of Botany, Rhodes University, P.O. Box 94, Grahamstown, 6140, South Africa*
12 *South African Environmental Observation Network, Fynbos Node, Private Bag X7, Claremont, 7735, South Africa*, and *Statistics in Ecology, Environment and Conservation, Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch, 7701, South Africa*

Author for correspondence: *M. Marianne le Roux, m.leroux@sanbi.org.za*

ORCID MMR, http://orcid.org/0000-0003-4705-2513 [for ORCID IDs of other authors, see end of paper]

DOI https://doi.org/10.12705/662.9

Abstract Taxonomy provides a universal method to classify biodiversity at different scales locally and globally. Currently, existing taxonomic treatments are scattered, limiting their accessibility and utility. The Convention on Biological Diversity has responded to this challenge by setting the goal of compiling a World Flora Online (Global Strategy for Plant Conservation Target 1, 2011–2020). This can be done by aggregating electronically available information provided by each country, region or specialist group. Developing a Flora or a high-level monographic product requires time and input from a large pool of taxonomic specialists. Completing a Flora may be difficult to accomplish for phytodiverse countries, such as South Africa, if the 2020 target is to be met. Fortunately, a large number of taxonomic contributions and many electronic tools exist that can enhance progress. Where these are available, efforts have to be made to access and digitise the literature. Here we describe a pragmatic approach to developing an online Flora, involving taking floristic information from multiple, previously published sources, digitising the legacy literature where needed and aggregating the required information into a single portal. South Africa is committed to producing an online Flora (the e-Flora of South Africa) and contributing the information to the World Flora Online initiative following the aggregator portal approach, a method described here that might be useful for other countries with high phytodiversity.

Keywords aggregator portal; big data; BODATSA; BRAHMS; data mining; e-Flora of South Africa; taxonomy; World Flora Online

Supplementary Material The Electronic Supplement (Table S1) is available in the Suplementary Data section of the online version of this article at http://www.ingentaconnect.com/content/iapt/tax

INTRODUCTION

The primary function of biological taxonomic science (in this sense including systematics) is to provide a universal means of identifying the components of biodiversity, to link them to information on their relationships and properties, and to create predictive information systems (Cracraft, 2002). Ultimately taxonomy aims to provide a complete treatment of all organisms at a range of scales (from local to global), to address the needs of various users. At present, information on most plant groups is scattered, being a geographically dispersed ill-defined sum of the many papers and books, often

Received: 25 Aug 2016 | returned for (first) revision: 2 Nov 2016 | (last) revision received: 25 Jan 2017 | accepted: 25 Jan 2017 || publication date(s): online fast track, 14 Apr 2017; in print and online issues, 4 May 2017 ||

published in different languages, supplemented by material conserved in museums and herbaria. There can be considerable costs attached to obtaining and accessing all the appropriate books and the necessary papers. Thus, for the majority of users, taxonomic information can be difficult to access, particularly in developing countries where the most diverse floras occur.

The 21st century has already seen significant efforts to address these issues technologically. The development of taxonomic resources on the internet has followed five principal routes: (1) Checklists: aggregation and dissemination of taxonomic names from multiple sources without taxonomic opinion (e.g., IPNI, 2015) to checklists and catalogues with accepted and synonymised names (e.g., African Plant Database, 2012; The Plant List, 2013; Catalogue of Life, 2015; World Checklist of Selected Plant Families, 2015; Tropicos, 2016). (2) Provision of virtual herbaria: digitisation of specimen collections in museums and herbaria, and disseminating them online (e.g., Reflora Virtual Herbarium, 2013; JSTOR Global Plants, 2015; Muséum National d'Histoire Naturelle, 2015; Naturalis Biodiversity Center, 2016). (3) Make specimen label information available in databases, e.g., BRAHMS (2015), and dissemination on a website. (4) Compile e-Floras or e-Monographs: projects that seek to compile legacy taxonomic (or floristic) information in structured characterisations (e.g., Encyclopedia of Life, 2016; GoldenGATE, 2016) or unstructured textual descriptions on the Web, ranging from family-scale resources (e.g., CATE-Araceae, 2015; Solanaceae Source, 2015) to those with regional foci (e.g., Flora of Nepal, 2014; Atlas of Living Australia, 2015; Flora of China, 2015a; Flora Iberica, 2016; Flora Zambesiaca, 2016) and projects that develop technologies to link and aggregate various biodiversity resources from across the Web (e.g., eMonocot, 2015; Encyclopedia of Life, 2015). (5) Development and dissemination of electronic biodiversity tools: biodiversity informatics projects that seek to provide the Information and Communication Technology (ICT) resources to help taxonomists publish their work on the Web by primarily managing specimen metadata (e.g., BRAHMS Online, BRAHMS, 2015; JACQ, 2016; Specify, 2016), taxonomic or nomenclatural data (EDIT Platform for Cybertaxonomy, Ciardelli & al., 2009 and Berendsohn, 2010; Scratchpads, 2015) and portals that provide content on actual taxonomic groups by implementing these tools.

The need for a world Flora was recognized by the Convention on Biological Diversity in its Target 1 of the Global Strategy for Plant Conservation (GSPC) 2011–2020 (GSPC, 2015) as the strategy acknowledged that sound taxonomic information is needed for efficient and effective conservation planning and management. The goal of Target 1 is to collect taxonomic information for all plants on Earth and present it as a global information service (WFO, 2016). The World Flora Online (WFO) project (WFO, 2016) follows the fourth principal route and aims to create a Flora of all known plants by 2020. In the case of a WFO, a species-level taxonomic backbone has to be established. Two challenges for the WFO initiative are that much of the information does not exist in electronic format and that species circumscriptions from different sources will be incongruent (a problem that has been outlined and discussed by Berendsohn & Geoffrey, 2007). These challenges may be addressed by taxonomic expert networks or individuals where networks are lacking.

The number of plant species in the world is estimated at around 400,000 (CBD, 2012; Paton, 2013; State of the World's Plants, 2016), with some estimating a further 50,000 that are as yet undescribed (Pimm & al., 2014; Pimm & Joppa, 2015). Collecting information for this number of taxa is an immense task that will only be achievable if consortia of individual researchers and research organisations work together. The WFO project is coordinated by the WFO Council that was formed in 2012 (Miller & al., 2014). The Council consists of 35 members (institutions or projects) that signed the WFO Memorandum of Understanding and agreed to work towards a global Flora (WFO, 2016). These institutions will participate in a global taxonomic information data mining initiative, with the South African National Biodiversity Institute (SANBI) mandated by the South African government to coordinate efforts in South Africa through the e-Flora of South Africa (e-Flora SA) project (Victor & al., 2014, 2015a). The project aims to collect taxonomic information for all species within the country, an estimated 20,500 species, 6% of the global Flora with a unique contribution of ca. 11,700 species (ca. 3% species of the world's flora) endemic to SA (Germishuizen & al., 2006), to make it available online, and feed the information into the WFO.

The purpose of this paper is to put South Africa's progress into a global perspective. The methodology that will be followed in the e-Flora SA project is presented to illustrate to other phytodiverse countries or regions how a similar outcome and contribution towards Target 1 of the GSPC may be achieved. We present an overview of progress of the e-Flora production in South Africa and compare this to other phytodiverse countries.

