

Weekly Notes - Bat Eco-Interactions Working Group

Meeting zoom: <https://deic.zoom.us/j/7588563077>

Meeting time: 5pm CET / 11am ET / 8am PT

Whimsical board for brainstorming:

<https://whimsical.com/biodiversity-knowledge-hub-WyqgfhD1V4rMg5W7d6CnYY>

GloBI Project page - <https://globalbioticinteractions.org/gbatnet>

23 July 2025	3
16 July 2025	4
9 July 2025	6
2 July 2025	8
25 June 2025	9
18 June 2025	10
11 June 2025	11
4 June 2025	13
Present: Aja, Nancy, Cecilia, Jorrit, Quentin, Anna, Jelle	13
28 May 2025	14
Present: Aja, Nancy, Kendra, Jorrit, Donat, Nate, Connor, Anna, Hernani	14
Agenda -	14
21 May 2025	20
Present: Aja, Nancy, Cecilia, Kendra, Jorrit, Jelle, Donat (and Vanessa), Ariadna, Hernani, Anna, Nate, Cullen, Quentin, Connor	20
Agenda -	20
14 May 2025	23
Present: Aja, Nancy, Cecilia, Kendra, Jorrit, Jelle, Donat (and Vanessa and Patrick), Ariadna, Hernani, Anna, Nate, Cullen, Quentin, Connor	23
Agenda:	23
7 May 2025	26
Present: DeeAnn, Aja, Nancy, Cecilia, Kendra, Jorrit, Jelle, Donat, Ariadna, Hernani, Anna	26
Aja - suggests fuzzy matching -	28
30 Apr 2025	28
Present: Donat, Aja, Anna, Hernani, Nate, Cullen, Kendra, Jorrit, Jelle	28
Agenda:	28
23 Apr 2025	31
Present: Donat, Iroro, DeeAnn, Aja, Anna, Hernani, Quentin, Iroro	31
Agenda:	31
• announcements	31
• review IUCN abstracts are due in one week! LINK to abstract draft.	31
16 Apr 2025	32
Present: Donat, Nancy, Jelle, Nate, Iroro, Kendra, Aja, Anna, Ariadna, Cecilia, Hernani	32

Agenda: IUCN abstracts are due in two weeks! LINK to abstract draft.	32
9 Apr 2025	34
2 Apr 2025	1
MDD taxonomic subcommittees	1
Text mining for interactions aspects	1
26 Mar 2025	1
19 Mar 2025	1
12 Mar 2025	1
5 Mar 2025	1
26 Feb 2025	1
19 Feb 2025	1
12 Feb 2025	1
5 Feb 2025	1
29 January 2025	1
22 January 2025	1
15 January 2025	1
18 December 2024	1
11 December 2024	1
4 December 2024	1
20 November 2024	1
13 November 2024	1
6 November 2024	1
30 October 2024	1
23 October 2024	1
16 October 2024	1
9 October 2024	1
2 October 2024	1
25 September 2024	1
18 September 2024	1
11 September 2024	1
4 September 2024	1
28 Aug 2024	1
21 Aug 2024	1
14 Aug 2024	1
7 Aug 2024	1
31 Jul 2024	1
24 Jul 2024	1
17 Jul 2024	1
10 Jul 2024	1
3 Jul 2024	1
26 June 2024	1

19 June 2024	1
12 June 2024	1
05 June 2024	1
10 May 2024	1
8 May 2024	1
1 May 2024	1
24 April 2024	1
17 April 2024	1
3 April 2024	1
27 March 2024	1
20 March 2024	1
6 March 2024	1
28 Feb 2024	1
21 Feb 2024	1

20 August 2025

Present: Nancy, DeeAnn, Donat, Nate, Aja, Jorrit, Cullen

Agenda:

- IBRC recap
- others?

Ant-eating bats

- Rhinopoma that do termite homes
 - Tie reproduction to termite emergences
- Maybe some flying ants
- <https://www.globalbioticinteractions.org/?interactionType=interactsWith&sourceTaxon=Chiroptera&targetTaxon=Formicidae> for examples of Chiroptera interacting with Formicidae (Ants)
- <https://batbase.org/interaction/15096> for example

BHL pdfs

- Jorrit worked on the upload for this to Zenodo – then Biodiversity PMC
 - <https://biodiversitypmc.sibils.org/>
 - <https://zenodo.org/communities/bhl-blr> 342,436 deposits as of 2025-08-20

LLM usage for data extraction

- Nate Layman will join in two weeks (Sep 3rd)
- Substantial progress, it sounds like
- Doing 1 paper at a time, checking against hand-extracted papers – vetting this process
- Fixing glitches by adjusting the prompts for LLM

IBRC recap

- Everyone very interested – gave out postcards with info
 - No one has yet filled out form to join
- Posters on BatLit
- 650 people attended – fewer Americans, many had lost grant funding
- Cullen was part of symposium on seed dispersal (incl Latin American folks)
- GBatNet community feeling / collaboration
 - Clearly present at the meeting
- Some folks are still traveling – email address should work
- Now ready to focus on the BTA-to-IUCN taxonomic alignment paper going

NASBR in October

- Edmonton, Canada
- DeeAnn and Nancy are going
- Have 220+ people registered

1500th bat species

- *Pipistrellus etula*
- Link to PDF: <https://slack-files.com/T014Y4DLN90-F092AHNBTC0-577f14458f>
- BatNames and MDD coordinated for this
 - Toward transition
 - Once the MDD brings subspecies into the fold
 - Junior synonyms of all the subspecies >> would be lost if we just flattened all to synonym level
 - Nancy and Rob retiring in June 2026 – moving to Colorado

Wanting to have infrastructure to connect MDD to BatLit / Zenodo

- Citing the machine-actionable version of the taxonomic treatment
-

23 July 2025

Present: Jelle, Aja, Cecilia, Hernani, Jorrit, Quentin, Nate, Nancy, Anna, Cullen

Agenda:

- IBRC planning
 - Tigga wants 1 slide for achievements of this group
 - Giving a talk about activities of different working groups in GBatNet
- BatLit update

- growing BatLit curatorial backlog (see <https://github.com/bat-literature/bat-literature.github.io/issues>) and setting expectations for them getting handled by BatLit curator

BatLit website

- Changing the navigation bar to be more consistent, allow for navigation
- Using the “edit” page to do these changes directly
- Using Github issues in a more direct and repeatable way – documenting clearly – asking for resources
 - >> Jorrit helping to guide Aja through this process

BTA manuscript

- Targeting PLOS Biology – make it broader than just bat taxonomy – but can use the example of the IUCN RedList for showing the impact
- Need figures at the beginning to broaden it out
 - E.g., Problem figures – taxonomic splits
- Section in discussion – about WHY this hasn’t been solved before
 - Tracking inputs and outputs
 - Exact files – inputs of data
 - Load into Github – get the exact digital signature of a given file
 - >> describing data, where you got it
 - Possible problems – IUCN versions – can be transcription errors that have been interpreted subsequently
 - >> editing is part of the workflow
 - Establishing a Github of input and outputs – big part of what’s needed
- What’s the audience – main audience
 - Bat scientists
 - Data scientists
 - Should that be 2 papers?
 - If so, one gets lost
 - Bring attention to the problem + the solution
 - Including what can be automated vs not

Barriers to access → access to literature is one (BatLit addresses) → expert knowledge to interpret the literature (BTA helps lower that bar) → Use of biodiversity knowledge to questions

16 July 2025

Present: Jelle, Donat, Aja, Kendra, Cecilia, Hermani, Jorrit, Iroro, Quentin

Agenda:

- BatLit update
- BTA
- ...

Progress on the batlit.org website. Nate and Aja have administrative access to batlit.org domain name registration. Please check out and comment - I guess to Aja

(Jorrit: note that content (e.g., text/images) is in, navigation and styles are under development. If you are super excited about skinning/theming websites, please feel free to contribute at <https://github.com/bat-literature/bat-literature.github.io> , a vanilla jekyll/github pages website)

Aja shared that BTA is still updating, making changes. Aja shared that it currently takes her 3 days to add a new version of a taxonomic authority into BTA . She hopes to work with Nate Layman to automate the process. Nate is currently still working on this bat co-roosting text extraction. Some mention of putting BTA in a relational database.

Manuscript [A Standardized Review of Names Across Multiple Taxonomic Authorities: an Example from Bats \(Mammalia, Chiroptera\)](#): how to proceed: use the Zoom whiteboard

Preparation of the Australia meeting

Donat talk about their work with Biodiversity Heritage Library.

BHL used Rod Page's BioSTOR 380k pdfs. 150k have cross ref doi, 100-150k pdf with no doi, others include 60k SciELO (south america: Brazil, Chile, Colombia, Mexico, Argentina).

BHL-BLR aims to take all of BHL into Zenodo

Mapping of all the license in BHL to Zenodo vocabulary based on standardized [SPDX licences](#) (Jorrit: for context and discussion, see <https://github.com/bio-quoda/preston/issues/325>).

Questions/Concerns:

1. Which DOI do we use? Original DOIs or Zenodo DOIs? We decided to use the original DOIs for social / political reasons.
2. PDF Metadata Quality - BHL evolved on ways to add metadata into the PDFs they made available. PDFs have metadata embedded into the pdf file.
3. Get BHL content into a separate platform to ensure continued access, making BHL content more resilient to host/funding issues.

Long term goals - extract treatments, and find ways to extract the knowledge they contain in machine readable formats. Hackathon.

```

},
"links" : [
    "self" : "https://zenodo.org/api/records?allversions=False&page=1&q=alternate.identifier%3A%22urn%3Alsid%3Abiodiversitylibrary.org%3Apart%3A405078%22&size=25&sort=bestmatch&type=publication"
]
}]
[main] INFO bio.guoda.preston.ZenodoUtils - executing query [<https://zenodo.org/api/records?all_versions=false&q=alternate.identifier=biodiversitylibrary.org%3Apart%3A405078%22&type=publication>] done.
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - POST https://zenodo.org/api/deposit/depositions HTTP/1.1
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - PUT https://zenodo.org/api/deposit/depositions/15981506 HTTP/1.1
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - upload [<hash://md5/f7e6176f68da79326248c5071be2de39>] as [bhlpart405078.pdf]
.
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - PUT https://zenodo.org/api/files/f8f67848-a8d8-4f58-bcc8-79407041d14b/bhlpart405078.pdf HTTP/1.1
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - calculating content length for [hash://md5/f7e6176f68da79326248c5071be2de39]...
counting bytes: [hash://md5/f7e6176f68da79326248c5071be2de39] 1 MB at 16.86 MB/s [main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - calculating content length for [hash://md5/f7e6176f68da79326248c5071be2de39] done.
[https://linker.bio/hash:...8da79326248c5071be2de39] 1 MB at 16.87 MB/s completed in < 1 minute
[main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - attempting to dereference [hash://md5/f7e6176f68da79326248c5071be2de39]
uploading: [hash://md5/f7e6176f68da79326248c5071be2de39] 98.3% of 1 MB at 22.59 MB/s ETA: < 1 minute [main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - copying [hash://md5/f7e6176f68da79326248c5071be2de39] complete.
[https://linker.bio/hash:...8da79326248c5071be2de39] 1 MB at 22.98 MB/s completed in < 1 minute
[main] INFO bio.guoda.preston.Zenodo.ZenodoMetadataFileStreamHandler - upload [<hash://md5/f7e6176f68da79326248c5071be2de39>] as [bhlpart405078.pdf] finished.
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - POST https://zenodo.org/api/deposit/depositions/15981506/actions/publish HTTP/1.1
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - executing query [<https://zenodo.org/api/records?all_versions=false&q=alternate.identifier:%22urn%3Alsid%3Abiodiversitylibrary.org%3Apart%3A194482%22&type=publication>]...

```

```

[main] INFO bio.guoda.preston.ZenodoUtils - PUT https://zenodo.org/api/files/1710-48b8-9ab2-ea9fd738c466/bhlpart406872.pdf HTTP/1.1
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - calculating content length for [hash://md5/8a809e2371de423df6187a229eea729e]...
counting bytes: [hash://md5/8a809e2371de423df6187a229eea729e] 1 MB at 19.45 MB/s [main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - calculating content length for [hash://md5/8a809e2371de423df6187a229eea729e] done.
[https://linker.bio/hash:...1de423df6187a229eea729e] 1 MB at 19.76 MB/s completed in < 1 minute
[main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - attempting to dereference [hash://md5/8a809e2371de423df6187a229eea729e]
uploading: [hash://md5/8a809e2371de423df6187a229eea729e] 98.1% of 1 MB at 23.09 MB/s ETA: < 1 minute [main] INFO bio.guoda.preston.Zenodo.DereferencingEntity - copying [hash://md5/8a809e2371de423df6187a229eea729e] complete.
[https://linker.bio/hash:...1de423df6187a229eea729e] 1 MB at 23.05 MB/s completed in < 1 minute
[main] INFO bio.guoda.preston.Zenodo.ZenodoMetadataFileStreamHandler - upload [<hash://md5/8a809e2371de423df6187a229eea729e>] as [bhlpart406872.pdf]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - POST https://zenodo.org/api/deposit/depositions/15981530/actions/publish HTTP/1.1
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo endpoint: [https://zenodo.org]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - zenodo token : [*****]
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - executing query [<https://zenodo.org/api/records?all_versions=false&q=alternate.identifier=biodiversitylibrary.org%3Apart%3A405717%22&type=publication>]...
[main] INFO bio.guoda.preston.Zenodo.ZenodoUtils - got query result [{ "hits" : { "hits" : [ ], "total" : 0 }, "aggregations" : { "access_status" : { "buckets" : [ ], "label" : "Access status" }, "resource_type" : { "buckets" : [ ], "label" : "Resource types" }, "subject" : { "buckets" : [ ], "label" : "Subjects" }, "file_type" : { ...

```

9 July 2025

Present: Quentin, Nancy, Jelle, Aja, Anna, Hernani, Cecilia, Jelle, Nate, Jorrit

Agenda ideas

- BTA alignment of BatNames and MDD – trying to get that matchup to zero

- Bat Lit update
- <https://batlit.org> domain name management ownership

Australia meeting – folks are leaving on Jul 28 – meeting is Aug 3-8: <https://www.ibrc.org/>

BTA >>> found that MDD v2.2 mistakenly contains subspecies column – should not have been released with that version as not yet curated

From Donat: "May be Jorrit can give a brief update of the starting upload of the ca 350000 BHL PDFs to BLR, and with it the further upload of them to BiodiversityPMC."

BatLit update:

- How to do the domain handoff
 - Can be tricky – virtual real estate – companies licensed to sell domains
 - namecheap.com is what Jorrit uses – rented batlit.org through that service
 - \$15 per year – to be transferred to batbase.org group
 - DNS servers are a chokepoint in the whole internet
 - Volunteers to help with DNS transfer
 - Cullen, Aja, Nate
 - Also potential to add American Society of Mammalogists as owner / manager / hoster of the service
 - Keeping design as own separate thing
- Search page: <https://batlit.org/refs>
 - Has all data loaded as JSON on the browser side – makes for easy search
- Someone requested to add a publication to BatLit

BHL update (Biodiversity Heritage Library)

- Announced that they are looking for a new funding source
- > 300k articles / books
- Donat has decided to put all in a Corpus similar to as for BatLit
- Decided to do the entire corpus
- All is CC-BY or other open license
- They publish an archive of references in RIS format
- Showing sandbox of the Zenodo upload
 - Explicit about the licenses
 - Links to the BHL landing pages (same data presented in a different way)
 - <https://sandbox.zenodo.org/records/281899>
-

2 July 2025

Present: Quentin, Nancy, Jelle, Aja, Anna, Hernani, Cecilia

- Summer heat!
- Aja updates:
 - Bat Literature
 - Coordinating with Jorrit about domain/server
 - Landing Page
 - Subheadings:
 - Frontpage
 - About
 - Search BatLit
 - Zenodo Collection
 - New Members
 - How to Contribute
 - Will have third party colleague to test drive
 - Edits from group?
 - Cover image:
 - Bat picture vs. fossil
 - Nancy to circulate a picture that we have permission to use
 - Authorship: need to confirm
 - Need to notify all authors that they are included
 - What do we think is “published literature”?
 - We will drop to be more liberal
 - White papers, dissertations,
 - Taxonomists/ICZN are more stringent criteria
 - Botanists have
 - Data paper
 - Confusing term: methods/protocol?
 - Conversation with Jorrit
 - Can keep then replace
 - Combine New Members + Contribute papers subheadings
 - “Get Involved”
 - First request for access Zenodo community and submission of literature
 - “We have an opportunity to improve our BatLit ingestion process with a new request. Can we add this to our agenda on Wednesday? The paper below was provided without any context. I can reach out to her, search for the publication, etc. Let's make these decisions and add them to our protocols together.
 - The record 'EchoVision: Translating Bat Echolocation into Assistive Technologies for the Visually Impaired' was submitted to your community 'Bat Literature Project' by '@PAVITHRA576'
 - Bat Taxonomy
 - May need separate meeting with Nancy/Jelle

- Ready to be published as a new version (v5)
 - Will update here: <https://zenodo.org/records/11193643>
- BatNames edits in internal version
 - 1911 differences between MDD and BatNames
 - Subspecies shouldn't be different
 - We should highlight the species-level differences
 - Subspecies will never not include the species as first two words
 - Quentin to help with re-running
 - Aja to email with links
- Current Manuscript:
<https://docs.google.com/document/d/1sWTh-GQqqJI95xdVbZC1ryZKlpNr7T2vUFHac3dlySE/edit?tab=t.0>
 - Figures
 - Map or Case Study Nate's example of *Miniopterus* splits
 - Quentin, Anna able to help
 - Breaking up sections into new "Tabs" feature of Google Docs
- NASBR <https://nasbr.org/>
 - October 8 -11th
 - Abstracts Due
 - Nancy, Aja planning to go
 - Aja to present on bat-lit
- Do we have a group list/contact information generally?
 - Slack group has 50+ members
 - Google invite list for this meeting lists email (~18 people)
 - Anna to start some internal contact list for our slack

25 June 2025

Present: Aja, Jelle, Anna, Cecilia, Hernani, Nate

BTA dynamics, looking at synonym matchup table

- Comparing BatNames and MDD
- Should have that equal 0 differences at the end of the table
- Want that species-level comparison to be the main one in the main text
 - Can have a subspecies-level comparison in the supplement

18 June 2025

Present: Nancy, Aja, Kendra, Jelle, Anna, Cecilia, Donat, Quentin

[Batlit.org](#) website presentation:

- What is website for editing?
 - Aja to connect with Jorrit
 - Probably:
<https://github.com/bat-literature/bat-literature.github.io/blob/main/index.md>
- Citation: DeeAnna + Kendra should be included
 - What is this for? Website may be different than data paper/Zenodo
 - Anna sent pull request
 - <https://github.com/bat-literature/bat-literature.github.io/blob/main/index.md>
- Layout:
 - Question for group from Aja: One page or multiple tabs?
 - Likely will need a hybrid of two
 - Mobile viewing is a priority
- Automated record counting
 - We currently do this some in Jorrit's data paper:
 - Pages, references, attachments
 - What do these values mean?
 - Example: <https://zenodo.org/communities/biosyslit/>
 - These are Zenodo stats, which require separate script
 - Records, Views, Downloads
 - Contact Donat for connect with Fatima who generates that for BLR
 - Whatever stats, somewhere there should be clear definition
 - Definitions don't need to be on front page, but should be somewhere
 - What is a "published work"
 - "Views" to Zenodo? To Batlit website?
 - An example is here



- Using images from the literature like BHL
 - Donat shared "ocellus" to find images
 - <https://ocellus.info/>
 - Don't use pictures of people/bats
 - Focus on pre-1950 publications
- BTA
 - Aja working on updates as well as tutorial video
- IBRC
 - Bat Lit Poster
 - Stickers!
 - QR code?

- Where do we get this
- Updating Bat Handling project by Quentin + summer students
 - Outlets for small publications suitable for student timeline?
 - Make an iNaturalist project and use the “Journal” option
 - e.g, <https://www.inaturalist.org/projects/project-porchlight/journal>
 - University of Kansas has a One Health Newsletter
 - <https://olathe.k-state.edu/research/one-health-newsletter/>
- Living Data Updates
 - Quentin + Donat finalizing sessions
- BLR updates
- Tail length
 - Prehensile tails?
 - Pangolins, monkeys, marine mammals
 - For species that overlap with ringtail in habitat range - rats?
 - Bats would be bad outgroup
 - bat tails are related to foraging behaviors
 - Its a major portion of the tail membrane
 - Large variation; standard measurement along with total length
 - Mouse bat tails: use as whisker for crevice navigation
 - <https://www.alamy.com/stock-photo-long-tailed-bat-23979816.html>
 - Functions?
 - Movement
 - Flying
 - Positioning during roosting - focus on crevice-dwelling, not social roosts
 - Prop during tree climbing
 - Anti-predation?
 - Primitive: when has it been lost? re-evolved ?
 - Not used for ectoparasite deterrent; not actively used when handling
 - Classic review:
 - <https://paperpile.com/shared/svisy3OPjTqaTXGEwzK46Tg>
 - New: https://paperpile.com/shared/sHYbbLhViROUxRkDX4i_irQ

11 June 2025

Present: Cecilia, Quentin, Anna, Nate, Connor, Cullen, Kendra, Aja

Thesis updates from Anna and Cecilia

IBRC happening in August

- Organized session on seed dispersal (Cullen)
- Aja part of a normal session
- Cecilia in a bats beneficial to agriculture
- Kendra was planning on it – but now not –

EcoHealth journal also shut down

- Authors had their manuscripts held up >> only just learned about it

ASM meeting in Purdue

- Kendra got a hardship grant from ASM to attend (awesome)
- Nate going to ASM also
- Connor too
- Anna not going – crunching on thesis stuff

Living Data in Colombia

- TDWG conference + GEOBON group + GBIF
- More data than policy centric

BatLit landing page

- Aja working in Google Pages for that
- Aiming to get this ready for the IBRC
 - PR push at that meeting – distributing cards with information + QR codes
 - Engaging the community

Getting the Taxonomic Subcommittees for the MDD going

- Also a PR push of sorts at the ASM meeting
- “We could send an open call out along with the form in general for the subcommittees. A QR code is probably easiest as long as we can make it widely accessible.”
- Planning to post to Mammal-L before ASM as well

Anna Ringtail work – citation diving

- https://docs.google.com/presentation/d/1VNX_k8_dmZ6uWgyPYrz0G-HXfCcQPzzMjvCKoGbzorc/edit?usp=sharing
- Snowballing
 - Going from one paper to all relevant citation – to all those papers and their citations – etc
- Could do this as a cumulative curve – accumulation over levels of snowballing
- Imaging this as two different papers
- Side note: State mammal of Arizona – some discussion of sequencing its genome

Aja

- Working on updating some for the BTA, to then be able to finish the Results section
- Found some things with BatNames back end – will talk to Nancy and Andrea
- IUCN
 - Updates on the bat taxonomic updates

4 June 2025

Present: Aja, Nancy, Cecilia, Jorrit, Quentin, Anna, Jelle

Agenda (as suggested by Nate): Agenda: keeping things rolling with the BatLit webpage, BTA manuscript, co-roosting -- potential to join as a Disentis use case

1. BatLit Webpage
2. BTA manuscript
3. existing co-roosting database /
4. nate layman collaboration work on
- 5..potential to join as a Disentis use case (needs Donat/Nate to talk about)

Notes:

Aja - focusing on BTA, holding off on documenting the workflow of Zotero -> Zenodo used for BatLit until BTA work is

Aja - still working on the BTA manuscript, "downloading" the latest version of BatNames, early next week, updating the agreement index table

Nancy - suggest to be specific on requesting comments on document drafts: what section, which document, etc.

Aja - indicates that June 16th will be another of BTA review with specific asks to specific people

Aja - is working on the BatLit Landing page, ask input through slack

Anna -

Aja - no update on the bat-co roosting dataset

Aja - continuing to work with Nate to run the workflows to extract interactions for known co-roosting literature. (<https://github.com/n8layman/bat-interactions>)

Aja - currently only the OCR bit is being handled

Quentin - are you using the bat-co roosting

Aja - write script to pull information that we are interested in, then compare to training dataset. In addition, we are planning to

Quentin - two approaches (1) let LLM extract interactions independent of the bat-co-roosting (2) let LLM use bat-corosting data as a training dataset so that the LLM can get better

Aja - Nate is testing extracting graphs and charts and tables using Mistral Services API facilitated by a R wrapper script.

Anna - offering to review methods that Nate and Aja are working on

(note by Jorrit): for reference, the bat-corosting database is indexed by GloBI at <https://github.com/globalbioticinteractions/bat-co-roosting-database>, pointing to a google sheet. So, no data paper associated with bat-co-roosting database as of yet.

