

Improving the standardization of wild bee occurrence data: Towards a formal wild bee data standard

Brianne Du Clos

US National Native Bee Monitoring Research Coordination Network

University of California, Riverside

Co-authors

- Katja Seltmann, University of California, Santa Barbara
- Nash Turley, The Pennsylvania State University
- Clare Maffei, U.S. Fish and Wildlife Service
- Erika Tucker, Biodiversity Outreach Network
- Ian Lane, U.S. Fish and Wildlife Service
- Hannah Levenson, North Carolina State University
- Hollis Woodard, University of California, Riverside

Manuscript overview

- Manuscript
 - Definitions, examples, use, credit, conduct, implementation and management going forward
 - Pathways to data sharing
 - Case studies with standardized protocols
- Appendix 1: *The Wild Bee Data Standard* Term List, v. 1.0.0
- Appendix 2: Glossary of Terms
- Appendix 3: Additional Resources

Online resources

- GitHub repository
- Templates for standardizing data
- Special issue page for protocols that demonstrate using the standard

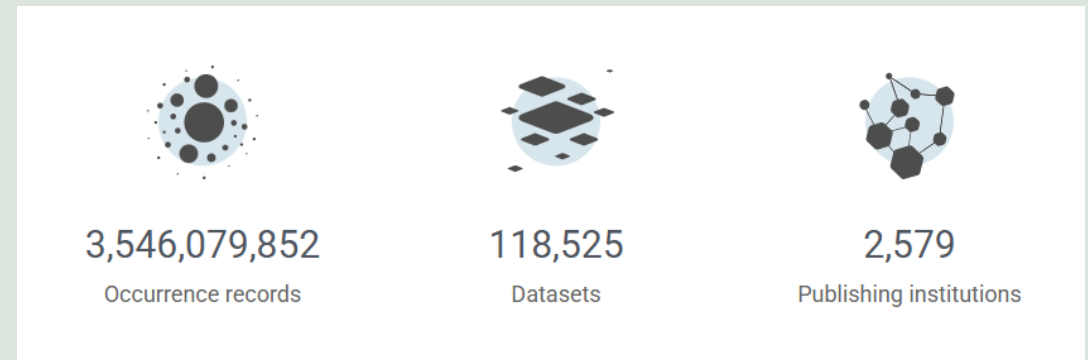
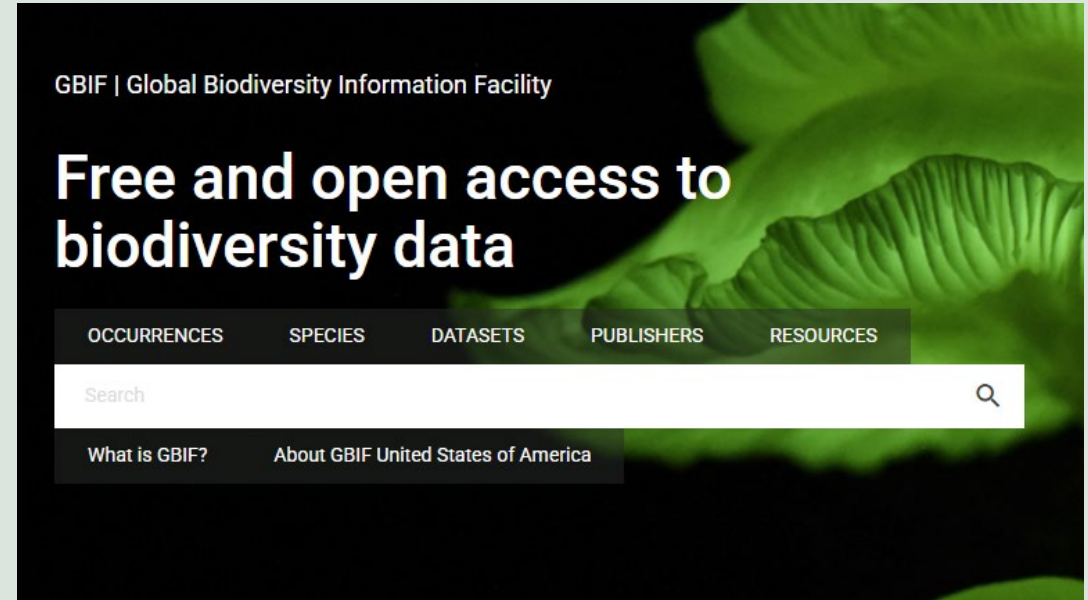


<https://github.com/Big-Bee-Network/wild-bee-data-standard>

“Wild bee occurrence data are generated when bees are collected or observed in the field, and include the date collected or observed, collector or observer identity, geographic location, and other information.”

What is occurrence data?

- An organism
- In a place
- At a time
- ...and other information!



What's the goal?