PROGRESS ON FLORA PRODUCTION FOR PHYTODIVERSE COUNTRIES

To compare the status of Flora compilation in phytodiverse areas across the world, a list of the top 11 phytodiverse countries was selected from Forzza & al. (2012) and includes Australia, Brazil, China, Colombia, India, Indonesia, Madagascar, Mexico, Papua New Guinea, Peru and South Africa (all with at least 9700 species of which more than 5600 species are endemic). A literature survey was conducted to determine and record the size of the country, the number of taxa and number of endemic species, the title of the checklist (if available), the most recent product that was delivered or being compiled, and the anticipated completion dates of projects or published products (Table 1).

Of the 11 countries with high plant diversity (Table 1), eight have complete checklists available, but only China has published a complete Flora, which is available in hard copy and electronic format. Eight countries are still working on their Floras: Australia, Brazil, Colombia, Indonesia, Madagascar, Mexico, Papua New Guinea and South Africa, but the Flora of Peru is currently dormant. No published checklist exists for three of the areas (Table 1).

The Flora of China project was started in 1994 and is now complete. It covers 31,362 species, had 478 contributors (Flora of China, 2015b) and was informed by the baseline of the Flora Reipublicae Popularis Sinicae project that commenced in the mid-1950s with the first volume published in 1959 (Ma & Clemants, 2006). The Flora Malesiana project (which includes two of the phytodiverse countries identified in Table 1, Indonesia and Papua New Guinea) commenced 19 years earlier, covers ca. 49,900 species and has had approximately 50 contributors up to now. It is still in progress after more than 60 years and, at the current rate, requires another 150 years before it will be complete (Roos, 2003). The number of recognised

Table 1. A comparison of electronic outputs of Flora projects in large, biodiverse countries for land plants (vascular plants and bryophytes).

Country	Area1 (km^2)	# vascular species	# endemic vascular species	Published checklist	Flora approach	Project or publication title(s)	Project or publication dates
Australia	7,741,220	15,638[1]	14,013[1]	Australian Plant Census[4]	Traditional	Flora of Australia[12]	1981–ongoing
					Aggregator	Atlas of Living Australia[13]	2010–ongoing
Brazil	8,514,880	32,364[1]	18,082[1]	The Brazilian catalogue of plants and fungi[5] (also available online: Lista de Espécies da Flora do Brasil[6])	Traditional	*Flora brasiliensis*[14]	1833–1906
					Traditional	Flora do Brasil 2020[15]	2010–ongoing
China	9,598,088	31,362[2]	15,623[2]	Flora of China Checklist[7]	Traditional	*Flora Reipublicae Popularis Sinicae*[16]	1959–2004
					Traditional	*Flora of China*[2]	1994–2004
Colombia	1,141,750	24,500[1]	10,500[1]	Not available	Traditional	*Prodromus florae novo-granatensis*[17,18]	1862–1873
						Flora de Colombia[19]	1983–ongoing
India	3,287,260	17,832[1]	6,113[1]	Not available	Traditional	*Flora of British India*[20]	1872–1897
						Fascicles of Flora of India[21]	1978–1996
Indonesia	1,904,570	29,375[1]	13,750[1]	Not available	Traditional	Flora Malesiana[22]	±1950–ongoing
Madagascar	587,040	9,753[1]	7,250[1]	Catalogue of the vascular plants of Madagascar[8]	N/A	Checklist to be expanded into an e-Flora in future[8]	2004–ongoing
Mexico	1,964,380	25,036[1]	11,250[1]	Flora Mesoamericana[9]	Traditional	Flora de México[21]	1993–ongoing
						Flora Mesoamericana[9]	1994–ongoing
Papua New Guinea	462,840	14,522[1]	13,250[1]	Census of vascular plants of Papua New Guinea[10]	Traditional	Flora Malesiana[22]	±1950–ongoing
Peru	1,285,220	18,055[1]	5,676[1]	Catalogue of the flowering plants and gymnosperms of Peru[11]	Traditional	Flora of Peru[21]	1936–dormant
South Africa	1,219,090	20,491[3]	11,700[3]	A checklist of South African plants[3]	Traditional	*Flora of southern Africa*[23]; *Bothalia*	1958–ongoing
					Aggregator	e-Flora of South Africa	2013–ongoing

1 Forzza & al. (2012); — 2 Flora of China (2015a, b); — 3 Germishuizen & al. (2006); — 4 Australian Plant Census (2016); — 5 Forzza & al. (2010); — 6 Lista de Espécies da Flora do Brasil (2010); — 7 Flora of China Checklist (2015, 2016); — 8 Madagascar Catalogue (2013); — 9 Flora Mesoamericana (2015); — 10 Conn (2008); — 11 Brako & Zarucchi (1993); — 12 Flora of Australia Online (2016); — 13 Atlas of Living Australia (2015); — 14 Martius & al. (1840–1906); — 15 Reflora Flora do Brasil 2020 (2016); — 16 Ma & Clemants (2006); — 17 Triana Silva & Planchon (1862); — 18 Triana Silva & Planchon (1863–1867); — 19 Bentancur & al. (1983–2014); — 20 Hooker (1875–1897); — 21 Frodin (2001); — 22 Flora Malesiana (2016); — 23 Marais (1958)

species in China is fewer than in the Flora Malesiana region, but the contributions of a community approximately nine times the quantity (ca. 450 contributors) have aided progress considerably (M. Roos., pers. comm.). From these two projects, it can be seen that to complete a Flora project (especially for a phytodiverse country or region), the scientific community needs to be heavily involved if it is to be finished in a reasonable timeframe.

■ THE STATUS OF THE FLORA PRODUCTION IN SOUTH AFRICA

The first major work published on the South African flora was Thunberg's *Flora Capensis* (1823), which treated ca. 2800 species known from the Cape (Glen & Germishuizen, 2010; Crouch & al., 2013). Harvey and Sonder published the first true Flora as *Flora Capensis* between 1859 and 1865. The project was concluded in 1933 with contributions by numerous authors in seven volumes (Hill, 1933). In total, 11,731 species were treated (Crouch & al., 2013). Twenty years later, in 1955, work began on the *Flora of southern Africa* (*FSA*) to replace the outdated *Flora Capensis*. The project was aimed at creating a complement to the *Flora Zambesiaca* series (Crouch & al., 2013). All known species across Botswana, Lesotho, Namibia, South Africa and Swaziland (Marais, 1958) were to be treated in the *FSA* series over the next 40 years. The first volume of the series was published in 1963, but progress was slow and only ca. 17%–18% (Crouch & al., 2013; Von Staden & al., 2013) of