Elton, Nomer, & Preston. (2025). A Review of Biotic Interactions and Taxon Names Found in globalbioticinteractions/bat-co-roosting-database
hash://md5/9c419345bc2ee8811bad5326bdcaa150. Zenodo.
<https://doi.org/10.5281/zenodo.15190815>

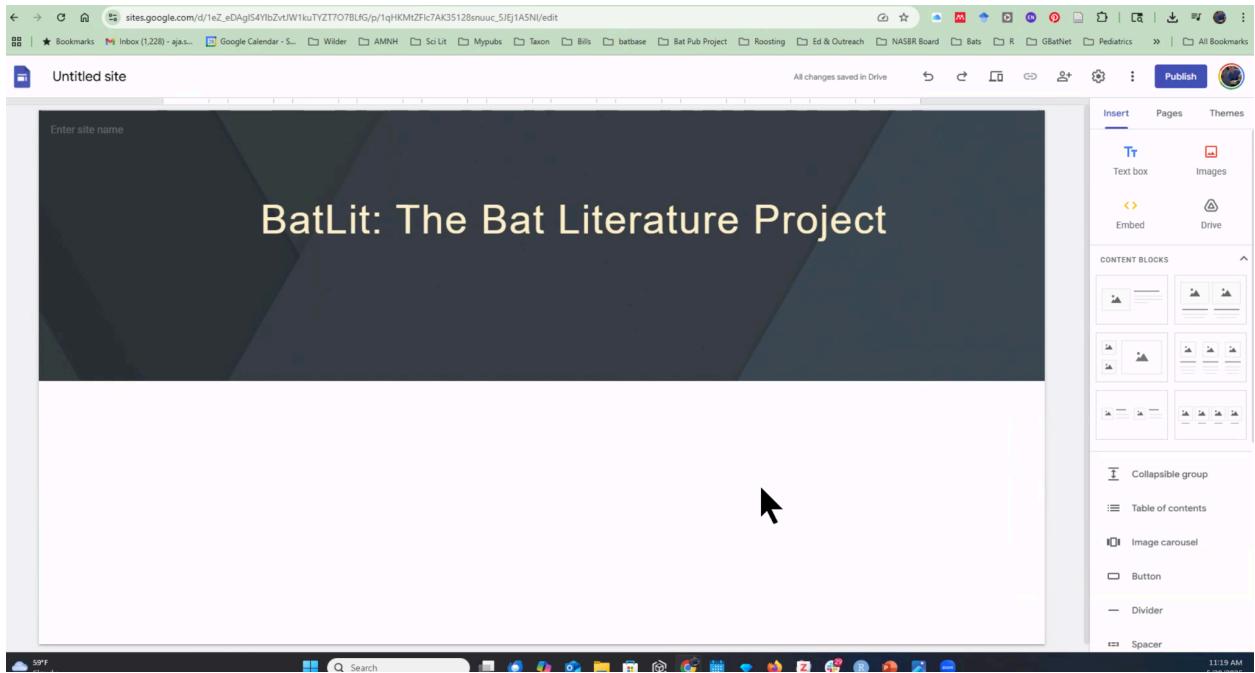
28 May 2025

Present: Aja, Nancy, Kendra, Jorrit, Donat, Nate, Connor, Anna, Hernani

Agenda -

1. Aja's BTA manuscript
2. Nate abstract at Living Data:
<https://docs.google.com/document/d/1vbqKUb9MaaCdNAmyEKhZOmETvbvJftjZRqu30wZmWn8/edit?tab=t.0>
3. linking bta / bat-corosting / batlit
4. and Anna's batlit comments
 - a. <https://github.com/bat-literature/bat-literature.github.io/issues/69>
 - b. <https://github.com/bat-literature/bat-literature.github.io/issues/70>

Aja is working on a frontpage template / user friendly design



Jorrit planning a Zotero upload from BatLit backend

- Could be a larger recorded session for sharing with others

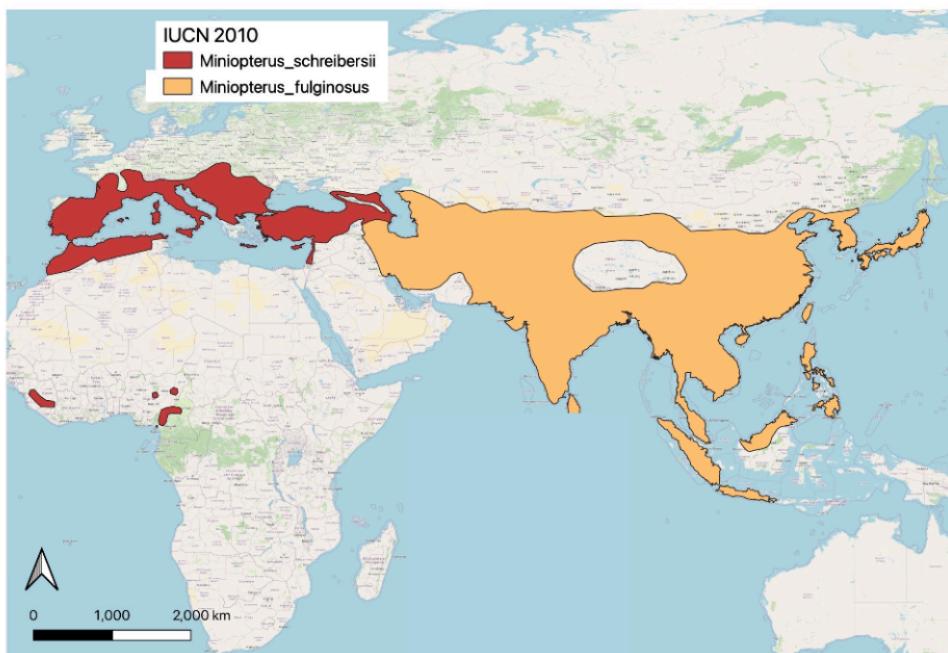
Living Data

- The deadline for abstract submissions has closed on May 25, 2025 (extended deadline)
- Nate abstract:
<https://docs.google.com/document/d/1vbqKUb9MaaCdNAmyEKhZOmETvbvJftjZRqu30wZmWn8/edit?tab=t.0>
- Many sessions – <https://www.livingdata2025.com/sessions.html>
 - Donat has sessions on Disentis, cases for re-use of data
 - Quentin has OneHealth session

BTA manuscript

- Ready for comments now! Only the Intro and Methods are ready for comments
- Still re-doing the Results
- Then the Discussion

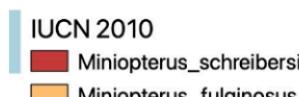
Example: *Miniopterus schreibersii* / *fulginosus*



Example: *Miniopterus schreibersii* / *fulginosus*

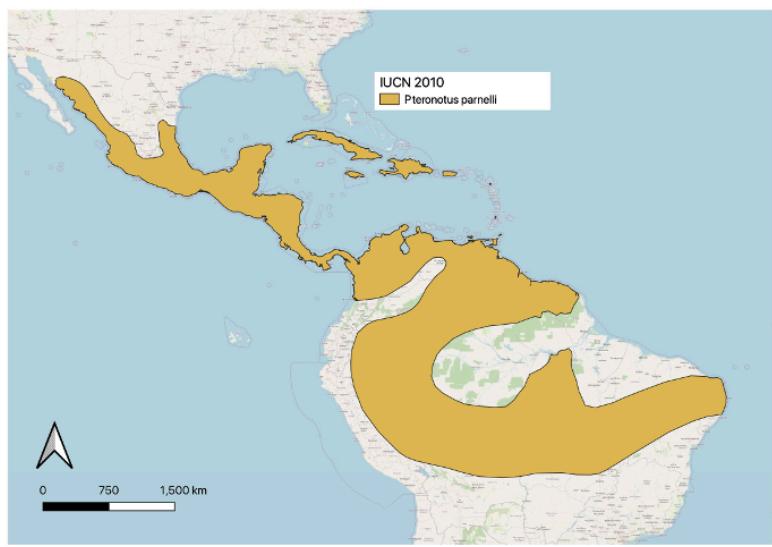


Example: *Miniopterus schreibersii* / *fulginosus*

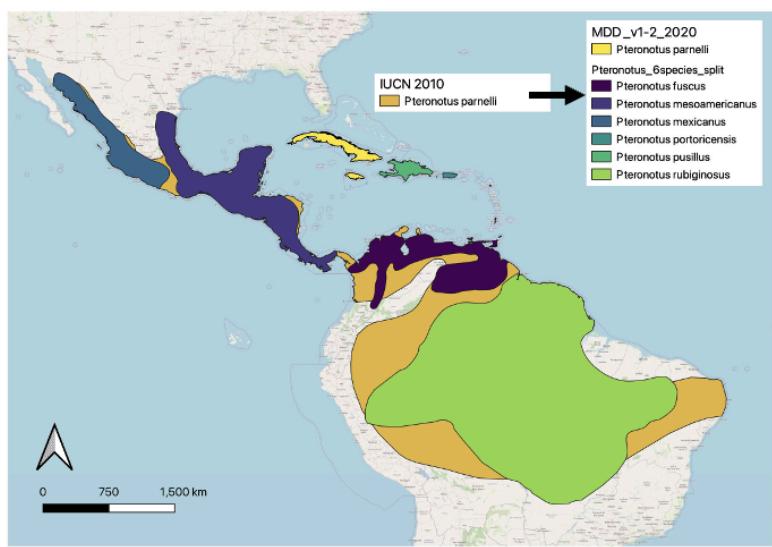


- Example we can add for a figure in the BTA manuscript
- Nate can work on the *Miniopterus schreibersii* example – could do one for *Pteronotus parnellii* too – made one previously but didn't show in NASBR talk:

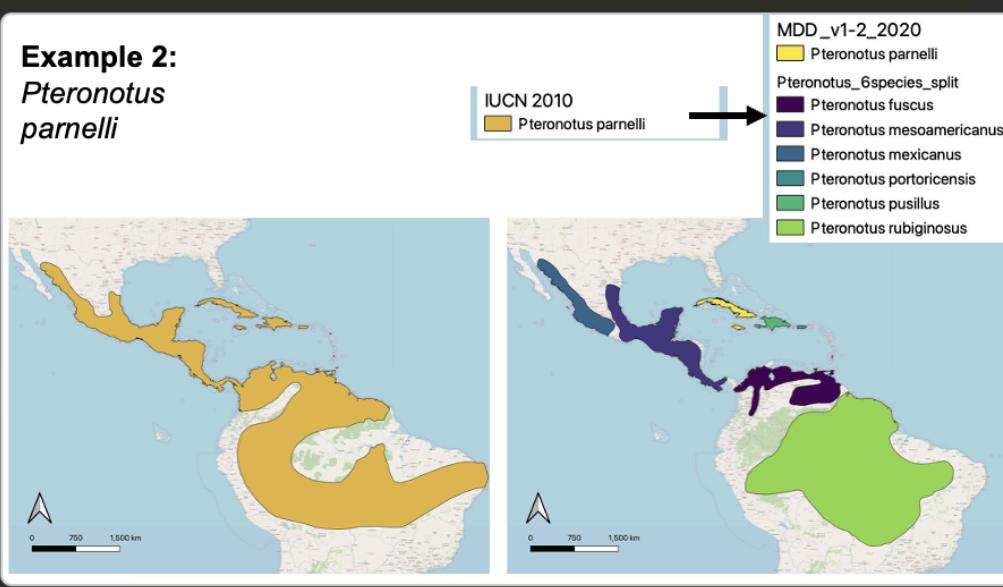
Example 2:
Pteronotus
parnelli



Example 2:
Pteronotus
parnelli



Example 2:
Pteronotus
parnelli



How to cite Lynx Handbook of the Mammals of the World dynamics...

- All the citation dynamics are confusing
- You could use the link to BLR on Zenodo <https://doi.org/10.5281/zenodo.3748525> and from there you could also site each species individually: Rhinolophus rhodesiae Roberts 1946 <https://zenodo.org/records/3808862>
- Jorrit: advise to include all of the data sources – link to a repository with those digital data in a way
 - Aja included all of the accounts verbatim as part of the BTA
 - But reconstructing that from primary sources is not direct/easy to do

Not easy to cite a literature piece / opinion – or many ways to do it

- Similarly many ways to refer to a given taxon
 - E.g., “Gen sp_A sensu lato” or “Gen sp_A sec Smith et al. 2020 HASH” (with the HASH part referring to a digital object that includes the range map, DNA seq, taxonomic treatment, etc – or a packing list thereof)
 - See Upham, N., & Poelen, J. (2024). Taxonomic Data Objects for Communicating the Meaning of Species Names. *Biodiversity Information Science and Standards*, 8, e139413.
<https://doi.org/10.3897/biss.8.139413>

Tanalgo, K. C., Tabora, J. A. G., de Oliveira, H. F. M., Haelewaters, D., Beranek, C. T., Otálora-Ardila, A., ... & Hughes, A. C. (2022). DarkCideS 1.0, a global database for bats in karsts and caves. *Scientific Data*, 9(1), 155.

<https://www.nature.com/articles/s41597-022-01234-4>

- You need to include the data from this paper for co-roosting
- It is a review about bat species co-occurring in caves worldwide
-

21 May 2025

Present: Aja, Nancy, Cecilia, Kendra, Jorrit, Jelle, Donat (and Vanessa), Ariadna, Hernani, Anna, Nate, Cullen, Quentin, Connor

Agenda -

5. Aja's BTA manuscript
6. Living Data biodiversity abstract Sunday 25 May
7. Hernani's prioritizing of bat species for sequencing
8. Linking Bat co-roosting literature with BatLit
9. Nate Layman Demonstration Impressions/ next steps
10. Status of building batlit
11. Cecilia / Hernani exchange re: [Fuzessy et al. NeoFrugivory](#)
<https://github.com/globalbioticinteractions/globalbioticinteractions/issues/1071> and associated

Notes

Aja making progress on BTA manuscript ([Manuscript](#)). Next week, I am hoping to share the draft and hand off pieces to co-authors. . Expect to cut a new release of BTA on Zenodo.

Anna - agreement index - what is it? (Jorrit: see <https://jhpoelen.nl/bat-taxonomic-alignment/> for example calculation of agreement index, e.g. table 3

Table 3. BTA@87c862c1 treatments, agreement index, and their associated names. **The agreement index is ratio of the number of pairwise agreements for a concept versus the total number of possible pairwise agreements.** Yellow/light colors indicate more agreement, green/dark shades indicate less agreement. Download table (minus agreement index) as tsv, csv, or json.

Hernani suggests Venn diagram or some other way to indicate the overlap between.

Some example of ways to visualize alignment below.

Quentin asking about how homonyms were accounted for in BTA.

Datos Vivos Oct 2025 Bogota - Nate's [Abstract - Living Data 2025](#), Donat will lead workshop on liberating data.

Hernani - planning to getting back to work on Big Bat Database (e.g., dental formulas) after finishing up some other project, the prioritization of bat sequencing. Hernani talking about "Bat

species sequencing prioritizing" and ways to quantify this as a priority score using categories (e.g., endemism, threatened, low protected area coverage).

Jorrit: how does the IUCN conservation status related to the "bat species sequencing prioritizing" method you and your collaborators used?

Hernani - some IUCN "least concern" are of high priority, because some criteria are not considered by IUCN (e.g., % in protected area, land use change). Among other things, he's sharing maps showing countries (or islands) being unique "hosts" of a specific bat species.

Figure 1. BTA Agreement Index using *Viridis scale*. Yellow/light colors indicate more agreement, green/dark shades indicate less agreement.

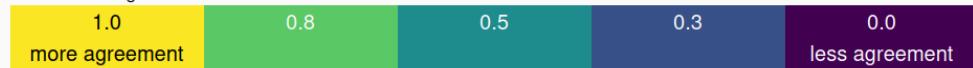


Figure 2. A clickable heatmap containing all BTA concepts and their associated agreement index organized in no particular order. Yellow/light colors indicate more agreement, green/dark shades indicate less agreement.

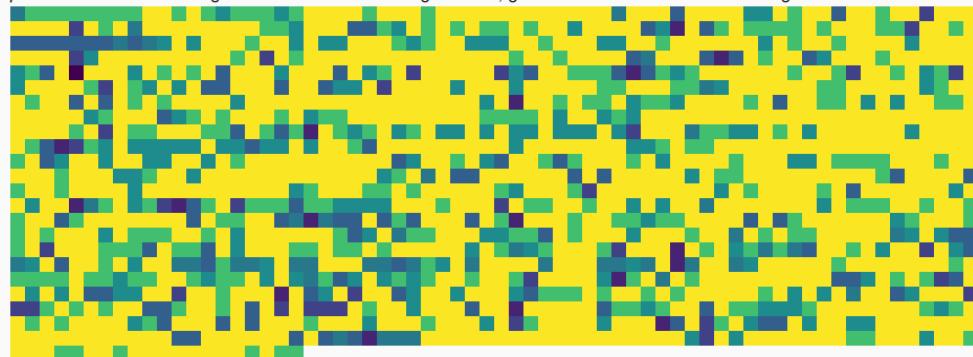


Table 2. BTA@87c862c1 agreement matrix. Each cell indicates the number of concept **disagreements** across catalog pairs. Total number of concepts in BTA@87c862c1 is 1538. Yellow/light colors indicate more agreement, green/dark shades indicate less agreement.

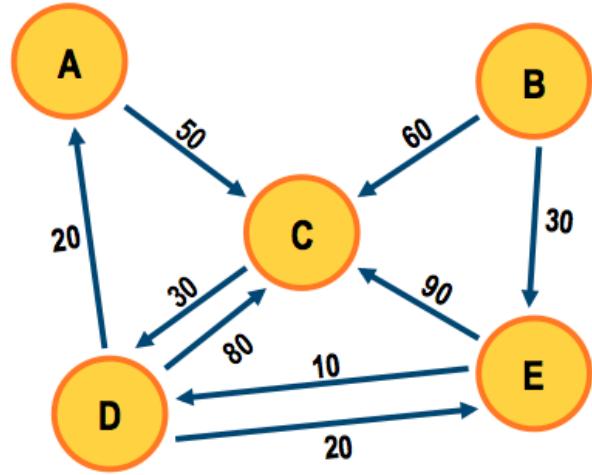
	BatNames	BatNames_2023	HMW	IUCN	MDD	MDD_2023	MSW3
BatNames	-	81	209	312	105	93	580
BatNames_2023	-	-	217	325	98	41	583
HMW	-	-	-	268	195	225	520
IUCN	-	-	-	-	330	324	494
MDD	-	-	-	-	-	89	595
MDD_2023	-	-	-	-	-	-	582
MSW3	-	-	-	-	-	-	-

Table 3. BTA@87c862c1 treatments, agreement index, and their associated names. The agreement index is the ratio of the number of pairwise agreements for a concept versus the total number of possible pairwise agreements. Yellow/light colors indicate more agreement, green/dark shades indicate less agreement.

Download table (minus agreement index) as [tsv](#), [csv](#), or [json](#).

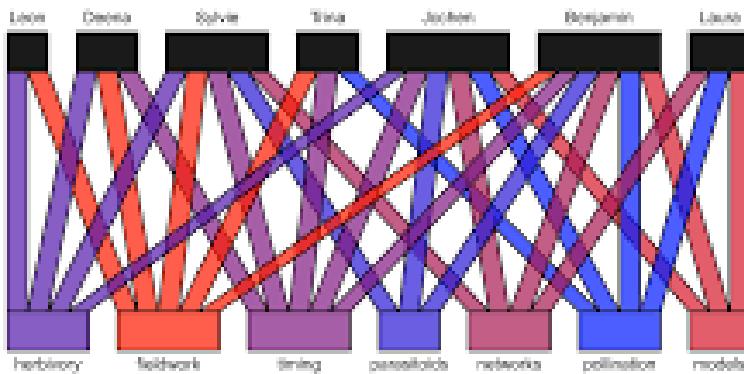
treatmentId	agreementIndex	name	accordingTo
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_MSW3
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_HMW
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_BatNames
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_MDD
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_IUCN
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_BatNames_2023
BTA_3@87c862c1	1.0	Acerodon celebensis	Name_MDD_2023
BTA_4@87c862c1	1.0	Acerodon humilis	Name_MSW3
BTA_4@87c862c1	1.0	Acerodon humilis	Name_HMW
BTA_4@87c862c1	1.0	Acerodon humilis	Name_BatNames

Spatial Interactions



Traffic Matrix

	A	B	C	D	E	T _i
A	0	0	50	0	0	50
B	0	0	60	0	30	90
C	0	0	0	30	0	30
D	20	0	80	0	20	120
E	0	0	90	10	0	100
T _j	20	0	280	40	50	390



14 May 2025

Present: Aja, Nancy, Cecilia, Kendra, Jorrit, Jelle, Donat (and Vanessa and Patrick), Ariadna, Hernani, Anna, Nate, Cullen, Quentin, Connor

Agenda:

- Nate Layman joining to update on data extraction for bat co-roosting
 - Just started a new job and may not be joining actually
- BTA manuscript
- Issues arising

Metabar-coding of plant species - with Venessa of the university of Porto (Portugal) working with Donat/Quentin

- Workshop where Quentin and Donat are currently working
- Vanessa provides a use case using metabarcoding of Bat Feces
- Biodiversa Plus - Paris - <https://www.biodiversa.eu/2024/04/15/metaplantcode/>
- workshop still going until morning of May 15, 2025
 - 105 Boulevard de l'Hôpital, 75013 Paris
 - Google Maps Link

Nancy, Connor, Cecilia were in Belize netting bats

- 1100 bats in two weeks, exported 4400 samples, safely back at AMNH now

BTA

- Alignment of bibliographic resources – similar to taxonomic alignment
 - Figuring out gaps in BatLit relative to the BTA
 - Which papers are cited in the BTA – are they indexed in the Bat Lit?
 - Gap analysis – or Use analysis depending on perspective
 - what is the overlap in the references in batlit and the references mentioned in some other resource (like BTA, MDD etc).

Nate Layman joined and shares on extracting data from ...

- OCR - "OhSeeR" pipeline
 - <https://github.com/n8layman/ohseer>
- Mistral a OCR resources -> markdown
 - <https://mistral.ai/news/mistral-ocr>
- LLM focusing on English at the beginning
 - Knowing which languages we have would be useful too
 - <https://cran.r-project.org/web/packages/ellmer/vignettes/ellmer.html>
 - <https://github.com/n8layman/ehallm>
- Example literature in the demo
 - https://zenodo.org/communities/batlit/records?q=metadata.publication_date%3A1994%20Gibbons&l=list&p=1&s=10&sort=bestmatch
 - Gibbons, P. (1994). Sustaining key old-growth characteristics in native forests used for wood production: retention of trees with hollows. *Ecology and Sustainability of Southern Temperate Ecosystems*, 59–84.
<https://doi.org/10.5281/zenodo.13528407>
- Jorrit: Q- is there a way to only do OCR, not the LLM?
 - yes, you can do OCR only
- Jorrit: Q - what is the price per OCR? What is their pricing model? How do you go about the processed/OCR-ed text and associate these with the origin of these texts? (would be nice to archive the results not just to document the results, but also to avoid re-work).
- How big an issue is needing to OCR things for the co-roosting dataset?
 - Aja: ~330 papers in the test corpus – ~100 papers need OCR out of that

- Getting the Table data out in a consistent way is an important step
- Tokenizing the document
 - Get turned into vectors in multidimensional space for the LLM to work with
- ellmer R package
 - wrapper around various Web API provided by LLM folks
 - From Hadley Wickham
 - <https://cran.r-project.org/web/packages/ellmer/vignettes/ellmer.htm>
 - Q (Jorrit) - what is the price associated with interacting with these folks? How much money have you spent so far while working on this prototype?
 - Nate - Millions of tokens (words submitted, word received) - \$3.75 per million tokens - related to Claude (?)

The screenshot shows the Anthropic Pricing page. The left sidebar has a navigation menu with items like 'Home', 'User Guides' (which is selected), 'API Reference', 'Cloud Code', 'Prompt Library', and 'Release Notes'. Below this are sections for 'Developer Console', 'Developer Discord', 'Support', 'started', 'review', 'ai setup', 'o to Claude', 'about Claude', 'cases', 'lens', 'ing', 'urity and compliance', 'd with Claude', 'ne success criteria', and 'elop test cases'. The main content area is titled 'LEARN ABOUT CLAUDE' and 'Pricing'. It says 'Learn about Anthropic's pricing structure for models and features'. A note states 'This page provides detailed pricing information for Anthropic's models and features. All prices are in USD.' Another note says 'For the most current pricing information, please visit anthropic.com/pricing'. A section titled 'Model pricing' contains a table with the following data:

Model	Base Input Tokens	Cache Writes	Cache Hits	Output Tokens
Claude 3.7 Sonnet	\$3 / MTok	\$3.75 / MTok	\$0.30 / MTok	\$15 / MTok
Claude 3.5 Sonnet	\$3 / MTok	\$3.75 / MTok	\$0.30 / MTok	\$15 / MTok
Claude 3.5 Haiku	\$0.80 / MTok	\$1 / MTok	\$0.08 / MTok	\$4 / MTok
Claude 3 Opus	\$15 / MTok	\$18.75 / MTok	\$1.50 / MTok	\$75 / MTok
Claude 3 Haiku	\$0.25 / MTok	\$0.30 / MTok	\$0.03 / MTok	\$1.25 / MTok

① MTok = Million tokens. The "Base Input Tokens" column shows standard input.

- Q (Jorrit) - which LLM can be run offline against a versioned copy of a specific LLM?
 - Yes, there are some (Q: which?) models that can be run offline. You'd have to provide your own hardware. (Q: What kind of hardware? What kind of hardware is needed to run LLM suited for this kind of work?)
- (Donat) can you use JATS input instead of markdown?
 - Used by PubMed Central
 - Would be convenient to use if could – markdown is an additional step
 - JATS format:
 - <https://jats.nlm.nih.gov/archiving/tag-library/1.2/chapter/nfd-journal-meta.htm>
- Current approach
 - Creating efficient prompts
 - NOT – using the training data from Aja to train a given model
 - Not enough data to fully fine-tune a model
 - Can instead
 - Use Aja's hand-extracted data as VETTING for the model – to be able to assess accuracy dynamics

(Q: Jorrit) If you ask the same question many times to some LLM, do you get the same answer?
(Q: Jorrit) If you ask the same question many times to some LLM 6 months later, do you get the same answer?
(Q: Jorrit) How do you keep track of the questions/answers question some LLM?

Related work done with Pensoft on extracting data from tables - as seen on
<https://www.globalbioticinteractions.org/pensoft/> in 2025-05-14

[...]

17 July 2020 - Mariya published blog post “Pensoft – GloBI workflow for FAIR data exchange and indexing of biotic interactions locked within scholarly articles” at
<https://blog.pensoft.net/2020/07/17/pensoft-globi-workflow-for-fair-data-exchange-and-indexing-of-biotic-interactions-locked-within-scholarly-articles/>.

11 Aug 2020 - Mariya submitted abstract “Semantic Publishing Enables Text Mining of Biotic Interactions” to virtual TDWG 2020 annual conference Oct 19-23 2020.
<https://www.globalbioticinteractions.org/pensoft/tdwg2020-semantic-publishing.pdf>
[...]

7 May 2025

Present: DeeAnn, Aja, Nancy, Cecilia, Kendra, Jorrit, Jelle, Donat, Ariadna, Hernani, Anna

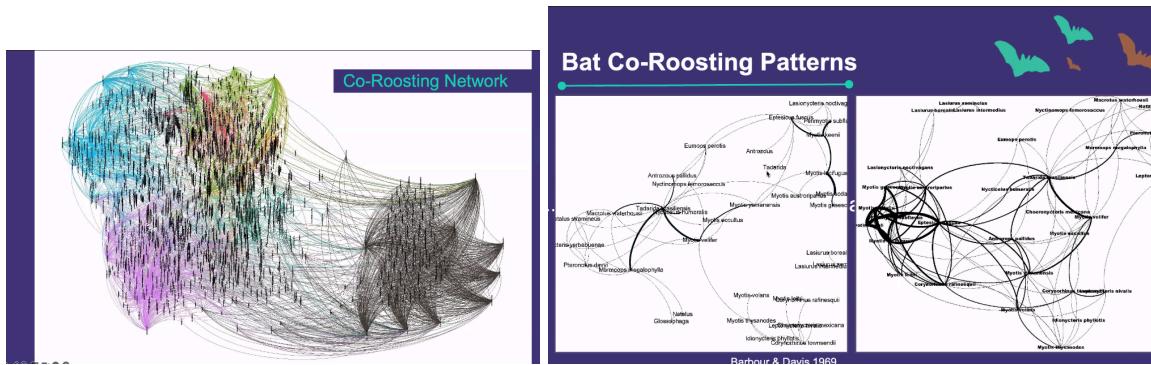
Notes:

Talking about bat books! Lots of them getting published. The importance of the “proper” childhood reflected in DeeAnn’s acknowledgements in here new book

Aja submitted the abstracts for IBRC

Aja asking for baby related bat pictures (bat reproduction) for baby shower

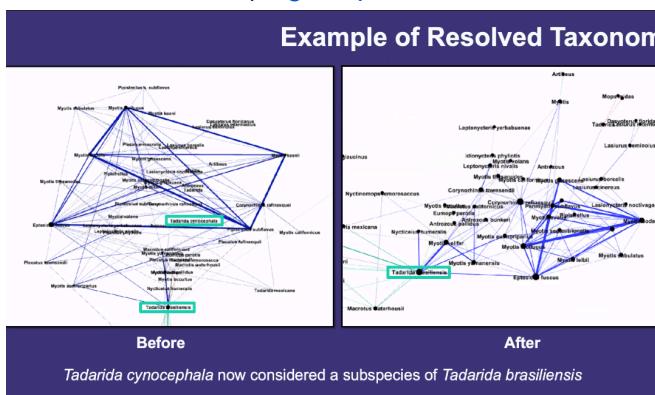
Kendra is giving a presentation in a few hours today at [MEPA](#) (Museums and Emerging Pathogens in the Americas). Question about a slide related to GBatNet group - linkages between bat hosts - viruses, digging deeper into ecological integration (e.g., bat co-roosting dataset), Bat Taxonomic Alignment and importance of that.



what does the second slide demonstrate?

Aja - correcting for valid species names alters the constructed species interactions - better representation of actual interactions between species. Consolidates data points of all synonymized names.

With similar slides ([original presentation with notes](#)) NASBR/IBRC 2022- Austin, TX)



Hernani - How BatLit and BTA would be combined in a useful way?

Nancy - you have to understand the taxon concept the author was using . Needs manual (and subjective?) review by taxonomists. No fixed timeline.

Donat - how can we educate the researchers to cite the specific concepts in their publications?

DeeAnn - considers to give a talk at the infectious disease on the importance of citing taxonomic concepts.

Kendra - presentation at the American Society of Virologists (last June) about the importance of identifying host species to a crowd of virologists, urging them to spend more time on viral hosts to make their outcomes re-usable beyond their own research. Some audience members were interested in vouchering the hosts and connecting to museums. Kendra also suggested collecting a variety of other data (echolocation, occurrence data) while already spending time/money on being in the field. Argues that it doesn't take that much more work to do so.

Museums often pitch in resources to help make this happen.

