Lepidozia cupressina.
Andrew Branson

Considerations for the future of the BBS database

Oli Pescott and **Sam Amy** discuss our future database options

In the first volume of the *Atlas of British & Irish Bryophytes*, Chris Preston (2014) reviewed the history of bryophyte recording as practised by the British Bryological Society (BBS) and its predecessors. His section on 'Changes in the techniques of bryophyte recording' makes it clear that, despite the fundamentals of field activity remaining largely unchanged, numerous adjustments to both tools and the Society's structure have taken place in recent decades in response to technological shifts and opportunities. In our recent summary of the BBS database holdings since the last atlas, we touched on the process by which records are compiled and submitted via Regional Recorders to the Recording Secretary (Preston *et al*., 2012; Amy & Pescott, 2022). These are first subject to a number of validation and verification checks within an MS Access database (designed by Mark Hill during his tenure as Recording Secretary), with a degree of manual checking where potential errors are flagged. Records are then uploaded to a relational database (provided by the company Oracle Database) managed by staff at the Biological Records Centre (BRC) within the UK Centre for Ecology & Hydrology in Wallingford, Oxfordshire.

The need for a move

There are two main reasons for the BBS to consider the need to move away from this system. First, the Biological Records Centre has been investing in web-based recording technology for some years as an alternative to the Oracle database (Amy & Pescott, 2023), and it is not efficient to continue to support two separate

systems in perpetuity. Second, the current system where Regional Recorders submit data in a variety of formats to a Recording Secretary (Preston, 2014), who is then faced with the task of validating all content and converting it into a standardised format for database upload, is both time-consuming and increasingly inefficient in an age of web-based technology. Inefficient because, in such a world, all validation (checking spellings, dates, grid references etc.) and standardisation requirements could be instantly handled by a centralised online system to which all Regional (or other) Recorders submit data directly (and potentially use for their own data management requirements). Other advantages could include:

- instant access to centralised BBS data for all members;
- automated data publishing to the NBN Atlas and elsewhere (removing the need to maintain two taxon registers and actively produce such exports; Amy & Pescott, 2023);
- customised data visualisation and analysis;
- instant feedback on things such as altitude records or new vice-county records (assuming that such data are also codified in the system); and,
- direct integration with field recording tools such as mobile apps.

This note outlines the options that should be considered for moving BBS recording systems forward from this point. We attempt to broadly consider all options for the Society, without being swayed by any existing option with which we have corporate links or experience; the reader can judge how successful we are in this. Inevitably our proximity to some solutions gives us more insight into these than others, and the BBS is encouraged to seek alternative opinions to supplement those given here. The mission of the Biological Records Centre has always been to support national recording schemes and societies in achieving their own goals, and not to make their decisions for them.

Introducing Indicia

We first briefly introduce the Indicia system, as this open-source software toolkit underpins a number of the following options. Indicia is a set of 'modules' that can be used to construct digital biological recording forms (http://www.indicia.org.uk); these can be as simple or advanced as required for collecting and storing biodiversity observations. This includes standard functionality for such things as the collection, management, reporting, mapping and verification of records, as well as options such as photo upload and more advanced data visualisation capacities.

When online biological recording first became a reality over a decade ago, there was a danger that a large number of competing systems would be set up, and that records would become increasingly fragmented. To help avoid this problem, iRecord (https://irecord.org.uk), a website and mobile-based application built using Indicia, was also developed. This is now one of the most widely used online biological recording platforms in the UK, and also allows records from many other websites and apps to be brought together in one place to utilise its verification functionality (Amy & Pescott, 2023).

Where do Indicia systems store data?

The Indicia toolkit provides a 'data warehouse' (essentially just a database run on a remote server), the primary purpose of which is to store records from biological recording websites, including species observations, sites, recorders, lists of terms used in the data, and geographic objects such as site boundaries. It is a 'spatially enabled' database, meaning that it can handle objects defined in geographic space and perform

spatial operations on them (such as intersections of grid square-based records with vice-county or other boundaries). A data warehouse has its own administration interface, designed for use by people whose role it is to set up and configure the surveys that are feeding data to it.