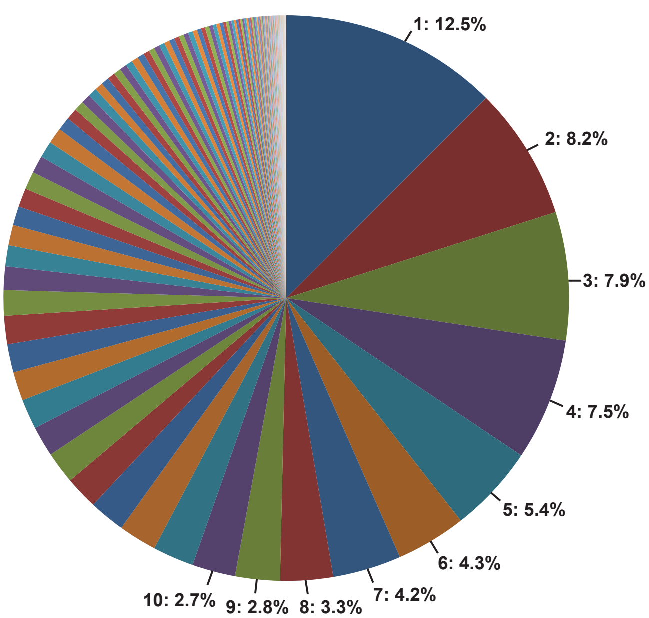

Fig. 1. A total of 132 journal or book titles in which taxonomic information are available for ca. 17,500 South African taxa. **1,** *Flora of southern Africa*; **2,** *Flora Zambesiaca*; **3,** *Illustrated handbook of succulent plants: Aizoaceae, A–E, F–Z*; **4,** *Flora Capensis*; **5,** *Journal of South African Botany*; **6,** *Annalen des Naturhistorischen Museums in Wien*; **7,** *Bothalia* (including *ABC Journal*); **8,** *Contributions from the Bolus Herbarium*; **9,** *Memoirs of the Botanical Survey of South Africa*; **10,** *Journal of South African Botany Supplement*. — Additional information is included in the Electr. Suppl.: Table S1.

the ca. 20,500 species known in southern Africa (Germishuizen & al., 2006) have been treated. A major achievement for South Africa was the checklist entitled *A checklist of South African plants* published by Germishuizen & al. in 2006.

Since the *FSA* project, a regional conspectus programme (ecological areas, e.g., the core Cape Flora and the extra Cape Flora that constitute the Greater Cape Floristic Region; or areas defined by provincial boundaries, e.g., plants of the northern provinces) for South Africa was initiated. The first three volumes have been published in the *Strelitzia* series, treating a total of ca. 14,400 taxa (ca. 67% of South Africa's flora; Retief & Herman, 1997; Manning & Goldblatt, 2012; Snijman, 2013).

In addition to the works cited above, an assortment of publications containing floristic information exists. Many taxonomic revisions of South African plants have been published across a wide range of national and international journals. Themed publications, e.g., bulb, succulent, tree and plant use books, also contain useful information. Thus the information required for the local Flora exists, but is scattered among many different sources. To assess how much floristic and taxonomic information is readily available (with copyright permission easily obtainable) for incorporation into the e-Flora SA, a list of publications and datasets was compiled (Table 2). A second list (Fig. 1; Electr. Suppl.: Table S1), in which three or more South African species have been revised, shows how wide the coverage is on the one hand, but how scattered taxonomic information is on the other. The total number of species was calculated for each publication (Fig. 1) and just over 50% of the treatments are in 7 sources, but treatments of the other species are spread among 125 different publications. In these calculations, multiple papers in the same journal (regardless of year of publication) are counted as being from the same source.

Information from collectors' labels of the herbarium specimens lodged in SANBI herbaria is recorded in the Botanical Database of Southern Africa (BODATSA) stored in the BRAHMS platform (BRAHMS, 2015). Additional databases of information include the Threatened Species Programme (Red List of South African Plants, 2015), the medicinal and magical plants (Arnold & al., 2002) and the common names (compiled from multiple sources) datasets. These are available for incorporation into the e-Flora without having to acquire permission from external copyright holders. There are also numerous type specimens and authoritatively determined specimens of taxa in South African herbaria that were digitised and made available under the auspices of the African Plants Initiative (Smith, 2004; Walters & al., 2010). These include 40,879 scanned specimens from local herbaria (Table 2). Numerous images of type specimens are also available from other herbaria world-wide through JSTOR Global Plants (2015). Furthermore, a total of 9983 items of artwork, slides and images are available on JSTOR Global Plants. SANBI retains the copyright of 30,327 specimens and images and these can therefore be used in the e-Flora. The remaining 16,861 specimens and 3674 botanical illustrations (or other images available online) can be included via links to the host herbaria or websites to avoid having to gain permission and having the additional responsibility of curating the images of specimens.

In 2013, the e-Flora SA project commenced in which SANBI is aiming to collect floristic information for all species within the country, to present the data online and ultimately contribute a national Flora towards the online world Flora initiative.

■ THE AGGREGATOR PORTAL APPROACH TO COMPILING AN E-FLORA

There is an urgency to promote and stimulate the production and dissemination of foundational biodiversity information (e.g., species names, descriptions and distributions), especially in electronic format (online checklists and Floras). The approach to developing online Floras is influenced by the legacy of botanical knowledge resources for the country or region, the availability of technological resources and skills and the available human capacity. These approaches broadly fall into two categories: (1) a traditional Flora approach, or (2) an aggregator portal approach.

The first approach, traditional Flora, involves developing an electronic Flora following traditional methods, where experts are invited to contribute and create accounts for genera, families or an entire region, following a set format determined by the editor(s). Usually new taxonomic work will be carried out by the scientific community and the contributions are coordinated and edited by one or more editors. Individual chapters may be published online while remaining chapters are still being compiled, e.g., Flora of Brazil (Brazilian Flora 2020 in Construction, 2016). Once all chapters are complete, a volume can be published and printed in hard copy. Alternatively, a complete Flora can be published in hardcopy, and then digitised and made available online after the completion of the project, e.g., Flora of China (2015a).

An advantage of this first (traditional Flora) approach is that the entire Flora is in a consistent format where each taxon treatment includes a single description and a set of standardised additional attributes such as habitat or distribution. This allows for clear communication of information where the user is advised of the taxonomic circumscription and associated

Table 2. Summary of the major sources of existing floristic information for South African plants.

Publication title/series/journal	Attribution	Date of publication	Number of items
Literature			
Flowering plants of Africa (including *Flowering plants of South Africa*)	Multiple authors	1921–ongoing	2,280 species
Flora of southern Africa	Multiple authors	1963–ongoing	ca. 4,390 species
Plants of the northern provinces of South Africa: Keys and diagnostic characters (Strelitzia 6)	Retief & Herman; multiple authors	1997	5,768 species
Plants of the Greater Cape Floristic Region: 1. The Core Cape Flora (Strelitzia 29)	Manning & Goldblatt; multiple authors	2012	9,251 species
Plants of the Greater Cape Floristic Region: 2. The Extra Cape Flora (Strelitzia 30)	Snijman; multiple authors	2013	3,715 species
Bothalia	Multiple authors	1923–ongoing as the *African Biodiversity and Conservation Journal*	Uncertain
Specimens			
Bews Herbarium, University of KwaZulu-Natal Herbarium (NU)	N/A	N/A	3,967 specimens
Bolus Herbarium, University of Cape Town (BOL)	N/A	N/A	12,106 specimens
Buffelskloof Nature Reserve Herbarium (BNRH)	N/A	N/A	727 specimens
H.G.W.J. Schweickerdt Herbarium, University of Pretoria (PRU)	N/A	N/A	61 specimens
SANBI, Compton Herbarium (NBG)	N/A	N/A	4,894 specimens
SANBI, KwaZulu-Natal Herbarium (NH)	N/A	N/A	1,260 specimens
SANBI, Pretoria National Herbarium (PRE)	N/A	N/A	15,059 specimens
SANBI, South African Museum in the Compton Herbarium (SAM)	N/A	N/A	2,805 specimens
Images			
Artwork (Flowering Plants of Africa)	Multiple artists	1868–2006	2,542 objects
Illustrations of southern Africa flora	Multiple artists	1929–2006	1,132 objects
SANBI, Slides	Multiple photographers	1920–2006	3,766 slides
SANBI, Photographic images of southern Africa flora (Compton and KwaZulu-Natal herbaria)	Multiple photographers	N/A	2,543 images

information. However, this method is time consuming, especially if the flora of the area is relatively unknown or diverse, and before the final volumes are available, the earlier ones are already out-of-date. When little taxonomic information is available for the specific area and time is no limitation, this method is appropriate, but it is likely to require many dedicated taxonomists conducting extensive research to achieve a complete floristic account, especially in highly phytodiverse countries or regions.

