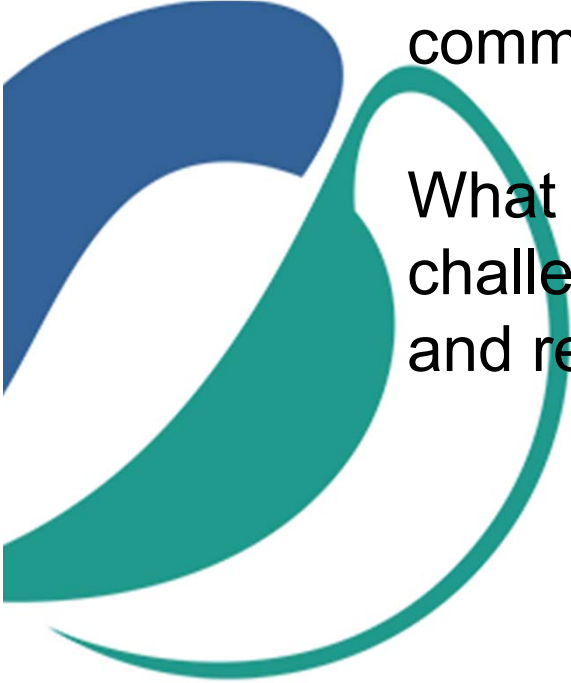




Whose data are they anyway?

Tom Hunt & Veerle Van den Eynden

NBN Conference
13 November 2019
Nottingham



Interactive session on data ownership, copyright, fair dealing and licensing of biodiversity records and other data that may be collected by citizens, professionals and commercial companies.

What are the opportunities and challenges when publishing, sharing and reusing them?

Workshop structure

- Background and context:
 - data ownership
 - copyright
 - licensing
 - citizen science
- Examples and attitudes from literature
- Case studies discussion
- Best practice guidance



Society for the Preservation of Natural History Collections Licensing And Rights Around Collections Data and symposium 25th - 31st May, Chicago



ALERC NC @ALERC_NC · May 29

This strikes me as a very important topic for a symposium or discussion, something we should have in the UK with input from experts. Whose data is it anyway? Would you agree @NBNTTrust @NaturalEngland @nature_scot @NatResWales @_NFBR ?



Laura Russell @pagodarose · May 29

Kicking off the licensing and rights symposium is @SomeSarahE @FieldMuseum. She's decoding legalise and discussing copyright in 15 minutes...just right. #SPNHC2019 @creativecommons



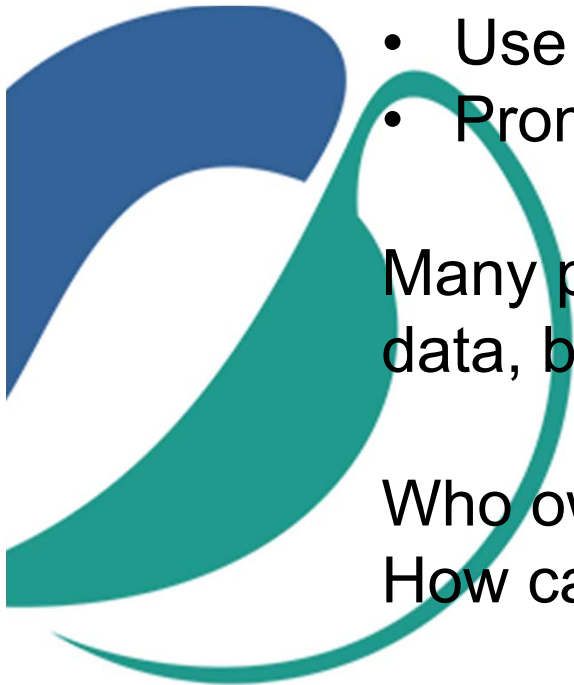
Why is this an issue?

- Greater amount of digitised data
- Wider range of sources
- Wider range of uses and users
- Wider range of data mediators
- Data linkage options
- Use of Creative Commons licences
- Promotion of particular licences

Many people making more decisions about data, but are they fair and legal?

Who owns biodiversity data?

How can they be used / shared?