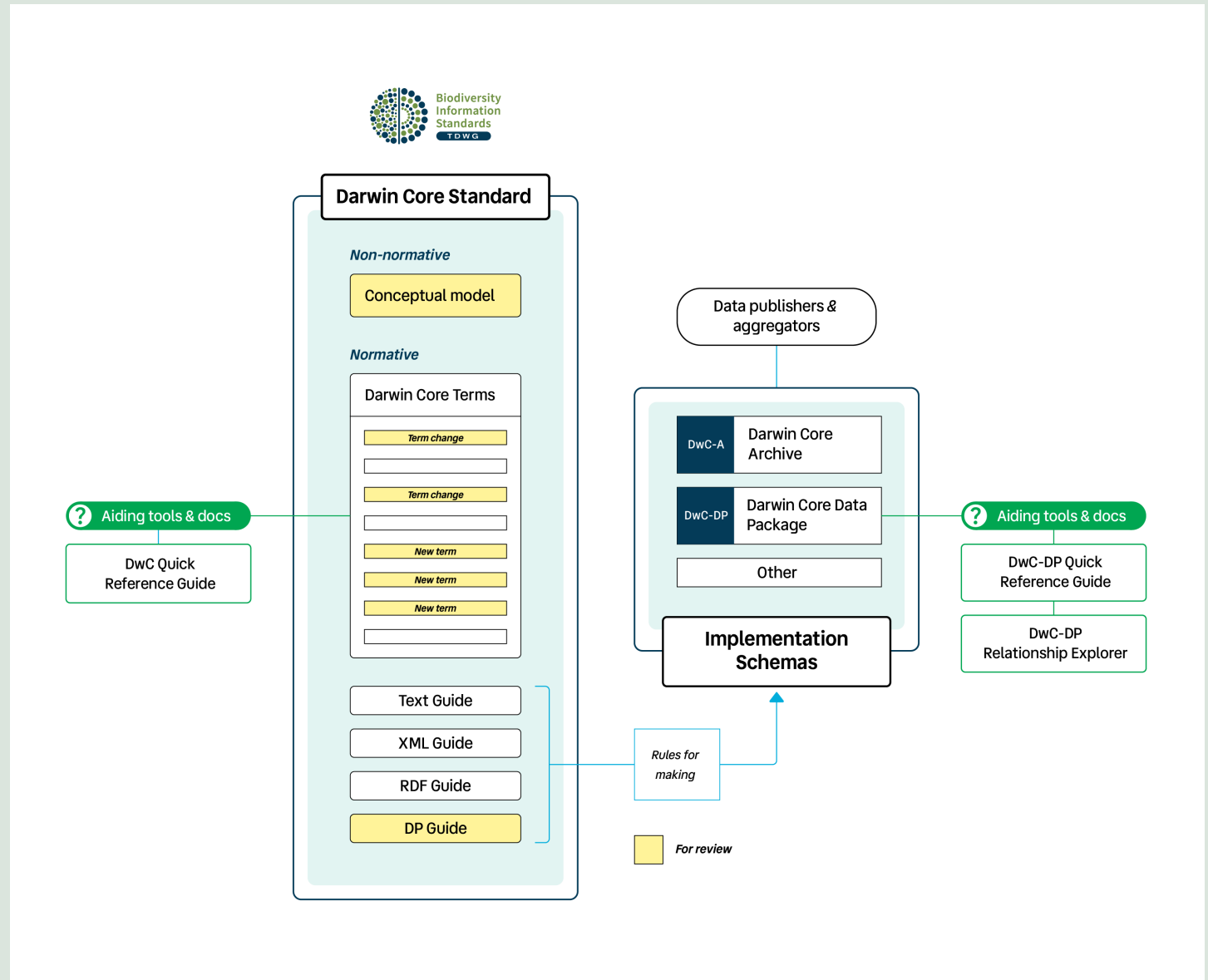
To assess species status and trends over time

How can we get there?

1. We need more data
2. Those data need to be standardized

How can we standardize occurrence data?

<https://dwc.tdwg.org/terms/>



How can we standardize wild bee occurrence data?

- *The Wild Bee Data Standard*
 - Not a new standard – just an adaptation of Darwin Core written specifically for wild bee occurrence data
 - The terms are the same, but they are presented differently
 - 75 terms in 3 tiers of requirement:
 - Core (26) – must be provided and collected
 - Recommended (22) – should be provided if collected
 - Optional (27) – can be provided if collected

26 core terms:

RECORD-LEVEL: institutionCode, basisOfRecord, informationWithheld

OCCURRENCE: occurrenceID, catalogNumber, recordedBy, individualCount,

occurrenceStatus

EVENT: eventDate, eventTime, year, month, day, samplingProtocol, sampleSizeValue,

sampleSizeUnit, samplingEffort

LOCATION: country, stateProvince, decimalLatitude, decimalLongitude,

coordinateUncertaintyInMeters

IDENTIFICATION: identifiedBy

TAXON: scientificName, genus, specificEpithet

Where to share?

GBIF

- One portal for all global taxa
- Publish data through an accredited publisher
- Data mapped to Darwin Core standard as part of publication
- Accepts physical specimen information
- Accepts photo observation information, but not image files -- photo data must provide links to where images are hosted
- Data downloads either as a simple CSV or a Darwin Core Archive

Symbiota

- Multiple taxa-focused portals
- Anyone can publish data to a Symbiota portal
- Data mapped to Darwin Core standard as part of publication
- Accepts physical specimen information
- Accepts photo observation information and hosts image files
- Data downloads either as a Symbiota Archive or a Darwin Core Archive
- Data uploaded to Symbiota portals can be served to GBIF for access and download
- Data can be live-managed in Symbiota portals

iNaturalist

- One portal for all global taxa
- Anyone can publish data to iNaturalist
- Data mapped to Darwin Core standard as part of publication
- Accepts photo observation information and hosts image files
- If observations become research-grade, data can be served to GBIF for access and download
- Can download research-grade observations through iNaturalist or GBIF

You can do this!

- Use the provided templates and the case studies in the protocols to develop in-house sampling methods that align with proposed standards
- By entering your data and your metadata into our template, you've created everything you need to share your data!
- Symbiota portals do not have barriers to entry!
 - They do a lot of the work for you – mapping terms, creating unique identifiers for each occurrence record
 - ...and data there can end up on GBIF!

Thank you!

- Participants in the RCN Data Management Workshop
- Co-authors
- Reviewers
- Journal editors
- National Institute of Food and Agriculture



Questions?



**NATIONAL NATIVE
BEE MONITORING
NETWORK**

bduclos@lumcon.edu