The second approach (aggregator portal approach) involves taking information from multiple published sources, often requiring digitising if they do not exist in electronic format, then mining data and aggregating it into a single portal, e.g. eMonocot (2015). This approach is largely made possible by the availability of technology that did not exist previously.

In the aggregator portal approach, there may be multiple descriptions available for each taxon. The format of descriptions and scope of the accounts may be inconsistent within and among species and there may be descriptive variation among taxa through time due to the lack of a standardised vocabulary, format of different sources and the change in authors. However, with the availability of a large body of existing taxonomic information, and where time is of essence, following this method will be more efficient. Some of the disadvantages may temporarily be overcome by presenting the data in a standard template or by presenting the most recent or appropriate description first. The most appropriate description would be based on set criteria, which include whether the description fits the current species circumscription (i.e., latest revision) as recognised within the country, and comprehensiveness of the data. The approach adopted by Flora projects in botanically rich countries such as Brazil (Reflora Flora do Brasil 2020, 2016), Malaysia (M. Roos, pers. comm.) and China (Flora of China, 2015b) suggests that time and large numbers of taxonomists are required to follow a traditional Flora approach. Sufficiently up-to-date published information exists for ca. 85% of South African species (Fig. 1). With less than three years left to meet Target 1 of the GSPC, an aggregator portal approach is the most feasible course of action.

The e-Flora SA two year project progress. — The initial time-frame for the e-Flora SA (Victor & al., 2014) was published based on a preliminary evaluation. New insights have been gained from lessons learned, markup and image management tools that became available and the database system (BRAHMS) that was further developed to accommodate the necessary descriptive information. All these factors have led to a revised methodology for the project. The interim project deadline was set and achieved for July 2016: descriptive data for 40% of species (8196 species according to the total number of taxa listed in Germishuizen & al., 2006) captured and deposited in BODATSA. The e-Flora team has worked beyond the target and collected descriptions for 13,800 species. The revised methodology to complete the remaining ca. 6700 species within three years includes the following steps (Fig. 2):

a. Checklist. – Before any floristic information can be collected, it is essential to know what the composition of the plant biodiversity is. A best-practice guide was developed by Hamer & al. (2012) for procedures for compiling and maintaining a national checklist. In South Africa, a committee, the South African National Plant Checklist Committee, was established to deal with competing or alternative taxonomies in order to create a consensus checklist.

1. Create an electronic list with all the names and author citations of the taxa that occur within the specific region.
2. Incorporate the data from the checklist into a database or make sure that the existing information in the database is accurate and up-to-date (in the case of South Africa, all information will be stored in BODATSA, hosted in BRAHMS, 2015).
3. Assign a unique identifier (UI) to each taxon name in the database (the IPNI UI can be included where available).
4. Develop mechanisms to track changes in circumscription of taxa (e.g., include notes of concept changes).

b. Select and convert literature. – Once the checklist is compiled, it can be expanded by adding floristic information as text blocks from at least one source for every taxon. Information is converted from the original published version (from hardcopy to digital copy) into a format that can be imported by the database, e.g., Microsoft (MS) Excel spreadsheet, DataBase (DBF) File. The format may be specific, depending on the database used.

1. Select the relevant literature to be incorporated for each taxon and, where necessary, obtain permission from the copyright holders to re-use it.
2. Hardcopy sources are digitised and converted to a format that can be edited in a text editor such as MS Notepad or MS Word using optical character recognition (OCR) software (e.g., ABBYY or Adobe Acrobat); if literature is ready and available in electronic format, this step is omitted.
3. Create an MS Excel template that will temporarily host floristic information and facilitate the transport thereof from the literature source into the database.
4. While the document is open in a text editor, e.g., MS Word, relevant bits of the information are carefully selected from the publication and copied into the MS Excel template; this process is known as markup.

Step four is largely automated by using a Markup Tool Add-In for MS Word (Von Staden, 2016). This software allows one to select information in MS Word, which is then automatically exported to an MS Excel document that contains column headings matching the field names in the database. It enables a quicker and more accurate transport of information. Some amending of errors, introduced during the OCR process, is largely done by the tool, but additional checking is required in MS Word. Other specialised biodiversity-orientated tools are available for markup as described in a pro-iBiosphere report (Scholz & Mietchen, 2013), such as GoldenGATE (2016; developed by Plazi), FlorML (developed by Hamann & al. 2014) or CharaParser (developed by Phenoscape, 2016) that allows automation of significant parts of the markup and data extraction process.

Identification keys form one of the central components of Floras (Scholz & Mietchen, 2013). Several tools are available to create electronic, interactive keys (Burkmar, 2014) and are generated by building a character matrix and coding the character states or by converting dichotomous keys into an electronic format using software such as CharaParser, Delta (developed at the Institute of Botany, Chinese Academy of Sciences, 2017), Lucid and Lucid Phoenix (developed by Lucidcentral, 2017), and Xper (developed by the Laboratory of Informatics and Systematics of the University Pierre et Marie Curie, 2017). Incorporating identification keys into the e-Flora SA project is considered a priority for the second phase, once all the correct names and descriptions are in place (Victor & al., 2014). Meanwhile, available electronic keys relevant to the South African flora have been linked to a central webpage and include keys to family, genus and in some cases species level (Identification Keys, 2017).

Images can also be incorporated into the database by storing them in a repository and saving the links to the images in the database.

c. Import data. – Floristic information, temporarily hosted in the MS Excel document, has to be imported into the database. In order to ensure that errors are avoided during transfer, it is recommended that the field names in the MS Excel document are matched with the field names in the database. It is also advisable to use UIs, allocated to taxon names, during the transfer process instead of actual scientific names, which could potentially create problems in the instance of spelling mistakes or inconsistent use of author names.

1. Information in the MS Excel template is transferred into the database; use UIs as the link between the taxon name and the information (the UIs are added to the MS Excel document by matching the names to a taxon dictionary created from the database).
2. The reference list in the database is updated and the necessary links are made between the floristic information and the literature source.
3. The quality and accuracy of the data are checked by specialists or proof-readers.