About UK Data Service

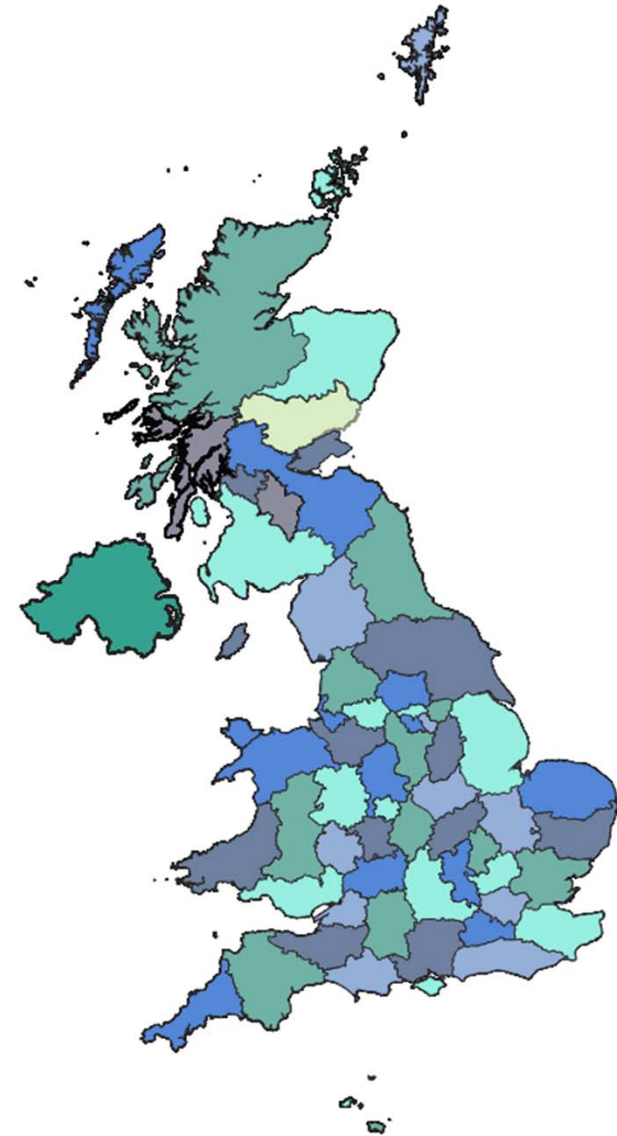
- Curate, preserve, provide access to social science data for reuse
- Data producers: academia, government, commercial, charities
- Funded by UK Research and Innovation – ESRC
- Data management advice, guidance, training for data creators
- Support for data users of the service
- Information about the use to which data are put

ukdataservice.ac.uk



About ALER

- Associative Records (
- Represent
- “To encourage between (ideas, col joint tools



Some terms

- Creative Commons
 - A not for profit organisation that creates easy to use and flexible licences based on copyright law
 - Use common language and standard terms to replace bespoke T&Cs.
 - Available in human-readable and machine-readable forms
 - Appealing for data sharers, as rights are well-clarified
 - Used by GBIF (Veerle to give more details)



Some terms

- Intellectual Property
 - Things that you own that are intangible
- Open Data (ODI)
 - Accessible (ideally via the internet) at no more than the cost of reproduction, without limitations based on user identity or intent
 - In a digital, machine readable format for interoperation with other data
 - Free of restriction on use or redistribution in its licencing conditions.
 - OGL, CC-0, CC-BY



Does copyright apply?

- Under the Copyright, Designs and Patents Act 1988, copyright applies to:
 - original literary, dramatic, musical or artistic works
 - sound recordings, films, broadcasts or cable programmes
 - the typographical arrangement of publications
- Copyright is an intellectual property right assigned automatically to the creator.
- It prevents un-authorised copying and publishing of an original work.
- Copyright will not cover the underlying facts, ideas or concepts, but only the particular way in which they have been expressed.
- Many research outputs such as spreadsheets, publications, reports and computer programs fall under literary work and are therefore protected by copyright.
- Information being in the public domain (e.g. online) does not mean copyright does not apply

Are biodiversity data / records facts or works?

Crawley, D. (2018) Derek Crawley Images. NBN Atlas.
https://records.nbnatlas.org/occurrences/search?q=data_resource_uid:dr1751#tab_record_images

Human observation of *Halichoerus grypus* | Grey Seal recorded on 2015-12-04

Flag an issue Contact data provider

Licence: CC-BY-NC

Date loaded: 2019-01-19

Date last processed: 2019-10-09



Overview

Occurrence ID	2792648
Basis of record	Human observation Supplied basis: "Occurrence"
Scientific name	<i>Halichoerus grypus</i> (Fabricius, 1791) - Grey Seal
Licence	CC-BY-NC
Rights holder	The Mammal Society and Biological Records Centre
Record Date	2015-12-04 (Day)
Locality	Invereshie And Inshriach
Location	Grid Reference: NJ018652 Latitude: 57.66639 Longitude: -3.64692
Recorded by	Ford, Andy
Identified by	Ford, Andy
Identification verification status	Accepted

Dataset

Data partner	The Mammal Society
Data resource	National Mammal Atlas Project, online recording
Collection code	iRecord Cairngorms NP

Abundance

Individual count	5
	2015-12-04

NBN Atlas (2019)

