

SCAN-LepNet

1. SCAN & LepNet to iDigBees

2. SCAN as a North American Arthropod Portal

3. LightningBug & Symbiota2

4. iDigBees & Research-Ready Data



Seeding Sustainable Digitization



SCAN TCN 2012-2016 (Active PENs continue)

1. **10** Collections, **7** funded PEN Projects (Still sponsoring PENs)
2. Original focus on Southwest Ground-Dwelling Arthropods
3. **2,252,066** records to date (2X expected)
4. **65** non-ADBC funded collections, **1,040,293** digitized records
5. **36** pubs on SCAN or using data
6. **Current focus = North American Arthropods, 298 collections, 30 million records and 5.2 million images**

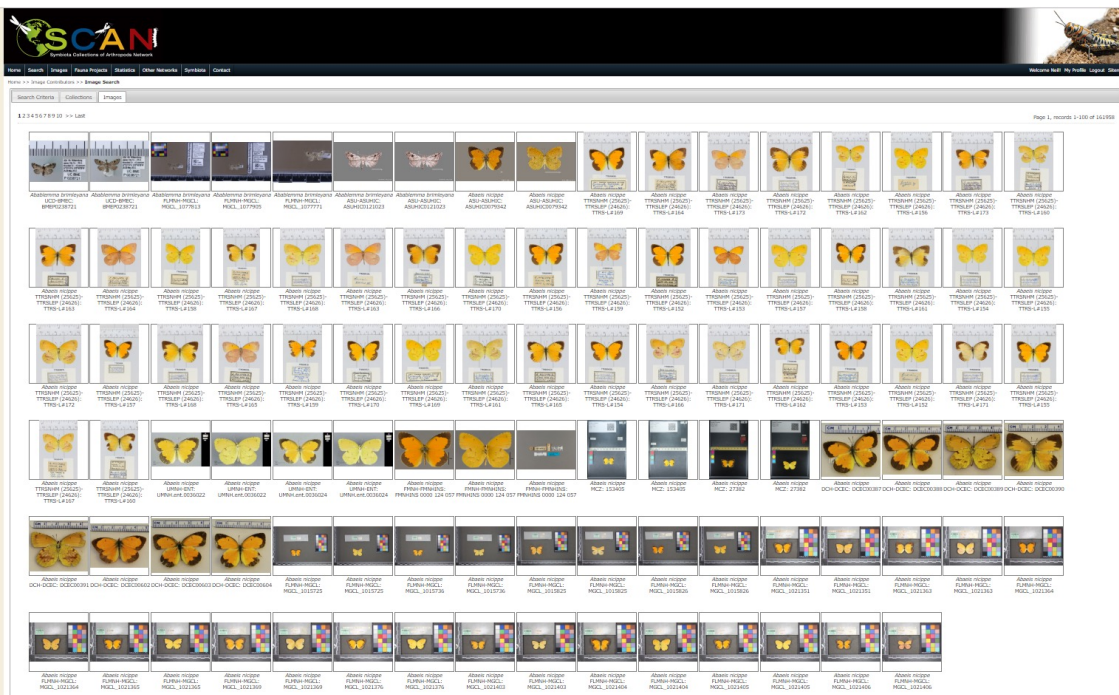


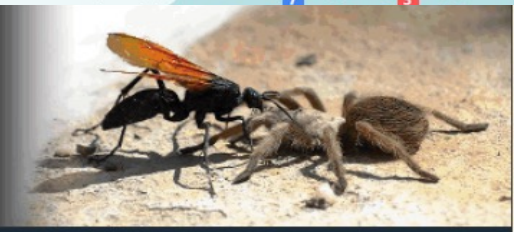
Core TCN Collections 2016-2020-1

1. Focus on North American Lepidoptera - ~142,000 species
2. 26 Collections, 4 PEN grants (Still sponsoring PENs)
3. 1.5 million records, 172k images

Broader Impacts

1. 132 collections contribute > 2.2 million Lepidoptera records
2. 148 families represented
3. 93% of records from North America, but 358 countries/regions represented





Symbiota Collections of Arthropods Network (SCAN): A Data Portal Built to Visualize, Manipulate, and Occurrences