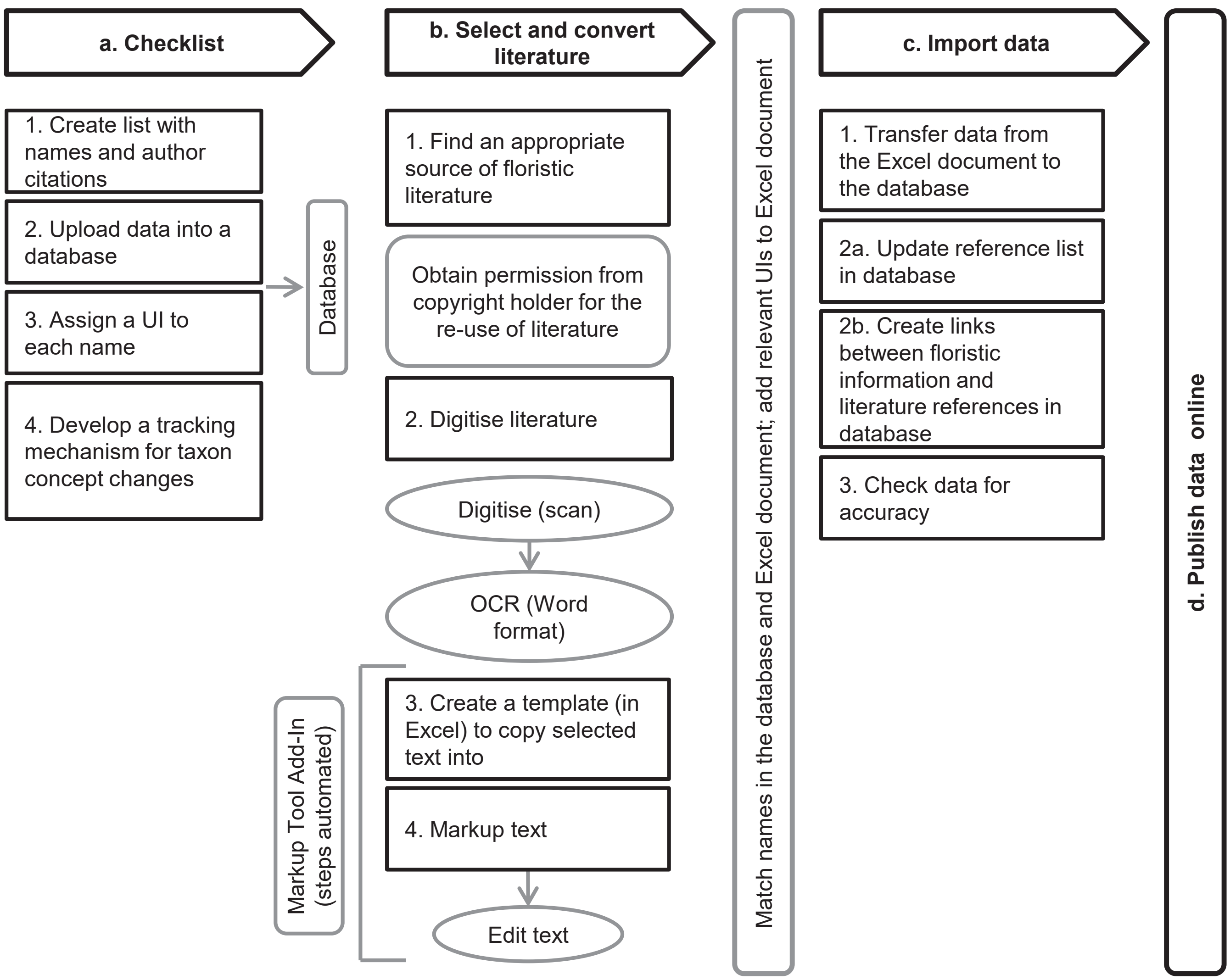

Fig. 2. Steps involved in collecting and storing floristic information to create an e-Flora using the aggregator portal approach. OCR, Optical Character Recognition; UI, Unique Identifier.

d. Publish data online. – The final step in the process will be to make floristic information accessible on the Web (Victor & al., 2014). BRAHMS Online enables one to publish information directly online (BRAHMS, 2015).

e. Biodiversity information standards. – After the data of the e-Flora SA project are published online the content will be shared with the WFO initiative. In order to share biodiversity information effectively between two data sources, the content has to be packaged into a format that corresponds with the structure and definitions of content in the second source. Integration of data can only take place effectively when the data structure of the two systems is homogeneous. A set of terms that is commonly used for this purpose is Darwin Core (Wieczorek & al., 2012) (used by Pensoft, 2015; the Global Biodiversity Information Facility – GBIF Tools, 2013; and eMonocot, 2015) and therefore this standard will be utilised in the e-Flora SA project.

f. Gap identification. – Successful compilation of information from published regional conspectuses (Retief & Herman, 1997; Manning & Goldblatt, 2012; Snijman, 2013) and electronic aggregation of existing published literature has resulted in descriptions for ca. 13,800 species (ca. 67%) (Gemishuizen & al., 2006) collected and captured in the database by the first project deadline. Three main tasks remain: (1) collecting information for the other 33% of species; (2) collecting comprehensive descriptions for those species where only a short description has been uploaded; and (3) encouraging revision of poorly treated or taxonomically out-dated groups. In South Africa this is being done by developing a strategy for plant taxonomic research (Victor & al., 2015a). Conspectuses are currently being compiled for the Nama-Karoo region, the Eastern Cape, Free State and KwaZulu-Natal Provinces, which, when added to the database, should cover most of the taxa within South Africa.

Although the collecting of descriptive data is currently the main focus, ideally images should be included with the aim of making at least one image (photo or scan of a herbarium specimen) available for every species by 2020. Existing images (preferably linked to voucher specimens) are being sourced from taxonomic experts. Where no image is available for a species, a scan of a representative herbarium specimen can be used.

There will potentially be shortcomings to the approach implemented by South Africa. Distribution maps created from databased specimen information will not always reflect accurate data and might include points from misidentified herbarium material, especially in groups that are poorly known (e.g., Aizoaceae, Victor & al., 2015a). Distribution patterns might also be incomplete for certain taxa as large regions of the country are underrepresented in the SANBI herbaria (Hamer & al., 2012). In future, the network of herbarium specimen data for South Africa will be expanded by serving specimen data online for all major herbaria in the country, which will lead to more comprehensive distribution maps. A commenting function will provide the opportunity for users to send feedback and allow for the improvement of data quality by its users.

Global benefits and utility. — The WFO will provide all users of plant information with a comprehensive world resource. It will be useful at a range of scales providing datasets and information based on state of the art taxonomic expertise in forms appropriate to the needs of users. The key global benefit of this resource will lie in the facilitation of regional to global biodiversity science. It will permit novel research programmes in evolutionary biology, ecology, climate change science, conservation science and agriculture from a biodiversity standpoint via a global dataset providing "big data".

Beyond 2016. — The e-Flora SA project will continue, beyond 2016, to collect basic floristic information for the remaining (33%) of South African species. The e-Flora SA project will continue to exist beyond 2020, with updates made and data expanded in accordance with the checklist being continually refined, and growing user needs.

Authors of taxonomic work are invited to contribute their publications to the e-Flora SA project to build and expand our knowledge base. Participation in the project will provide an opportunity to broaden exposure of researchers' taxonomic work to the national and international arena. This could lead to improved citations (often a concern for taxonomists) and exposure to new groups of potential users from both scientific and non-scientific backgrounds (Krell, 2002; Victor & al., 2015b). Tracking online usage will also provide a valuable additional measure of impact for taxonomic outputs.

Leadership. — The success and sustainability of large biodiversity projects, such as the e-Flora SA project, are greatly facilitated if a single organisation can play a leadership role, coordinating the project, providing sustainable resources to maintain the data, and to ensure that the data remain accessible. In South Africa, SANBI is in a position to fulfil a leadership role, but countries that lack such a government-funded organisation will need to depend on a team or partnership of experts who are committed to the project to take on this responsibility, as well as external funding.