<https://covid19taf.slack.com/archives/C011W4GPXDX/p1746222455888919>

@Cullen Geiselman @Aja Sherman @anna For instance, who should be responsible for solidify the link between the citation string:

Arendt, W. J. 1986. Bat predation by the St. Lucia boa (Boa constrictor orophias). Caribbean Journal of Science, 22: 219- 220.
listed in the Bat-co-roosting dataset and their available copy in BatLit at

Arendt, W. J. (1986). Bat predation by the St. Lucia boa (Boa constrictor orophias). Caribbean Journal of Science, 22, 219–220. <https://doi.org/10.5281/zenodo.13528299>

? (edited)

Aja - suggests fuzzy matching -

Nancy - different journals will have different citation styles - many different versions of the same citation.

Jorrit - how to curate a collection of fuzzily matched citations and their associated

Nancy - suggests that the title should always be the same.

Jelle - begs to differ

Anna - aren't DOIs supposed to solve this? For the co-roosting datasets introducing citation id
(e.g., Arendt_1965.pdf to link between citation table and interaction table)

Donat - How to make sure that it links to the same publication?

Anna - co-roosting citation is connected to Arendt_1965.pdf in the co-roosting datasets. That pdf has a digital md5 signature (independent of pdf name), and this signature can be used to link to BatLit. BatLit keeps a list of md5 signatures of all the pdfs.

Aja - ocr is getting better - do fuzzy search on content of the pdf in addition to the citation.

Both Anna and Aja are proposing to use the pdf as the anchor.

two methods -

1. fuzzy matching by citation string
2. matching by pdfs (fuzzy matching, compare md5 digital signature file, filename)

30 Apr 2025

Present: Donat, Aja, Anna, Hernani, Nate, Cullen, Kendra, Jorrit, Jelle

Agenda:

- review batlit poster abstract by Cullen [BatLit Poster IBRC 2025 Abstract](#)
 - review AI voice for BTA video draft
- Abstract for BTA poster
<https://docs.google.com/document/d/1evO-NtV9vOKATfPqmRbD1wtf32M8VeQyVuxLtU-uGOM/edit?tab=t.0>

- Abstract for BatLit Poster
https://docs.google.com/document/d/1Cw3tSBKBvjD1AVyAd_orEKu8jHvjiL3Rfk2GI8G7a4g/edit?tab=t.0

BTA paper

- Aja working on the methods

BatLit workflows

- Some need for multiple people to be able to upload PDFs from Zotero to Zenodo
 - Jorrit has a generic workflow: (1) translate data from RIS / Zotero to Zenodo format for the metadata; (2) then integration is the same
 - Translation is to match up with Zenodo API
- General purpose batch uploads to Zenodo
 - Sometimes the existing DOIs are not found directly – but even if not found, they can later be associated – why add DOI to Zenodo metadata so can search with this
 - Batch scripting is simple process
 - Want more tools to process abstracts to get keywords etc
 - Zotero can help
 - Either will reference the DOI and get metadata from CrossRef, or else do OCR and pull out of the PDF directly
 - There will always be mistakes
 - Goal: produce a referenceable corpus, then can go back and fix errors / edit, produce a new version with a hash etc – can then make improvements in a controlled way
 - Keep the option to fix errors
 - Unlike in NCBI / GenBank and sometimes GBIF
- Jelle has built his own scripts
 - Dealing with old literature that doesn't have DOIs
 - Own format for Hesperomys references
 - Aja has his PDF collection for bats
 - Has already uploaded aspects of this to Zenodo – only the references that already had metadata

BHL lost support from the Smithsonian

- Jorrit and Donat working to backup / upload to Zenodo
- BHL in a transition, looking for someone to take it over – it will still be online
- What portion of things does BHL host on its own servers?
 - Whole volumes
 - But more useful as parts, so not separately indexed with metadata
 - Some on Internet Archive
 - But mostly the large volumes
 - They used to do the scanning for BHL

- Most metadata of the parts appears only on the BHL website
- Goal: to move BHL into the Biodiversity PMC
- Many item parts are not hosted on BHL
 - About 10% – Scielo different countries
- Core value of BHL
 - (1) They have the scientific names attached to each page
 - Something trying to reproduce in Biodiversity PMC
 - Were indexed by the GlobalNames project – but includes some errors
 - (2) Good bibliographic metadata, can be extracted
 - Useful for completing full references of authorities
 - BHL put lots of energy to this – wanted the DOIs to be associated with this directly
- One Problem w BHL
 - Started as a library project, not a science project
 - Article and volume-level data – rather than extracted information that you could actually analyze
 - LINKING is what's needed
 - Corpora – but then to make them useful, need to link content together meaningfully – by species names / concepts, references, authors
 - This is a strength of our Bat-Eco-Evo group as well
- Purposes of links – WHY do this
 - Can compare names
 - Once have name, can link to vast array of other data
 - Internal sanity checks
 - On data integrity
 - Ecological claims
 - Are different authors using the same level of data
 - What tests, assays
 - Tracing back references to claims
 - References
 - Actual article context
 - Full text
 - Metadata of article

Kendra – giving a MEPA talk on different projects of the Bat Eco-Interactions group

All the activities of the Bat Eco-Interactions group

- Where is the public face of all this?
- <https://globalbioticinteractions.org/gbatnet>
 - Has good index there
- Connecting out to resources
 - Global Mountain Biodiversity Assessment (GMBA)
 - IPBES
- We could have more of a landing page to document this

- BatLit to BTA connections
 - Are a major breakthrough
 - Combining those initiatives makes it a big deal

Co-roosting

- Dataset has references, names, links to BatLit
- Could use this to show the whole web of knowledge that is related here
- Use case
- Example to push and publicize
- This is also MASSIVE in terms of the amount of data
 - Want to make something digestible enough
 - **Smaller** use case should really be the thing to push
 - Automated workflow – manual checking
 - GloBI would be a great venue for this
 - Do this use case with every name linked to a taxonomic treatment
 - Make this the way forward
 - Simple, but immediate effect

Major key take-home message from meeting today:

- Let's pull out a digestible case study from the Bat Co-roosting Dataset to showcase to broader audiences -- displaying the full depth of linking for names to taxonomic treatments, ecological claims to evidence in literature, linking from names to other ecological interactions (host-virus, predator-prey, animal-habitat, etc) -- and then use that to illustrate the power of LINKING as an approach to doing biodiversity science
- Question to ponder: Which case study would be best for that goal?

23 Apr 2025

Present: Donat, Iroro, DeeAnn, Aja, Anna, Hernani, Quentin, Iroro

Agenda:

- announcements
- review IUCN abstracts are due in one week! [LINK](#) to abstract draft.
- review batlit poster abstract by Cullen [🔗 BatLit Poster IBRC 2025 Abstract](#)
- review AI voice for BTA video draft
- question about batlit website

Donat solicited bat related pdfs that have not yet been in Treatment Bank / BatLit / BLR, especially 19th and 20th century bat literature. We have already completed 173 publications (see the [list](#) already, and are looking for additional publications (reference?) .

Hernani asks a question about BatLit - can I make a list of publications that mention (insert bat species) that were published before some publication year (2013). Hernani: please add your specific question, ideally with specific examples, as feedback on <https://batlit.org> (e.g., github issue).

Aja mentioned that the (BTA?) manuscript is planned to include a section on the relation between name changes and the historic literature that may have been affected.

Action item - include specific example of Hernani's idea into the upcoming manuscript. Aja's is going to work on that.

Quentin - mentioned an example Pepistrel ? Please clarify examples.
https://en.wikipedia.org/wiki/Soprano_pipistrelle *Pipistrellus pygmaeus*

DeeAnn suggests including a name split/lump example that appeals to non-bat specialists (e.g., viral reservoirs). Popular, news worthy example. She also offers to contribute some of those examples to the manuscript.

Iroro - suggests to also include more complicated examples (e.g., west africa bat genus changes) . How to translate tracking the name changes of existing reports (e.g.,
<https://www.google.com/search?client=safari&rls=en&q=african+chiroptera+report&ie=UTF-8&oe=UTF-8>) <https://www.africanbats-npc.org/acr/african-chiroptera-report-2024> also include csv version .

16 Apr 2025

Present: Donat, Nancy, Jelle, Nate, Iroro, Kendra, Aja, Anna, Ariadna, Cecilia, Hernani

Agenda: IUCN abstracts are due in two weeks! [LINK](#) to abstract draft.

Linking the BatLit Zenodo to MDD

- Could be directly linking within the MDD species pages
- BatLit articles are not indexed based on species names yet – but we could develop this
- MDD cited papers are not yet in BatLit, not parsed for bib info

OCRing old bat literature

- Plazi has worked on this some
- Some things are already on BHL
 - But OCR is not good
 - How do you measure this? Absence of the taxonomic names is the main criterion
 - Flagging what standard of OCR that a PDF meets — quantifying
 - Extracting the treatment for indexing on TreatmentBank, GBIF

Which DOI to use from Plazi treatments?

- Assigning DOIs to old literature makes sense
- But sometimes conflicts – JSTOR, Cross-ref members (eg Oxford U Press, Taylor and Francis) and BHL have the same article
 - Then it ends up resolving to the publisher (fee for use) article
 - Can screw up the bibliometrics if you have these resolving to different article versions – if you cite the Zenodo and not the publisher, harder to track

IBRC meeting abstracts

- Due on April 30th: <https://www.ibrc.org/call-for-abstracts.html>
- Who is going?
 - Aja, Ariadna, Iroro, maybe Kendra, maybe Cecilia
- Aja abstract
 - BTA centric, influence on the IUCN – then also plug BatLit for accessing the literature
 - Maybe also a talk and a poster type dynamic

IUCN integration more generally

- Craig Hilton Taylor follow up – how are the BTA recommendations being used by IUCN?
- We can reach out more further

New jobs

- Kendra at Univ Minnesota (Dept of Vet & Biomedical Science, College of Vet Medicine)
 - Focused on vector-borne diseases until more federal funding opportunities become available to support bat-virus biosurveillance program
 - Chronic wasting disease at Minnesota Center for Prion Research & Outreach (MNPROM)
 - Infectious proteins (prions), stay active in the environment for years
 - Mainly studied in deer, elk, moose – slow disease progression
 - Developed rapid diagnostics (as a start-up company—Prionex)
 - EHA dissolved as of April 1, 2025
- Ariadna in France

Iroro website setup for BatLit

- Landing page

- Ownership of website – use a general webmaster email
- batLit.org

Hosted portals

- E.g., <https://data.europeanjournaloftaxonomy.eu/>
- Could do this based on MDD or Batnames taxonomies
- Updated landing page for BLR too:
 - <https://zenodo.org/communities/biosyslit/>

9 Apr 2025

Present: Jelle, Jorrit, Iroro, Aja, DeeAnn, Cullen, Anna, Connor

idea - find ways to connect GBatNet related projects through their literature references.
Literature references are the one thing that all the projects have in common, in addition to
taxonomic names . . .

How to shape this?

DeeAnn - join a GBatNet meeting and discuss

e.g., - next Tuesday 15 April / 16 April 2025 .

MEETING I 19:00 US Central Time (CT) APRIL 15th = 1:00 AM GMT APRIL 16th

Tigga Kingston is inviting you to a scheduled Zoom meeting.

Join Zoom Meeting

<https://texastech.zoom.us/j/91463939233?pwd=QHqbP2gtS6oaXwwC7DFL9NK58VBgTa.1>

Meeting ID: 914 6393 9233

Passcode: 932065

MEETING II 07:00 US Central Time (CT) APRIL 16th = 13:00 GMT APRIL 16th

Tigga Kingston is inviting you to a scheduled Zoom meeting.

Join Zoom Meeting

<https://texastech.zoom.us/j/97236805373?pwd=KmygeqJ1wYIPTda9F1QbDGcAwkGbZ1.1>

Meeting ID: 972 3680 5373

Passcode: 250687

Ideas on linking between MDD and BatLit

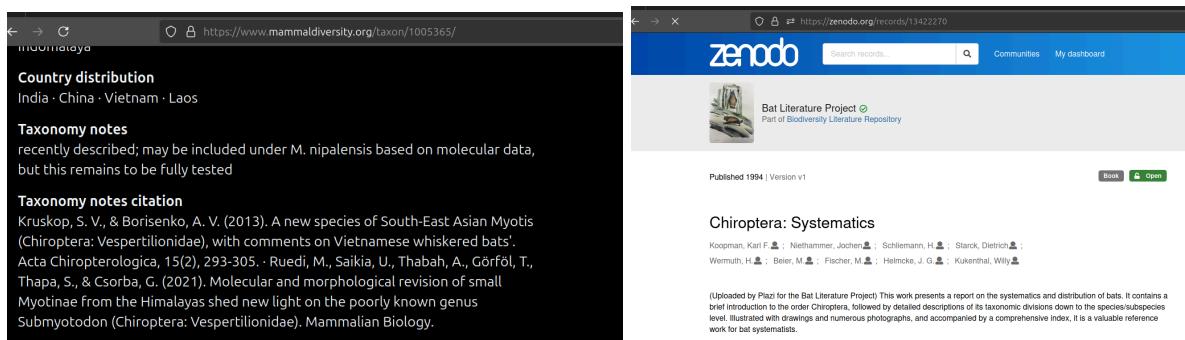
DeeAnn - what would it look like to integration MDD and BatLit

DeeAnn wants to find the paper associated with the taxonomic notes citation, get the pdf right away. DeeAnn also wants to import the citation into her Zotero library. DeeAnn want to hover over the citation and see where this literature can be retrieved, and include a link to the DOI and link to batlit, and provide context to other MDD taxonomic pages that reference this publication.

Aja wants a paper link to find the paper pdf and import into her Zotero library. On visiting the pdf landing page, I'd like to see which other MDD taxonomic pages link to this publication. The publication may impact other taxonomic concepts. Aja would like to be able to "tag" the paper with tags "species description", "genera change", "family change", "elevation", "split", "merge", "lump". Also, tags like "from MSW" to help group literature records in meaningful ways, like "find all the literature associated with "splits".

Cullen wants to links on a link that shows a webpage with the pdf in it. Then, I'd like to download the pdf and import them into Zotero library.

Conner wants to link literature to their batlit "record" (e.g., DOI, zenodo deposit id) as he is curating MDD.



Next topic - Abstracts for upcoming conferences (IBRC Australia, NASBR Canada, Living Data Columbia Oct 2025)

Cullen - **IBRC** international bat research conference abstracts due April 30th, 2025. Who is going? Who is interested in presenting?

DeeAnn is not planning to go. Aja is planning to go. Iroro planning to go pending visa logistics. Kendra is going (in absentia), Cullen is going but presenting on seed dispersal. Aja - idea to present on ways to link MDD, BatLit, BTA etc. and ways to connect to other projects (e.g., GBatNet projects BigBatDB). Aja to figure out who's going (via slack), share an abstract draft on Slack and stimulate folks to contribute.

Quentin - **Living Data Conference Oct 2025**. Symposium was accepted - "Re: Congratulations! Your Session 6796861 has been accepted into the Call for Abstracts round for Living Data 2025. Dear Quentin Groom,

The Academic Committee is pleased to inform you that your proposed session 6796861 titled "Biodiversity data for One Health: Integrating open data to address global health challenges" has been accepted for inclusion at the Living Data 2025 conference as a 1. However, because there was significant overlap among topics, either one or two other sessions are being merged with yours.

The impact of these mergers should be the potential for extra help from organizers of sessions being merged with yours. We will send you details separately, so that you can work with the other organizers to coordinate your session."

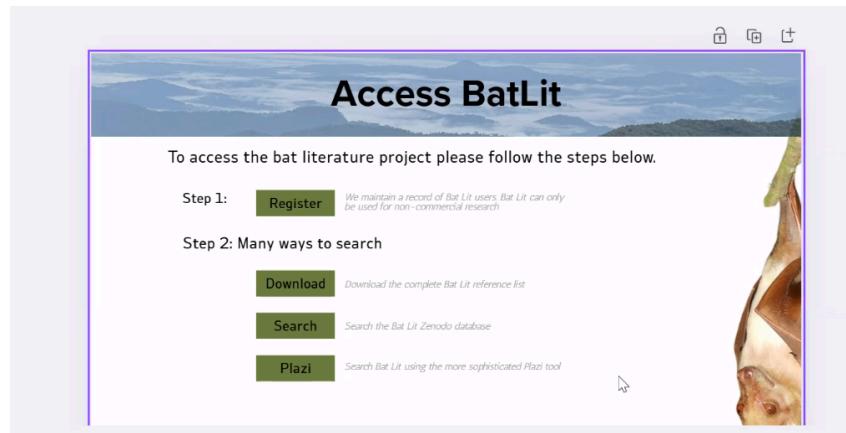
NASBR 2025 in Edmonton - Aja is going, Kendra maybe, Iroro pending visa logistics,

2 Apr 2025

Present: Nancy, Jelle, Jorrit, Iroro, Aja, Nate, Cecilia, Cullen, Anna
Jelle is new father - awesome

BatLit – think about the front end, how to make this more user friendly

- Softer landing page
 - Simpler
- Focus has been on corpus – pouring to digital concrete
- Journal publication possible
- Could be as simple as just making a single page to land on
 - Shift current front page info to an “About” page
- Likely other person needs to help – not Jorrit
- Timeline
 - IBRC meeting in Aug 2025 in Australia
 - Can be hard launch at that time
- Iroro
 - Has a simple draft landing page – using Canva



early draft

- Simple, has the general functionality
- Could either:

- Integrate with current <https://batlit.org/>
- Swap out the current page – and just add links
- Current BatLit.org backend is on Github — just add to the Github Pages context:
<https://github.com/bat-literature/bat-literature.github.io/tree/main>
- Could be nested within BatBase
 - E.g., batbase.org/batlit
- Doing an actual journal publication
 - Would help CVs of those in group
 - More formal process
 - Peer review vetting could help refine explanations of the process for others
- BatLit membership procedure
 - Draft doc:
https://docs.google.com/document/d/1hs3sVor9Jse7BkEgrvVV_dHlo9KHGE-olhFV7Hpcqbk/edit?tab=t.0
 - Sign up form:
https://docs.google.com/forms/d/e/1FAIpQLScTFfM3qNvhwTuMY_9xhvwyJpcccxAIR4UqEdp4dOlqNyPdVA/viewform
 - Seems pretty solid!
 - Ping general email to BatLit – curated by several people (rather than only Aja)
 - Then manual input of access to Zenodo
 - In future, if needed – could likely automate via Zenodo API – Donat will check on this
 - But we don't want to let this slow us down
 - Discussion of who should be owner / curator / manager / reader – need to decide at some point – core group of people curating only

Donat: BHL whole articles are now going into Biodiversity PMC

- 30k articles in BHL are all now searchable
- Bat articles – included
 - Searching text – OCR of old stuff has already been done

MDD taxonomic subcommittees

- Make the role of building MamLit database similarly part of this
 - As modeled after the batlit.org work
- In addition to providing oversight on taxonomic decisions
- MDD v2.1 version coming soon

GBIF Checklist Bank

- Need to get MDD on this – especially now that we have the synonym list

Text mining for interactions aspects

- Nate Layman – finishing up the EcoHealth internal documents

- Potential to host some of these on Zenodo

26 Mar 2025

Present: Nancy, Anna, DeeAnn, Connor, Jorrit, Iroro, Aja

Iroro - talks about a student book plan - Nigeria non-profit Small Mammal Conservation - Ecoevo Book Program <https://www.smacon-africa.org/ecoево> - sending books to students all over Afrika and Afrikan Islands. Planning to buy 100 copies of DeeAnn's Bat Book , students can sign up.

DeeAnn - Did Mammals of Africa get distributed by the Ecoevo Book Program ? 8 volume pdf is available, and we need to ask for permission to share. Fully searchable pdfs.

Anna - are eBooks suitable for the <https://www.smacon-africa.org/ecoево> program?

Iroro - we did do pdf electronic distributions, but not (yet?) coordinated through the book program.

Anna - UGA (University of Georgia) is now preferring eBooks over physical copies.

DeeAnn - mentioned that she felt a moment of shame when seeing the 150 requests for papers on research gate.

Jorrit: is there a way to redirect folks to BatLit.org so they can access your papers.

Anna shares some links of DeeAnn's papers currently

https://zenodo.org/communities/batlit/records?q=metadata.creators.person_or_org.name%3A%22Reeder%2C%20DeeAnn%22&l=list&p=1&s=10&sort=bestmatch

Iroro / Nancy suggests the put the word out and getting <https://batlit.org> .

Big question - are we ready to publicize batlit.org beyond our group?

Nancy et al. - make a splash page with a easy to understand mission statement / burb.

Silly example: "If you are looking for bat literature, look no further. We have it all!"

Goal - prettify and make more accessible batlit . Shopping list:

1. nice photo
2. burb (2 sentences)
3. QR code
4. contact information

5. button/link to sign-up page

6. code of conduct

Iroro - can help with designing the landing page.

Jorrit - please send the html / css etc. of the landing page, and I'll integrate it into <https://batlit.org>

Aja made some decisions about BTA –

Current BTA status - name alignment across all name authorities.

<https://docs.google.com/spreadsheets/d/1oR0L4BisisXHQuGNyMjBi-uTVkgwgi2tbXK2qwMvUul/edit?qid=1089845149#gid=1089845149>

Jorrit and Nancy suggest allocating some specific time to work through specific BTA use cases.

19 Mar 2025

Present: Cecilia, Nancy, Anna, DeeAnn, Ariadna, Kendra, Jorrit, Iroro,

discussing link rot - and ways to counter them, using examples

<https://github.com/bio-quoda/preston/issues/321>

Anna mentioned a digital library conference in St Louis last year. This year 2025 the meeting will be in Denver. Perhaps an idea to submit an abstract to describe our BatLit methods?

A missing service across research communities - ways to preserve, access and reuse digital data just like we've been doing with book etc. for centuries. Who's going to take this on systematically, we are trying to do our best to keep our stuff available (e.g., BatLit).

Nancy's strategy - migrate to BatNames to MDD . MDD has a professional society (sustainability) behind it. Tip - always have to have an exit strategy (for both digital and physical sample).

DeeAnn - which institution is going to preserve my physical samples?

Nancy - all about priorities.

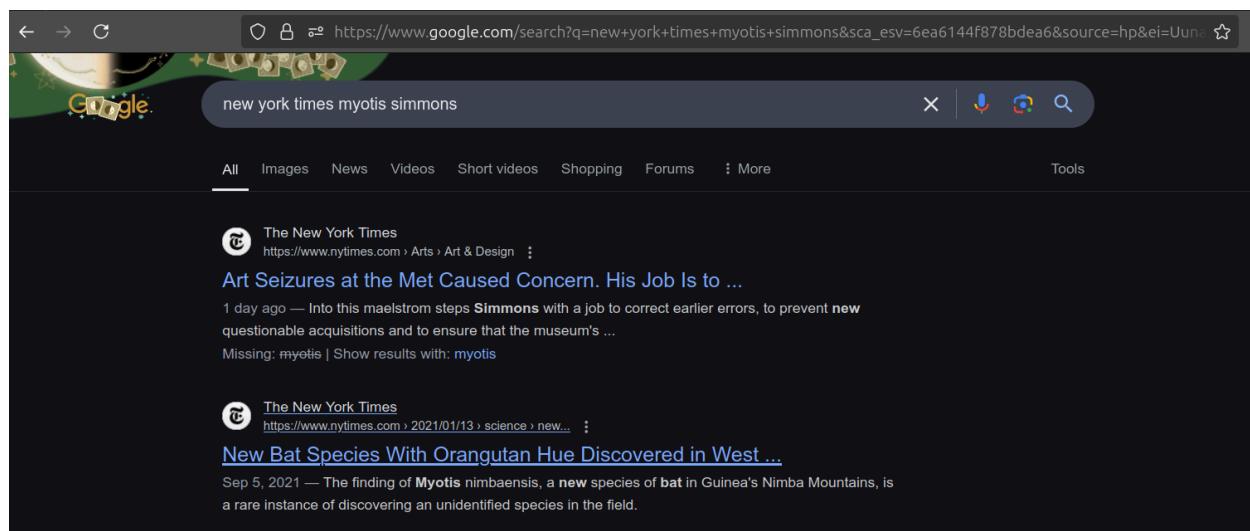
Iroro - idea - have a bot send up to 5 emails once in a while, even when there's 1000 of broken links, or things that need attention.

Harnani - we are good at describing species - but need to get better at linking associated literature and datasets.

Nancy - shares story on discovering a new Myotis species (*Myotis nimbaensis*)* along with her new life goal to never study Myotis again.

* featured in the New York Times

<https://www.nytimes.com/2021/01/13/science/new-bat-species.html> . Jorrit found the article after following Ariadna's suggestion to just search for "Myotis New York Times"



12 Mar 2025

Present: Jelle, Ariadna, Nancy, Cullen, Cecilia, Jorrit, Aja, Kendra, Hernani, Anna

- New MDD!!! <https://www.mammaldiversity.org/explore/>
- Aja is working on BTA name matching for additional authorities and agreement index tables for paper
- Discussed who is going to Australia for IUCN, many not sure
- Ariadna looking for the admission form for batlit community on Zenodo for a student
 - Cullen and Donat are making progress on the pipeline, it will be available soon
 - Suggested that the student email us and will be given individual access

- Short meeting

5 Mar 2025

Present: DeeAnn, Nate, Quentin, Connor, Donat, Jorrit, Aja, Cullen, Kendra

Slack for our working group

- Long term solution here?
- Migrate to a non-profit most ideally
- ASM Mammal Diversity Database transfer could be possible
 - Quentin will connect to Ana Casino – Dimitry Koureaus

Old notes did get lost at some point

- Earlier notes doc disappeared

Donat– press release came out

- Disentis Roadmap declaration: <https://www.bouchoutdeclaration.org/background-2024/>
- Extract all the data that can be reused
 - 70 people signed up, 13 institutions
 - Can still sign this
- Followup of the Bouchout Declaration in 2014
- Getting the literature to be more open access
 - In line with the DNA
- Only 20-30% is open access for taxonomy in particular
 - Biodiversity literature, as used for IPBES – 35% are open access
- Demonstrating whole workflow
 - Biodiversity Data Journal – <https://data.bdj.pensoft.net/>
 - EJT - <https://data.europeanjournaloftaxonomy.eu/>
 - All the raw data is coming through these directions too
 - Plazi:
<https://www.gbif.org/publisher/7ce8aef0-9e92-11dc-8738-b8a03c50a862>
- Hosted portal via GBIF: [bdj](#), [ejt](#)
 - Thin web layer on top of the data
 - More user friendly way to dive into this
 - Potential way to understand the power of data in publications
 - Can build things NEXT TO this GBIF digital concrete
 - Tools to provide folks with data
 - But also potentially annotations?
 - Can point to concrete, add changes
 - Jorrit is already doing this for BEE data -see e.g.,
<https://github.com/big-bee-network/bif>
 - Want the annotations to then *feedback* into GBIF more directly – already within hat ecosystem, DarwinCore, etc

- DeeAnn
 - Abstract for the Bat ID meeting in Chicago in July:
https://docs.google.com/document/d/13_SZ-SszpsH74k_rO-6ddJWLRSkHsXxeBp6dDX-sM1A/edit?usp=sharing
 - <https://brooklab.org/programs/bat-id-2025>
 - Not in much in there yet, placeholder for building this
 - Nate's PPTX of slides at NASBR in October:
https://www.dropbox.com/scl/fi/8w7txztmkhn1r2yd2ihgj/UphamEtAl_NASBR_BatTaxoViral_24Oct2024_final.pptx?rlkey=9xme8n6lefjvf6p9p3ji3i0ku&dl=0
 - For an (incomplete) timeline of our activities as a group, please see also
<https://www.globalbioticinteractions.org/gbatnet#timeline> .
- News in general
 - Too much going on to parse almost
- WW2, apparently had “bat bombs”: https://en.wikipedia.org/wiki/Bat_bomb

26 Feb 2025

Present: DeeAnn, Jelle, Quentin, Connor, Donat, Jorrit, Cecilia, Nancy, Aja, Hernani, Cecilia, Cullen

- NIH status quo and how funded projects are going on
- Bat infectious disease meeting University of Chicago, July 9-11, 2025
 - DeeAnn is attending ; will present on the tools that we have built (BTA, co-roosting, lit)... will have an abstract draft soon
 - <https://brooklab.org/programs/bat-id-2025>
- Australian Meeting <https://www.ibrc.org/> 3 – 8 August 2025 in Cairns, Australia
 - Too expensive to go?!
 - Many people from SE Asia is attending
- Living Data 2025 conference Bogotá October 21-25
 - 99 submission for symposia...
 - <https://www.livingdata2025.com/>
- DeeAnn’s new book is coming out in March 25, 2025 ([Amazon](#))
 - DeeAnn Reeder 2025. The Lives of Bats: A Natural History. Princeton University Press. ISBN 9780691257624
<https://press.princeton.edu/books/hardcover/9780691257624/the-lives-of-bats>
 - DeeAnn might have a code to get it cheaper [EXH25 is the 30% discount code - to use directly at Princeton University Press site;
https://press.princeton.edu/books/hardcover/9780691257624/the-lives-of-bats?srsltid=AfmBOoo5owTMGf-0kPOXY80BsN7HFw_NKiNtIQ0bSf3wypKayIAXGDob#preview - but then shipping is \$8 - so... maybe easier to just buy on Amazon/your favorite bookseller].