The BRC hosts a specific instance of an Indicia data warehouse (the 'BRC Data Warehouse') to support organisations capturing biological records online, thereby minimising the infrastructure requirements of other organisations (this is the warehouse receiving records from the iRecord website, as well as numerous other projects; Amy & Pescott, 2023). The BBS, if it chose an Indicia-based system, could opt to have its records stored within the BRC Data Warehouse. Here, records from different schemes can be kept separate, or shared as required. This approach allows automated processes to be used to share records; for example, quicker and more frequent updates to the NBN Atlas are possible via Indicia than they are currently from the BRC Oracle database, and Local Environmental Records Centres can also be given direct access to data through existing mechanisms. The model of having data from multiple schemes in a single database is not new, and is actually how the current BBS database also works. Although it is not advertised as such, BBS data are stored in tables alongside data from other schemes and projects, the only distinguishing factor is simply whether the record is of a bryophyte or not: the current 'BBS database' is really just the bryophyte component of the larger BRC Oracle database.

Future options

The following options outline some of the future possible configurations for the BBS database, both built using Indicia and without. Note that we do not discuss the import of the existing BBS data holdings to any of these hypothetical new systems, as this would just entail the relatively minor (although not necessarily rapid) technical issue of matching existing columns and formats between the old and new systems, and applies equally to all (apart from Option 5, which is essentially stasis).

Option 1: A bespoke BBS recording platform using its own data warehouse

Under this route, a developer would use the Indicia toolkit to build a totally customised recording platform for the BBS. This would entail not only the design of the system (although of course existing database 'models' implied by the use of Indicia would have to be used), but also ongoing costs for its hosting and maintenance. New functionality not already available within the Indicia system would have to either be funded, or the Society would have to wait until such functionality was included in the core toolkit (although it is difficult to imagine what such functionality might be, given the fact that biological recording schemes typically demand similar options, and that these almost certainly already exist within Indicia). Under this model, integration with other platforms (e.g. data exchange with iRecord, data publishing to the NBN Atlas) would also be straightforward. One example of this model is the EU-funded 'Fungi Without Borders' project (http://www.fungi-without-borders.eu/en/home).

Option 2: A bespoke BBS recording platform using the BRC Data Warehouse

This option would entail a customised recording website, but one which stores its data in the existing BRC Data Warehouse already used by other projects such as iRecord. This option is often used by projects which need specialised recording forms and project documentation; the National Plant Monitoring Scheme website is

an example of such (www.npms.org.uk). Here, the need for an independent scheme identity, and the requirement for more complex, plot- and sample-based recording forms, meant that an independent website was designed for the NPMS using Indicia and other tools. The data, however, are ultimately stored in the BRC Data Warehouse. In this instance the BRC is able to support the running and maintenance of the database, but the scheme pays for, and organises, the upkeep and development of the website. For the NPMS this includes both the deployment of recording forms based on the Indicia toolkit, and the web content management system used (Drupal). In a similar vein, novel recording forms could potentially be embedded into other website content management systems, such as WordPress (as now used by the BBS website). There are numerous other examples of this model, including CEDaR online (http://www2.habitas.org.uk/records/home) – the records centre for Northern Ireland; NatureSpot (https://www.naturespot.org.uk/) – a charity recording wildlife in the Leicestershire area; the National Moth Recording Scheme (https://mothrecording.org); and the database of the British Mycological Society (http://www.frdbi.info/). Note that using the BRC Data Warehouse with a customised recording website means that data received via this website can be completely separated from other bryophyte data in the Warehouse (cf. Option 3).

Option 3: Direct use of iRecord with a bespoke recording form (already in use)

This would essentially be a move to using the bryophyte recording form described elsewhere in this issue by Amy & Pescott (2023). This could be customised further, but essentially it would mean that all 'BBS' bryophyte data submitted through this route would be contained within the BRC Data Warehouse alongside all other bryophyte submissions. Data submitted via various routes (e.g. the casual occurrences page, the mobile app, and the bryophyte recording form) are distinguishable in the database, but this would only relate to the data entry form or method used, and would not offer the ability to specifically ring-fence BBS bryophyte data from other bryophyte data in the BRC Warehouse (cf. Option 2).