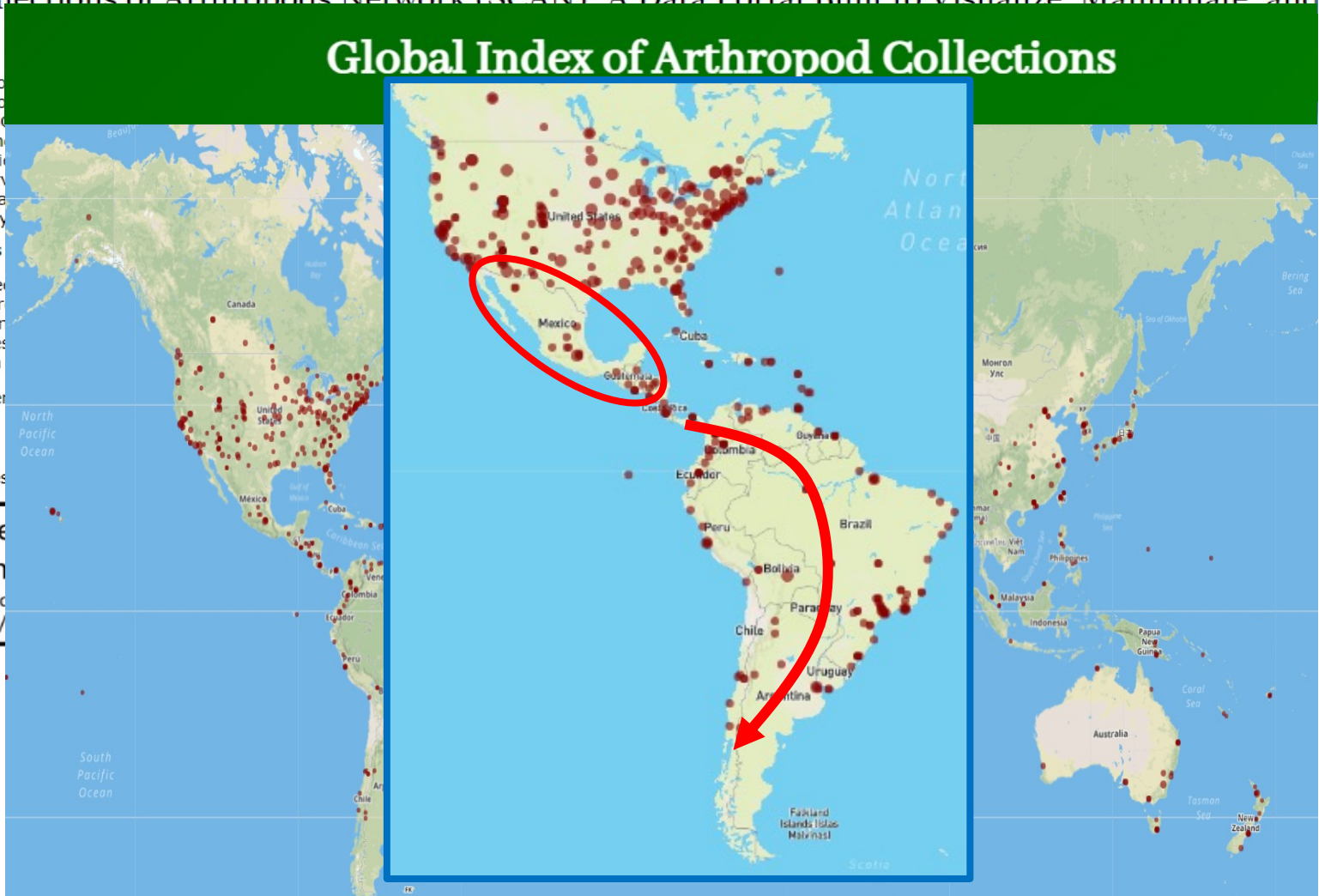
SCAN serves arthropod data for North America but globally. It serves data to GBIF. SCAN is the primary repository and the primary TCN. InvertEBase serves bees (iDigBees). We allow users to easily query data.

- Important features**
1. Easy web-based interface
 2. Download entire datasets
 3. Map georeferenced data
 4. Upload high-resolution images
 5. Design custom maps

This website is the central hub for network updates.

SCAN currently serves

Get a free Den... Manage your collec... with a free skype/...