- - Nancy wants to ship some copies as gifts to IBRC
 - Images databases used to get images are terrible in term of identification of bats
- What is the most cited publication? Probably not the one you want, see e.g. [Nancy Simmons](#)
- Quentin shared a list of bats he recently found on Montserrat to complement his work on endangered trees, for which he will put in a proposal:
 - *Sturnira thomasi vulcanensis*: plant visiting, roosting high up in trees; [BiodiversityPMC](#)
 - *Chiroderma improvisum* plant visiting, cave roosting?; [BiodiversityPMC](#)
 - *Ardops nichollsi montserratensis* plant visiting, cave roosting? [BiodiversityPMC](#)
<https://www.globalbioticinteractions.org/?interactionType=interactsWith&sourceTaxon=Ardops%20nichollsi>
 - *Monophyllus plethodon* plant visiting, cave roosting? [BiodiversityPMC](#)
<https://www.globalbioticinteractions.org/?interactionType=interactsWith&sourceTaxon=Monophyllus%20plethodon>
 - *Tadarida brasiliensis*: high flying insect eating, cave roosting?, builds on of the biggest gathering of mammals in the world; Mexican free tailed bat - might be called American free tailed bat... [BiodiversityPMC](#)
<https://www.globalbioticinteractions.org/?interactionType=interactsWith&sourceTaxon=Tadarida%20brasiliensis>
- There is an open call for publication of data related to datasets of vectors of human disease for GBIF
<https://www.gbif.org/news/70AuWTs68FiGEZEyALB1TV/third-call-for-data-papers-describing-datasets-on-vectors-of-human-diseases>
-



Brazilian Free-tailed Bat, Mexican Free-tailed Bat
(*Tadarida brasiliensis*)



Supported by:

  <https://pmc.ncbi.nlm.nih.gov/articles/PMC5852992> Provider:  @article {Keck2025.01.24.634685, author = {Keck, Fran{\c c}ois and Broadbent, Henry and Altermatt, Florian}, title = {Extracting massive ecological data on state and interactions of species using large language models}, year = {2025}, doi = {10.1101/2025.01.24.634685}, journal = {bioRxiv}} Accessed via <https://github.com/fkeck/gpt_interactions/archive/24c1a83b16441b223073b19064b3a5e769aeb348.zip> at 2025-02-17T09:12:16.225Z. 

Refuted by:

None.

interacts with



Western Barn Owl
(*Tyto alba*)



Example of interaction claim indexed by GloBI using Keck et al. 2025 preprint

<https://doi.org/10.1101/2025.01.24.634685> Keck2025.01.24.634685, author = {Keck, Fran{\c c}ois and Broadbent, Henry and Altermatt, Florian}, title = {Extracting massive ecological data on state and interactions of species using large language models}, year = {2025}, doi = {10.1101/2025.01.24.634685}, journal = {bioRxiv}}

<https://www.globalbioticinteractions.org/?interactionType=interactsWith&sourceTaxon=Tadarida%20brasiliensis&targetTaxon=Tyto%20alba>

This preprint shares data extracted via some large language model and aligned names with ncbi taxonomy. See also https://github.com/fkeck/gpt_interactions

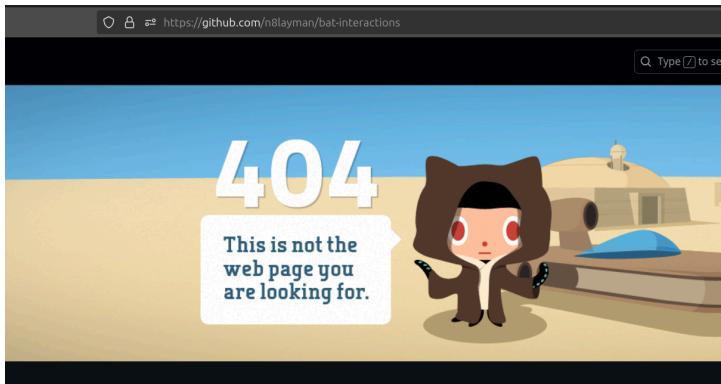
The interesting point is the data they actually used:

"The text corpus was sourced from the PubMed Central (PMC) Open Access database, which contains millions of publications spanning various scientific disciplines⁵¹. To narrow the scope of our analysis, we first refined our corpus to focus specifically on publications related to ecological sciences. Using the OpenAlex database⁵² and its topic classification system, we identified 99,555 publications tagged under the subfield "Ecology" from an initial pool of 6,159,719 PMC publications. We downloaded the corresponding 99,555 XML files from PMC and parsed them to extract 3,552,030 paragraphs for further analysis."

That means there are only a few journals covering biodiversity included and most of the data we assemble in our publications is not part of it.

- Nate Layman will be joining us soon to share his progress on the automation of our co-roosting (eco-interactions) data mining pipeline. If you have any questions for him, would you add them here? He is curious as to what we are interested in.

<https://github.com/n8layman/bat-interactions> -- Jorrit: 2025-02-26 can't access the repo, is it public?



19 Feb 2025

Present: DeeAnn, Jelle, Kendra, Anna, Nate, Jorrit, Donat

- NIH concerns
 - Federal register is blocked – so cannot announce new meetings
 - Bits of it are working
 - Fired all the at-will employees
 - But is this the tip of the iceberg?
 - NIAID potentially in the cross-hairs
 - Yearly money seems okay
 - View from Germany – surprised, upset, Ukraine
 - Europe having to raise defense budgets beyond current 2%
 - Elections next week in Germany
 - AfD dynamics
 - “Firewall” to not talk to the Nazi rightwing
 - Kendra – western Asia countries are filling into Putin’s crosshairs
 - Syria, Jordan, Turkey
 - Climate crisis, intersection of many things here
- Jelle at OpenAI now
 - Optimistic, lots of competitors
- BHL annual meeting in Berlin – IBPES will be speaking
 - Donat is organizing, Plazi
 - Bat work is relevant here
- DeeAnn – upcoming meetings
 - BatID meeting in Chicago in July

- Every 3-4 years, fourth time holding it
- Cara Brook is hosting
- Used to lack bat biologists – ecology and taxonomy was not part of it originally –
 - Meeting: <https://brooklab.org/programs/bat-id-2025>
- BTA especially relevant
 - DeeAnn could present on this
- Kendra also attending
 - Western Asia work, or just networking
- Nate can send slides to DeeAnn
- Got pushback on the Weber et al. paper
 - Weber, N., Nagy, M., Markotter, W., Schaer, J., Puechmaille, S. J., Sutton, J., Dávalos, L. M., Dusabe, M.-C., Ejotre, I., Fenton, M. B., Knörnschild, M., López-Baucells, A., Medellin, R. A., Metz, M., Mubareka, S., Nsengimana, O., O'Mara, M. T., Racey, P. A., Tuttle, M., ... Reeder, D. M. (2023). Robust evidence for bats as reservoir hosts is lacking in most African virus studies: A review and call to optimize sampling and conserve bats. *Biology Letters*, 19(11), 20230358. <https://doi.org/10.1098/rsbl.2023.0358>
 - Can talk to clarify this
- Other recent paper
 - Fearing, A., Faulkner, K., Smith, P., Humbrecht, E., Kyne, P. M., Feldheim, K. A., Moore, A. B. M., Brink, A. A., McDavitt, M. T., Smith, K. L., Whitty, J. M., Wiley, T. R., Wueringer, B. E., & Phillips, N. M. (2025). Assessing confidence in zoological specimen collection metadata for use in scientific studies. *Journal for Nature Conservation*, 84, 126815. <https://doi.org/10.1016/j.jnc.2024.126815>
- How to best use the poor quality data?
 - Power law type approach
- Jorrit is indexing a pre-print
 - https://github.com/fkeck/gpt_interactions/pull/1
 - Also BatNames –
 - Had published a PDF to Zenodo
 - But needed a machine-readable
- BatLit progressing
 - Cullen and Donat and moving that forward
 - Version 0.7 on Zenodo
- BTA
 - Connecting the input to the output
 - Hard to reconstruct – hand curation – not automated
 - Releases will become harder to do, trace errors,
 - Thinking about BTA as training data

- Input are more unclear than outputs

12 Feb 2025

Present: Aja, Nancy, Connor, Jelle, Kendra, Nate, Jorrit

- NIH concerns
 - Virus research could be targeted
 - Limiting research on bats is possible – infectious diseases
 - Bird flu
 - <https://www.nytimes.com/2025/02/06/health/cdc-bird-flu-cats-people.html>
- NSF concerns
 - Broader impacts aspect
 - Looking grant by grant
- Live work session with BTA manuscript?
 - Better as a subgroup – also Cecilia
- BatLit-to-MDD
 - Start with the bat literature linked in MDD
 - Find those papers, clean up, link to BatLit and MDD
 - Establish the workflow as based on bat portion
 - Then expand to rodents
 - All mammals
 - Question about needed specificity of the database — would this work across all mammals?
 - 50,000 PDFs from Jelle would definitely help
 - We could write best practices
 - Share resources
 - Re-use workflows
 - BLR on Zenodo –
- Taxonomic subcommittees
 - Better to have fewer larger groups
 - Not split bats up
 - Not split Carnivora
 - Benefit from consistency – helpful to have more people with a broader perspective
 - New vs Old world should be consistent
 - Can have structure within the groups
 - Larger group, allow for people step up, make things happen
 - Keep Ungulates together

- Marsup + monotremes
- Rodents all as one group?
 - Huge amount of overlap in methods and perspectives
 - Similar size to bats
 - Easier to invite involvement –
- Having all subcommittees in a single Slack could be good
 - Start channels with same few letters for each group
 - Main Google Sheet >> discussions
 - Use new Slack tools
- Do a virtual workshop every 2 years
 - Subcommittees discuss things
 - Would go against the regular updates to the MDD list though
- Formal communication vs informal on the Slack
 - Decisions, proposals
 - Zenodo white papers
 - Distill the conversations in Slack
 - Especially for more tricky taxonomic decisions
- BTA dynamics
 - Jelle helped update
 - Added Corbett and Hill 1 and 3 – and MDD v2.0
 - Also new BatNames version
- Subspecies
 - Could add to MDD for v2.1? Likely not
 - Larger process of curation – will need to work on this through time
 - Jelle has added his info to MDD currently
 - MSW3 subspecies – all valid are supposed to be in bold
 - BatNames
 - Only recognizes subspecies for the species that have “long accounts” – some species do not have this –
- Nate Layman
 - Automating the co-roosting dataset
 - Just pulling out text
 - Code: <https://github.com/n8layman/eco-interaction-extraction>
 - Now employed by Cullen – contract for all the way through publication
 - Awesome –
 - Can bring him to this call once gets some results
 - Finishing up at EcoHealth Alliance
 - Pushing out publications, aiming to
 - Spending down their accounts.

5 Feb 2025

Present: Aja, Hernani, Nancy, Adriana, Jorrit, Donat, Anna, Quentin, Connor, Jelle

- ~~Live work session with the BTA manuscript~~ - Last week had discussed using time to review and develop figures; did not pursue
- Bat Taxonomic Alignment
 - IUCN report has been sent off
 - Subspecies discussion (Connor, Aja, Jelle, Nancy)
 - New sources to integrate
 - MDD v2 in hand
 - Koopman 1994
 - Jelle to prioritize autonomously extracting data
 - Plazi to perform treatment
- BatLit
 - General confusion on Zotero-Zenodo connection
 - Currently metadata-only updates do not get pushed to Zenodo or new BatLit versions
 - Aja requests that metadata-only updates in Zotero transfer to Zenodo with new version updates too; Aja to make github issue
 - Jorrit to update pipeline
- Quintin
 - In Montersatt (?) for:
 - groundtruthing new species for the island
 - Encouraging establishment of GBIF portal
- Living Data conference in October
 - <https://www.livingdata2025.com/calls.html>
 - Quintin + Jorrit going/organizing Symposium on Vectorborne Diseases
 - Other abstracts/proposals will soon be accepted

29 January 2025

Present: Aja, Kendra, Anna, Jorrit, Nancy, Hernani, Nathan, Cecilia

BatLit Zenodo permissions / google form / user agreement

- possible citation - BatLit Community. Bat Literature Project (BatLit, <https://batlit.org>) v0.7 hash://md5/ce1e7618f237ff56e74d51f496e811a3 - Reference List, Metadata, and their Provenance (0.7.2) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.14662695> (Jorrit: this links to a Zenodo deposit describing the BatLit index).
- Request to cite, but no way to enforce it

- Idea of BatLit is to enable (i) access to literature; (ii) collective analysis of that literature
 - Broad agreement about this
 - But also some concern about how widely re-distributed the information could be – Language on the last question:
 - *Use of this literature collection is limited to research, educational, non-commercial "fair use". Colleagues are welcome to download, print, or use material from the catalog, for their individual research purposes, so long as the sources are cited appropriately. Users may not, however, copy material from the collection into other databases or onto other websites, or otherwise disseminate the information, without permission from the copyright holder.*
- Kendra suggests to include how-to videos in addition to textual instructions
 - Jorrit shared an example: Here's a how-to video
<https://opentraits.org/members.html> with Kermit the Frog !
<https://youtu.be/bgVd-OHdZH4>
 - https://docs.google.com/forms/d/e/1FAIpQLScTFfM3qNvhwTuMY_9xhvwyJpcccxAIR4UgEdp4dOlqNyPdVA/viewform

Update on the BatLit search

- Plazi search tool is now linked
- We could embed on the same website:
 - Embed tags:
`<embed src="http://www.example.com" style="width:500px; height: 300px;">`
 - Also iframes:
`<iframe src="/othersiteurl" width="100%" height="300">`
`<p>Your browser does not support iframes.</p>`
`</iframe>`

Manuscripts

- BTA – Cecilia has updated some of the document – the OUTLINE in particular
- BTA document:
<https://docs.google.com/document/d/1sWTh-GQqqJI95xdVbZC1ryZKlpNr7T2vUFHac3dySE/edit?tab=t.0>
 - Just the outline part for now
 - Aiming for PLOS Biology
 - Using bats as a case study for a broader tool
 - BTA
 - But couched more broadly
- Goals for future work
 - Wanting this to be more automated
- Focusing on Figures and Tables
 - related to <https://jhpoelen.nl/bat-taxonomic-alignment/> - get a sense of how disagreement

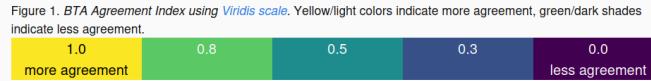
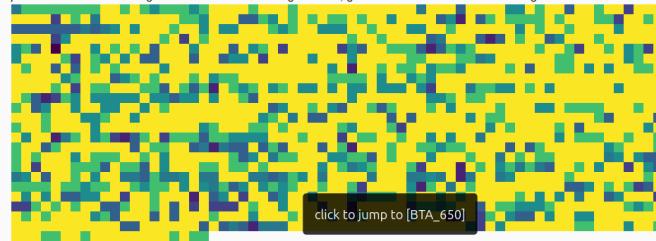


Figure 2. A clickable heatmap containing all BTA concepts and their associated agreement Index organized in no particular order. Yellow/light colors indicate more agreement, green/dark shades indicate less agreement.



- The Figures would be the narrative part
- Coworking times could be good
-

22 January 2025

Present: Donat, Hernani, Nancy, Jorrit, DeeAnn, Conner, Cecilia, Aja, Anna, Cullen

Coroosting data sets

- Nate Layman starts working on the co-roosting data set
- Compare this with Patrick SIBiLS approach

Paper publishing

- Decide in the topic and audience, one or two papers? Impact way beyond bats thus a more widely circulated journal
 - Pursue more general audience than just bat taxonomists
- Publishing where? <https://academic.oup.com/sysbio/article/65/4/561/1753624>
 - Criteria: the Virus community needs to read this journal, that it should be in PubMed;
- Journals:
 - Nature scientific data
 - “publish data descriptors” - not a good fit I think (DR)
 - Biodiversity Data Journal
 - Biological Reviews
 - [Systematic Biology](#) (DR)
 - Nature Reviews Biodiversity - new journal from Nature (HO)
 - Nature Methods
 - PLoS One (may be as followup on PLoS Biology refusal)

- [PLoS Biology](#) first target; methods and resources section
- Ecological Applications
- Methods in Ecology and Evolution
- [Scientific Reports](#)

How to govern the batlit community

- Policy paper

[Batlit membership](#)

https://docs.google.com/document/d/1hs3sVor9Jse7BkEgrvVV_dHlo9KHGE-olhFV7Hpcqbk/edit?tab=t.0

- How to start
 - Create a google form to sign including that they tick a box “i agree with the code of conduct”
 - How stringent should the policy be? Rather loose to enable as many people as possible to participate
 - Needs to have a bat research focus, so for each paper in batlit we can explain why it is there.
- How to get people hooked up
 - Allow to request not digitally accessible papers
 - Create a bibliography of bat literature online where for those references that have no PDF the PDF can be uploaded
- Additional streams
 - Publishers: talk to them

15 January 2025

Present: DeeAnn, Donat, Cecilia, Jorrit, Kendra, Anna, Nate, Aja,

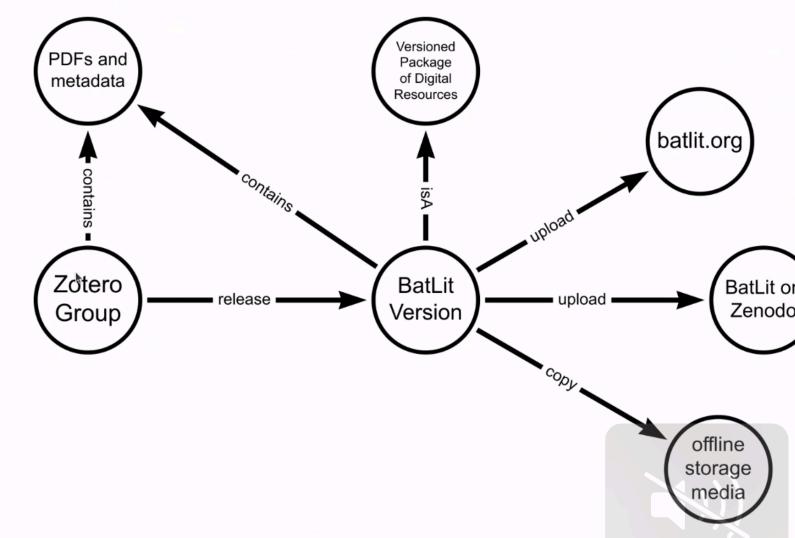
Kendra to move to Minnesota – to work with Peter Larsen – expand bat biosurveillance (One Health) program using field-adapted laboratory diagnostics

- Moving the first week of April – but starting in two weeks remotely
- Peter developed field-adapted BSL2 workflows
 - Metabarcoding bat species in the Amazon
 - Prion work on deer – spongiform encephalitis
- Univ Minnesota
 - Vet school

Bat Lit

- Some confusion on what the goal of the project is
 - Donat: Goal should be to make all these publications accessible – e.g., the Africa Bat Report – make everyone able to access
 - Lots of emphasis on reproducible workflows

- Dona ist working on a manuscript that outlines his vision on workflows that involve batlit and other projects:
 - https://docs.google.com/document/d/1pVhXhRDTRJPgp8om30_RoV5rpONr0lrjGYQniqICqWM/edit?usp=sharing
- Overall criteria for how what we want BatLit to be / become
- Bat Lit is involved in large ecosystem
- But Bat Lit Corpus is a well-defined digital work
- Donat makes the point that Batlit sensu Jorrit is part of the vision we should agree in terms of building the batlit community in Zenodo.
- Jorrit
 - Reviewing <https://batlit.org/>
 - Goal is to have easier access to bat literature
 - Suggest to have distinction between BatLit Corpus and BatLit Project
 - May be disagreement here, should resolve



- Zotero as the reference organization part of this
- To upload publications not in zotero and baltic community, Aja explains how to clean up the metadata, for example to remove html tags or urls . The Zotero group is part of the community, not open
- Aja manages the batliterature group in Zotero, so far 9 members. For access ask Aja.
- Definitions of Users vs. Curators
 - Cullen/Donat working on this: [BatLit Governance Document](#)

The screenshot shows a Zotero library window with several tabs open at the top: "Systematics of big-eyed bats", "Harmonizing taxon names", "The taxonomic name resolution", "Bats in the Anthropocene", "Inferring the mammal tree", "mhZapph", and "Recommendations for us". The main pane displays a list of articles under the heading "My Library". One article is selected, showing its full bibliographic details in a right-hand panel:

Morphogenetic Effect of L-cysteine on *Pseudogymnoascus destructans* and Related Species

Info

Item Type: Journal Article
 Title: Morphogenetic Effect of L-cysteine on *Pseudogymnoascus destructans* and Related Species
 Author: Raudabaugh, D B
 Author: Miller, A N
 Publication: Mycosphere
 Volume: 5
 Issue: 6
 Pages: 737-746
 Date: 2014
 Series:
 Series Title:
 Series Text:
 Journal Abbr:
 Language:
 DOI: 10.5943/mycosphere
 /5/6/4
 ISSN:
 Short Title:
 URL: http://ezproxy.stir.ac.uk/
 /login?url=http://
 /edsweb.wustl.edu/login.aspx?direct=true
 &db=edsweb&
 AN=000363227200004&
 site=eds-live

Accessed: Archive
 Loc. in Archive

Proposed agenda item:

1. group review of <https://batlit.org> and work through examples.

 **Jorrit Poelen (GloBI)** 10:27 AM
 @anna suggested to take some time in our Wednesday meeting to go over <https://batlit.org> and work through some of the examples. I think this may help to figure out ways to facilitate reuse of this rich, carefully curated, versioned dataset of Bat Literature. Curious to see whether there's animo for this.

bat-literature
BatLit: The Bat Literature Project
 The Bat Literature Project aims to facilitate discovery of

- Established that Zotero is the internal editing platform and Zenodo is the platform where the larger community can access pdfs
- Anna has been struggling with sharing the Zenodo platform, how to connect them to the collection, and how they can add to the corpus
- A huge need to establish protocol for vetting and including members to the Zenodo community,
https://docs.google.com/document/d/1kuDALvMlfQrlh9mhHnzOdsEm88_Xjzvl/edit
- Anna wants to access v.1, can she do that without Jorrit's instruction? No. Users cannot download a copy. Jorrit can, but this workflow would need to be established in the future. (Jorrit notes on 2025-01-22: users can download copies as long as they have access to Zenodo and/or Zotero).
- A file like this would allow to get access to all the articles.
<https://docs.google.com/spreadsheets/d/1i1qV-OuWo5-UhAN1DdalLYuxDylZnZDd/edit?usp=sharing&ouid=117441245162738210165&rtpof=true&sd=true>

- ◆ From Anna - How does this list give access to articles? What is different on this list vs. there several ref lists on github/the “version description”: see <https://github.com/bat-literature/bat-literature.github.io/issues/53>
- Backup SD storage is not allowed

Present: Cecilia, Donat, Aja, Jorrit, Ariadna, Connor, Kendra, Anna, Nate

The screenshot shows a search interface with the query "make up a description of a high desert myotis bat species that" entered. Below the search bar, there are filters: "All" (which is selected), "Images", "News", "Videos", and "Goggles". The main result is an article titled "Bat Species Loses Wings Gains Tail". The article's text discusses the High Desert Myotis (Myotis desertus caudatus) and its unique adaptations to arid environments.

Bat Species Loses Wings Gains Tail

The **High Desert Myotis** (*Myotis desertus caudatus*) is a unique and adapted species of bat that has evolved to thrive in the arid high desert environments of North America. This species is characterized by its remarkable loss of wings and concomitant development of a long, slender tail.

Physical Characteristics

- Body length: 7-8 cm (2.8-3.1 in)
- Tail length: 15-18 cm (5.9-7 in), making up approximately 60% of the total body length
- Weight: 5-7 grams (0.18-0.25 oz)
- Fur: Thick, woolly, and grayish-brown in color, providing excellent insulation against the harsh desert climate

Jorrit posts idea that ChatGPT is making up species descriptions

- Similar to <https://en.wikipedia.org/wiki/Rhinogradentia> idea (Donat)
- Cautionary tale
- Also articles that are intentionally clowning, e.g.: <https://www.chm.bris.ac.uk/sillymolecules/birds.pdf>
- (Jorrit: what is the litmus test to (a) detect material originating from trusted sources (b) have not been altered after publication/citation)

Aja work on IUCN taxonomy changes

- Taxonomy changes tracked

- https://www.dropbox.com/scl/fi/1di9xwajardyn7ha3t0o1/Taxonomy-changes-in-SI_S_Chiroptera_2025.xlsx?rlkey=rt4q4yvri4dh9lizrie5wbfm&e=1&dl=0
- SIS sheet
 - https://www.dropbox.com/scl/fi/p87hpa4oxsix9u86e08ya/SIS_Add_Taxonomy_Request_2025.xlsx?rlkey=2xkng6iqc2fjkswe71iwtu9a&e=1&dl=0
- General work on figuring out how this SIS system works
- SIS folks are all pro bono
 - If we overthink this, it won't work so well
 - Goal is to upload things ASAP
 - Don't want to slow everything down necessarily
- Two part goal here really
 - Improve bat taxonomy
 - But also improve the SIS process

Manuscripts for the Bat Eco-Interactions group

- BTA manuscript

Zenodo DOI discussion / suggestions

- <https://batlit.org/refs>
- Reference page that lists everything
 - Helpful to have all info in one place
- But don't need have directly link to the publisher DOI, and do not misuse the publisher DOI

we should complement or provide the references in Batlit

- Krishna (06/2018) [Adiposity associated changes in serum glucose and adiponectin levels modulate ovarian steroidogenesis during delayed embryonic development in the fruit bat, Cynopterus sphinx](#) *General and Comparative Endocrinology* hash://md5/66bf44b07c49bdcb47c6ce598dda8746 doi: [10.1016/j.ygcen.2018.02.016](https://doi.org/10.1016/j.ygcen.2018.02.016)
- similar to the reference in the Zenodo Batlit citation
- Krishna, A. (2014). Role of Adiponectin in Delayed Embryonic Development of the Short-Nosed Fruit Bat, *Cynopterus sphinx*. *Molecular Reproduction and Development*, 81(12), 1086–1102. <https://doi.org/10.1002/mrd.24620> and add the hash hash://md5/66bf44b07c49bdcb47c6ce598dda8746
- Add similar to Taxodros a dataset that includes the batlit database <https://zenodo.org/records/13841002> in Zenodo, including a user friendly CSV or similar file, including at least the Zenodo ID so users can install a bibliographic database on their local system. BTW the huge number of downloads of this DB in taxodros (>5K) indicated an interest in the community for this.