■ CONCLUSIONS

A global Flora has to be created by 2020 to meet Target 1 of the GSPC (2015). This task involving an estimated 400,000–450,000 taxa is too large for a single institution to undertake and requires international collaboration. The WFO project will build a global product using contributions from smaller individual projects, whether monographic, family treatments or regional Floras such as the e-Flora SA. The traditional method used to compile Floras is time-consuming and will be unrealistic to complete before 2020 unless a very large number of taxonomists participates (e.g., Brazilian Flora in Construction, 2016). The alternative method outlined here (aggregator portal) would still require capacity and resources to carry out digitisation of hardcopy information, markup and improvement of data, but would be more feasible to achieve in a shorter space of time.

South Africa, a highly plant biodiverse country, will complete a Flora using an aggregator portal approach. The method outlined in this paper can serve as an example for other highly phytodiverse regions to complete their Floras in time for the 2020 goal. If institutions or partnerships in the ten phytodiverse countries listed in Table 1 together with South Africa can take a leadership role in compiling electronic Floras by 2020, more than 50% of the WFO will be complete. It is clear that no single world organisation could successfully shoulder the responsibility to produce a world Flora by 2020 and so national projects, especially in the most phytodiverse nations, will have to participate in the global project in order for the WFO to meet the targets that were set by the GSPC for 2011–2020.

Major advantages of an electronic Flora include the possibilities for expansion, maintenance and growth inherent in a dynamic system. The WFO, to which South Africa will contribute a significant unique portion through the e-Flora SA project, will revolutionise plant taxonomy from a technological standpoint and produce a change in its relationship with its users, who will be the greatest beneficiaries of an online world Flora system.

■ ACKNOWLEDGEMENTS

The National Research Foundation (grant reference number: UID92629) and Royal Society (grant reference number: SA140038) are thanked for the funding that was awarded to host an e-Flora workshop in South Africa.

■ LITERATURE CITED

African Plant Database. 2012. http://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php (accessed 11 Mar 2015).

Arnold, T.H., Prentice, C.A., Hawker, L.C., Snyman, E.E., Tomalin, M., Crouch, N.R. & Pottas-Bircher, C. 2002. *Medicinal and magical plants of southern Africa: An annotated checklist.* Strelitzia 13. Pretoria: South African National Biodiversity Institute.

Atlas of Living Australia. 2015. http://www.ala.org.au/ (accessed 10 Mar 2015).

Australian Plant Census. 2016. https://www.anbg.gov.au/chah/apc/about-APC.html (accessed 23 Aug 2016).

Bentancur, J., Galeano, G., Aguirre-C., J. & Amat-G., G. (eds.) 1983–2014. *Flora de Colombia*, vols. 1–31. Bogotá: Impreso en Colombia.

Berendsohn, W.G. 2010. Devising the EDIT platform for cybertaxonomy. Pp. 1–6 in: Nimis, P.L. & Vignes-Lebbe, R. (eds.), *Tools for identifying biodiversity: Progress and problems*. Trieste: EUT Edizioni Università di Trieste.

Berendsohn, W.G. & Geoffroy, M. 2007. Networking taxonomic concepts—uniting without "unitary-ism". Pp. 13–22 in: Curry, G. & Humphries, C. (eds.), *Biodiversity databases – Techniques, politics, and applications*. Baton Rouge: CRC Taylor & Francis. https://doi.org/10.1201/9781439832547.ch3

BRAHMS. 2015. [Botanical Research and Herbarium Management System, version 7]. http://herbaria.plants.ox.ac.uk/bol/ (accessed 10 Mar 2015).

Brako, L. & Zarucchi, J.L. 1993. Catalogue of the flowering plants and gymnosperms of Peru. *Monogr. Syst. Bot. Missouri Bot. Gard.* 45: 1–1286.

Brazilian Flora 2020 in Construction (Brazilian Flora). 2016. [Rio de Janeiro Botanical Garden]. http://floradobrasil.jbrj.gov.br/ (accessed 25 May 2016).

Burkmar, R. 2014. *The shifting paradigm of biological identification: A report for the tomorrow's biodiversity project*. http://www.tombio.uk/sites/default/files/TomBio-IDResources-22082014.pdf (accessed 17 Jan 2017).

Catalogue of Life. 2015. http://www.catalogueoflife.org/ (accessed 9 Mar 2015).

CATE-Araceae. 2015. [eMonocot Team CATE-Araceae]. http://araceae.e-monocot.org (accessed 10 Mar 2015).

CBD (Convention on Biological Diversity). 2012. [Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA). Global strategy for plant conservation: World Flora Online by 2020, 16th Meeting, Montreal, Canada. In: Convention on Biological Diversity, UNEP/CBD/SBSTTA/16/INF/38]. http://www.cbd.int/doc/meetings/sbstta/sbstta-16/information/sbstta-16-inf-38-en.doc (accessed 20 Jun 2015).

Ciardelli, P., Kelbert, P., Kohlbecker, A., Hoffmann, N., Güntsch, A. & Berendsohn, W.G. 2009. The EDIT platform for cybertaxonomy and the taxonomic workflow: Selected components. *Lecture Notes Informatics* 154: 625–638.

Conn, B.J. 2008. Census of Vascular Plants of Papua New Guinea. http://www.pngplants.org/PNGCensus (accessed 5 May 2015).

Cracraft, J. 2002. The seven great questions of systematic biology: An essential foundation for conservation and the sustainable use of biodiversity. *Ann. Missouri Bot. Gard*. 89: 127–144. https://doi.org/10.2307/3298558

Crouch, N.R., Smith, G.F. & Figueiredo, E. 2013. From checklists to an e-Flora for southern Africa: Past experiences and future prospects for meeting Target 1 of the 2020 Global Strategy for Plant Conservation. *Ann. Missouri Bot. Gard.* 99: 153–160. https://doi.org/10.3417/2011107

eMonocot. 2015. [Version 1.0.2]. http://e-monocot.org/ (accessed 9 Mar 2015).

Encyclopedia of Life. 2015. http://eol.org/ (accessed 18 Jun 2015).

Encyclopedia of Life. 2016. TraitBank. http://eol.org/info/516 (accessed 22 Aug 2016).

Flora Iberica. 2016. http://www.floraiberica.es/eng/ (accessed 9 Jun 2016).

Flora Malesiana. 2016. http://floramalesiana.org/ (accessed 24 Aug 2016).

Flora Mesoamericana. 2015. http://www.mobot.org/mobot/fm/intro.html (accessed 5 May 2015).

Flora of Australia Online. 2016. http://www.environment.gov.au/biodiversity/abrs/online-resources/flora/main/ (accessed 24 Aug 2016).

Flora of China. 2015a. http://www.efloras.org/flora_page.aspx?flora_id=2 (accessed 10 Mar 2015).

Flora of China. 2015b. [Vol. 1]. http://www.efloras.org/volume_page.aspx?volume_id=2001&flora_id=2 (accessed 11 Mar 2015).

Flora of China Checklist. 2015. http://www.tropicos.org/projectwebportal.aspx?pagename=Home&projectid=8 (accessed 5 May 2015).

Flora of China Checklist. 2016. http://flora.huh.harvard.edu/FloraData/002/Vol01/FOC-01-03-Authors-and-Illustrators.pdf (accessed 26 May 2016).