SCAN Portal

1. Primary Data Aggregator
Most Complete arthropod occurrence data for North America
(30 million records)
(5.2 million images)
(201 Institutions)
(298 collections)
2. The Arthropod Index
Database of 938 arthropod Collections >
Promote SCAN Model

Contributing to the Global Pipeline: GBIF

54 GBIF “Collection” datasets (NOT Project datasets)

INSTALLATION

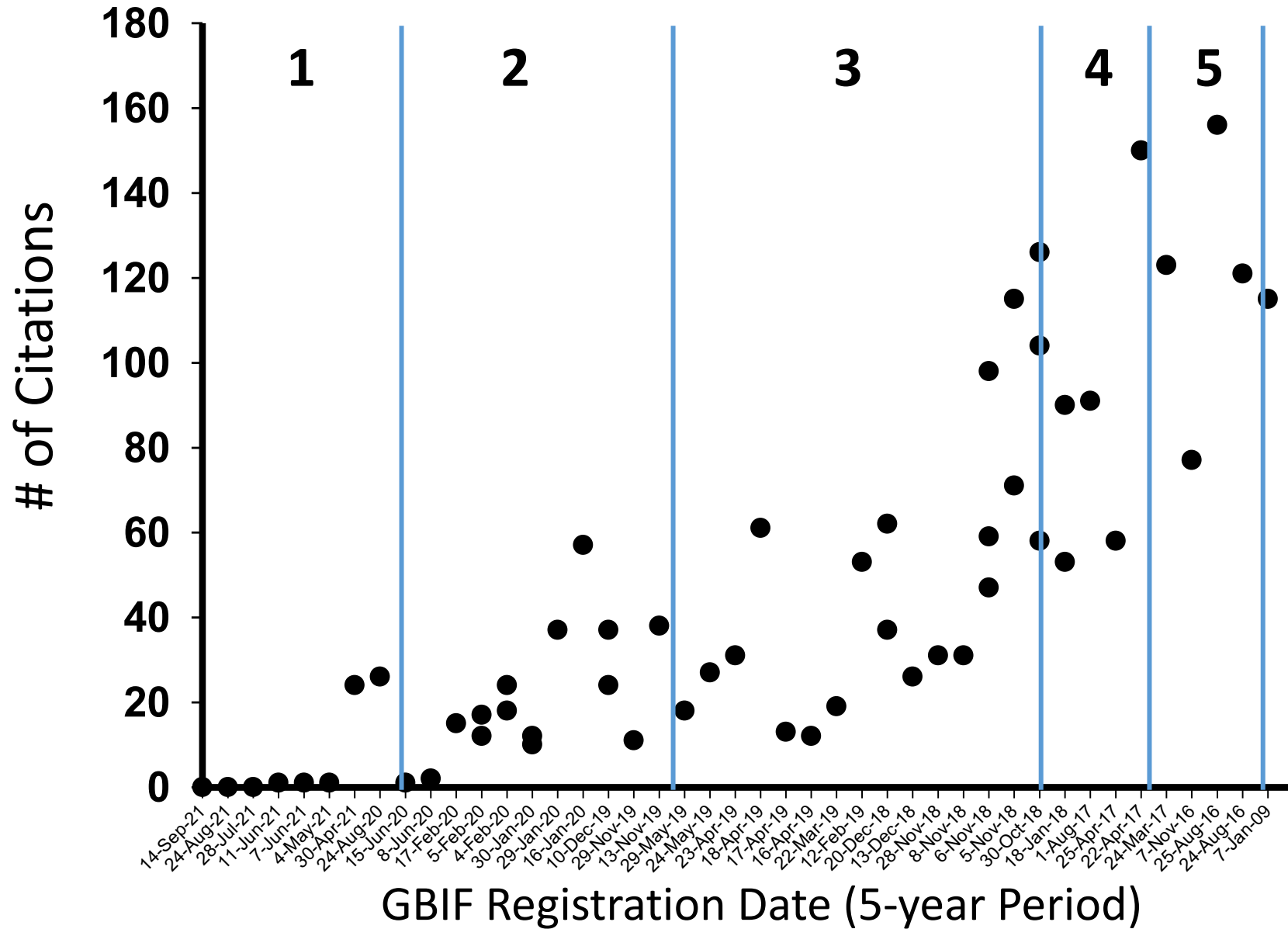
Symbiota Collections of Arthropods Network (SCAN)

The Symbiota Collections of Arthropods Network (SCAN: <https://scan-bugs.org>) serves specimen occurrence records and images from over 100 North American arthropod collections for all arthropod taxa. The focus is on North America but global in scope. SCAN is built on Symbiota (<http://symbiota.org/docs/>), a web-based collections database system that is used for other taxonomic data portals, including (Symbiota Portals). SCAN is the primary repository for occurrence data produced by the four continuing Thematic Collections Networks (TCNs), the Southwest Collections of Arthropods Network (SCAN TCN), the Lepidoptera of North America Network (LepNet TCN), Terrestrial Parasite Tracker (TPT: <http://parasitetracker.org/>), and arthropod data produced by InvertEBase TCN (<http://invertebase.org/>).