232 results for Data resource: Derek Crawley Images

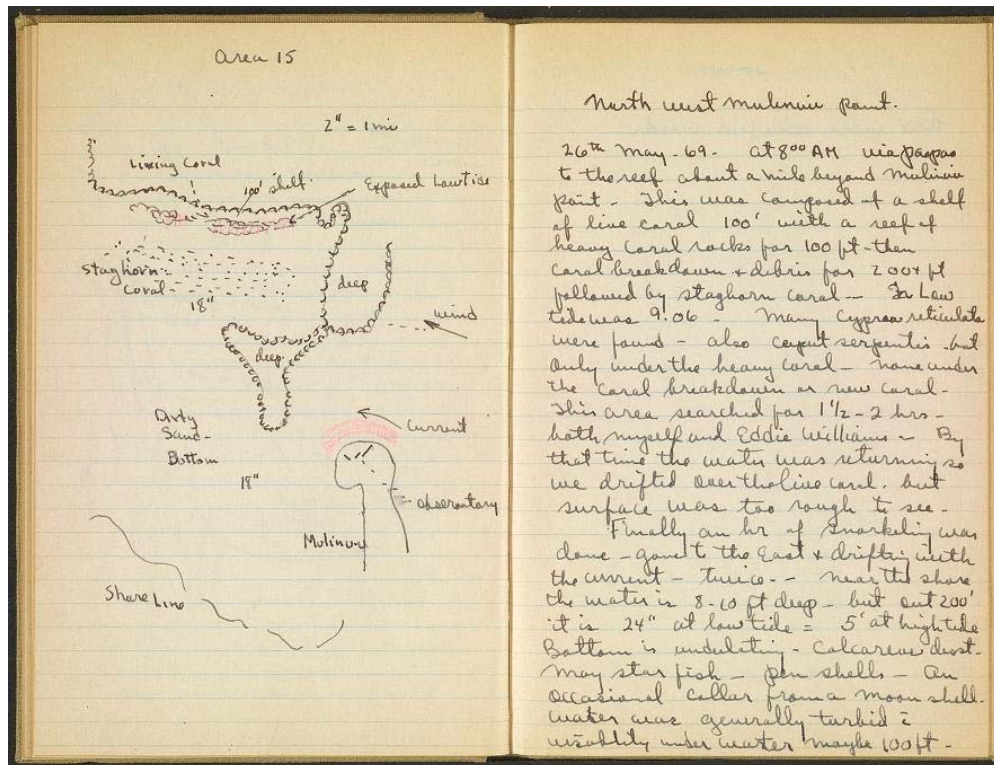
Some of the displayed records may not be available for commercial use. Please check the licence conditions and non-commercial use guidance [here](#)

Records Map Charts Record images

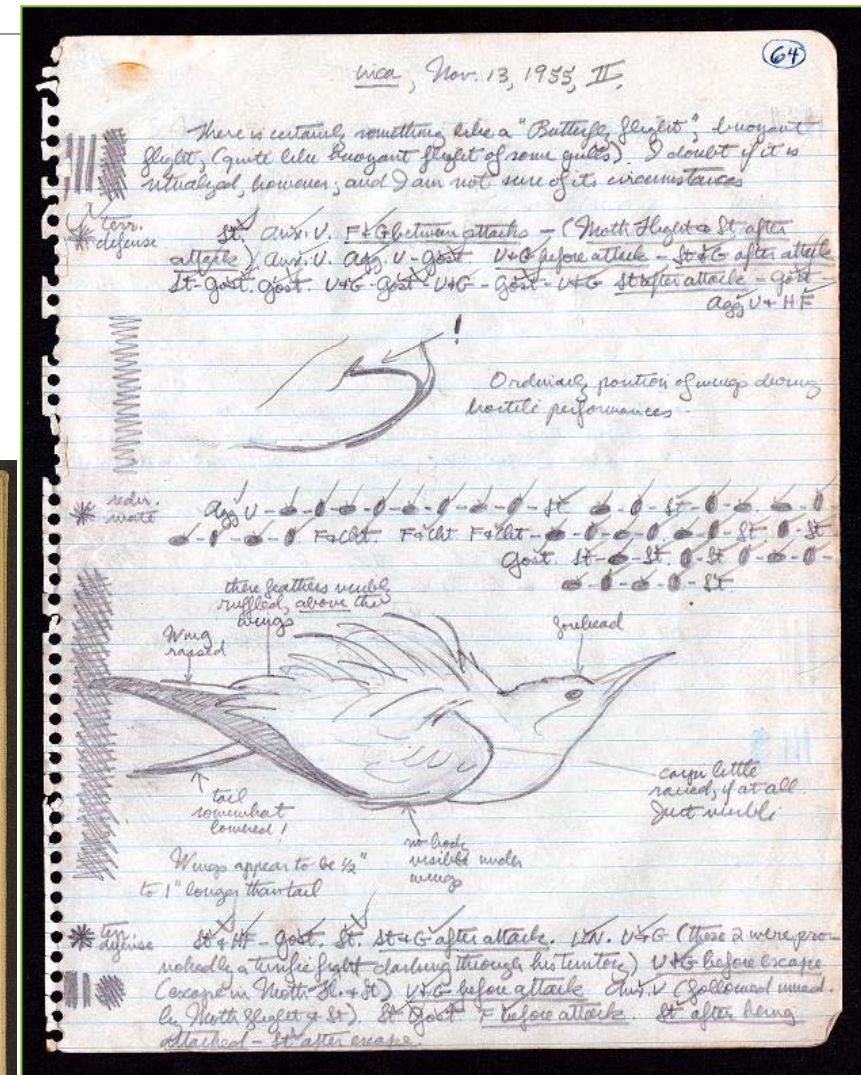
Images from occurrence records



Are biodiversity data / records facts or works?