Option 4: A third-party web database

All the preceding options start with the assumption that the BRC-supported Indicia infrastructure is a sensible platform for the BBS to use. Indicia is underpinned by support to BRC (both from the Joint Nature Conservation Committee, and from the Natural Environment Research Council via UKCEH core funding), and, more generally, the BRC is funded to support recording schemes and societies. Clearly this situation reduces costs for a society. Using Indicia also means that the BBS could indirectly benefit from ongoing improvements to the Indicia toolkit, including bug fixes, security updates, and new functionality added to the core systems for other projects (as well as requesting their own features under some of the above options). However, commissioning a third-party to design and implement a web-based recording system is also an option. The Botanical Society of Britain and Ireland, for example, maintain their own database system independently of the Indicia framework, although their database designer is a permanent member of that Society's paid staff. The Fungus Conservation Trust also has its own bespoke online database (https://cate.fungustrust.org.uk/public/). Third-party systems can still easily exchange data with Indicia systems, an opportunity that is taken advantage of by the BSBI for importing National Plant

Monitoring Scheme data to their 'distribution database'. The British Trust for Ornithology's BirdTrack system has also been configured to share data with iRecord (see https://tinyurl.com/birdTrackiRecord).

Option 5: Do nothing

Doing nothing is always an option that should be considered. In the current case there is as yet no specific date on which the BRC has determined that it will withdraw support for the Oracle database used to store BBS data (although loss of skilled staff capable of maintaining this system could unexpectedly precipitate this event, or at the least introduce large delays into the process of loading or amending data). Moreover, the recent increase in BRC botanical personnel has meant that BBS data have been processed more rapidly than has been the case in recent years. However, alongside these considerations, Recording Secretary terms have generally been around ten years in the recent past, and handing over the current, rather complex, system to a new Recording Secretary is something that is probably best avoided, especially if a new system is required within a few years. On the other hand, a new Secretary existing outside of the auspices of BRC may wish to invent their own wheels.

Conclusion

A BBS move towards any web-based data capture and storage system need not preclude people from carrying on with their current methods of data capture and storage, be that on paper, personal databases or spreadsheet (although there would be no need to continue updating the current BBS Excel-based recording forms were a web-based equivalent available). Data from these systems could still be uploaded to the new website, or passed on to a nominated compiler such as a Regional Recorder or the Recording Secretary. All recorders would enjoy benefits in terms of time saved, reductions in formatting errors and access to shared data that the more streamlined web-based approaches to data management outlined here could bring.

Acknowledgements

Thank you to Martin Harvey, David Roy and Chris Preston for comments which improved the paper. Oli Pescott and Sam Amy were supported by the NERC award number NE/R016429/1 as part of the UK Status, Change and Projections of the Environment (UK-SCAPE) programme delivering National Capability. This work was also supported through the Natural Capital Ecosystem Assessment programme, a partnership programme managed by Defra.

References

Amy, S. & Pescott, O.L. (2022). Patterns of bryophyte recording in Britain and Ireland since the 2014 Atlas. *Field Bryology* 128: 2–11.

Amy, S. & Pescott, O.L. (2023). Bryophyte distribution data in iRecord: an introduction to the resource. *Field Bryology* 129: 36–46.

Preston, C.D., Blackstock, T.H., Bosanquet, S.D.S., Godfrey, M.F., Hill, M.O., Holyoak D.T. & Rothero, G.P. (2012). *Bryophyte recording handbook*. British Bryological Society.

Preston, C.D. (2014). Recording bryophytes in Britain and Ireland, 1990–2013, in T.L. Blockeel, S.D.S. Bosanquet, M.O. Hill & C.D. Preston (eds), *Atlas of British and Irish bryophytes*, vol. 1, pp. 2–11. Pisces Publications, Newbury.

Oli L. Pescott[1]
Sam Amy[1]
[1]UKCEH Wallingford
e olipes@ceh.ac.uk