Stefanie E. LaZerte

ROpenSci

ESA 2021

INS 6: **Open Data Resources During a Global Pandemic**

Live Discussion: Tuesday August 3 2021, 11am - 12pm Pacific

Using rOpenSci packages to access open data for ecology

🐦 @steffilazerte

⌨ steffilazerte

🌐 steffilazerte.ca

🐦 @rOpenSci

🌐 rOpenSci.org

R OpenSci

About　　Blog　　Projects　　Packages　　Community　　Resources

New!　rOpenSci at useR!2021 - Presentations from Staff and Community　➡

Transforming science through open data, software & reproducibility

We help develop R packages for the sciences via community driven learning, review and maintenance of contributed software in the R ecosystem
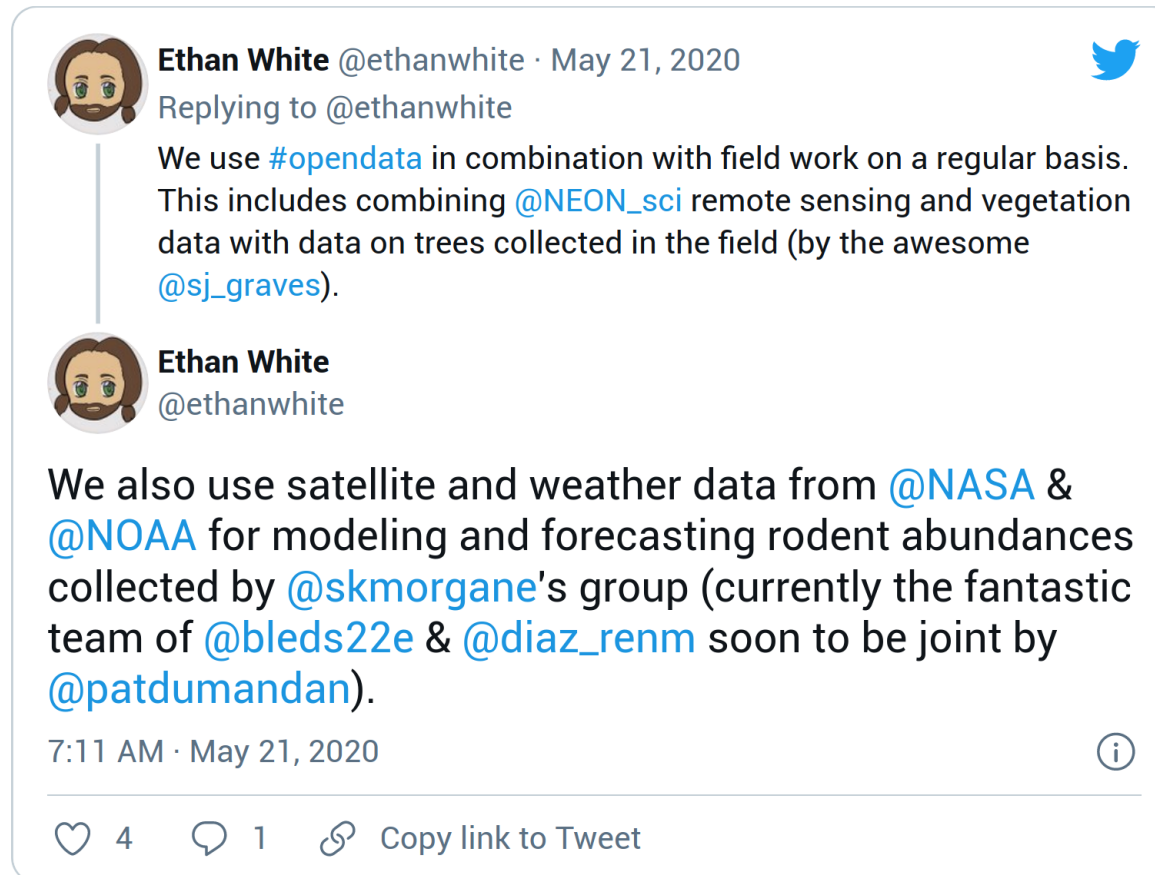
Browse Our Packages　　Our Mission