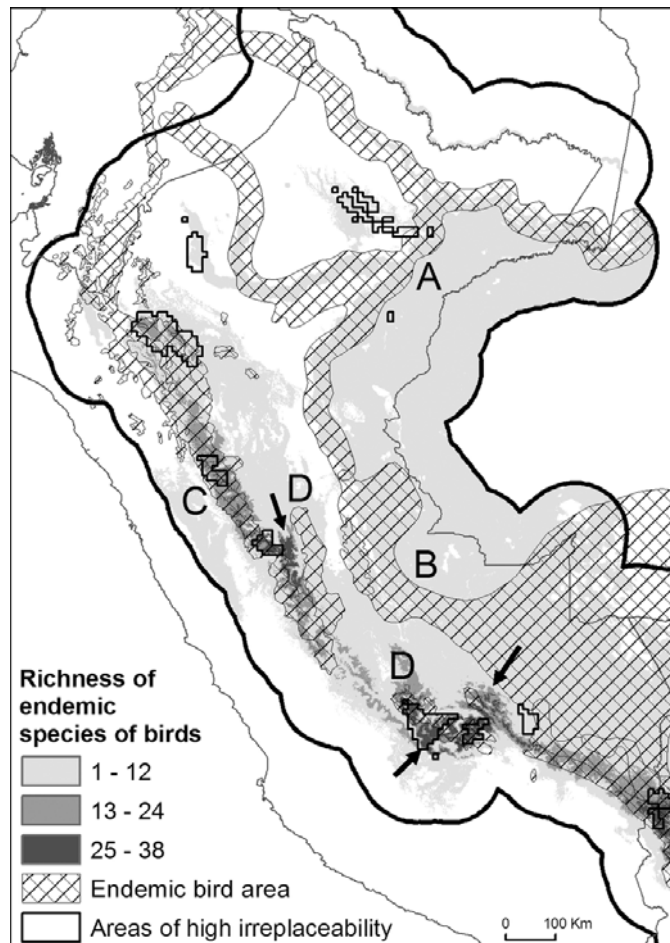


Lambrechts, A.E. (1969) Samoa, April-August 1969. Fieldbook. Smithsonian Field Books collection. Biodiversity Heritage Library. DOI 10.5962/bhl.title.142145



Moynihan, M.H. (1955). Gull Notes (2 of 3). Smithsonian Field Books collection. Biodiversity Heritage Library. DOI 10.5962/bhl.title.96923

Are biodiversity data / records facts or works?



country	site	treatment	hive	block	r0_date	bees_r0	male_brood_r0	egg_stage_r0	larvastage_r0	pupastage_r0	nectar_r0	pollen_r0
G	G1	C	65	1	17/04/2015	7425	0	3120	6960	12720	5760	960
G	G1	C	66	1	17/04/2015	10725	1200	3120	7680	18000	12000	8160
G	G1	C	67	1	17/04/2015	11625	0	5280	7200	17760	11520	3360
G	G1	C	68	1	17/04/2015	11850	1440	4080	5520	18000	4800	4560
G	G1	C	69	1	17/04/2015	12525	0	5280	8160	13920	17040	6480
G	G1	C	70	1	17/04/2015	12600	0	1440	2640	13200	12720	11520
G	G2	ctd	49	1	17/04/2015	6975	0	3600	6480	12240	15600	7680
G	G2	ctd	50	1	17/04/2015	9525	0	2400	5040	12960	12960	5040
G	G2	ctd	51	1	17/04/2015	9600	0	4800	6240	8880	11040	8160
G	G2	ctd	52	1	17/04/2015	10500	0	1920	5280	12000	14880	10080
G	G2	ctd	53	1	17/04/2015	12000	240	3120	8640	16560	6000	2640
G	G2	ctd	54	1	17/04/2015	16875	0	3600	10320	18000	13440	6240
G	G3	tmx	57	1	17/04/2015	6900	0	3840	6240	11040	14400	10080
G	G3	tmx	58	1	17/04/2015	9600	0	5040	5520	13440	9120	6480
G	G3	tmx	59	1	17/04/2015	11400	720	6000	5040	18000	3600	3360
G	G3	tmx	60	1	17/04/2015	12900	480	6240	4560	19680	13440	3120
G	G3	tmx	61	1	17/04/2015	13350	960	2400	7680	19200	12000	5760
G	G3	tmx	62	1	17/04/2015	13575	0	4080	4560	15120	16560	13920
G	G4	C	1	2	17/04/2015	10425	1680	2400	3120	9120	10560	10560
G	G4	C	2	2	17/04/2015	11025	240	2160	7200	17040	15840	9120
G	G4	C	3	2	17/04/2015	11325	0	3840	4320	11760	6240	3840
G	G4	C	4	2	17/04/2015	13425	960	4080	12000	16320	10800	6960