Living Data call

- <https://www.livingdata2025.com/calls.html>

18 December 2024

Present: Aja, Connor, Nancy, Quintin, DeeAnn, Jelle, Ariadna, Kendra, Anna, Jorrit (joined late), Cullen, Donat

- Aja update on IUCN 200+ taxonomic change recommendations
 - (Nancy) submitted files need to be in exact format- IUCN strict on this
 - Submitting to general group for broad taxonomic shifts, not order-specific groups
 - <https://www.iucnredlist.org/assessment/authorities>
 - Also providing detailed guidance to bat specialist group for once taxonomic changes are implemented
 - Need to get contact info from Dave

11 December 2024

Present: Cecilia, Nancy, Jelle, Donat, Aja, Jorrit, Hernani, Ariadna, Iroro, DeeAnn, Connor, Kendra, Anna

Agenda:

1. Australian Bat Symposium The 20th International Bat Research Conference (Aja)
2. BigBatDB update (Jorrit, Ariadna)

Ariadna is planning to go to <https://www.ausbats.org.au/ibrc2025.html> 3 - 8 August, 2025 – Cairns, Australia and is asking about some email.

Cecilia was contacted by organisers of a symposium inviting her as a speaker for a digital databases symposium. She forwarded email to Aja and Nate who are better suited to submit some titles for potential talks on behalf of Eco-interactions groups.

Ariadna on the BigBatDB - we decided to collect dental formulas and index it by Encyclopedia of Life (EOL) - (step 1) transcribe dental formula in a Google Sheet
<https://docs.google.com/spreadsheets/d/1xGPXuwEuqEao2ezya7kW2DWixp4pejqVEjKFD1sqG9g/edit?gid=0#gid=0> . Ariadna created a google form to help facilitate data entry
https://docs.google.com/forms/d/1w4_ttyMr5r1cfQQ-XqYBGZ-B5NEgapyqoaMsEVSQWaY/edit . Ariadna is planning to take the form responses and reformat them into the dental formula suited for EOL indexing and name alignment.

Hernani mentioned a reptile trait database that may be a source of inspiration.
<http://www.reptile-database.org/>

A link directly to a copy of Koopman
<https://linker.bio/hash://md5/70c3e0fba7379e09e95a38569fe29da7>

same, but with page 77
<https://linker.bio/pdf:hash://md5/70c3e0fba7379e09e95a38569fe29da7!p77> renders a single page

Similar derived pdf for koopman page 77-80 ranges.
<https://linker.bio/pdf:hash://md5/70c3e0fba7379e09e95a38569fe29da7!p77-p80>

Cecilia / Hernani suggests including the date and locality information as a separate field to help with the name alignment.

Arianda - how to best capture the locality information

Anna - suggest to include verbatim (what is in the text, e.g., "southern Mexico to Nicaragua"))

T. c. coffini (southern Mexico to Nicaragua), *T. c. cirrhosus* (Costa Rica to Amazonian Brazil), *T. c. ehrhardti* (eastern Brazil and perhaps Bolivia).

Anna - do keep track of the basisOfRecord for the dental formula

<https://github.com/BigBatDB/bigbatdb/issues> - contains all the issues

<https://github.com/BigBatDB/bigbatdb/issues/5> - Katja Schulz (of EOL) suggested to include Lifestage - Nancy says that (for bats) are only for adults due to their specific dental development, different from other mammals.

Discussion of extracting trait data using OCR or similar. No script for doing this from Koopman 1994 exists yet. There is structured data from the HMW series <https://hpoelen.nl/hmw/> (Jorrit: this work was extracted from already semi-structured data)

What do we get out of doing this inside BigBatDB? Idea is to build an expandable database that can be extended into more data, instead of a single spreadsheet that can be forgotten. This is the test case for BigBatDB.

Name alignment. Jorrit explains the data format for BigBatDB releases. Standardized (Darwin Core) format allows doing name alignment using standard tools.

github.com/bigbatDB/bigbatdb-name-alignment

Allows aligning with a variety of other sources. Run a GitHub Action that produces an alignment report summarizing the results.

Aja: Taxonomic alignment for Bat1K.

4 December 2024

Present: Cecilia (left early), Nancy, Nate, Jelle, Donat, Aja, Jorrit, Hernani, Ariadna, Cullen, Iroro

Aja action items for IUCN-MDD matchup

- <https://docs.google.com/spreadsheets/d/164fcpl7i36MC4ZUElrzd2Zj06fiQndlBO9k-wISX4aE/edit?gid=204177465#gid=204177465>
- To enable the IUCN update for bats
- Linking to BatLit at the same time
 - Good way to ensure that BatLit is complete
 - Can do this for the BTA more generally
- Wanting to mark down the per species assessment date + change date in the literature
- Will follow up further in a separate call

Cecilia wants titles for potential talks for a symposium at IBRC in Australia

- <https://www.ausbats.org.au/ibrc2025.html>
- 3 - 8 August, 2025 – Cairns, Australia
- suggests to come up with a title / abstract
 - Jorrit suggests: "IUCN Conservation as a driver to preserve and develop digital research infrastructure."
 - Pragmatic, practical means of implementing now
- Email notice:
 - "I am putting together a symposia proposal titled "Digital resources for bat researchers and managers" and was hoping I could include you or someone in your group to talk about your online resources."
 - The symposia is designed to advertise and discuss issues around developing these sorts of tools and databases"
- Could plan to have the BatLit be more advanced
 - Should apply to everyone – who uses literature
 - Live tools, with BiodiversityPMC
 - Data organized, links to publications
 - Not just ideas, but tools
- The Taxonomy work has a smaller user group and interest base

IUCN reassessment list

- 200+ species that are valid but not in the IUCN list
- Geographic ranges are not matched to these new lists – we do have some MDD maps, but not precise – could still be used
- Body size
- In 5 years or so we could do a much better job at this

- Then can look at predictive power to greater extent
- Iroro has a paper looking at bat extinction risk
 - Using IUCN data as current

Linking the BTA to BatLit more generally → taxonomic alignment tool, how could we use it for GBIF / CoL, which is more directly used for research

- Much more interested in alternative taxonomies at GBIF now
- Publications – but then taxonomy from the publications
- May need to get more specific when say “GBIF”
 - Not just one thing
 - Put datasets in one place
 - Tools are getting better so we can zoom in more
 - GBIF is accumulating data – not producing data – data model used to be only occurrences, but now it is broader
 - iDigBio, DiSSCo are doing lots of digitization, but don't care about the linking as much
 - Name system is how we do the linking
- How to use BTA to link datasets?
 - Can implement this in Nomer perhaps
 - Coming up with a super simple use case – that is the best place to get started
- We can document how Aja is doing this
- Nomer
 - A name alignment tool
 - <https://github.com/globalbioticinteractions/nomer>
 - Could be key for doing the name-based linking part of this
 - Use a versioned copy of the BTA for this – well-defined sandbox to test/play
- Name- vs concept-based matching
 - Taxonomic Data Objects for Communicating the Meaning of Species Names
 - Nathan S. Upham, Jorrit Poelen
 - <https://doi.org/10.3897/biss.8.139413>

Example

- Trachops cirrhosus
- <https://synospecies.plazi.org/#Trachops+cirrhosus>
- We could track this through manually through the BTA
 - From Jorrit to Ariadna - please send me the row in the BigBatDB you'd like to align (the name) to BTA

Jorrit wonders about our current strategy to coordinate activities of Bat1k, Big Bat DB with projects like BatLit, BTA, GloBI, Plazi's BLR, MDD etc.

20 November 2024

Present: Kendra, Cibele, Cecilia, Nancy, Nate, DeeAnn, Anna, Jelle, Donat, Connor, Iroro

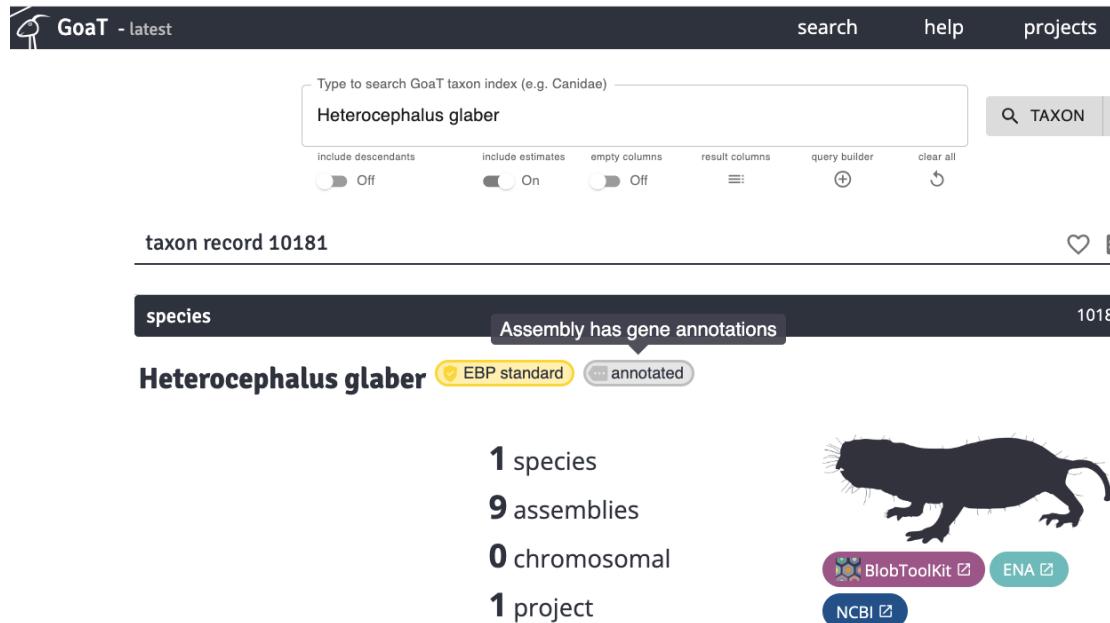
Cibele joining from Darwin Tree of Life / Wellcome Sanger Institute

- Part of time is on the Earth Biogenome Project
- Genomes on a Tree (GoAT)
 - <https://goat.genomehubs.org/>
 - Large data system that underlies
 - Runs on NCBI taxonomy (as backbone)
 - Stores lots of data related to genome assemblies
 - Chromosome number, genome size
 - Inferred ploidy
 - Runs on Bioproject IDs
 - Working closely with NCBI taxonomy curators
 - Now seem willing to change some aspects of taxonomy
 - But needs to be done at the *record* level rather than wholesale taxonomic changes
 - Would result in issues of taxonomic misIDs
- ToL IDs
 - Known to be a problematic issue
 - Not robust to taxonomic change
 - Moving to align these with NCBI taxon IDs
 - Works as well as the NCBI taxonomy does
 - Article on using the TaxIDs –
<https://wellcomeopenresearch.org/articles/9-591>
 - Developing a pipeline for submitting batch changes to the taxonomy of a given set of records
 - Could work for individual sequence records also
- Similar to issues encountering with the IUCN

Linking to taxonomy → but also want to link to taxonomic treatments

- Plazi linking to the genomes themselves: <http://plazi.org/treatmentbank/>
- Existing links with GBIF
- ENA – european NCBI
- Assigning a name to a genome – are referring to one concept of a given name – should be explicit about this >>
 - On the one hand a small thing – but becomes much larger when talk about AI linking of biodiversity knowledge – need explicit links of this specimen / genome is a representative of taxon A according to author B
- Potential to do for all EBP – but starting with the bats
 - BatLit project

- BatLit on Zenodo:
<https://zenodo.org/communities/batlit/records?q=&l=list&p=1&s=10>
- <https://biotxplorer.sibils.org/?species1=Pholidota&size=1>
- Exploring taxonomic names via Synospecies, a triple store based on taxonomic names in Catalogue of Life and TreatmentBank. See eg [Rhinolophus sinicus](#)
- Possible meeting to be arranged with Checklistbank to discuss their approach to alternative taxonomies and how this could be combined with the Bat-eco group members taxonomic names subprojects?
- CSC Post meeting thoughts: Ways GoaT could immediately help:
 - Workflow 1: a)The working group could double check bat names and ncbi taxIDs that have genomes. b) Associate specific BioSample used to produce the genome with most relevant treatments. c) If there is a validation for used name = ncbi taxID = specific treatment, goat can add a validation tag "valid Name" or "Taxonomy verified" in the record's page. (like the tags in the image).



The screenshot shows the GoaT taxon index search interface. At the top, there is a search bar with the query "Heterocephalus glaber". Below the search bar are several filter options: "include descendants" (Off), "include estimates" (On), "empty columns" (Off), "result columns" (dropdown menu), "query builder" (button), and "clear all" (button). A "TAXON" button is also present.

The main search result for "taxon record 10181" is displayed. The species "Heterocephalus glaber" is highlighted. A tooltip indicates that the assembly has gene annotations. Below the species name, there are statistics: 1 species, 9 assemblies, 0 chromosomal, and 1 project. To the right of the statistics is a silhouette of a mole rat. Buttons for "BlobToolKit" and "ENA" are shown, along with a "NCBI" link.

- Create a new column for the “number of treatments” that would contain a link-out to the equivalent search result for that name, like in the [Rhinolophus sinicus](#) example. That would at least inform the user that disambiguation might be needed.
- Current columns under “names” that can already be filled if needed:

scientific_name	taxon_id	synonym	tolid_prefix	common_name	≡
<i>Heterocephalus glaber</i>	10181	<i>Heterocephalus dunnii</i>	mHetGla	Naked mole-rat	>

- GoaT can produce a list of names not found on NCBI taxonomy that can be used to prioritize the names Bat1K needs to request from ENA/NCBI. This is publicly available in our S3 store, but currently outdated: [Not found](#) | [spellcheck suggested](#). Next update scheduled for Jan25.

13 November 2024

Present: Aja, Quentin, Cecilia, Nancy, Nate, Connor, Donat, Meike, Kendra, Anna, Ariadna
Big bat database

- https://docs.google.com/spreadsheets/d/1gtENmce4tv2AJWduWrj6dKnDFE0O-tLNZVA_G6BAn5r4/edit?gid=0#gid=0
- How to keep track of changes to taxonomy in the context of traits
 - Record locality, date
- Focus on primary versus secondary literature
- Donat: <https://zenodo.org/records/13505746> or <https://doi.org/10.5281/zenodo.13505746> for the publication
 - Also include the BatLit link to the Zenodo record – allows for linking to the extraction of trait via machine learning – subsequently
 -

Similar to dynamic of the Bat1K data – how to keep/organize the metadata

- Meike – tracking the metadata / sample coordination
- All genomes derived from only one organism
- Already collecting lots of metadata
 - Geographic location, not all with photos
- They use <https://goat.genomehubs.org/> – GoaT
 - Uses NCBI taxonomy likely – could have more taxonomic data maybe from MDD?
- Quentin: legacy of mess of thinking that have a genome of species not a specimen
- Backwards data curation is what's happening
 - Important to talk to the people who collected the specimens
 - Collector of the tissues / submitter needs to be the one to link the taxonomic treatment to the sample
- Need for a *review / perspective* article to clarify what types of metadata should be associated with a given new genome sequence
 - We could work on this as a group
 - MIXS: <https://github.com/GenomicsStandardsConsortium/mixs>
 - MIXS: Minimum Information about any (X) Sequence
 - <https://www.gencsc.org/> – Genomic Standards Consortium
 - Data nerds already know about this
 - But what's needed is a broader paper written in plain language
 - Science Forum: The critical importance of vouchers in genomics
 - <https://elifesciences.org/articles/68264>
- 3 letter codes for genome names
- Current input form for Bat1K:
https://docs.google.com/spreadsheets/d/10cz-E3aLc1vA_IqYFgQQKPdCdQwIdZTt/edit?usp=sharing&ouid=113418377508345956042&rtpof=true&sd=true

From Jorrit:

"We are currently visiting extended family in Totnes, Devon, UK: in my absence, I hope that you'll give Meike Mai and Cibele Ciao of Bat1K a warm welcome (should they decide to attend) and chat about ways to collaborate and help establish bidirectional links between https://bat1k.com genomic information with MDD/BatNames as well as BatLit and possibly other projects. For some ideas, please see <https://github.com/jhpoelen/bat1k-dwca> and <https://github.com/jhpoelen/bat1k-talk-2024-10-27>. Enjoy!"

Linking sequences to other data including the voucher

- <https://biodiversityknowledgehub.eu/wp-content/uploads/2024/07/ENA-Cross-Reference-Search-Factsheet.pdf>
- <https://www.ebi.ac.uk/ena/browser/xref>
- <https://ena-docs.readthedocs.io/en/latest/retrieval/programmatic-access/cross-reference.html>

6 November 2024

Present: Aja, Jorrit, Quentin, Cullen, DeeAnn, Cecilia, Nancy

Jorrit - Bat1k - Nancy on the steering committee

Bat1k -[: <https://bat1k.com/about/> refer

Bat1k -[:uses names from]-> MDD/ BatNames

Bat1k (genomes)

MDD / Batnames / BTA (taxonomic stuff)

BigBatDB (traits)

BatLit (literature)

BatBase/GloBI (interactions)

How do our activities connect to other related initiatives like: GBatNet working groups.

<https://jhpoelen.nl/bat1k-talk-2024-10-27/#/title-slide>

slides sources and data <https://github.com/jhpoelen/bat1k-talk-2024-10-27>

Jorrit explained the Bat1K interactions he had at NASBR and the scripts (Preston, Nomer) he used to compare names in Bat1K with MDD (see above). Cullen suggested including name and location. It was clarified that more data will be collected to update the dataset with needed location and other metadata, like how the data were obtained. Next steps to turn metadata into

archive so it can be indexed by GBIF to see what assemblies are in Bat1K without exposing them directly. Nancy said there is a lot of complexity with assemblies which is why they are using Globus data storage, including annotations. Time is an issue because they have a bandwidth problem while they are preparing the paper. Jorrit is seeing a lot of connections between MDD/BatNames and how he can help them link to other places and make the links more explicit. Nancy said they are now dealing with the issue that bat names will change overtime and these need to be included in their dataset. Jorrit planted the idea of data repackaging and wants to know what's next: Jorrit and Nate are on the Bat1K slack, invited Bat1K to come to one of our meetings.....

Aja pointed out that using automation to compare name lists have been problematic and it would be good to explore how we can improve it. Jorrit recommended adding BTA into the Nomer search (?). DeeAnn wants to use BTA in a paper. Group discussed publishing short paper on the tool. Nancy said we could publish in Biodiversity Data Journal and then DeeAnn could cite it, etc. We could use a few examples (*Pteronotus*, *Miniopterus*).

Note that <https://github.com/globalbioticinteractions/nomer/issues/157> already documented a desire to integrate BTA in Nomer. DeeAnn's list would be make for a good test case to implement this automation.

Next step - create diagrams to help show connection between the different projects

30 October 2024

Present: Aja, Jorrit, Nate, Connor, Hernani, Jelle, Kendra, Anna, Quentin, Jorrit, Anna, Cullen, Donat

Nate netted bats with Nancy near Tempe in a *Macrotus californicus* (insectivore, see also <https://www.mammaldiversity.org/taxon/1004950>).



Anna – bats that are eating hornets in Georgia

- Invasive hornets (*Vespa velutina*)
- <https://www.globalbioticinteractions.org/?interactionType=interactsWith&sourceTaxon=Vespa%20velutina&targetTaxon=Chiroptera> - no indexed interactions in GloBI Bat <> Vespa
-

Gary McCracken (<https://eeb.utk.edu/people/gary-mccracken/>) interaction at the airport

Asked where certain papers were in BatLit – he's writing a book right now, so close to access for references

- Idea to have all references freely available within the BatLit repo
- Broader access to BatLit
 - Few people requesting to become members, but not a ton
 - Select folks that we can slowly assemble

Kendra is working on a paper with Iroro on a trait paper for all IUCN assessed species

- Can integrate that effort with the Big Bat Database

BigBatDB v0.1 triggered ideas on structuring data by Katja Schulz of Encyclopedia of Life -

<https://github.com/jhpoelen/bigbatdb/issues/1>

GloBI <> Virion - <https://github.com/globalbioticinteractions/globalbioticinteractions/issues/665>
<https://github.com/viralemergence/virion/issues/67>

as of 2024-10-30 - (with weeks ago being 2024-10-17)

<https://github.com/viralemergence/virion/issues/67#issuecomment-1222619458>

The screenshot shows a GitHub issue comment from user **jhpoelen** dated Aug 22, 2022, at 11:42 AM CDT. The comment text is:

@tpoirot @cjcarlson what is the status of this issue?
Also, suggest to rephrase the title from "Recursion problem with GloBI" to "improve methods to de-duplicate VIRION records"

A smiley face emoji is present below the comment.

Below the comment, user **cjcarlson** has closed the issue as completed 2 weeks ago.

Lots and lots of databases in Verena: <https://www.viralemergence.org/data>

"October 17, 2024: The GLOBI dataset has been de-indexed from VIRION, due to a growing number of concerns about the reliability of text-mined data. This represents an important departure from the pipeline described in the *mBio* publication. The current pipeline is described at the top of this README." – bottom of <https://github.com/viralemergence/virion>

- Very strange
- GloBI does not focus solely on text-mined data
- It indexes existing datasets see <https://globalbioticinteractions.org/datasets> .
- suggestion to update GloBI's description/issues at
<https://github.com/viralemergence/virion/issues/93>

BHL and the internet archive going down

- Relative to Plazi converting BHL articles into Biodiversity PMC – current approach is using a professional vendor to do the scan and OCR – then Plazi is extracting from that
- Very few people support / use BHL (Jorrit: could you be more specific?)
- But it is clearly essential to the taxonomic enterprise
- BHL operates like a library – just digitizing whole library catalogs, but some of it is less valuable (Jorrit: less valuable because ...?)
- How does one really cite BHL? Tends to just be mentioned in the methods –
- For mammals
 - Have opportunity to get all the original descriptions
 - And then get all the modern revisions
 - And then we step through all of this – showing the value of reconstructing the conceptual and nomenclatural changes
 - Anna *did* this for ringtails
 - Citation method doubles the number of papers in the reviews
 - https://docs.google.com/spreadsheets/d/1smvZJh4V2Av_CVD55evzYBe_UWI7eyGsevEYMkdeUvQ8/edit?gid=2112296708#gid=2112296708
- BatLit
 - Has the value of showing how you turn a big collection of literature – into something that you can compute with
 - Count all the words that start with A, or something more relevant
 - Aja: goal to make sure that the abstracts are included in the metadata

Two members of the Bat Eco-Interactions group won awards at NASBR

COP16 in Colombia – NASBR folks were there

- <https://www.science.org/doi/10.1126/science.adu0409>
- <https://www.science.org/doi/10.1126/science.adg0344>

Jorrit: I am thinking about ways to have batlit <> bat1k <> MDD <> BTA linked through literature records and taxonomic names, genome sequences. Hoping to discuss next time.

- Mos def

23 October 2024

Present:

- In person! Aja, Jorrit, Cecilia, Cullen, Nancy, Ariadna, Nate, Connor, Hernani, DeeAnn (left early?)
- Online: Jelle, Kendra, Anna, Quentin

How to discuss the BatLit service while at NASBR

- Legal framework
 - Anything that is 2000 and before is open access
 - Anyone can jump into that by going to batlit.org
 - You can help curate
- Password access can be given to certain researchers
 - They can reach out to info@batbase.org
 - Contribute and use the repo
 - How do we decide this?
 - Communities of practice
 - We want to establish this
 -
- Reach out to people
 - Ask for theses, gray literature, reports – stuff that can actually disappear
 - Anna: theses in Mexican public university library could be a nice start.
 - This is a list of repositories I've put together:
 - <https://docs.google.com/spreadsheets/d/1k2Ob-OKEqaZogg5NDv8vEpfFe1fNofQhOv4bejR7Kss/edit?usp=sharing>
 - UNAM has 568+ bat theses
<https://repositorio.unam.mx/contenidos?&q=murcielago>
 - Creative Commons BY-NC-ND 4.0 Internacional
 - <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>
 - Membership at BatLit?
 - CV too much?
 - Tiered system, multiple ways to engage
 - A lot is already open
 - Example for liberal membership from UGA's Center for Ecology of Infectious Diseases
 - <https://www.ceid.uga.edu/about/information-for-members/>
 - Co-roosting talk
 - Extrapolation of the 300 publications to the expected full diversity of bat species interactions → 5200 total publications on co-roosting, many of which don't exist
 - AlterNate shared some slides of how to use the co-roosting archive

- Showing how we would extract these data
- Comparing the ML data extraction aspects with the rule-based vs GPT contexts – validating – human in the loop too
- Important to build in checks to the workflow
 - >> Education
 - Important to have an *intention* to understand the technology at every stage
 - Make sure that have off-ramps, backups
- Anna: can rerun summary scripts from many years ago:
Output:
<https://covid19taf.slack.com/files/U01P3UH96KA/F07TKAR1W81/status-of-lit-extractions.pdf>
- Script:
<https://covid19taf.slack.com/files/U01P3UH96KA/F07SS48S6KZ/status-of-lit-extractions.rmd>

16 October 2024

Present:

Jelle, Donat, Aja, Kendra, Jorrit, Cecilia, Cullen, Nancy, Anna, Ariadna, Quentin

Data review issues: it always needs more people to review and “clean” datasets

NASBR

- Where to share the files: [GoogleDrive](#)
- Our symposium is Friday morning
- Cullen will let everyone know where to meet before the actual symposium, may be for breakfast or another time before.

IUCN

- IUCN Species Specialist Group
 - Bat
 - Bat oneHealth
 - Redlist group
 - Taxonomic Backbone - Jorrit: is the IUCN taxonomic backbone openly available? Who is managing it? Should I make an effort to add IUCN taxonomic backbone to the Nomer Corpus of Taxonomic Resources? Answer: The Redlist is the taxonomic backbone and is available online.

Aja's presentation:

- Bat taxonomy under control by Nancy; there are no processes in place; IUCN is sometimes reassessing the taxa in the Redlist. An effort is on the way to create a pipeline to update names in the IUCN backbone taxonomy.
- IUCN versioned?
 - Each assessment has a date, which allows to get back
 - Yes, but the back-log isn't readily available

Reference local species/databases:

- <https://enciclovida.mx/especies/33942-anoura-geoffroyi>
- More complete language translations

AI perspectives

How to continue the discussion with Nate and Patrick

- Learn more about the black boxes
- Skin the cat

Proposal by Nate

- 1.1. Data Schema Design
- 2.1. OCR (Optical Character Recognition)
- 2.2. Output Structuring
- 2.3. Establish Performance Metrics

“Draft a grant” to progress?

e.g.