Flora of Nepal. 2014. http://padme.rbge.org.uk/floraofnepal/ (accessed 10 Mar 2015).

Flora Zambesiaca. 2016. http://apps.kew.org/efloras/fz/intro.html (accessed 25 May 2016).

Forzza, R.C., Leitman, P.M., Costa, A., Carvalho, A.A., Jr., Peixoto, A.L., Walter, B.M.T., Bicudo, C., Zappi, D., Costa, D.P., Lleras, E., Martinelli, G., Lima, H.C., Prado, J., Stehmann, J.R., Baumgratz, J.F.A., Pirani, J.R., Sylvestre, L.S., Maia, L.C., Lohmann, L.G., Paganucci, L., Silveira, M., Nadruz, M., Mamede, M.C.H., Bastos, M.N.C., Morim, M.P., Barbosa, M.R., Menezes, M., Hopkins, M., Secco, R., Cavalcanti, T. & Souza, V.C. 2010. *Catálogo das plantas e fungos do Brasil*, vols. 1–2. Rio de Janeiro: Andrea Jakobsson Estúdio and Rio de Janeiro Botanical Garden.

Forzza, R.C., Baumgratz, J.F.A., Bicudo, C.E.M., Canhos, D.A.L., Carvalho, A.A. Jr., Coelho, M.A.N., Costa, A.F., Costa, D.P., Hopkins, M.G., Leitman, P.M., Lohmann, L.G., Lughadha,

E.N., Maia, L.C., Martinelli, G., Menezes, M., Morim, M.P., Peixoto, A.L., Pirani, J.R., Prado, J., Queiroz, L.P., Souza, S., Souza, V.C., Stehmann, J.R., Sylvestre, L.S., Walter, B.M.T. & Zappi, D.C. 2012. New Brazilian floristic list highlights conservation challenges. *BioScience* 6: 39–45. https://doi.org/10.1525/bio.2012.62.1.8

Frodin, D.G. 2001. *Guide to Standard Floras of the World: An annotated, geographically arranged systematic bibliography of the principal floras, enumerations, checklists and chorological atlases of different areas*, ed. 2. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511541803

GBIF Tools. 2013. http://tools.gbif.org/ (accessed 18 May 2015).

Germishuizen, G., Meyer, N.L., Steenkamp, Y. & Keith, M. (eds.) 2006. *A checklist of South African plants.* Southern African Botanical Diversity Network Report 41. Pretoria: SABONET.

Glen, H.F. & Germishuizen, G. 2010. *Botanical exploration of southern Africa*, ed. 2. Strelitzia 26. Pretoria: South African National Biodiversity Institute.

GoldenGATE. 2016. http://plazi.org/resources/treatmentbank/goldengate-editor/ (accessed 22 Aug 2016).

Grey-Wilson, C. 1980. *Impatiens of Africa.* Rotterdam: A.A. Balkema.

GSPC (Global Strategy for Plant Conservation). 2015. The Targets 2011–2020. http://www.cbd.int/gspc/targets.shtml (accessed 10 Mar 2015).

Hamann, T., Müller, A., Roos, M.C., Sosef, M. & Smets, E. 2014. Detailed markup of semi-monographic legacy taxonomic works using FlorML. *Taxon* 63: 377–393. https://doi.org/10.12705/632.11

Hamer, M., Victor, J. & Smith, G.F. 2012. Best practice guide for compiling, maintaining and disseminating national species checklists, version 1.0. Global Biodiversity Information Facility, Copenhagen. http://www.gbif.org/orc/?doc_id=4752 (accessed 19 Aug 2016).

Harvey, W.H. & Sonder, O.W. 1859–1860. *Flora Capensis: Being a description of the plants of the Cape Colony, Caffraria and Port Natal (and neighbouring territories),* vol. 1, *Ranunculaceae to Connaraceae*. Dublin: Hodges, Smith & Co; Capetown: A.S. Robertson. https://doi.org/10.5962/bhl.title.821

Harvey, W.H. & Sonder, O.W. 1861–1862. *Flora Capensis: Being a description of the plants of the Cape Colony, Caffraria and Port Natal (and neighbouring territories),* vol. 2, *Leguminosae to Loranthaceae*. Dublin: Hodges, Smith & Co; Capetown: I.C. Juta. https://doi.org/10.5962/bhl.title.821

Harvey, W.H. & Sonder, O.W. 1864–1865. *Flora Capensis: Being a description of the plants of the Cape Colony, Caffraria and Port Natal (and neighbouring territories)* vol. 3, *Rubiaceae to Campanulaceae*. Dublin: Hodges, Smith & Co; Capetown: A.S. Robertson. https://doi.org/10.5962/bhl.title.821

Hill, A.W. (ed.) 1933. *Flora Capensis: Being a description of the plants of the Cape Colony, Caffraria and Port Natal (and neighbouring territories),* vol. 5(2, suppl.), *Gymnospermae*. Ashford: Reeve. http://dx.doi.org/10.5962/bhl.title.821

Hooker, J.D. 1875–1897. *Flora of British India*, vols. 1–7. London: Reeve. https://doi.org/10.5962/bhl.title.678

Identification Keys. 2017. http://biodiversityadvisor.sanbi.org/research-and-modelling/identification-keys/ (accessed 19 Jan 2017).

Institute of Botany, Chinese Academy of Sciences. 2017. DELTA – DEscripton Language for TAxonomy. http://delta-intkey.com/ (accessed 19 Jan 2017).

IPNI (International Plant Names Index). 2015. http://www.ipni.org/ (accessed 18 Jun 2015).

JACQ. 2016. Data input and manipulation in the JACQ system. http://jacq.nhm-wien.ac.at/dokuwiki/doku.php?id=export_documentation (accessed 21 Nov 2016).

JSTOR Global Plants. 2015. http://plants.jstor.org/ (accessed 11 Mar 2015).

Krell, F.-T. 2002. Why impact factors don't work for taxonomy. *Nature* 415: 957. https://doi.org/10.1038/415957a

Laboratory of Informatics and Systematics of University Pierre et Marie Curie. 2017. Xper³. http://www.xper3.fr (accessed 17 Jan 2017).

Lista de Espécies da Flora do Brasil. 2010. http://floradobrasil.jbrj.gov.br/2010 (accessed 5 May 2015).

Lucidcentral. 2017. Lucidcentral.org. http://www.lucidcentral.com/en-us/home.aspx (accessed 17 Jan 2017).

Ma, J. & Clemants, S. 2006. A history and overview of the *Flora Reipublicae Popularis Sinicae* (*FRPS*, *Flora of China*, Chinese edition, 1959–2004). *Taxon* 55: 451–460.

Madagascar Catalogue. 2013. Catalogue of the vascular plants of Madagascar. Missouri Botanical Garden, St. Louis, U.S.A. and Antananarivo, Madagascar. http://www.tropicos.org/project/mada (accessed 24 Aug 2016).

Manning, J.C. & Goldblatt, P. (eds.) 2012. *Plants of the Greater Cape Floristic Region*, vol. 1, *The core Cape Flora.* Strelitzia 29. Pretoria: South African National Biodiversity Institute.

Marais, W. 1958. The proposed Flora of South Africa. Proceedings of the 3rd AETFAT Congress, Coimbra, Portugal. *Mem. Soc. Brot.* 13: 51–52.