Woodcock, B.A.; Knäbe, S.; Jakel, K.; Scrimshaw, P.; Bullock, J.M.; Shore, R.F.; Heard, M.S.; Pereira, M.G.; Redhead, J.; Ridding, L.; Dean, H.; Sleep, D.; Henrys, P.A.; Peyton, J.; Hulmes, S.; Hulmes, L.; Genersch, E.; Beckmann, B.; Mitschunas, N.; Webb, J.A.; Pywell, R.F. (2017). Population responses of honeybees to oilseed rape neonicotinoid seed treatments in Hungary, Germany and the UK. NERC Environmental Information Data Centre. <https://doi.org/10.5285/eac530fe-54ad-4570-83d3-c59e70c0af9d>
BibTeXRIS

© Syngenta Ltd., Bayer CropScience, NERC (Centre for Ecology & Hydrology)

Young, Bruce & Franke, Irma & P.A.Hernandez, & Herzog, Sebastian & Paniagua, Lily & Tovar Ingar, Carolina & Valqui, Thomas. (2009). Using Spatial Models to Predict Areas of Endemism and Gaps in the Protection of Andean Slope Birds. The Auk. 126. 554-565. 10.1525/auk.2009.08155.

Are biodiversity records facts or works?

- Single observations
- Photographs
- Videos
- Sound recordings
- Written observations
- Maps

Do not ignore, but manage copyright



Copyright

- Joint copyright:
 - Data/works created by multiple people
 - Derived data/works
- Contracts
 - Depending on the employer, your contract may state that any works created during employment are the intellectual property of the employer
 - Even so, employers may be OK for employees to be listed as copyright holder, especially in research

Database right

- If information is structured in a database, the structure acquires a database right, alongside the copyright in the content of the database.
- The structure is protected, not the content.
- Legally, a database is a collection of independent works arranged in a systematic or methodical way.
- A database may be protected by both copyright and database right.
- The database must be the result of substantial intellectual investment in obtaining, verifying or presenting the content in an original manner. Simply entering facts into a spreadsheet does not count as substantial effort.
- The database right is an automatic right and protects databases against the un-authorised extraction and reuse of the contents.

Data repositories

- Most data repositories operate a system of not acquiring any copyright ownership in the data
- The creator (or funder) of the work/data is listed as the copyright owner
- The data repository then distributes the dataset through a licence agreement



Duration of rights

Type of work	Copyright duration
Literary and artistic works	70 years from the end of the year of the death of creator
Sound recordings	50 years from date of creation
Typographical arrangements	25 years from date of publication
Crown Copyright	50 years from date of publication or 125 years from date of creation
Database right	15 years from year of completion

Fair dealing

- Works/data can be copied, used, published etc. for non-commercial teaching or research purposes, private study, criticism or review without infringing copyright, provided that the owner of the work is sufficiently acknowledged
- This only applies to literary, dramatic, musical or artistic work, not to films or recordings
- Also applied to database right
- An acknowledgement should give credit to the source used, the distributor and the copyright holder

Also no copyright permission needed if:

- Copyright has expired
- Free-to-use licence applies, e.g. Creative Commons

Copyright permission

- Trace copyright owner
- Ask permission, specifying the use
- If copyright holder cannot be traced; this does not have to exclude reuse
- How other areas handle this:
 - Orphan works register for books, music, paintings, films; UK Intellectual Property Office can licence on behalf of absent copyright holder, if a diligent search has failed to trace/locate the copyright holder
 - Reuse of legacy medical data without consent: ethics committee can consider the case, weighing potential risks to people versus the benefits to society

Other ownership aspects

- Traditional knowledge rights
- Ethics of intellectual property in citizen science
 - Acknowledge effort of recorders, though citation, especially those spending considerable time
 - It may not be practical to name all contributors in a citation as 'authors' and it will be easier or more logical to cite through the organization
 - Individual contributions can still be acknowledged as contributors of a dataset.

Licensing data

- A licence agreement is a legal arrangement that sets out what a user can or cannot do with the data
- Variety of licence types: [Licence selector](#) available

Choose a License

Answer the questions or use the search to find the license you want

↺ Start again



What do you want to deposit?

Software

Data

Search for a license...

Public Domain Mark (PD)

The work identified as being free of known restrictions under copyright law, including all related and neighboring rights.

Publicly Available



Public Domain Dedication (CC Zero)

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

Publicly Available



OPEN

DATA

Creative Commons Attribution (CC-BY)

This is the standard creative commons license that gives others maximum freedom to do what they want with your work.