<https://new.nsf.gov/funding/opportunities/innovation-infrastructure-innovation-biological-research>

- First define the research question that you are trying to answer (not that you could imagine that could be answered - this because answering needs a commitment later, once a system is built)
 - Go back to the original question driving the groups:
 - outbreak of a pandemic
 - Prediction of a pandemic, using interactions, geography, behavior, roosting near human populations, reservoirs
 - Disease risk
 - What are the databases needed for above?
 - Build a system to get data out publications
 - Data integration
 - Create a data format that allows reuse of data extracted from publications

Next meeting:

Same time, just for some of use in the same room at NASBR

9 October 2024

Present: Aja, Nancy, Donat, Jelle, Nate, Cullen, Jorrit, Kendra, Ariadna, Anna, Cecilia, Nate L, Patrick Ruch, Quentin

Focus today on text-mining

- Patrick Ruch: <https://www.sib.swiss/patrick-ruch-group>
- Nate Layman: <https://www.ecohealthalliance.org/personnel/dr-nathan-layman>

BatLit community in Zenodo >> staging area for text mining

BiotXplorer, SIB approach to data extraction and interaction labeling

- <https://biotxplorer.sibils.org/>
- Two species and an interaction type
- Brings the passages from the text with it, and the PMID
- How good is the tool?
 - Did two benchmarks – relative to 100 manually curated triplets and 100 triplets from GloBI
 - GloBI data is about accessing existing species datasets, from very heterogenous sources – citing hundreds of datasets that are cited
 - All datasets in GloBI:
<https://www.globalbioticinteractions.org/datasets>
 - ML-derived dataset:
<https://depot.globalbioticinteractions.org/reviews/EOL/pseudonitzchia/>
 - Result: BiotXplorer has a lower accuracy with the manually curated data, as compared to the GloBI data (with known provenance)
 - Could scale this result to 100k triplets rather than 100, and then GloBI would be the most relevant
 - Positive side: all datasets in this comparison have known provenance
 - <https://github.com/globalbioticinteractions/globalbioticinteractions/issues/556>
 - Using the ROBI relationship ontology
 - Interaction types – synonyms of those interactions – some improved, but some increased false positives
 - Next step: looking at the agents of the interaction – species / taxonomy – but the *intersection* of the taxonomic category and the interaction type
 - Overall
 - Uses a traditional NLP approach – sentence splitting – regex matching – stemming of words / fuzzy matching – expecting 74% precision currently
 - Classifiers rather than LLMs – binary classification
 - Sensitivity – not estimating this currently – would require a human to go through a 20 page paper and list all the interactions
 - Tokenizing is the first step in a lot of LLMs – but BiotXplorer is not using this currently (using regex searching)

- Are using BERT – and tokenizing at that level
- Tokens:
 - Might be able to access these through the paywall >> but could also yield full sentences that would still be protected
- Would ideally be looking at the *network* aspect of how that is changing based on tweaks in the model – could measure differences
 - Adding visualization as well
- LLM approach – has advantage that the interaction could be 3 sentences later – not all info just in one sentence

Comparing across methods

- Possible to compare / contrast the strengths and weaknesses of these different methods? Should be yes
 - >> How combine, ideally
- Could use the different methods to check / verify each other
 - One approach is more traceable than another – scalability differs too – also with human-in-the-loop as well
- <https://github.com/globalbioticinteractions/globalbioticinteractions/issues/1021>
- https://en.wikipedia.org/wiki/Blackboard_system
- At some point we will be able to use fully read / tagged papers for use in calculating both sensitivity and precision
 - E.g., the rodent-virus interaction data from ASU / Beckett Sterners group
-

Broader

- you may be interested to <https://github.com/big-bee-network/bif> - we are compiling a giant list of interaction terms and their RO mapping from ~10-20 millions interaction candidate claims mined from GBIF indexed datasets. - see <https://github.com/Big-Bee-Network/bif/issues/6>

NASBR

- We can share our NASBR presentations for team edits using this shared file; <https://drive.google.com/drive/folders/1nwD1eafKZFdc82I9H4kVnzVPLNxVBV7B>
 - Can't wait to see everyone in real life
- Planning to have the Oct 23rd meeting in person

2 October 2024

Present: Aja, Nancy, Jelle, Nate, Cullen, Donat, Nathan Layman, Jorrit, Kendrai, Anna, Ariadna, Cecilia

Cullen: BatLit access to the DB

- need some onboarding process for getting into the Zenodo "batlit" group to get access to the restricted publications.
- Parallel with BatBase – everyone is approved for someone that can search the DB
- Nancy - we need a document for folks to agree with before they become a member. To make sure to do legal
- Terms of Agreement / Terms of Use
 - Example from the World Spider Catalog we could build from
 - Not easy to change the terms, although perhaps possible
- We made bridge between Zotero and Zenodo – and opened up large amounts of literature
 - We want to share this sooner rather than later
- 20k records in BatLit
 - >> this version 0.6 is the one that will be discussed from the

Aja working on IUCN bat RedList updates

- Wanting to host a taxonomy call to discuss in more detail

Nice search box, Jorrit! <https://batlit.org/>

- four lines of html code, thanks to Zenodo -
 - from
<https://github.com/bat-literature/bat-literature.github.io/blob/f43d595e61eb95f03f56bd598f2b60ed53791956/README.md?plain=1#L1>

```
<form class="searchbox" action="https://zenodo.org/communities/batlit">
  <input type="search" id="name" name="q" placeholder="Search BatLit..." />
  <button type="submit">Search</button>
</form>
```

NABR presentation guide - <https://nasbr.org/presentations24>

tentative schedule

https://docs.google.com/spreadsheets/d/1h-pFB6et0ryGUaKfkZDrpC_CCNpxYartpd8fl5TVxuc/edit?gid=0#gid=0

Paperpile (Ariadna), Zotero (Aja), Endnote (Cullen), Custom software (Jelle)



Nomen nescio. 2024. <https://api.openai.com/v1/>

accessed at 2024-09-27.

25 September 2024

Present: Aja, Nancy, Jelle, DeeAnn, Nate, Kendra, Cullen, Donat, Jorrit, Nate Layman
(alter-nate, EHA), Hernani, Anna

Introductions

- Nate Layman has done work on transferable / transmissible vaccines in wild animals – degree in evolutionary biology – worked with Scott Nuismer at UIdaho as a postdoc – <https://imci.uidaho.edu/all-of-our-people/name/nathan-layman/>
- Kendra now a research associate at AMNH

ML-based extraction of data from the bat co-roosting dataset

- Already started working some with Aja on this
- World Organization for Animal Health – the ML work is derived from that in a prior EHA contract >> goal was to estimate the number of cases / outbreaks that they missed
 - At the time ChatGPT / LLMs were developing
 - Grabbed 100k abstracts, fed to GPT, asked it to read the abstracts and pull out all the outbreaks of given viruses in given years
 - Similar in several ways the approach we are using
- Slides on ecological data modeling
 - Aiming for tabular data as the output from data extraction
 - Function calling schema
 - How you prompt ChatGPT – using a javascript function for generating prompts – ask for return to be tabular / structured

- Respond with co-roosting interactions + page number + snippet of text
- Context allows for still having the human check it
- Structured response engineering
 - This is critical for downstream usage
- How to evaluate performance?
 - Ranking metric – accuracy, specificity, precision, recall
 - Computer vision can now extract data from figures
- Work in progress for sure, lots of errors still
 - Now working on fine tuning – domain specific knowledge
 - Want to be able to re-use the pipeline – bat-plant interactions
- Jorrit: What data was used to train ChatGPT version you used?
 - Likely no
 - Entire internet used
 - ChatGPT is a black box – doesn't cite sources
 - Answers today could be different than answers tomorrow
 - Approaches to this
 - Ask multiple times – see if gives same result
 - Include humans in the workflow >> reducing work, not eliminating
 - Is ChatGPT repeatable and suitable for academic research?
 - Don't know who created, what data went in, paid service,
 - (Jorrit notes) Llama

<https://research.facebook.com/publications/llama-open-and-efficient-foundation-language-models/> opens the model, but **not** the data. ChatGPT offers to interact with the model in a pay-to-play "cloud" model, but not the model itself. Both, do not clearly define what data corpus (Jorrit: likely the provenance of the data is undefined because the LLaMA/ChatGPT folks don't know how or have no interest in citing their data sources, likely both) is used to train the models.

 - We could endeavor to build our own LLM / fine-tuned model for ecological data extraction
 - But not a trivial exercise
 - Need to scope how much work would indeed be involved
- Could bring in Patrick Ruch's group to discuss with Nate Layman – compare approaches
 - the accuracy metrics are a central part of this
 - <https://www.sib.swiss/patrick-ruch-group>
 - <https://sibils.text-analytics.ch/search/>
 - <https://biotxplorer.sibils.org/>
 - <https://biotxplorer.sibils.org/?species1=Rhinolophus&interaction=co-roosts%20with&size=1> here is an example
- Co-roosting of bat species >> couldn't a rule-based NLP model be used? (Anna)
 - Context could be missed – LLM will provide more nuance to the output (presumably), flexibility of LLM is the power – but NLP
 - Compare to BiotXplorer approach

- Live coding of pipeline in R Studio

18 September 2024

Present: Aja, Nancy, Jelle, DeeAnn, Nate, Kendra, Cullen, Quentin

NASBR plans for 23-26 Oct 20204

- Nancy already making her talk
 - Introducing the full symposium
 - We can send Nancy content, aim to set the context – to help set up the talks
- 5 weeks until the first day!
-

EcoHealth Alliance

- Kendra no longer works there, lots of people are leaving
- Wow, because have no money, no scientists
 - NIH is barring them from receiving funding
- World OneHealth Alliance in South Africa
 - Some EHA folks are still going
- Organization is folding
- Many folks looking for new jobs
 - Incl Kendra, directly contacting people

Aja – batLit update – de-duplicating 30k records that were found

- No more empty PDFs
- Near ready, will send to Jorrit >> previously had the Trash file indexed – now cleaned up
 - V0.5 of the Bat Lit Corpus
- Aim is for this all to “go live” while at NASBR
 - Need to think more about the governance structure – what are folks agreeing to with respect to becoming members / users of the repository
- Search features
 - Could just keep with Zenodo for now
 - Then advocate for a step-wise approach of using Google Scholar – then to BatLit to find the full text
- Versioning
 - Many reasons to want to be able to cite
 - E.g., “cited v3.2 of Bat Lit for doing my text mining routine, etc”
 - We want this to live on beyond this group of curators though too
 - Zenodo – each PDF has its own page, but then they come together as a corpus (Zorpus?)
 - <https://batlit.org/>

- Came to a consensus that we think as a group that versioning is important, including content-based versioning

IUCN is the next issue coming

- Aja will report on next week

EHA generated a pipeline for doing LLMs with the co-roosting dataset

- Nate Lehman at EHA
 - Still works there for now
 - <https://github.com/n8layman/ehallm>
 - Already looking at the data, can pull out different aspects of the dataset
 - Using OpenAlex also – for finding papers
- Aiming to understand what can be automated more easily >> for NASBR talk

Symposium planning

- Abstracts:
 - <https://docs.google.com/document/d/1TFPWLNxf5KrczCNGQzybiru93AwRtbx7l6NNWVgRcbY/edit>
- Having a meeting at Weds Oct 23

11 September 2024

Present: Aja, Nancy, Donat, Jelle,

How can we get most efficiently more colleagues be included

Strategies on how to deduplicate new paper collections

Jelle is looking into how to deduplicate the existing zotero file

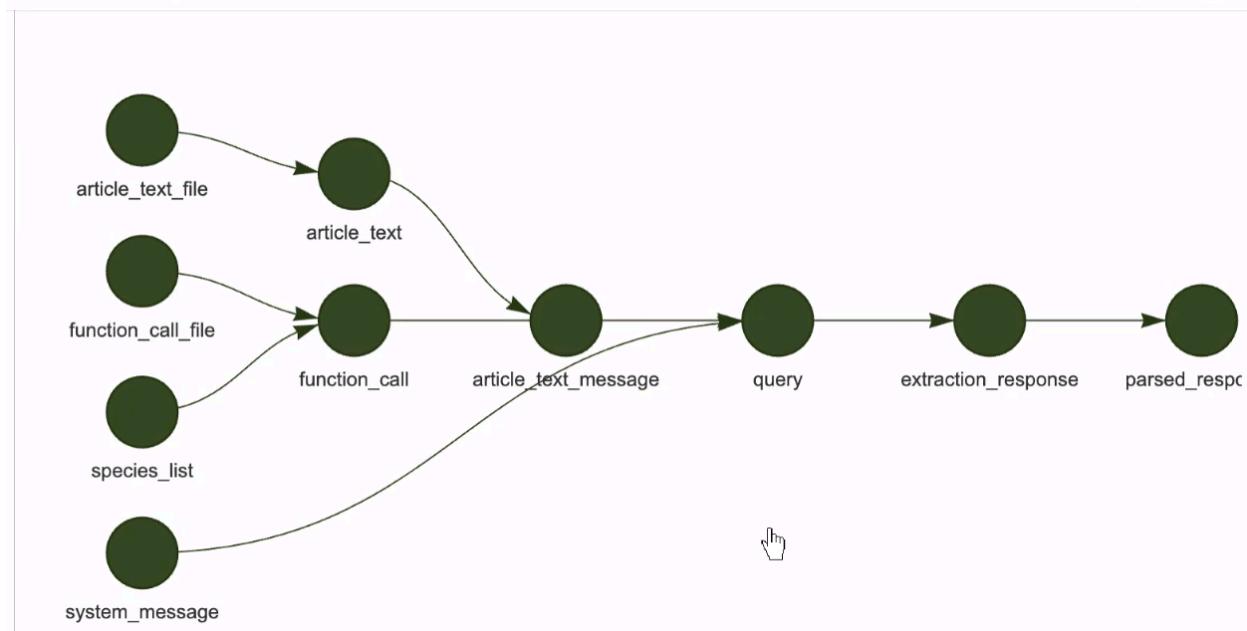
We need to include a column in Zotero that includes the Zenodo ID of the uploaded files.

Each Zotero that has been uploaded to Zenodo should be kept when deduplicating

Question arises what to do when an old bad PDF is replaced by a better one?

Value of openAlex for next generation of bat literature, and how much back is openAlex helpful to find old bat publications

Co-roosting: new tool by Nate Leman (?) <https://github.com/ecohealthalliance/ehallm> announced by Aja, each query running at 8 cts.



This could complement the development in biodiversityPMC

<https://biotxplorer.text-analytics.ch/?species1=Pholidota&size=1>

Add here is the list of terms used in biotxplorer for biotic interactions. Additional terms, or synonyms can be added.

4 September 2024

Present: Aja, Nancy, Jorrit, Kendra, Cullen, Donat, Jelle

Agenda:

Soft launch of the BatLit Corpus

Issues arising

BatLit -

Jorrit asks what to do with publications without a publication date
<https://github.com/bat-literature/bat-literature.github.io/issues/27>.

Soft launch result can be: <https://zenodo.org/communities/batlit>.

Open Question:

Donat: Should we keep versions of BatLit in both Zenodo and Zotero?

Cullen: For starters, we'll consider the Zenodo records be "readonly" and the Zotero records are going to be read-write.

Simplified Workflow

Aja curates records in Zotero -[:versioned_and_translated_into]* -> Zenodo BatLit community

<https://zenodo.org/records/13538164> - example of restricted batlit publication

<https://zenodo.org/records/13538144> - example of open batlit publication

"Amazon" BatLit package consists of three parts:

1. Packaging slip (describing the contents and their origin, e.g., I saw pdf x at location y on date z)
2. Zotero Metadata (structure data as retrieved through the Zotero Web API, for example, see <https://linker.bio/cut:hash://md5/03118e04a34ea876a0b48c511cd24a8d!/b300600-303154>)
3. PDFs download from the restricted Zotero Group using proper credentials.

Only 1. and 2. are open access, and are published on

<https://github.com/bat-literature/bat-literature.github.io> .

3. are stored offline and in restricted member access only locations. So, they are *not* open available.

On releasing a future BatLit version, the updated (meta)data would be added to the existing BatLit Archive and a new version "hash" would be created to help update the Zenodo records in a systematic and controlled manner.

Zotero -> BatLit "Amazon" Package -[:translated] -> Zenodo speak -[:update existing or create new] -> Zenodo

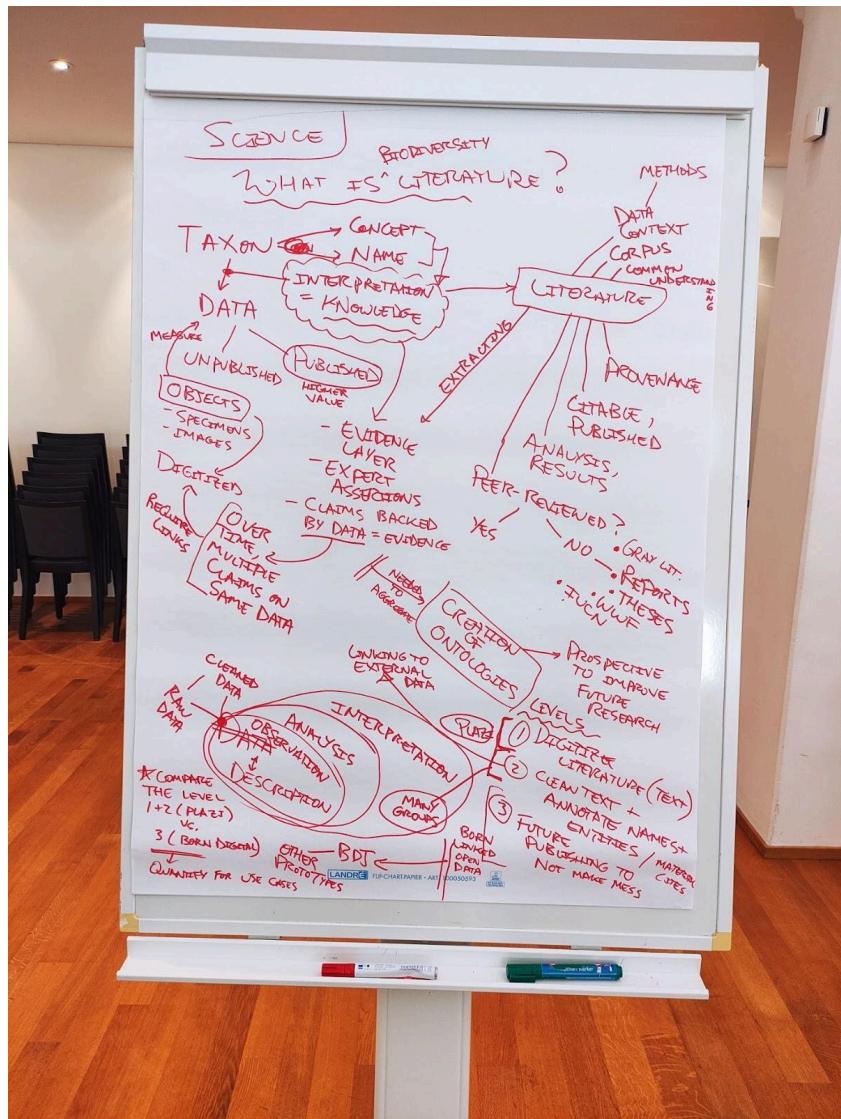
Next Steps: Aja to divvy up the work to correct the publication dates.

28 Aug 2024

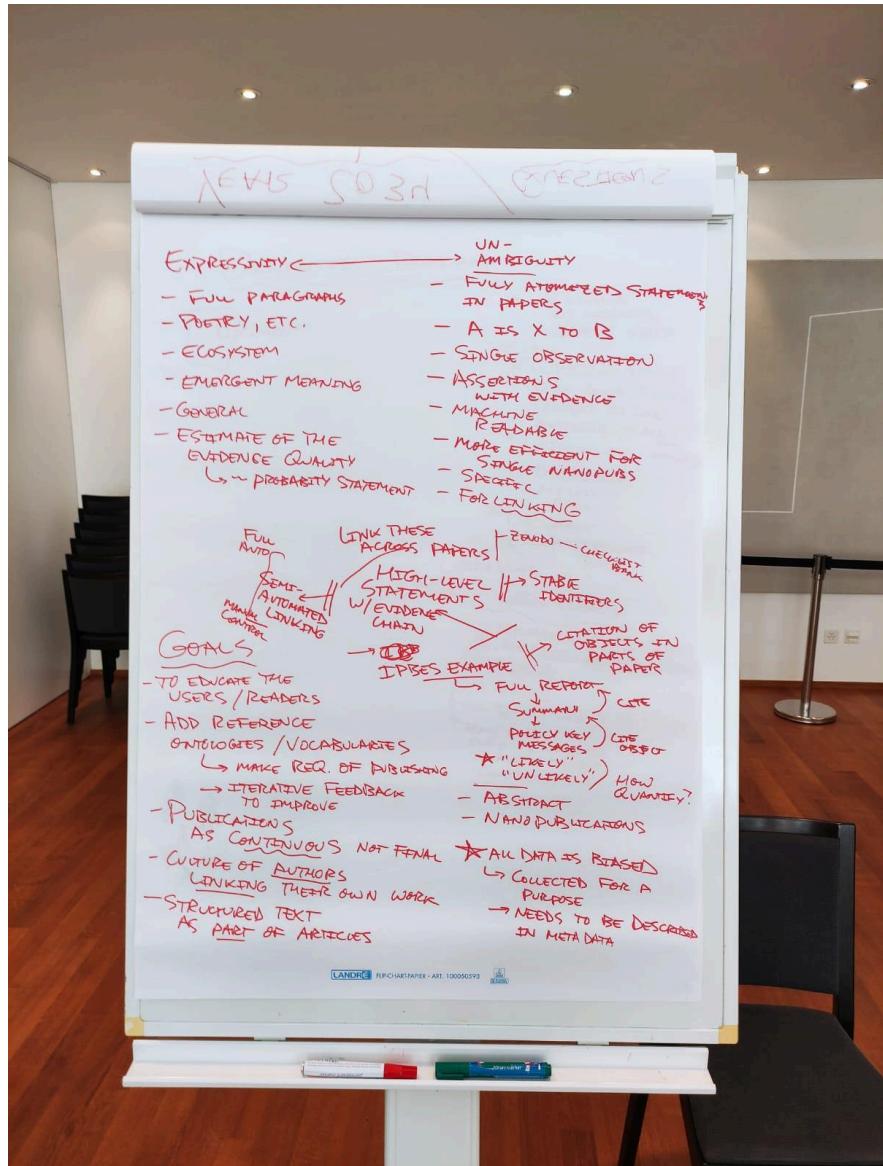
Present: Aja, Nancy, Ariadna, Nate, Jorrit, Kendra, Cullen, Donat

- Year of the Bat 2026
 - Bane or Blessing? Reviewing Cultural Values of Bats across the Asia-Pacific Region <https://journals.sagepub.com/doi/full/10.2993/0278-0771-41.1.18>
- Agenda

- ? BatLit v0.5 release
- Updates on the forthcoming 'Distentis Roadmap' that stem from the meeting Donat organized last week (the BatLit project was viewed as an excellent use case)
- BatLit data reviews, governance
- Issues arising
- shared by Kendra following Aja's announcement around "2026: Year of the Bat" - cultural significance of Low, M.-R., Hoong, W. Z., Shen, Z., Murugavel, B., Mariner, N., Paguntalan, L. M., Tanalgo, K., Aung, M. M., Sheherazade, ., Bansal, L. A., Sritongchuay, T., Preble, J. H., & Aziz, S. A. (2021). Bane or Blessing? Reviewing Cultural Values of Bats across the Asia-Pacific Region. *Journal of Ethnobiology*, 41(1), 18-34.
<https://doi.org/10.2993/0278-0771-41.1.18>
- Summary from the Disentis meeting
 - 10-year plan for the forthcoming path of liberating global biodiversity knowledge from literature
 - Day 1 - Group 1 (Scientific) Brainstorming Session



- Day 2



- Learned intuition from the use case of actually doing this work
 - Plazi process of extracting from literature
 - Initial work starting in 2002
 - Old papers – harder because they don't template well – start from the bibliographies
 - Zenodo repos
 - Storage and metadata
 - Biodiversity PMC
 - For use in additional annotation
 - Zotero use for gathering bibliographic metadata
 - GBIF folks integrating these data
 - Material citations
 - Data Futures

- Peter Cornwall
- “Green Field” tool for extracting citation metadata from publications
 - Just a prototype now – last year of Arcadia funding
- Main question:
 - Where do we invest?
 - Opening up as many new journals as possible
 - Use cases like BatLit
 - Working on old literature and the metadata of this
 - Likely depends on where / how the money comes from
 - Use cases can drive some of this funding and innovation
 - Ultimately want to open up 500 million pages
 - No one best way to do this

BatLit v0.5

- Data publication near ready to go
- Nate and Ariadna reviewed the administrative / machine-readable part of the BatLit
 - The digital concrete == the software architecture part of it
- <https://github.com/bat-literature/bat-literature.github.io/issues/23>
- Sandbox:

<https://sandbox.zenodo.org/communities/batlit-review-md5-26f7ce5dd404e33c6570edd4ba250d20/records?q=&l=list&p=1&s=10>

 - 17 items included as a pilot
 - Linker.bio output is coming from server in Germany – and/or Zenodo –
 - But there is some core dependency on Jorrit knowing where these things are – talk about how to keep these communities alive and sustainable
 - Bat Lit community, others
- Through Zenodo
 - We cannot make everything fully downloadable – but can cite things that are closed access in a way that maintains the privacy but still confirm that the dataset is complete – hiding portions of it
- Some concern about the keeping of extra copies outside of Zenodo
 - Does the server in Germany count as an extra copy?

21 Aug 2024

Present: Aja, Hernani, DeeAnn, Nancy, Ariadna, Nate, Jorrit, Anna

- Kendra and Aja will review Nate Layman’s recommendation for a workflow for our co-roosting dataset autom
- BatLit v0.5 was released <https://batlit.org> - request for review available at <https://github.com/bat-literature/bat-literature.github.io/issues/23>

- IUCN name matching for bats
 - Relative to the MDD, and Batnames
 - <https://docs.google.com/spreadsheets/d/164fcpl7i36MC4ZUEIrzd2Zj06fiQndlBO9k-wlSX4aE/edit?gid=0#gid=0>
 - With the purpose of updating the IUCN taxonomy for bats
 - (Jorrit) Which version of BTA did you use? How'd you like IUCN to cite this BTA version?
 - (jorrit) Still not sure which version of BTA is used for this IUCN comparison.
 - was it "Sherman, A., Geiselman, C., Poelen, J., Simmons, N., Reeder, D., Upham, N., Phelps, K., & Agosti, D. (2024). A Standardized Review of Bat Names Across Multiple Taxonomic Authorities [Data set]. Zenodo. (<https://doi.org/10.5281/zenodo.11193643>)" ?
 - Which MDD version? Aiming for the v1.13 MDD version, published Jul 13, 2024
 - IUCN has their own system
 - What we are sending – just the google sheet
 - Informally we'll send
 - But then that IUCN-to-MDD alignment would not get published
 - We COULD publish this on Zenodo or similar – why not do so?
 - What is IUCN's process?
 - Different sub authorities are in charge of assessments
 - But they cannot assess anything until the backbone taxonomy in the IUCN is changed
 - But that is a major bottleneck – someone has to notice the change, fill out a spreadsheet, then someone accepts, and then it is possible for a reassessment to occur
 - Solution - make more batch-level changes to the IUCN backbone taxonomy
 - What we need is a formal data publication that can be iteratively sent to the IUCN
 - >> and to cite the metadata resources involved in generating this
 - Hash-based versioning
 -

14 Aug 2024

Present: Jelle, Aja, Hernani, Kendra, Cullen, Jorrit, Donat, DeeAnn, Nancy, Nate

- Kendra and Aja will review Nate Layman's recommendation for a workflow for our co-roosting dataset automation (<https://github.com/ecohealthalliance/ehallm>)
- Jorrit's recommendation to start a new journal "BioKnowDar":
<https://pad.carpentries.org/bioknowdar>

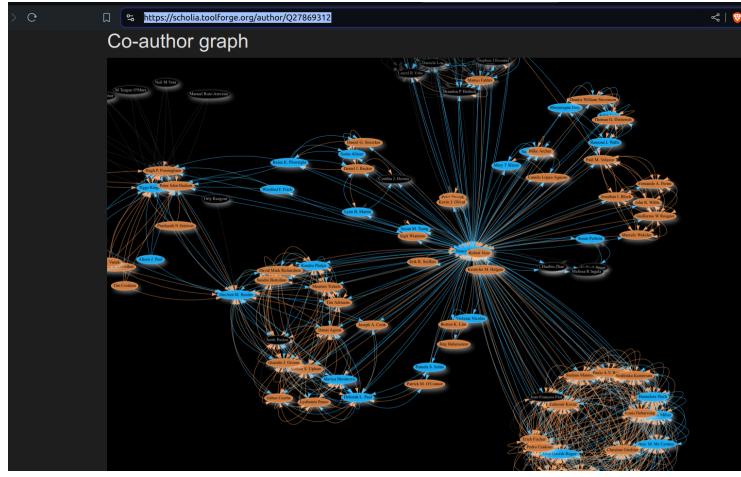
- Challenges of siloed projects
- Datasets are always coming online – how to keep aware of
- Digital scholarship initiatives
 - How to deal with tenure review boards?
 - Rigor of digital publications need to be acknowledged
 - Tension about academic credit
 - BUT tools for data review will also make it easier to write ‘traditional’ papers
 - Bot can publish naive reviews, but then humans can come along and write
- Peer review is big within this
 - Is what you are producing being checked by a peer review?
 - Dynamic digital publication – with a peer review system
 - But peer review is also a big time and
- Academic credit models
 - Versioning
 - Elaborate schemes of contribution half-life
- Hernani has an in-progress MS along similar lines:
 - https://docs.google.com/document/d/1FSsG2GQoclEsBX_f4RTU8_YI3x1GqXauG7cWTWrUFM/edit
 - “This is what we have been drafting so far...I was going to send an invitation to everyone when things were more structured”

BatLit – governance doc:

https://docs.google.com/document/d/1kuDALvMlfQrlh9mhHnzOdsEm88_Xjzvl/edit

- We are doing a 5 min brain dump regarding some of the highest value use cases for the project
- Accessibility
 - PDFs are difficult to access in the Global South, especially older literature, even smaller schools in the US
 - Instead of all doing this independently, we pool resources and create a “friends” community around the world
- Finding complete metadata for known papers
 - More accurate, faster, better tools, especially for old literature
 - Abbreviated journal names
 - Authors names with one or two initial
 - Author indexes as well – e.g., Nancy Simmons in Hesperomys: <https://hesperomys.com/h/3643>
 - Helpful for determining if these are the same individual

- Scholia also does this: <https://scholia.toolforge.org/author/Q27869312>



co-author graph according to wikidata visualized via Scholia, note the cluster of co-authors associated with the GBatNet/Covid19 group).