Open data

rOpenSci fosters a culture that values
open and **reproducible** research using
shared data and reusable software

We do this by

"Creating technical infrastructure in the form of carefully vetted,
staff and community-contributed R software tools that
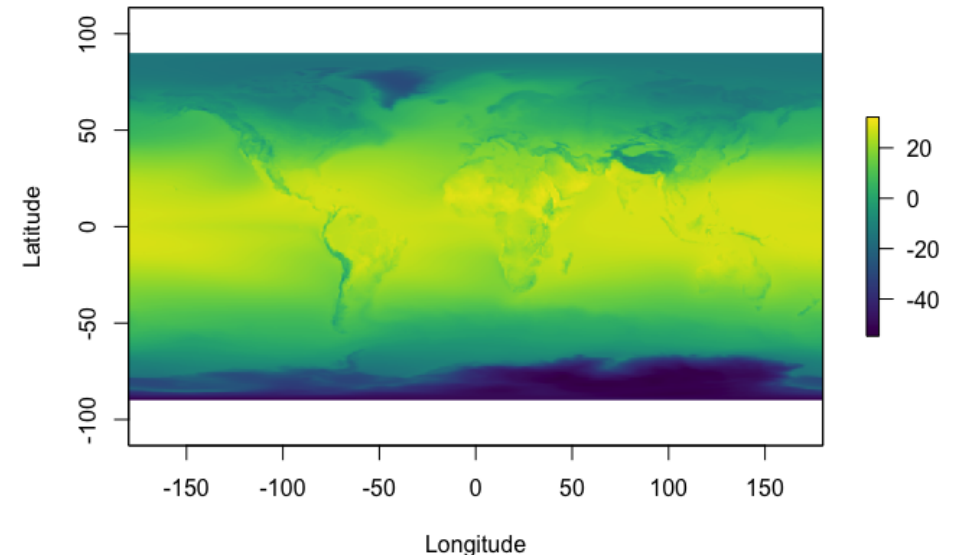**lower barriers to working with scientific data sources on the web**"

Open Data → Open Science → Better Science

Complement field data

Ethan White @ethanwhite · May 21, 2020
Replying to @ethanwhite
We use #opendata in combination with field work on a regular basis. This includes combining @NEON_sci remote sensing and vegetation data with data on trees collected in the field (by the awesome @sj_graves).

Ethan White
@ethanwhite

We also use satellite and weather data from @NASA & @NOAA for modeling and forecasting rodent abundances collected by @skmorgane's group (currently the fantastic team of @bleds22e & @diaz_renm soon to be joint by @patdumandan).

7:11 AM · May 21, 2020

♡ 4   💬 1   🔗 Copy link to Tweet

Open data to study ecology

Complement field data

- **rnoaa: [NOAA](#) weather, climate, tides, ice, etc.**

  (Scott Chamberlain, 2021)

- **nasapower: [NASA POWER](#) Global Meteorology, Surface Solar Energy and Climatology**

  (Adam Sparks, 2018)

- **riem: Weather data from international airports via Iowa Environment Mesonet**

  (Maëlle Salmon, 2016)

Global 30-year annual average temperature modelled from satellite data accessed via rnoaa*

*rOpenSci blog [POWER to the People](#) by Adam Sparks

When field or lab work is not an option*

- **rgbif: Download taxonomy and occurrences from [Global Biodiversity Information Facility](#)**