Publicly Available














































OPEN

DATA

Creative Commons Attribution-ShareAlike (CC-BY-SA)

<http://ufal.github.io/public-license-selector/>

Rights, licensing and usage conditions

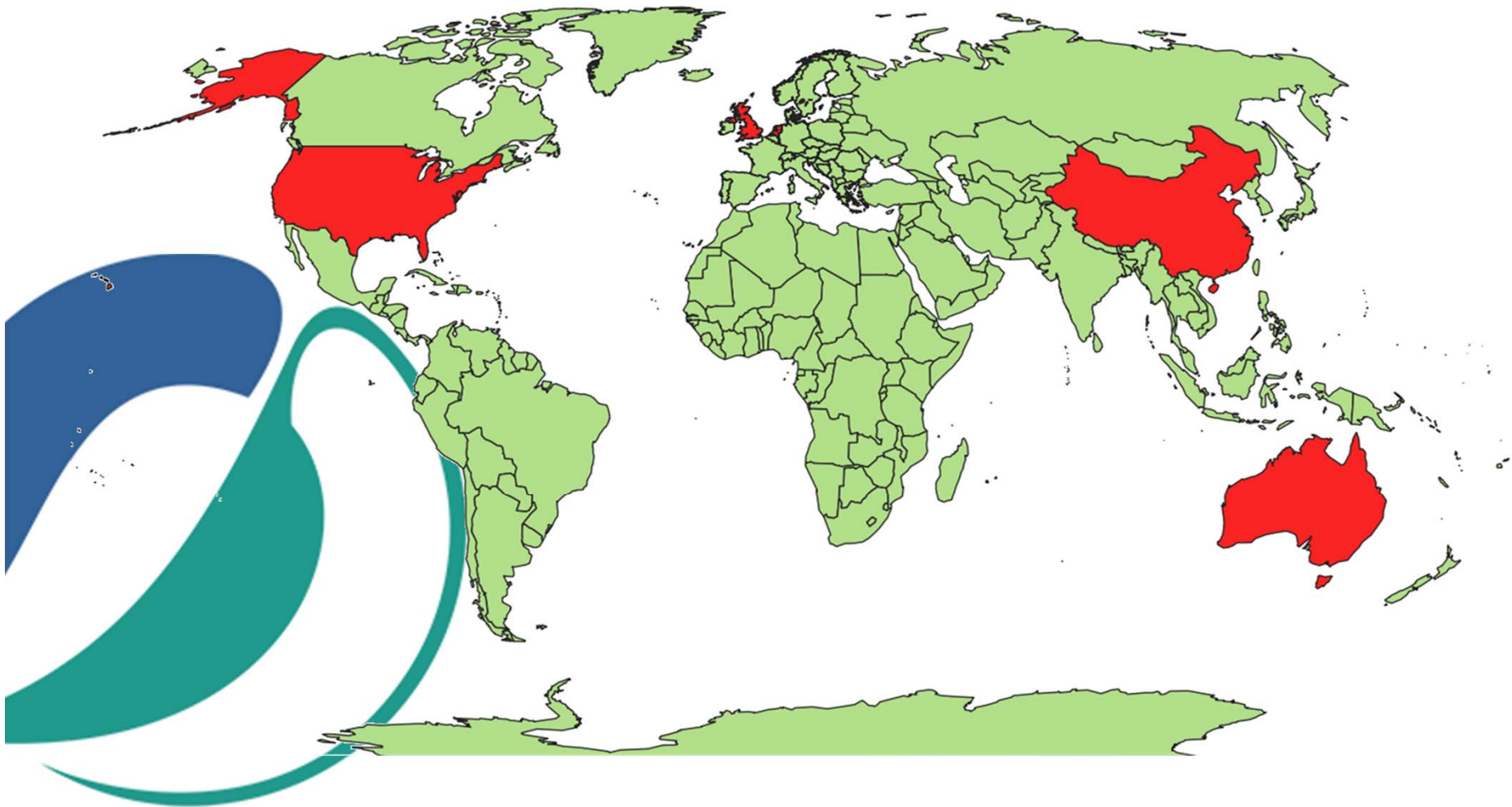
 creative commons LICENSES		Copy & Publish	Attribution Required	Commercial Use	Modify & Adapt	Change License
	Public Domain					
	BY Attribution					
	BY-SA Attribution ShareAlike					
	BY-ND Attribution NoDerivs					
	BY-NC Attribution NonCommercial					
	BY-NC-SA Attrib NonComm ShareAlike					
	BY-NC-ND Attrib NonComm NoDerivs					

Service

Creative Commons recommended by GBIF

- A CC licence cannot be revoked once it has been issued
- CC0
 - completely open CC licence
 - copyright owner waives all its rights, including the database right and the right to be identified as the creator
- CC BY
 - attribution
- CC BY NC
 - attribution
 - non-commercial

Attitudes around the World



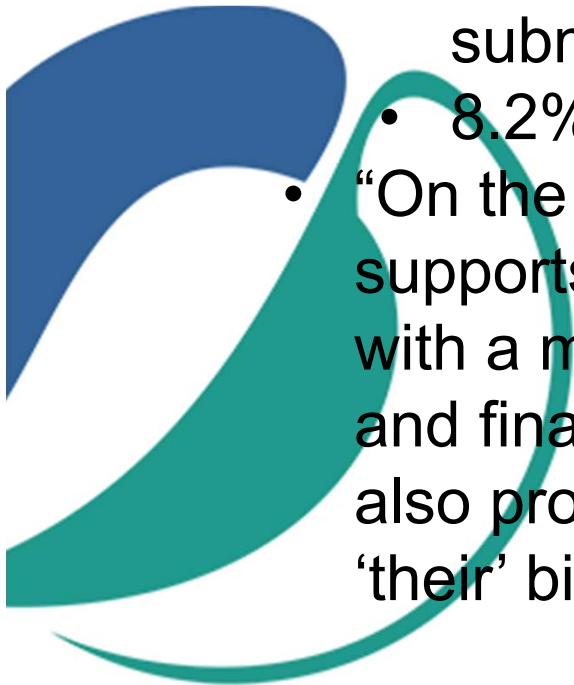
Yang and Chan (2015), China

- A study into geckos. Detailed location information not published because of risk of poaching for the pet trade