Martius, C.F.P. de, Eichler, A.W. & Urban, I. 1840–1906. *Flora Brasiliensis*, vols. 1–15. Monachii [Munich] et Lipsiae [Leipzig]: apud R. Oldenbourg. https://doi.org/10.5962/bhl.title.454

Miller, J.S., Thomas, W.W., Watson, W., Simpson, D. & Jackson, P.W. 2014. World Flora Online Council met in St. Petersburg. *Taxon* 63: 959. https://doi.org/10.12705/634.35

Muséum National d'Histoire Naturelle. 2015. [Vascular plants (P)]. https://science.mnhn.fr/institution/mnhn/collection/p/item/search/form (accessed 18 Jun 2015).

Naturalis Biodiversity Center. 2016. Digitization. https://science.naturalis.nl/en/collection/digitization/ (accessed 19 Aug 2016).

Paton, A.J. 2013. From working list to online flora of all known plants: Looking forward with hindsight 1. *Ann. Missouri Bot. Gard.* 99: 206–213. https://doi.org/10.3417/2011115

Pensoft. 2015. http://www.pensoft.net/journals/ (accessed 18 May 2015).

Phenoscape. 2016. http://phenoscape.org/wiki/Software (accessed 9 Jun 2016).

Pimm, S.L. & Joppa, L.N. 2015. How many plant species are there, where are they, and at what rate are they going extinct? *Ann. Missouri Bot. Gard.* 100: 170–176. https://doi.org/10.3417/2012018

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. & Sexton, J.O. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344 [12 pp.]. https://doi.org/10.1126/science.1246752

Red List of South African Plants. 2015. [Version 2014.1]. http://redlist.sanbi.org (accessed 11 Mar 2015).

Reflora Flora do Brasil 2020. 2016. http://reflora.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/ConsultaPublicaUC.do#CondicaoTaxonCP (accessed 24 Aug 2016).

Reflora Virtual Herbarium. 2013. http://www.herbariovirtualreflora.jbrj.gov.br/jabot/herbarioVirtual/ (accessed 18 Jun 2015).

Retief, E. & Herman, P.P.J. 1997. *Plants of the northern provinces of South Africa: Keys and diagnostic characters.* Strelitzia 6. Pretoria: National Botanical Institute.

Roos, M.C. 2003. Flora Malesiana 1991–2001. What has been achieved: Revitalisation, momentum? What next? *Telopea* 10: 1–10. https://doi.org/10.7751/telopea20035600

Scholz, H. & Mietchen, D. 2013. D3.3.1. Report on state-of the art and research horizons of semantic integration of biodiversity literature. The pro-iBiosphere Consortium. http://wiki.pro-ibiosphere.eu/w/media/2/2a/Pro-iBiosphere_WP3_MfN_D3.3.1_VFF20122013.pdf (accessed 9 Jun 2016).

Scratchpads. 2015. http://scratchpads.eu/ (accessed 10 Mar 2015).

Smith G.F. 2004. The African Plants Initiative: A big step for continental taxonomy. *Taxon* 53: 1023–1025. https://doi.org/10.2307/4135568

Snijman, D.A. (ed.) 2013. *Plants of the Greater Cape Floristic Region*, vol. 2, *The extra Cape Flora.* Strelitzia 30. Pretoria: South African National Biodiversity Institute.

Solanaceae Source. 2015. http://solanaceaesource.org/ (accessed 10 Mar 2015).

Specify. 2016. http://specifyx.specifysoftware.org/ (accessed 21 Nov 2016).

State of the World's Plants. 2016. https://stateoftheworldsplants.com/report/sotwp_2016.pdf (accessed 18 May 2016).

The Plant List. 2013. (Version 1.1). http://www.theplantlist.org/ (accessed 18 Jun 2015).

Thunberg, C.P. 1823. *Flora Capensis, sistens plantas Promontorii Bonae Spei Africes: Secundum systema sexuale emendatum, redactas ad classes, ordines, genera et species, cum differentiis specificis, synonymis et decriptionibus*, 2 vols. Stuttgardtiae [Stuttgart]: sumtibus J.G. Cottae. https://doi.org/10.5962/bhl.title.3650

Triana Silva, J. & Planchon, J.E. 1862. *Prodromus florae Novo-Granatensis*. Paris: Victor Masson & Fils. https://doi.org/10.5962/bhl.title.5385

Triana Silva, J. & Planchon, J.E. 1863–1867. *Prodromus florae Novo-Granatensis, Cryptogamie*. Paris: Victor Masson & Fils. http://bibdigital.rjb.csic.es/ing/Libro.php?Libro=301

Tropicos. 2016. http://www.tropicos.org/Project/FM (accessed 21 Nov 2016).

Victor, J.E., Smith, G.F., Turland, N.J., Le Roux, M., Paton, A., Figueiredo, E., Crouch, N.R., Van Wyk, A.E., Filer, D. & Van Wyk, E. 2014. Creating an Online World Flora by 2020: A perspective from South Africa. *Biodivers. & Conservation* 23: 251–263. https://doi.org/10.1007/s10531-013-0595-0

Victor, J.E., Smith, G.F. & Van Wyk, A.E. 2015a. *Strategy for plant taxonomic research in South Africa 2015–2020.* SANBI Biodiversity Series 26. Pretoria: South African National Biodiversity Institute.

Victor, J.E., Smith, G.F., Ribeiro, S. & Van Wyk, A.E. 2015b. Plant taxonomic capacity in South Africa. *Phytotaxa* 238(2): 149–162. https://doi.org/10.11646/phytotaxa.238.2.3

Von Staden, L., Raimondo, D. & Dayaram, A. 2013. Taxonomic research priorities for the conservation of the South African flora. *S. African J. Sci.* 109: Art. 1182 [10 pp.]. https://doi.org/10.1590/sajs.2013/1182

Von Staden, R. 2016. (Version 1.3.1). https://github.com/rudivs/SpeciesMarkupAddIn (accessed 18 May 2016).

Walters, M., Smith, G.F. & Crouch, N.R. 2010. The African Plants Initiative (API) in South Africa. *Taxon* 59: 1942–1948.

WFO (World Flora Online). 2016. http://www.worldfloraonline.org/ (accessed 18 May 2016).

Wieczorek, K., 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. https://doi.org/10.1371/journal.pone.0029715

World Checklist of Selected Plant Families. 2015. http://apps.kew.org/wcsp/ (accessed 18 Jun 2015).

ORCID Paul Wilkin: http://orcid.org/0000-0003-4982-7175; Kevin Balkwill: http://orcid.org/0000-0002-1181-5657; J. Stephen Boatwright: http://orcid.org/0000-0003-1250-4874; Benny Bytebier: http://orcid.org/0000-0002-4661-5727; Ronell R. Klopper: http://orcid.org/0000-0002-0948-5038; Marinda Koekemoer: http://orcid.org/0000-0003-0527-3545; Laurence Livermore: http://orcid.org/0000-0002-7341-1842; Roy Lubke: http://orcid.org/0000-0002-7563-5372; Anthony R. Magee: http://orcid.org/0000-0002-7063-7299; Alan Paton: http://orcid.org/0000-0002-6052-6675; Tim Pearce: http://orcid.org/0000-0003-1892-6824; Jasper Slingsby: http://orcid.org/0000-0003-1246-1181; Janine E. Victor: http://orcid.org/0000-0001-5703-1762; Lize von Staden: http://orcid.org/0000-0002-2652-7180