- ORCID is trying to do this – so is WikiData
- Broader goal of BatLit – trying to find + connect *all* the publications about bats, ever
 - Because of that, we should not allow for mission creep there
 - Cannot solve all the world's problems there
- Other databases can cite the BatLit papers
 - “Provide a durable literature resource for others to link to and find!”
 - Stop re-inventing the wheel in this context
 - Centralization – but in the sense of “let's use the same data and methods to refer to the same literature”
 - PDF repo that everyone can access and cite
 - Some risks of centralization too – if a few people control this, then those folks gain outsized power, and the power can lead to negative consequences
 - What we want is a place to start collecting these data

Need for a “terms of use” on the BatLit

- Example from the World Spider Catalog – basically saying that cannot replicate the entire data and serve elsewhere
- BUT – from a technical perspective, actually we DO want the ability to translate the entire corpus
- We want to show that the corpus can be re-used – more than just being “open” in terms of “allowed to download”
 - Ideally – have batches of 5k pdfs available for download, etc – but for own personal purposes

Broader thing – how can we ultimately unite these different literature repos?

-

7 Aug 2024

Present: Nancy, Aja, Jelle, Cullen, Hernani, Jorrit, Kendra, Nate, Donat, Anna

- Plantar fasciitis
 - Approaches – good shoes
- [BatLit Corpus Governance Doc:](#)
 - Reviewing this as a group, Cullen initially drafted
 - Steps
 - Zotero for organizing the PDFs and Metadata
 - Versioned archives whenever Aja puts up the BatSignal
 - Zenodo then reflects – the Zenodo is just an archiving service
 - If we make edits there, then they do not feed back to Zotero
 - Who can upload and from where?
 - Currently need to upload to Zotero then Zenodo reflects that
 - UI for exploring the PDFs, searching
 - Not going to be Zenodo
 - But Zenodo does have a search feature
 - Could be the batlit.org interface
 - Requires a relational database for the backend
 - If using Zenodo only as the file repository
 - Use the Zenodo web API
 - Re-use this API to create own webApp wrapper
 - Can use the web URL to be writing those queries
 - Build an index
 - Different projects
 - Curatorial workflow for creating / uploading PDFs to the BatLit
 - Web UI for users to explore the batLit Corpus
 - Hernani: Could Arctos be a potential good solution for this issue?
<https://arctosdb.org/about/>
 - Anna: paid service that operates from member fees
 - Jelle: built a custom search for Hesperomys that searches through PDF text, e.g. <http://hesperomys.com/search?q=trachops+coffini>. Definitely could be a lot better but as Jorrit says it wouldn't be too hard to build something similar for BatLit.

- BiodiversityPMC
 - <https://sibils.text-analytics.ch/search/>
 - Based on MedLine / Plazi TreatmentBank / Supplemental data
 - Accessibility?
 - Can you provide access to closed pubs
- Game: how to use the minimal effort to make these resources accessible?
 - Re-use Jelle's search tools? Uses Amazon web-services to host currently
 - But this does not require the PDFs to be on Zenodo – so they are somewhat less open/stable
 - Costs money to keep this running
 - BiodiversityPMC
 - Is this a separate upload? Or directly from Zenodo – ideally this happens automatically, but not yet in place
- What are the requirements for a particular use case?
 - Search a name – get PDFs for those
 - Well defined, finish that, then add more functionality after that
 - Next small problem
- Each of us can think of use cases >> can think of commonalities – for next time
- Parallel project with TaxoDros
 - They are taking a similar approach – Jorrit took PDFs and metadata that was custom – decoded – similar workflow as with BatLit – box it, export to Zenodo – they built a web application at <https://www.taxodros.uzh.ch/search/lit.php> for the search
- Here's another example from Donat: World Spider Catalogue
<https://wsc.nmbe.ch/>
 - Full community – requires a signup – they have a governance document – have to agree to not reproduce and sell the information
 - This is awesome, and would be a great target for the BatLit Corpus
- BatLit v0.4 release
 - Up to 20k papers now – does this include duplicates too?
 - <https://batlit.org/>
 - More attachments (29k) than references (20k) – can have multiple files per reference, sometimes supplemental materials
 - Metadata travels together with the PDF
 - Data review
 - Invites feedback, data errors can be caught

- Others

-

31 Jul 2024

Present: Jorrit, Jelle, Aja, Nate, Hernani, Cullen

IUCN updates relative to bat taxonomy

- Adding inform from <http://hesperomys.com/> – Jelle Zijlstra is joining
- Raised issue that Small Mammal Specialist group could also benefit from this type of update

1-page document for the GBatNet

- Drafted based on the images taken at the meeting
- Then collected feedback
- Start based on what they had emailed
- Focus on genome-based metadata standards
 - Will write this up
 - Eye on enabling future taxonomic changes for the published genomes

Jelle is working on importing classifications into Hesperomys

- Imported the HMW files
- Aligned with the MDD taxonomy as well
- And relative to IUCN
 - Aja had been doing that somewhat manually = Bat Taxonomic Alignment / BTA
 - She was also adding alignments to taxonomic meaning – considering the geographic locations / range maps to some extent
 - Now wanting to review all names again
 - Jelle helped with this matchup + the links to the new records
 - IUCN will have the original + current links to the info
- IUCN in general
 - They do not allow or practice versioning
 - They lack funding in a way
 - How could we better version these? Using the GloBI approach
 - Upload older versions of the taxonomy + range maps
 - Citation hygiene brings a lot of these issues to the front
 - IUCN has lots of these older data – but they are broadly unwilling to share it / feel it is outdated and not useful

BatLit work

- Aja is having to manually update metadata – due to some duplicates
 - Hard to know that have duplicates until enter the metadata
 - Some OCR is misidentifying authors
 - What can we do here
 - One solution: Tracking the metadata for each specific PDF
 - Treat the raw PDFs as one publication
 - Then next publication is v1.0 of the metadata on those PDFs
 - Versioned copies of Zotero basically
 - Helps understand how our knowledge of the metadata on these PDFs has changed through time
 - Can we help Aja compare the journal names / info relative to what Zotero is extracting from the PDFs
 - Use the DOIs → CrossRef as a check
 - Check which years the journal was published

Raw PDFs → automated annotations (e.g., OCR, DOI extraction) → manual annotations

About 500 away from 20k PDFs ingested

24 Jul 2024

Present: Jorrit, Anna, Nancy, Ariadna, Aja, Cullen, Connor, Kendra, Harnani

Review changes for GBatNet IUCN

https://docs.google.com/document/d/1au0_bSqZMn5b43qxKb0XEZYJca2PMcHF50pYLKLoPi0/edit

Project: Updating Bat Taxonomy for IUCN Red List

Participating Networks and Working Groups:

- Bat Eco-Interactions Working Group
- Global Bat Taxonomy Working Group
- IUCN SSC Bat Specialist Group

Rationale:

The IUCN Red List has not kept pace with current bat taxonomy, meaning that many currently recognized species are not listed and therefore their conservation status has not been appropriately assessed. Using the Bat Taxonomic Alignment Tool, we propose to provide an updated species list and associated data – and build a pipeline for future updates – to facilitate the inclusion of new and revised taxa in the IUCN assessment process.

Alignment with GBatNet Strategic Directions:

- 3.3 Help maintain accurate species assessments

- 1.6 Assist networks in prioritizing strategic actions

Approach:

The taxonomic boundaries of species and higher-level taxa frequently change as knowledge of population structure and evolutionary relationships improves, and older hypotheses are reframed. We have collected and aligned treatments across primary taxonomic authorities for Chiroptera to create a single dynamic and versioned name translation tool. The Bat Taxonomic Alignment (BTA) tool integrates data from three of the most utilized treatments: The Handbook of the Mammals of the World Volume 9 Bats (HMW), Bat Species of the World: A Taxonomic and Geographic Database (Batnames database), and Mammal Diversity Database (MDD). The BTA reveals that the IUCN Red List (version 2022) lacked over 180 bat species currently accepted by MDD and Batnames, along with over 120 superseded epithets. The BTA allows easy identification of missing species and will be used to propose updates to the Red List to facilitate more accurate species status assessments. We will align the IUCN Red List species list using the BTA to identify missing or outdated treatments and will determine the nature of the discrepancy (split, merge, new species described, etc.), the alternative names associated with the current binomial, and any journal publications supporting the change. This information will be provided to the IUCN in the format required by them for taxonomic updates to their lists. Once IUCN has ingested the new names, species can be prioritized for updated assessments. This project will benefit all GBatNet networks and working groups by ensuring the globally recognized list of species of conservation concern is accurate and useful.

Key targets and timeline:

- Obtain most recent list of bat species recognized by IUCN from Dave Waldien (done)
- Align Red List species with BTA and identify gaps (done)
- Obtain IUCN forms to propose species updates (waiting for response from Dave Waldien)
- Prepare IUCN forms for new species and submit to IUCN (Fall 2024)
- Prioritize species assessment updates through the Bat Specialist Group (Spring 2025)

Key Contacts: Aja Sherman aia@batbase.org; Nancy Simmons simmons@amnh.org; Ricardo Moratelli rimoratelli@gmail.com; Connor Burgin connorburgin@gmail.com; Julie Shapiro julie.teresa.shapiro@gmail.com; Nate Upham nathan.upham@asu.edu; Dave Waldien dwaldien@gmail.com

Nancy: There are no regular updates of the IUCN taxonomy. No mechanism is place across the IUCN taxonomic or bats.

Conner: In IUCN non-bat groups there have been "well done" (e.g., primate group).

other topics beyond the GBatNet project one pager -

1. How to keep track of all the activities / outcomes of this group (<https://globalbioticinteractions.org/gbatnet>)?

GBatNet has a quarterly activity reporting form - required by National Science Foundation (NSF) Nancy to share the form.

2. Cullen - batlit corpus update - we need some kind of governance document to share how this will live on in perpetuity.

3. Aja - > 300 hours uploading pdfs etc. lots of duplicates. Many of them were abstracts, deduplicated ~1/3 replaced. ~30s to process a pdf for automated process, with manual updates processing takes longer (2 minutes). Estimated release BatLit v0.4 <https://batlit.org> .

Aja priority is to get the basic metadata out (author names, date, publication, title).

Anna - Is the literature corpus addition protocol written down somewhere? Can I help review?

Suggest to setup a curation decision tree.

Connor - sent intro email to Aja with cc to Jelle Zijlstra. Jelle keeps a collection of pdfs related to taxonomies.

Aja goes over the batlit pdf addition protocol. Google drive -> download as zip -> import into Zotero

Jorrit: did you consider rclone ? This tool allows for access / syncing data in various "cloud" storage solutions, including google drive. I've been using it to sync files in dropbox .

Anna - suggests to create a journal reference lists .

Jorrit - suggests to have a dedicated review meeting when BatLit v0.4 has been released.

17 Jul 2024

Present: Hernani, Jorrit, Anna, Nancy, Donat, Ariadna, Cecilia

- Jorrit working with Brazilians on World Fair project: <https://www.globalbioticinteractions.org/worldfair/>

Recap of last week

- EHA funding changes
- BatNet domain name <https://batlit.org>
- BTA use cases <https://jhpoelen.nl/bat-taxonomic-alignment/>
 - People don't know where to look
 - Frequently Nancy can point to sources in 5 min
 - Will be putting together BTA tutorial videos
 - How to tract taxonomic changes? Eg phylogeny
 - Quentin has worked on this
 - Synospecies
 - Other taxa communities
- Tools can be hosted by GBAT net as long-term commitment for storage and dissemination.
 - Eurobat traits Shiny: <https://jasja.shinyapps.io/ClimBats/>
- publications vs. publish website - articles appear to be more static, less likely to include video or dynamic elements / links.
- Nancy
- Hernani - how to spot overlap between trait databases?
 - Jorrit - perhaps register your dataset with <https://opentraits.org> - see also <https://opentraits.org/best-practices.html> - "Ten (mostly) simple rules to

- future-proof trait data in ecological and evolutionary sciences" in Methods in Ecology and Evolution. doi:[10.1111/2041-210X.14033](https://doi.org/10.1111/2041-210X.14033)
- Jorrit - did you consider integrating with Encyclopedia of Life TraitBank ?
<https://eol.org/traitbank>
 -
 - ClimaBats data browser <https://jasja.shinyapps.io/ClimBats/>
 - Anna - suggest a way to list which traits have been extracted for specific publications. So, that others don't have to redo the work for those publications.
 - Anna - mentions - Stephens PR, Pappalardo P, Huang S, Byers JE, Farrell MJ, Gehman A, Ghai RR, Haas SE, Han B, Park AW, Schmidt JP, Altizer S, Ezenwa VO, Nunn CL. Global Mammal Parasite Database version 2.0. *Ecology*. 2017 May;98(5):1476. doi: [10.1002/ecy.1799](https://doi.org/10.1002/ecy.1799). PMID: 28273333.
 - <https://www.nature.com/articles/s41597-023-02157-4>
 - Linking traits to the sources, where is the corpus stored?
 - ClimaBats pdfs are likely with the first author:
 - jeremy.froidevaux(at).gmail.com
 - ClimaBats Viewer vs. the Publication
 - Viewer should have citation, methods
 - references are currently in the download (Traits_references.csv), but not through the viewer
 - https://jasja.shinyapps.io/ClimBats/_w_c316e7bb/EuroBaTrait_v1.0.zip
 - How do you retain the detail? Eg. individual with a morphometric trait
 - Aja's method of co-roosting
 - is to retain page number and text/sentence of trait export as a data cell in the database
 - Running list of edits
 - How to grow project
 - Short term goals, what is achievable
 - Work packages, can be cited - (e.g., WorldFAIR 10 had 3 pubs 10.1/10.2/10.3 see <https://globalbioticinteractions.org/worldfair>
<https://doi.org/10.5281/zenodo.10719265>)

10 Jul 2024

Present: Hernani, Kendra, Aja, Donat, Jorrit, Cecilia, Anna, Quentin

- EHA's debarment from federal funds
 - In the media:
 - <https://www.nytimes.com/2024/05/15/health/ecohealth-alliance-peter-daszak-nih-grants.html>
 - <https://www.nytimes.com/2024/05/01/science/covid-lab-leak-ecohealth-nih.html>
 - Unclear timeline (officially 3-5 years; possibly longer or investigator specific)
 - Kendra transitioning after August to somewhere new
- Aja got 20k Bat literature reference, working towards v0.4 of <https://batlit.org> .

- Aja mentioned that the IUCN project is a little delayed - bat names review of IUCN RedList to point out missing names, or name updates needed. Cullen and Aja are figuring out how to help IUCN to update their names. Reusing method
- Kendra - re: IUCN batnames list - how did Aja select the names from IUCN. Aja used the name agreement index <https://jhpoelen.nl/bat-taxonomic-alignment/> .

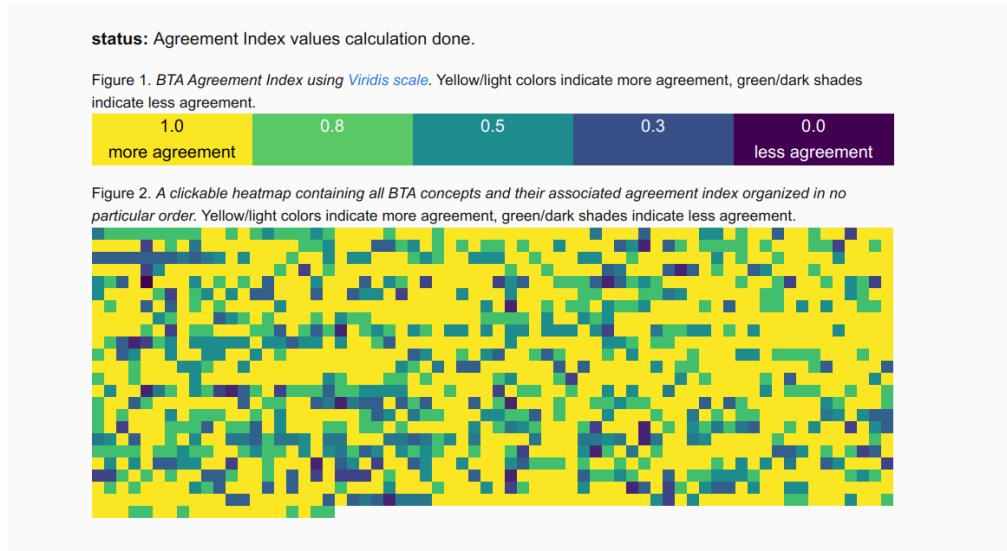


Table 2. BTA@87c862c1 agreement matrix. Each cell indicates the number of concept **disagreements** across catalog pairs. Total number of concepts in BTA@87c862c1 is 1538. Yellow/light colors indicate more agreement, green/dark shades indicate less agreement.

	BatNames	BatNames_2023	HMW	IUCN	MDD	MDD_2023	MSW3
BatNames	-	81	209	312	105	93	580
BatNames_2023	-	-	217	325	98	41	583
HMW	-	-	-	268	195	225	520
IUCN	-	-	-	-	330	324	494
MDD	-	-	-	-	-	89	595
MDD_2023	-	-	-	-	-	-	582
MSW3	-	-	-	-	-	-	-

- Kendra asked a question about *Pteropus giganteus* - *Pteropus medius/vampyrus*, from [MDD](#) “the name medius is used instead of giganteus because giganteus is now considered a synonym of P. vampyrum”. Reference to record - <https://linker.bio/line:hash://sha256/87c862c149f6a33c7d8c344568a9a7f80310cbeed324395aa94fff2d5248a3e6!/L2.L1182.tsv> csv version: <https://linker.bio/line:hash://sha256/029b12ab3a7d5c2cef01df411ea3ecb4a1fa5c199acc3bee4544a6ee59691c7!/L1.L1182.csv>

See reference for *P. giganteus* change to *P. medius*: Mlíkovský, J. (2012). Correct name for the Indian flying fox (Pteropodidae). Vespertilio, 16, 203-4.

https://www.ceson.org/vespertilio/16/203_204_Mlikovsky.pdf

- Donat asks about input for the selection of the BHL publications to be imported into BLR / BiodiversityPMC
<https://docs.google.com/spreadsheets/d/1fxi-ndcZJymKiAJx0Xillx29UuIO5Ndzo0vo9uEZxg/edit?gid=0#gid=0>
- Aja/Quentin - re: co-roosting database - no progress yet
- Bat Literature Project:
<https://sandbox.zenodo.org/communities/batlit/records?q=&l=list&p=1&s=10&sort=publication-desc>
 -
 - What is the Purpose?
 - Anna: a corpus should help provide a way to link a copy of an article that has a DOI. Document the workflow, which paper have been contributed in what way
 - How do we bridge in trait/content-based data to the corpus
 - Aja/Cullen:
 - Use keywords to build subcorpor. Identifies: What is a bat publication: taxonomy, just mentioning of a .
 - Donat:
 - Get immediate access to papers from any bibliography we are publishing in the future
 - Aja:
 - Provide access to students
 - Currently lacks definition/scope
 - Create a policy of what is a bat publications
 - Nancy's current goal is Koopman paper digitization (taxonomy)
 - Drosophila
 - What is the scope of the community: publications, supplements, data sets, PhD thesis, grey literature, reports, etc.?
 -
 - Currently lacks procedure:
 - Who can contribute?
 - What is the metadata? How complete do you need to be? Does curator need to do a check?
 - How do you attach content-derived metadata? How complete do you need to be?
 - Is this a hosting platform or will it point to already accessible platforms (e.g., BHL)?
 - Donat: hosting
 - Legal issues?
 - Pre-2000 open, the rest according to the existing open / closed access
 -

3 Jul 2024

Present: Hernani, Nancy, Nate, Cullen, Cecilia, Donat, Kendra, Aja

NASBR symposium abstracts

- Deadline now to Jul 10th
- We currently have 8 symposium abstracts
- Impact of taxonomic splits on bat-viral data
- <https://docs.google.com/document/d/1TFPWLNxf5KrczCNGQzybiru93AwRtbx7I6NNWVgRcbY/edit> – Current doc
- BatLit abstract
 - Could be edited more
 - Have the website: <https://batlit.org/>
 - Now 5.5k of the 20k+ PDFs from Van Cakenburge
 - All now uploaded via the cloud
 - AfricaBat Report, e.g., 2016:
https://www.researchgate.net/publication/311888337_African_Chiroptera_Report_2016
 - Lots of data here, difficult to access
 - I think this is the latest one:
https://www.researchgate.net/publication/364943183_African_Chiroptera_Report_2022_Complete_Report
 - Perhaps there is a structured database that underlies this?

Jorrit is additionally processing the BHL bat publications

- 25k articles that cover bats, but also some other taxa (Jorrit: 2024-07-05 an initial assessment suggests about 6k works that mention "Chiroptera".
<https://github.com/bat-literature/bat-literature.github.io/issues/15> in BHL . Please cite sources for the 25k articles so that I can understand the discrepancies)
- Part of the Plazi workflow – but Jorrit is helping to optimize in some ways
- Data in supplements in PubMedCentral – OCR and using the figures and tables
- Could import these via JATS etc
- OR import the articles as supplements – into Biodiversity PMC
 - Shortcut for getting data in – still searchable and parseable
 - Jorrit helping with versioning of those data
 - Donat wanting to make the data accessible for search
- Cullen and Donat are sponsoring for now – until we get further funding (R01?)
- Targets:
<https://docs.google.com/spreadsheets/d/1fxi-ndcZJymKiAJx0Xillx29UuIO5Ndzo0vo9uEZxg/edit?gid=0#gid=0>

-

Initial BatLit uploads

- <https://github.com/bat-literature/bat-literature.github.io/issues/14>
- In Zenodo Sandbox, with metadata – for checking now
- Could we also add the Supplementary Files?
 - E.g., scan documents for “supplement” and then have a part of the pipeline that brings in those files as well
 - Currently not being dealt with – but it is a big issue
- Could we pull out full text of all these articles in a batch way – if OCRed
-

Cullen linked to – <https://www.nature.com/articles/s41597-023-02472-w>

- A dataset on African bats’ functional traits
- “Here, we present the AfroBaT dataset, a compilation of trait data on 320 African bat species containing 76,914 values for 86 traits focusing on morphology, reproduction, life-history, trophic ecology, and species distributions.”

26 June 2024

Present: Hernani, Aja, Nancy, Nate, Ariadna, Anna, Cullen, Cecilia, Donat, Jorrit

Bat literature

- Data from Cakenburghe
 - 1400 articles just starting with E
- OSF library is having difficulty with the size of the data – 50 Gb, likely more than that, of PDFs
 - Jorrit has 10 TB for 20 euros a month
<https://www.hetzner.com/de/storage/storage-box/>.
- Zotero shared library
 - Jorrit is indexing from here: <https://bat-literature.github.io/#version-history>
 - https://bat-literature.github.io v0.3 just released today.
 - refs.csv has limited metadata of ~1000 paper
- Swiss-based bat literature repository
 - These were all file cabinets of non-digitized literature
 - Similar to what is at major museums

- From Manuel Ruedi
 - “Well, as I said, I don't have so many pdfs, mostly from the last 15 years (which you certainly already have). Before that, had my private collection of reprints, but not digitalized, and not so many (if you want them all, you can get them; most will be destroyed anyway upon my retirement, in 3 years from now...). But the most complete library of paper reprints of bats of the World (former private collection of Villy Aellen + many more since Pascal's rule), are in our Museum library; for this, contact our chief librarian Laura.Zbinden@ville-ge.ch. As our museum is now under repair for the next few months, she will also tell you if and when those bat reprints will be again available (currently, they are in boxes and inaccessible). Hope it helps. Manuel Ruedi”
 - “Dear Manuel— Do you happen to have Pascal Möschler's email – or is he not anymore involved in bat and the global bat library? His email seems not to work anymore. pascal.moeschler@ville-ge.ch Thanks for a hint Donat”
- Searching DOIs
 - Looking up open access or not:
 - R Package Open Alex: <https://docs.ropensci.org/openalexR/>
 - Also the RCrossRef package
 - <https://ciakovx.github.io/rcrossref.html>
 - Jorrit already did this query in OpenAlex:

<https://github.com/bat-literature/bat-literature.github.io/issues/2>

 - Not 100% reliable though
 - Donat helped follow this up:
 - <https://github.com/bat-literature/bat-literature.github.io/issues/2#issuecomment-2153428532>
 - Unpaywall is another option
 - <https://unpaywall.org/products/simple-query-tool>
 - Up to 1000 queries at a time
 - “Reference-extract sheet that is incomplete and using Zotero tools/platform to fill gaps. This open access column is not in the Zotero references.csv

<https://docs.google.com/spreadsheets/d/1y5uBKvyzDQgQUtyRHvn5f20IOA7LoMho/edit?usp=sharing&ouid=111971835458035348757&rtpof=true&sd=true> ”
- Aim is to emulate this TaxoDros:

- <https://zenodo.org/communities/taxodros/records?q=&l=list&p=1&s=10&sort=newest>
- <https://taxodros.github.io/> << view on how to query these things
- Donat: “cakenberg has 34480 bibliographic references” !