Ganzevoort et al (2017), Netherlands

- Surveyed citizen scientists' attitudes to data sharing.
- “Who owns submitted records?”
 - 48.7% “the data are nobody’s property (public good)”
 - 27.4% owned by the organisation they were submitted to,
 - 8.2% consider these data as personal property
- “On the other hand, only a small minority (12.3%) supports completely unconditional use of this data, with a majority specifying rules around attribution and financial gain, and a majority of the volunteers also professed an interest in tracking the use of ‘their’ biodiversity data.”



Martin, Christidis and Pecl (2016), Australia

- Survey of volunteer marine surveyors
- Surveyors demonstrated a very high willingness to share data with research organisations (although not 100%)
- “Private research companies or consultants” ranked lowest in this research.



Bowser et al (2017), USA

- “Volunteers value open data, and even find bragging rights in broad information dissemination: “I’d like to know if my data is being used by other projects. In fact I’d tell my wife and kids.” For many volunteers, personal motivations for sharing data outweigh the risk associated with ceding their privacy”



Groom et al. (2017), GBIF

- “Many volunteers retain a **sense** of ownership of data they have submitted, particularly if significant investment was required, and will withhold them if they perceive those data may be used inappropriately.”
- “The provision of data by citizen scientists to organizations undertaking biological recording or monitoring is underpinned by **trust**, which requires those organizations to take account of volunteer perspectives when making decisions about data sharing, and to be open about potential data uses.”



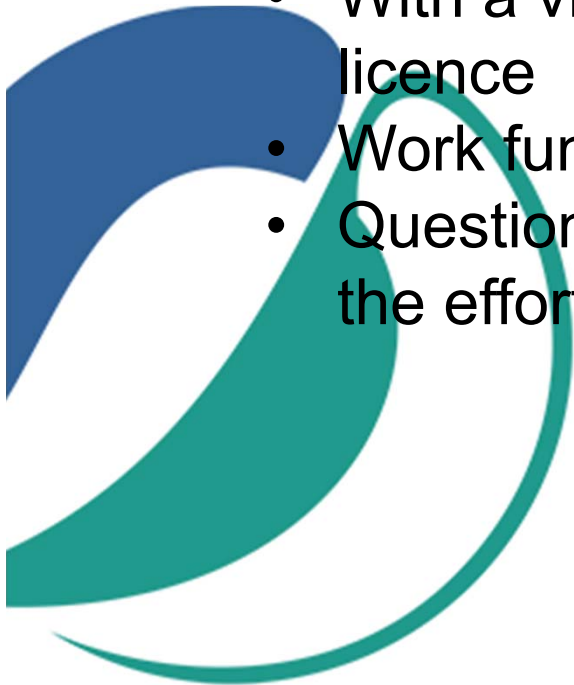
Ellis and Waterton (2005), UK

- “What we find interesting in the light of this exchange of knowledge objects is that knowledge circulation and representation both engender and are framed and enabled by a series of **imagined contracts** which are presently in a state of flux as amateurs and professionals negotiate their contours in a time of change and experiment introduced by participation.”



Sharing Sussex Beetle Records

- Digitised historic coleoptera records
- Digitised contemporary records
- Contacted all the recorders they could, to seek permission to share records
- With a view to placing on the Atlas under an open licence
- Work funded by DEFRA
- Question, should it be officially recorded somewhere, the efforts made to contact recorders?



An 'as-complete-as-possible' Sussex Beetle Records dataset covering the period to March 2017

SUSSEX BIODIVERSITY RECORD CENTRE

DESCRIPTION

Beetle records from Sussex shared by various recorders, most notably Peter Hodge - the Sussex County Recorder for Beetles - who has contributed over 65,000 records to this dataset and invested huge amounts of time in verifying the Sussex beetle records. Grant funding from Natural England enabled SxBRC to conduct a large-scale consultation with Sussex beetle recorders and secure the necessary permissions to publish beetle recorders' records at full resolution under a CC-BY licence. This dataset does not include ALL Sussex beetle records held by SxBRC, only those records where we have appropriate permission to share the records under a CC-BY licence, in accordance with the NBN guidance for

corded beetles in Sussex and would like your records to appear in

: Sussex Biodiversity Record Centre: bobforeman@sussexwt.org.uk
, see www.sxbrc.org.uk

N
he City of Brighton & Hove.

s who will have been recording for a range of different purposes



92,465 records

Data access

- View records
- Download usage stats
- Alert me about new records
- Alert me about annotations

100% records have verified identifications

Citations

- <https://doi.org/10.15468/nabhqz>
- 7 citations for these data

Licence

Creative Commons with Attribution 4.0 4.0



Temporal scope

Conclusions

- The terms IPR and copyright are hardly ever used (if at all).
- Ownership tends to be “sense of ownership”
- Recorders’ own attitudes seem to differ depending on circumstances.
- Potentially a difficult environment to navigate when making decisions on management.