Installation type: Symbiota installation

<https://www.gbif.org/installation/2c733a9d-363d-4d66-9aef-3e0f7bc44bec>

Collections GBIF Registered by SCAN (N=54)



Key Stats

- ❖ Data are increasing incrementally
- ❖ Citations are increasing exponentially
- ❖ **2,501** Citations (~500 unique citations)



Complete restructuring and modularization of Symbiota code

Open-modular code will allow for developers to easily create Symbiota modules or API linkages to 3rd party products

Symbiota2 is now available to developers <https://gitlab.com/symbiota2> and will be available to end users by January, 2022

Demo site allows anyone to see progress <https://demo.symbiota2.org/sitemap>

Everything Symbiota2 can be viewed at <http://symbiota2.org/>

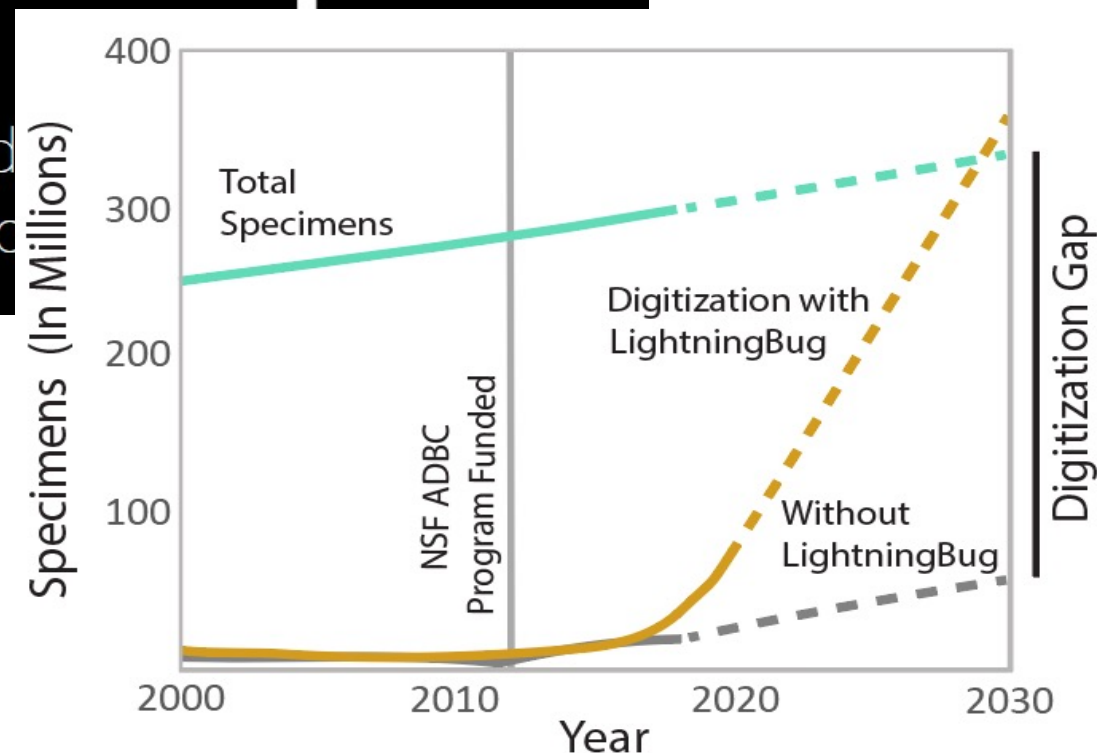
An Integrated Pipeline to Overcome The Biodiversity Digitization Gap

lightningbug.tech

through semi-automated
and 3D photogrammetric

- ❑ Prototypes at Yale and Harvard by June, 2022
- ❑ Models ready for general testing by January, 2023

1. Image with labels on specimens
2. Creates virtual label, OCR, parse to database
3. Creates 360 suite of images for photogrammetry



Develop existing data sets – Example Bees

- Define distributions for all ~5,000 species of bees from Alaska to Costa Rica (and beyond to Chile)
- Determine ecological factors mediating past, present and future distributions
- Inform conservation efforts



Any target group or area

SCAN –Ground-dwelling

Carabidae

Acrididae

Spiders

LepNet

Macromoths

SUMMARY

1. Both SCAN & LepNet TCNs successful in digitizing specimen labels and images, both still receiving PENs
2. Continue to add a new collection to SCAN every month and every other month to GBIF
3. Technology focus: **LightningBug** technology, **Symbiota2** will greatly increase data mobilization for data providers
4. Biological focus: Publishing data sets and connecting with the world