Big Bat Database work

- They have 100 databases that they are extracting from
- Taking all the verbatim names and locations from the databases
- Not as broad as PanTHERIA, etc
- More primary syntheses of individual traits from the literature
- @Hernani (he/him) @Ariadna Morales (she/her) @Cecilia Montauban et al. do you consider species interactions a trait? (e.g., parasite, fruit plant, prey etc.)
 - Yes, to some extent – not an entire focus, but does include this
 - Initially related to the genomes of the bats – for the Bat1K folks

Abstracts for the symposium

- <https://docs.google.com/document/d/1TFPWLNxf5KrczCNGQzybiru93AwRtbx7I6NNWVgRcbY/edit>
- Guidelines doc:
<https://cdn.wildapricot.com/214854/resources/docs/meetings/year/2017/2017%20NASBR%20Abstract%20Checklist.pdf?version=1495429744000&Policy=eyJTdGF0ZW1lbnQiOiBbeyJSZXNvdXJjZSI6Imh0dHBzOi8vY2RuLndpbGRhcHJpY290LmNvbS8yMTQ4NTQvcmVzb3VyY2VzL2RvY3MvbWVIdGluZ3MveWVhci8yMD E3LzlwMTclMjBOQVNcUiUyMEFic3RyYWN0JTlwQ2hIY2tsaXN0LnBkZj92ZXJzaW9uPTE0OTU0Mjk3NDQwMDAiLCJDb25kaXRpb24iOnsiRGF0ZUxl3NUaGFuljp7IkFXUzpFcG9jaFRpbWUiOjE3MTk0MTQ5NDd9LCJJcEFkZHJlc3MiOnsiQVdTOINvdXJjZUlwljoiMC4wLjAuMC8wln19fV19&Signature=YODu-S8zff4028Kom s3ZRbZlw8WfSVUSTa0SNLJJaYL4O-Hk9k8a~Qh2yZi8GhsQTYcZ4-VR7fgRfsMP63bcaF3WLGIDy7zWfCrmQH-6MQbPb7RpP1kJQC-foxCrjyZj4MGlbdrcqvpuexA23qrkksTWnPUR~Vz87jSfeQNCMoitX9F-k4jSub3kPvNCrwgYst3UB1Y9OE6cgdFc5CM9~IdccoGYKu1MNeH-6OP0sx~RahYD3RiTGXxHQeD0ZM853X8Oz9he1e01DoFAryQL4CzcPVyKi1VpuhfJTsYgrgrwr5m-6Fo7~QPpf2V9gLMKMRLg3DHvuDBFVreSNx~1w &Key-Pair-Id=K27MGQSHTHAGGE>

19 June 2024

Present: Quentin, Hernani, Aja, Nancy, Nate, Ariadna, Anna, Cullen, Hernani, Cecilia, Kendra, Donat

NASBR Symposium abstracts

- Cullen to submit collation in two weeks (July 3 – but abstracts to her by July 1)
- Most pressing

Jorrit is open to going to NASBR

- But we could also re-invest those funds into more data extraction
- Work has been to get all those publications up on Zenodo
 - Then following the Drosophila workflow
- We could have a pre-meeting workday
- Oct 24-26 are the main meeting days (<https://www.nasbr.org/annual-meetings>)
 - Oct 23 we could have a meeting everyone together
- Ask the group about to what extent want to interact in person versus keep that into the repository
 - If he goes – maximize his impact – advertise the initiative
 - Could make a poster for the purposes
 - Recruit PDFs from meeting attendees

Bat literature digitization work

- Uploading to BiodiversityPMC
- Not only taxonomic papers
 - Ecological papers
 - Shared archive of Bat Cons International
- OSF folder where collecting PDFs: <https://osf.io/28cvq/files/googledrive>
 - Want to start uploading documents
 - Do we have a quality control pre-check? Could be above the technical capabilities of the folks with the PDF libraries
- Journals:
 - From web of science journal list:
 - African Bat Conservation News
<https://www.africanbats-npc.org/abcn>
 - AFRICAN CHIROPTERA REPORT
<https://www.africanbats-npc.org/acr>
 - Bat Research News
-
- Donat is at Davos meeting - <https://www.worldbiodiversityforum.org/en/wbf-2024>
- Market out there for environmental assessments

- Map of Life folks are there
 - But not currently sending data back to GBIF
- GBIF, CoL folks are at the meeting – Tim Robertson
- Plazi aims to get new names into CoL
- Question
 - How to resolve all the names
 - Need to cite taxonomic treatments
- COI as possible ‘type’ for names / taxa

Bouchout Declaration meeting

- Hackathon-like meeting
- Integrating publications into CoL
- <https://bouchoutdeclaration2024.org/>
- Tim Robertson, Marcus Deiring, Olaf Banki
- Goals
 - Get more names in CoL
 - Make names databases more FAIR
 - Trying to get nomenclature folks to talk to each other
- Plant folks working on this too
 - Schellenberger Costa, D., Boehnisch, G., Freiberg, M., Govaerts, R., Grenié, M., Hassler, M., Kattge, J., Muellner-Riehl, A.N., Rojas Andrés, B.M., Winter, M., Watson, M., Zizka, A. and Wirth, C. (2023), The big four of plant taxonomy – a comparison of global checklists of vascular plant names. *New Phytol*, 240: 1687-1702. <https://doi.org/10.1111/nph.18961>
- CoL is trying to enable cross-linking / translation across taxonomies
 - Not going for “one taxonomy” to rule them all
- Don;t want just another wishful thinking list
 - Want actual needs from stakeholders
 - What is needed, how to achieve
 - Time to actually make an impact

12 June 2024

Present: Kendra, Quentin, Aja, Cullen, Ariadna, Hernani

Quentin scored \$6 million grant to study invasive species for 3.5 years

Dave Waldien shared a spreadsheet of IUCN bat species - Aja has updated the taxonomic alignment

Ariadna possibly to attend NASBR, would present on the integration of Bat1K and other genomic data with the Big Bat Database for data mining

05 June 2024

Present: Nancy, DeeAnn, Kendra, Quentin, Aja, Ariadna Morales, Hernani Oliveira, Cecilia Montauban, Nate (late)

- Introductions for guests
- Discussing NASBR 2024 Symposium:

Potential Presenters and Proposed Titles:

Presenter: Nancy B. Simmons – American Museum of Natural History (USA)

Proposed Title: Key links in the global bat data chain

Presenter: Hernani F. M. Oliveira - Universidade de Brasília (Brazil).

Proposed title: Challenges and perspectives in collating bat trait data worldwide.

Presenter: Winifred Frick - Bat Conservation International (USA)

Proposed Title: Bat data for conservation planning: the What, How, and Why

Presenter: Nathan Upham - Arizona State University (USA)

Proposed Title: When do taxonomic splits matter for bat biodiversity and viral risk analyses?

Presenter: Aja Sherman - Bat Eco-Interactions Project (USA)

Proposed title: The Bat Taxonomic Alignment (BTA), A Standardized Review of Bat Names Across Multiple Taxonomic Authorities

Presenter: Cullen Geiselman - Bat Eco-Interactions Project (USA)

Proposed Title: Liberating literature and the data within: the Bat Literature Repository and evolution of data extraction and curation

- Cullen suggested offline “Perhaps we should consider the main points we'd like participants that attend the whole thing will take away instead of talking about BTA, Co-roosting dataset, etc in separate talks.”
- Aja suggested one talk on liberating inaccessible literature and dark data (as well as linking) and another on taxonomy for this process, the BTA, and struggles with keeping up with taxonomic changes.
- Hernani will present on Big Bat Database, traits that colleagues are measuring, and how to link.

- Nancy was concerned that Nate and Aja's talks will be too similar. Let's discuss this with Nate when he returns. Also concern that 3 taxonomy talks might be redundant. Agreed that Nancy's talk will still be relevant since some friends don't know about existing databases, darwin core, or other available resources.
- Kendra asked about the status of the Big Bat Database. Members gave us background and informed us that changes are happening with new management...
- Nancy recommends Hernani presents on the Big Bat Database, the history, and current challenges
- Aja introduced the co-roosting
- Hernani will share literature associated with best practices for data mining
- Hernani asked if we document unpublished anecdotal interactions, Nancy recommends to leave these out unless published evidence
- Cecilia interested in how we keep track of where the interactions are documented from, Aja shared citation lists and internal identifiers
- <https://www.nature.com/articles/s44185-024-00043-9>
- <https://libguides.derby.ac.uk/literature-reviews/prisma-lr>
- <https://biotxplorer.text-analytics.ch/?species1=Rhinolophus&size=1>
- <https://zenodo.org/communities/covihc/records?q=&l=list&p=1&s=10&sort=newest>
- DeeAnn emphasizing gray literature from conservation organizations
 - How to access this?
 - E.g., interactions noted from non-peer reviewed literature
 - Could flag this in the data matrix
 - <https://smartconservationtools.org/en-us/>
 - “SMART platform consists of a set of software and analysis tools designed to help conservationists manage and protect wildlife and wild places. SMART can help standardize and streamline data collection, analysis, and reporting, making it easier for key information to get from the field to decision-makers.”
- Nate parsing his talk relative to Aja's
 - Will be a use case of the BTA on bat-virus interaction data – so will follow logically after Aja's BTA talk
 - Interfacing with Kendra too
- Nancy and Aja mentioned the IUCN integration of the Batnames.org / MDD taxonomy for bats – feeding this into conservation action

10 May 2024

Present: Nate

Brainstorming for the GBatNet meeting – name of group and notes on the “short blurb” that we need prior to Weds to be included in the printed materials:

- Name ideas
 - Bat Biotic Interactions for Organizing Mammal Ecology (BatBIOME)
 - Bio Biotic Interactions Organized by Species (BatBIOS)
- Blurb [sent to Susan Tsang on 16 May]
 - Our mission is to digitally liberate and organize all knowledge of bat biotic interactions – including that contained in taxonomic and ecological literature, databases, biocollections, and other archives – for studying the multifaceted ecological roles of bats worldwide. We are a group of biodiversity collections-related scientists that have been meeting weekly since April 2020. The Covid-19 pandemic reminded us that much of the knowledge of bats interacting with other bats, as well as with parasites, predators, prey, and habitats, are locked in offline books or paywalled PDFs and thus constitute ‘dark data’ that are digitally unavailable for synthetic analyses. These data are central to building robust models of viral spillover risk and other complex ecological dynamics, the predictive abilities of which are increasingly critical to society. Additionally, the pace at which taxonomic changes to bats and other taxa are occurring has brought considerable instability to sources of genetic and ecological data that are linked by species names and inform these models. Therefore, our group endeavors to both improve the digital access and taxonomic linking of biotic interaction data for bat species globally.

8 May 2024

Present: Aja, Kendra, DeeAnn, Cullen, Quentin

Aja caught DeeAnn up on what happened last week

Kendra notes that she will be attending the iDigBio meeting - will give a presentation at the “Digital Collections Data and Tracking Disease Workshop” - really interested in bat synonymy and in co-species data [so, this is us]

Kendra asked for slides to help - Aja shared prior slides, Cullen also has some...

DeeAnn - we really need to settle on our working group name.

- Bat Biotic Interactions and Taxonomy Working Group [or are we a network - GBatNet categorization] - Cullen will ask Tigga for advice on working group or network. Or **consortium**?
- Bat Biotic Interactions and Taxonomic Alignment Consortium?
- Need greater group input; should have this settled by next week's meeting

Aja/ChatGPT suggestions:

Global Bat Taxonomy and Ecology Network (GB-TEN)

Bat Taxonomic Interactions Initiative (BTII)

International Bat Biodiversity and Interaction Consortium (IB-BIC)

Bat Taxonomy and Ecological Network (BTEN)

Global Bat Diversity and Interaction Project (GBDIP)

Bat Taxonomy and Ecological Dynamics Consortium (B-TEDC)

International Bat Taxonomy and Ecological Nexus (IB-TEN)

Bat Taxonomy and Interaction Collaborative Hub (BATICH)

Worldwide Bat Taxonomy and Ecological Exchange (W-BTEE)

Bat Taxonomy and Interaction Assessment Network (BTIAN)

As a group, we reviewed the current GBatNet networks and "Interdisciplinary Projects" [AKA working groups]. We are akin to the "Bat Phenotypes and Evolution Network (BPEN)"??

Aja news: BTA done (will step aside for a bit, then review); will work on making it a relational database as next steps. Co-roosting and BTA manuscripts are in the works. She is also creating how-to videos and tutorials for the BTA.

Aja invited the team to review both datasets while she is writing, but plans to review one more time before publication.

We are still tackling the landing site for the searchable platform for the literature collection and BTA (probably batbase).

Co-roosting data and BTA are done!

Zotero goals: 20K more by end of summer

Aja - to DeeAnn and Nancy - need more examples of lineages with lots of change/unusual scenarios (e.g., *Pteronotus*, *Miniopterus*)

Aja will work with Quentin on Gephi interaction models, species accumulation curve, and possible data mining automation resources.

Aja asked the group: When making co-roosting dataset - was there a plan for automating it? Cullen thought we would be able to use our data as a training set...

Cullen - provided some funds for Jorrit to work with Zotero collection - pdfs to zenodo etc....

Discussed NASBR symposium proposal:

<https://docs.google.com/document/d/1TFPWLNxf5KrczCNGQzybiru93AwRtbx7I6NNWVgRcbY/edit>

Will submit the 14th

Quentin new business: recently had a meeting with a group that wants to push forward a IUCN working group on parasites. Interesting to think about what IUCN status might look like for parasites. Endo and ectoparasites etc... will keep our group apprised as to how it progresses.

1 May 2024

Present: Aja, Anna, Nate, Kendra, Cullen

Anna planning to defend her PhD in early Fall

- Still doing some data collection
- Looking at isotope signatures of processed vs natural food

Kendra taking notes through this OtterPop app:

https://otter.ai/u/haFT-sl_mGZPjC_o7IJJUg8SxqA?utm_source=va_chat_link_1

AI translation / transcripts have gotten very good lately

- <https://sonix.ai/> this is the website I was using
- <https://www.linguee.com/>
 - Seems much better than Google Translate
 - Helpful for her Zooniverse project: Canyon Critters

NASBR meetings in October

- Cullen has been organizing
- We have 6 talks, could be ready to submit already
- Potential for Jorrit to be invited too
 - Digitization of bat literature
 - Lots of PDFs and metadata available now
 - Bringing the publications to Zenodo → Plazi method for this
 - But then additional steps needed to index and search
 - Modeling it off of Drosophila:
 - <https://zenodo.org/communities/taxodros/records?q=&l=list&p=1&s=10&sort=newest>
 - <https://www.taxodros.uzh.ch/>

- Building the search feature is the part that Jorrit will help with – also creating the digital signatures from that
- NA bat people at doing their own symposium
- EuroBat people don't have their own funding
 - <https://www.nature.com/articles/s41597-023-02157-4>

24 April 2024

Present: Aja, Nancy, Nate, DeeAnn

- DeeAnn is back from fieldwork in NW corner of Uganda
 - Flight cages, molecular lab – now just need bats to eat /
- We will all be in Houston for the GbatNet meetings in ~1 month
 - Prep?
 - Everyone is representing a working group here
 - Bringing these together with the networks
- Symposium at NASBR
 - Bringing bat datasets together
 - Projects
 - Aja
 - BTA
 - Cullen:
 - Co-roosting dataset
 - Jorrit:
 - Bat Literature group
 - Other talks
 - Nancy
 - Intro to symposium, progress on BatNames and MDD alignment
 - Nate
 - When do taxonomic splits matter / range maps changing through time and intersecting that with observational data
 - Outside folks
 - BCI – Fred Frick – Teague O'Mara

- Ash Malmov; Anna Fagre – ash@bathealthfoundation.org;
anna.fagre@gmail.com
- BTA updates
 - Making a website for the BTA –
 - Setting up a peer review system for that
 - Flagging the content by Nancy, DeeAnn, Nate – Connor B too
 - Making subsets of the BTA too
 - Automate this process
 - Use SQL to select fields – the underlying database would always be the same
 -

17 April 2024

Present: Aja, Nancy, Cullen, Donat, Nate,

- Nancy doing fieldwork in Belize in May
 - 85 people going over the course of 2-3 weeks
 - Becker lab, Melissa Ingala, Kelly Spear, all taking samples at different time points – aiming for recaptures to get data through time
 - How many re-captures are they getting?
 - PIT tagging started in 2021 – 13% recapture rate – arm banding the vampire bats
 - Bats tend to have 5-6 years life span in some species, or 40+ years in Myotis in the wild
 - Collecting animals for whole genome work – 23 genomes being targeted
 - Air samplers for collecting eDNA in the caves / hollow trees / human structures

Other agenda items:

- Potential symposium at NASBR in Oct 2024
 - Guadalajara, Jalisco between 23-26rd October 2024
 - Due May 15th: <https://form.jotform.com/231825735583160>
 - Ideas
 - Considering how the databases can be used in planning and decision making
 - Databases in bat biodiversity science

- Inviting Jorrit
- Topic – bats and databases – linking networking databasing – what questions are uniquely enabled by aggregating data
 - Uses for informing bat research priorities
 - Liberating, linking, leveraging bat data for research and conservation priorities
 - Leveraging = taking advantage of something existing for a new purpose / scale of questions
 - Liberating side – how could I contribute to this – how could I use it?
- Roundtable as part of this
 - Symposium + workshop
 - Workshop could focus more on “how to” of liberating and linking
 - Panel discussion – bringing ideas together
- Hernani Oliveira

3 April 2024

Present: Aja, Nancy, Cullen, Kendra

- GBatNet meeting
 - Possibly include NABat, Hernani Oliveira (oliveira.hfm@gmail.com), Orly Razgour, and/or Ariadna Morales (ariadna.biologia@gmail.com)
 - There will be a questionnaire soon for travel details. Reservations will be made through GBatNet, not individuals
 - Need catchy title
 - Big Bat Databases: Easier Said Than Done
 - Speaking with facilitators
- Bat Literature Project
 - If Aja finds manuscripts, she will omit
 - Kendra will add literature
- BTA
 - Spoke about species country lists and if it was possible to solve island only issues - maybe revisit this topic, maybe not
- Kendra brought to our attention this grant opportunity
 - <https://www.nsf.gov/pubs/2024/nsf24069/nsf24069.jsp>
 - She also apologized for using the saying “negative Nancy” during previous meeting

27 March 2024

Present: Quentin, Aja, Nancy, DeeAnn, Nate, Anna

GBatNet connections

- Meeting in Houston – Week of the 20th of May – 21-24/25
- New name needed
 - Bat-Pathogen Working Group?
 - Other groups: <https://www.gbatnet.org/interdisciplinary-projects/>
 - Big Bat Database is taking off
 - Hernani Olvera, and Ariadna Morales leading this
 - How old is this bat?
 - Evo-devo group
 - Population sizes of bats
 - Bats in habitats – parasites/microbiomes
 - Stress in bats (DeeAnn)
 - Immune systems in bats (Dan Becker)
 - Key habitats
 - SciComm of bat science
 - Socioeconomics group
 - 11 of original 15 –
- Our group
 - Co-roosting dataset
 - Bat taxonomic alignment
 - Ecological interaction
 - Bat Biotic Interactions Working Group
 - Biodiversity Exchange on Bat Biotic Interactions
 - Chirocosmos!
 - Bat Interactions and Digital Knowledge
 - Bat Biotic Interactions and Taxonomy Exchange (Bat BITE)
- Asdell's Patterns of Mammalian Reproduction
 - Punch cards burned in a fire
 - Digitized?
 - Nate has mammae number and litter size data digitized – and collaborator Tom Stewart does have a scanned copy of the volume – planning to upload to Zenodo in the near future

Moth defenses against bats

- New employee at AMNH: Jesse Barber
 - Barber, J.R., Plotkin, D., Rubin, J.J., Homziak, N.T., Leavell, B.C., Houlihan, P.R., Miner, K.A., Breinholt, J.W., Quirk-Royal, B., Padrón, P.S. and Nunez, M., 2022. Anti-bat ultrasound production in moths is globally and phylogenetically widespread. *Proceedings of the National Academy of Sciences*, 119(25), p.e2117485119.
<https://www.pnas.org/doi/10.1073/pnas.2117485119>

DeeAnn wrote a book:

- "Lives of" Bats; series Princeton University Press
- "Lives of Bats" book
 - Out in October 2024!

Co-roosting dataset

- <https://docs.google.com/spreadsheets/d/1y5uBKvyzDQgQUtyRHvn5f20IOA7LoMho/edit#gid=691336520>
- Looking at the GBIF matched names vs others
 - Record secondary authorities in "interpreted_name" column
 - Migrate (*) in GBIF
 - Plants: Quentin
 - Macroparasites: Anna
 - Ants: Donat?
- Aja working on manuscripts:
 - Methods for co-roosting
 - Methods for bax taxonomic alignment

20 March 2024

Present: Quentin, Aja, Cullen, Nate, Donat at the end

IPBES reports – species assessment

- Figuring out baselines vs monitoring of species

Co-roosting data – Aja is cleaning up the names relative to the GBIF backbone

iDigBio meeting 2024 in Kansas:

<https://www.idigbio.org/content/digital-data-biodiversity-research-conference-2024>

TDWG in Japan in Sept: <https://www.tdwg.org/conferences/2024/call-for-abstracts/>

NASBR in Mexico / Guadalajara in Oct: <https://www.nasbr.org/>

- We could do a symposium? <https://www.nasbr.org/symposia24>
 - NA bat people
 - Occurrence data
 - Data aggregating – movement data aggregation
 - Ask Nancy who all could be giving talks
 - Way for our talks to not be lost
 - Roundtable possibly
 - Bat morphometric data possibly

- Bat data aggregation and synthesis generally
- May 15 – deadline – time to assemble

Oscar project

- <https://oscars-project.eu/>
- Covers all disciplines (not just biodiversity) but there is a life sciences cluster
- Could be good for Jorrit –

iNaturalist data on degree of establishment of invasive species in a given country

- Current issue with iNaturalist not pushing these data to GBIF

Donat talked to Jorrit – going to get all of these bat publications in Zenodo

- Scope for upload:
 - Taxonomy and ecology initially
 - Nancy, DeeAnn, Nate (MDD), Victor Van Cakenburgh
 - Gabor Scroba and Sandro Bertolino would be good to contact too

6 March 2024

Present: Quenting, DeeAnn, Aja, Nancy, Kendra, Donat

Aja explaining

- the [biotic interactions table](#) and what she needs from the team
- Cleaning up the mammal database synonyms, priority
- Starts writing the ms biotic and coroosting
- Andrea has not change document for [BTA](#)
- Writing the BTA paper by summer

Bat literature list

- Zootero add keywords automatically? If not possible do it later

28 Feb 2024

Present: Quentin, Donat, Aja, Cullen, Kendra

Reviewed top priorities with the group that wasn't present last week

Aja compared taxonomy (of interaction dataset) against Checklist Bank & GBIF for agreement

BiCIKL+ submission by Donat & Quentin (will know if funded by July)

Donat to send template to Aja for organizing citations stored on Zotero

21 Feb 2024

Present: Quentin, Aja, Nate

Camping gear

- <https://www.outdoorgearlab.com/>
- <https://www.bioliteenergy.com/products/campstove-2-plus>

Top priorities right now

- Taxonomy for the co-roosting dataset
 - Did a check against the GBIF backbone – using this to verify
 - Bat-bat interactions
 - We can dig into this part of the dataset to a greater extent
 - Plot the two phylogenies facing each other
 - Make line thickness proportional to the frequency of interactions
 - Rarefaction analysis as well
 - Are these sampling curves saturating? Or will it go on and on depending on increased sampling
- BTA manuscript
- How-to for the BTA

14 Feb 2024

Present: DeeAnn, Nancy, Aja, Nate

AirTags – helpful for tagging the field supplies when traveling

Ebola virus-like particles – just have the protein on the outside to be recognized by the bats –

Using DNA/RNA shield – also BAMM Banker for tissue culture

- Slowly make way into freezing
- Needs new -80C

Aja update

- BTA versions – batNames –

7 Feb 2024

Present: Cullen, Kendra, Nate, Nancy, Anna, Donat, Quentin

Anna presenting on ringtail omnivory at the University of Georgia, Odum School of Ecology annual student symposium

- Rank abundance curves of dietary items in scat
- Being an omnivore is on a big spectrum
- https://www.inaturalist.org/observations?place_id=any&taxon_id=41676
 - Includes lots of feces images
- Early ASM poster about this work ([Spear et al. 2022](#))

Cullen / Donat / Aja / Nancy

- Digitization of bat literature – all literature is there –
- First working on 20k Drosophila publications
 - Will figure out full workflow there
- Bat literature library in Switzerland – in Geneva – Can order PDFs from them
 - Muséum in Geneva
<https://www.ville-ge.ch/musinfo/bd/mhng/cco/accueil.php>, and was founded by the now retired Pascal Moeschler.
- Lots of low hanging fruit here
 - Victor Van Calkenburgh has a massive PDF library that he gave access to
- What are the goals?
 - Every publication ever on a bat?
 - Can be a parallel bibliographic effort and PDF ect
 - Multiple goals
 - 1. Get full bibliography reference for all publications
 - 2. Get PDFs for those papers
 - 3. Upload PDFs to Zenodo
 - 4. Tag / annotate / extract PDFs for forming digital knowledge
 - For 2-3-4 steps, can target a specific topic / clade
- What metadata do we need?
 - Some columns already developed on the “[reference-extract” sheet](#) of the co-roosting data.
- Nate / MDD can send literature on bats from the Zotero of the MDD
 - Will send to Aja

Bat Libraries:

- Bat bibliographies generated from the General Index subset:
 - Special Index to Species: <https://archive.org/details/IndexGenerium>

Other Internet Archive scans:

- BioStor <https://archive.org/details/biostor-127588/page/438/mode/2up>
- Big issue with bioheritage library is the lack of separated articles. Volumes are scanned and not parsed by article metadata.
 - Lack of funding/staff to do this. Priority is Smithsonians own stuff

24 Jan 2024

Present: Donat, Cullen, Kendra, Nate, Nancy, Aja, Quentin

Our collective bodily injuries and aging

- Nancy retiring – getting rid of administrative things

Liberating large amounts of bat literature

- Could we take all the PDFs from Nancy's computer – plus the files from batBase
- Possible to have a drag and drop?
 - Nancy has a combined folder for the first time – just joined things across computers
- Aja can help to create the Zotero record – BibTex or other –
- Then Donat can work with Plazi team to automate the uploads to Zenodo
- Also add the MDD references in this context
 - 196 "bat" references included – not updated for the v1.12 version

Philippines project work – aiming to get more funding via EcoHealth

- <https://www.washingtonpost.com/national-security/2023/09/07/deep-vzn-virus-wild-usaid-program/>
- <https://stopspillover.org/> (program that is funded by USAID)
- Lots of in country work being done –
 - Capacity building ultimately

10 Jan 2024

Present: Quentin, Cullen, Kendra, Aja, Nate, Donat

Updating the taxonomy of the bat co-roosting data

- Quentin took all the names included + matched it against the GBIF taxonomy backbone
- <https://docs.google.com/spreadsheets/d/1y5uBKvyzDQgQUtyRHvn5f20IOA7LoMho/edit#gid=691336520>
 - WorkingDoc_Raw-Data-Extraction_ExpertLiterature_8_8
- Steps were to use the GBIF taxonomy matcher tool
 - Species lookup: <https://www.gbif.org/tools/species-lookup>
 - ChecklistBank has something different – provides stable identifiers –
 - <https://www.checklistbank.org/tools/name-match>
 - Will return “valid” species names – but the species could still definitely be split
 - Kendra example with Miniopterus schreibersii – used to be widespread species, now restricted to Europe – so all the records in the Philippines are inaccurate
- Sources for digitizing all the bat literature
 - Try different sources
 - Private funding
 - Swiss Govt – RDF version of checklist bank – \$300k for these conversion services – just for the Swiss bats for the starter
 - Proof of what can be done for the swiss species
 - BiCikL – Arcadia
 - Working with Nancy – to get all the bat literature – leverage to get more funds to expand further
 - Gathering PDFs from her, batnames, MDD
 - Some of the bibliographic data contain errors currently
 - Knowledge Hub with the BiCiKL
 - <https://bicikl-project.eu/biodiversity-knowledge-hub>
 - Pensoft is involved because of the JATS conversion needed for PMC – Biodiversity PMC
 - <https://sibils.text-analytics.ch/search/> (Sibls portal)
 - Also this “Knowledge Centre for Biodiversity”:
https://knowledge4policy.ec.europa.eu/biodiversity_en
- Locality information for the bat co-roosting information

- There is a need to consider the “localityCountry” field in comparing to the BTA for taxonomic translation
 - Compare this to the “countryDistribution” field in the MDD for example – create a pairwise country matrix for doing this
 - >> Then have logic for the subsequent comparisons
 - E.g., if nameA exists in countryX in taxonomyY, then we have the possibility of properly translating taxonomic splits

03 Jan 2024

Present: Kendra, Nancy, Quintin, Anna

- Hellos from the holidays
- Quintin working on
 - flora online : no global one available that's great; need for more specific areas (e.g., Congo)
 - Difficult to acquire funding for the projects
 - IUCN invasive species as threat
 - Data available through redlist R package
- Role of reference books
 - ICTV virus taxonomy still heavily referenced book
 - Walker's large bat book: almost 25 lbs
 - Walker's Mammals of the World may have had last version in 2018
 - Interesting/Harsh Review broaching the topic
<https://www.mammalwatching.com/2018/07/17/a-review-a-long-review-of-the-latest-version-of-walkers-mammals-of-the-world/>
 - Out of date by the time they're published?
 - IUCN
 - embedded citations + species-specific authors can give idea of when updated
 - 2000 was when all animals started to be assessed
 - Many bats that were listed covered since the '80s
 - Taxonomy always an issue in any of the records
- How to deal with rejection/controversy