Conclusions

- Historical records / data
 - some searching/tracing of ownership (copyright) unavoidable
 - ask permission to share/use; this is not impossible
 - some people you can find, others you can't
 - more complex the more you go back in time
- Many recorders will happily agree to their data being shared; some will disagree; some will not respond
- Concerns are not necessarily about the ownership of the records (possessiveness), but rather about **having a say in who can use data and how they can be used**, e.g. excluding commercial use, excluding public access to rare/sensitive species data, fear of damages being done to species
- Anticipate this by making provisions to enable such conditions or exclusions

Recommendations

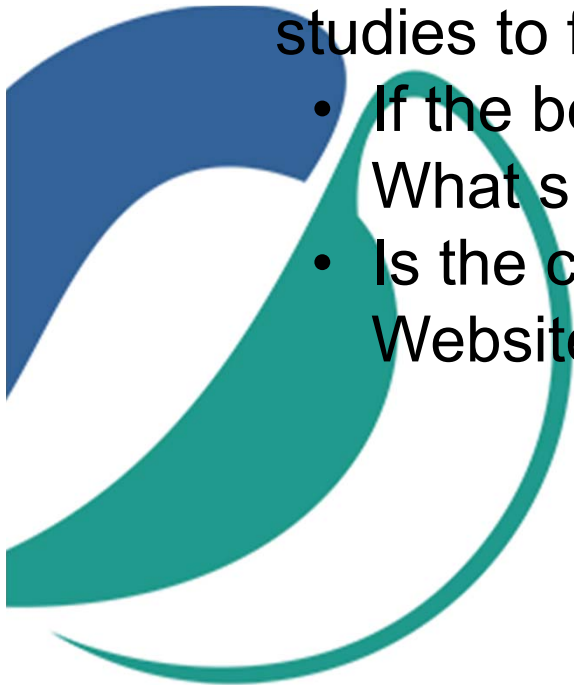
- At times difficult to determine whether or not copyright applies to data/records
- Do not ignore, but manage copyright
- Irrespective of copyright / ownership, recorders will most likely want to have their efforts acknowledged
- Use common sense in citation: recorders/contributors vs organisation as 'authors'
- Individual contributions can still be acknowledged as contributors of a dataset.
- A clear understanding how records/data can be shared/used, i.e. terms and conditions recorders can agree/disagree to, makes things easier
- Clear terms and conditions are very helpful for data contributors and data users
- Use explicit licences to make it clear how data can be used
- Risks need to be weighed up against benefits
- Have best practice guidance for data / records

References

- Yang, J. H., & Chan, B. P. (2015). Two new species of the genus *Goniurosaurus* (Squamata: Sauria: Eublepharidae) from southern China. *Zootaxa*, 3980, 67–80. <https://doi.org/10.11646/zootaxa.3980.1.4>
- Ganzevoort, W., van den Born, R. J. G., Halffman, W., & Turnhout, S. (2017). Sharing biodiversity data: Citizen scientists' concerns and motivations. *Biodiversity & Conservation*, 26, 2821–2837. <https://doi.org/10.1007/s10531-017-1391-z>
- Martin, V. Y., Christidis, L., & Pecl, G. T. (2016). Public interest in marine citizen science: Is there potential for growth? *BioScience*, 66, 683– 692. <https://doi.org/10.1093/biosci/biw070>
- Bowser, A., Shilton, K., Preece, J., & Warrick, E. (2017). Accounting for privacy in citizen science: Ethical research in a context of openness. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing (CSCW '17)*. ACM, New York, NY, USA, 2124–2136. <https://doi.org/10.1145/2998181.2998305>
- Groom, Q., Weatherdon, L., & Geijzendorffer, I. R. (2016). Is citizen science an open science in the case of biodiversity observations? *Journal of Applied Ecology*, 54, 612–617. <https://doi.org/10.1111/1365-2664.12767>
- Ellis, R., & Waterton, C. (2005). Caught between the cartographic and the ethnographic imagination: The whereabouts of amateurs, professionals, and nature in knowing biodiversity. *Environment and Planning D: Society and Space*, 23, 673–693. <https://doi.org/10.1068/d353t>

Discussion

- We have three case studies for discussion, based on real scenarios.
- Please feel free to introduce your own case study. We can sort this out over the tea break.
- After tea we will divide into groups and assess the case studies to find out:
 - If the best practice was followed in each in case. What should have been done differently, if anything?
 - Is the current best practice guidance on the NBN Website sufficient or does it need amending?



Questions

Veerle Van den Eynden

veerle@essex.ac.uk

Tom Hunt

Tom.Hunt@alerc.org.uk