  (Scott Chamberlain & Damiano Oldoni, 2021)
    - all taxonomic groups

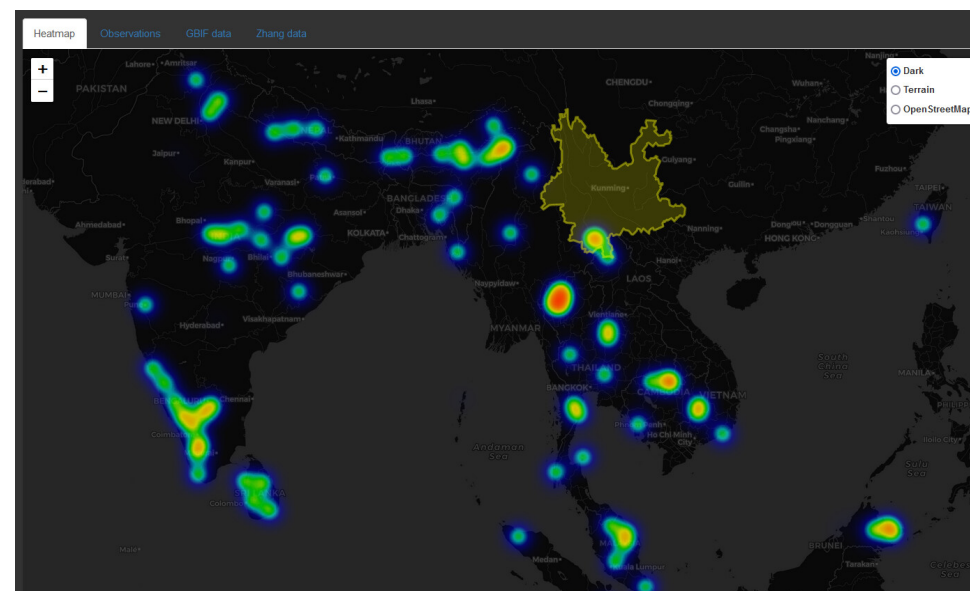- **rinat: Download observations from ❧ [iNaturalist](#)**

  (Vijay Barve, Edmund Hart & Stéphane Guillou, 2021)
    - all taxonomic groups

- **auk: Manage observations from [eBird](#)**

  (Matthew Strimas-Mackey, Eliot Miller & Wesley Hochachka, 2018)
    - avian observations

Heatmap of asian elephant observations using rgbif**

*rOpenSci blog [When Field or Lab Work is not an Option - Leveraging Open Data Resources for Remote Research](#) by The rOpenSci Team *et al.*

**rOpenSci use cases [Mapping asian elephant observations with rgbif](#) by Tuija Sonkkila

Open data to study ecology **of a pandemic**

Explore how birds change their use of urban landscapes during a pandemic[*]

- Pandemic resulted in declines in human activity

- How does this affect use of urban areas by birds?

- Open data from eBird

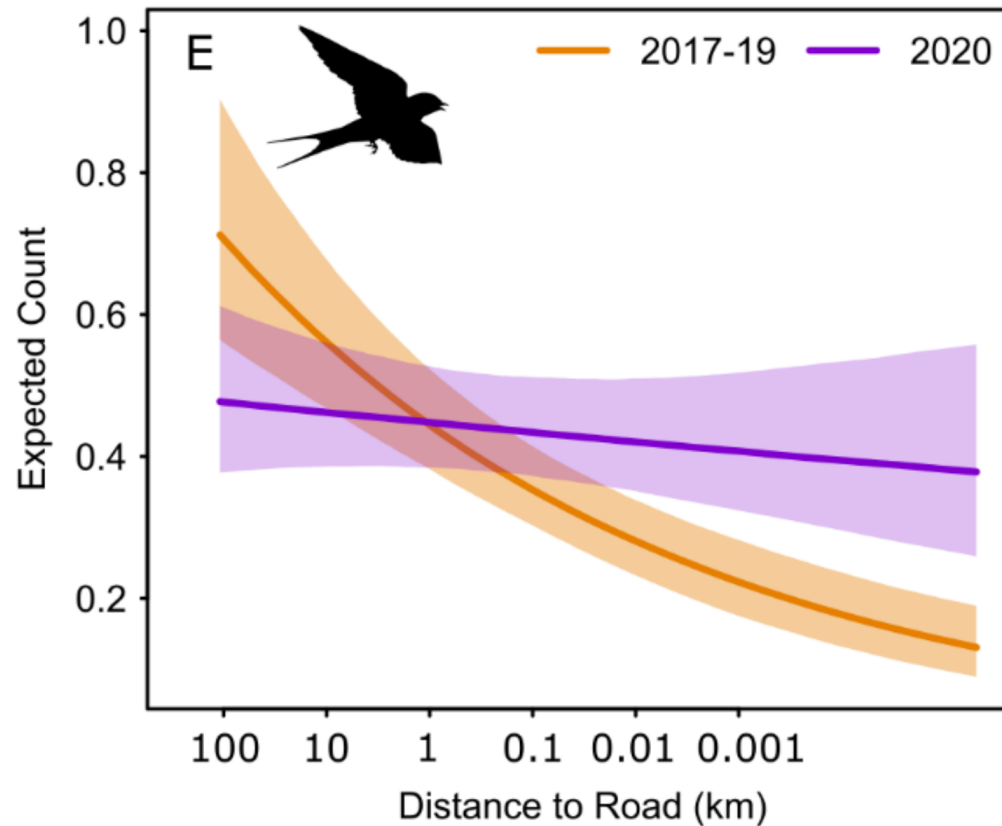- rOpenSci **auk** package to manage huge eBird data sets

Science Advances Reduced human activity during COVID-19 alters avian land use across North America
(accepted pending revisions) by Michael Schrimpf *et al.*

Explore how birds change their use of urban landscapes during a pandemic*

*Science Advances** Reduced human activity during COVID-19 alters avian land use across North America
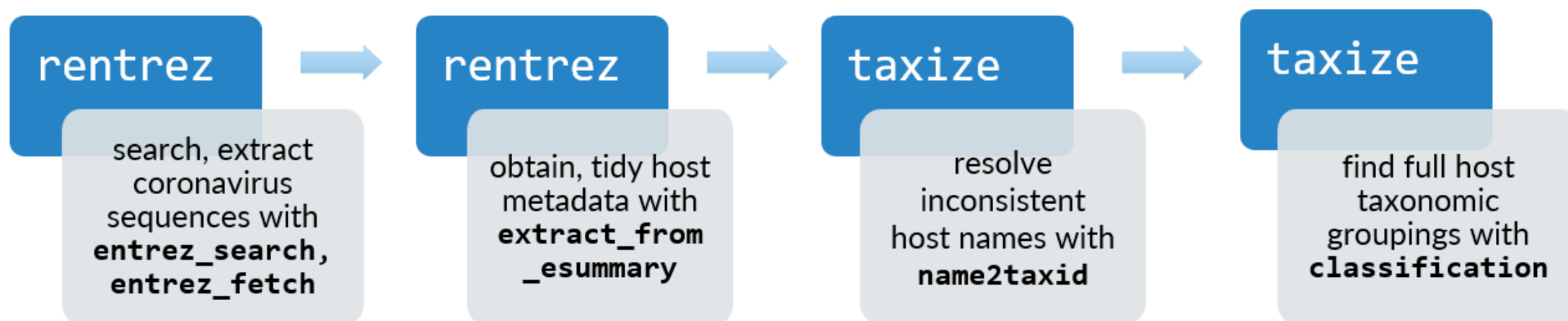(accepted pending revisions) by Michael Schrimpf *et al.*

Open data to study **a virus**

Data on the coronaviruses*

- **rentrez: Search, download data from, and otherwise interact with NCBI databases**

    (David Winter, 2017)

- **taxize: Search by species names for taxonomic hierarchical information**

    (Scott Chamberlain & Eduard Szocs, 2013)

rentrez → search, extract coronavirus sequences with `entrez_search, entrez_fetch`

rentrez → obtain, tidy host metadata with `extract_from _esummary`

taxize → resolve inconsistent host names with `name2taxid`

taxize → find full host taxonomic groupings with `classification`

*rOpenSci blog [Using Open-Access Tools (rentrez, taxize) to Find Coronaviruses, Their Genetic Sequences, and Their Hosts](#) by Liam Brierley

Open Data → Open Science → Better Science

Thank you!

Dr. Steffi LaZerte
Analysis and Data Tools for Science

🐦 @steffilazerte

⌨ steffilazerte

🌐 steffilazerte.ca

R▸OpenSci

🐦 @rOpenSci

🌐 rOpenSci.org