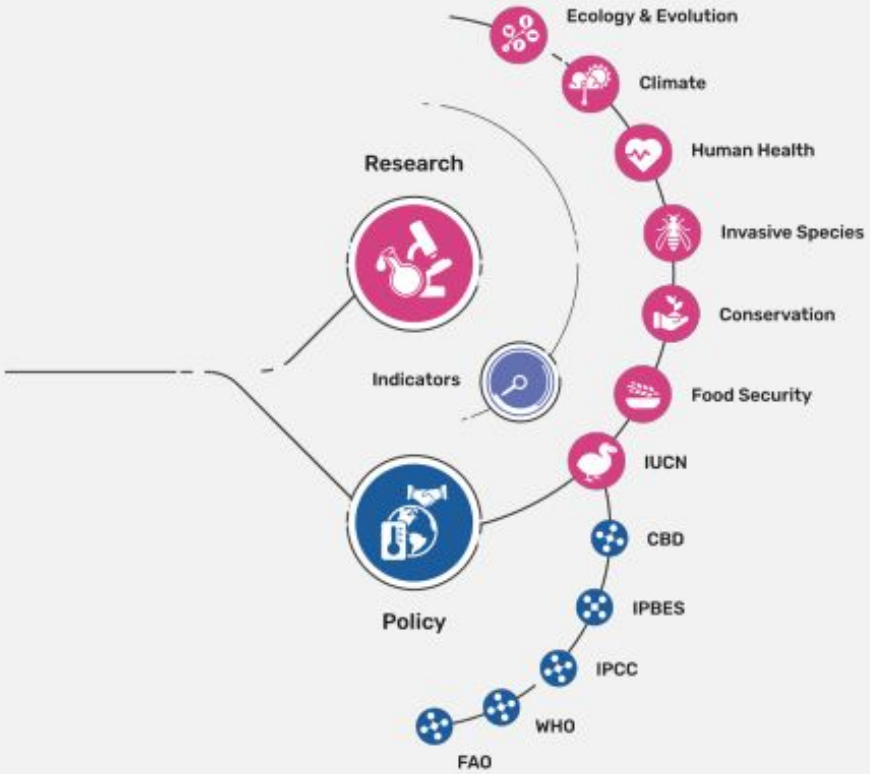
Why use GBIF data?

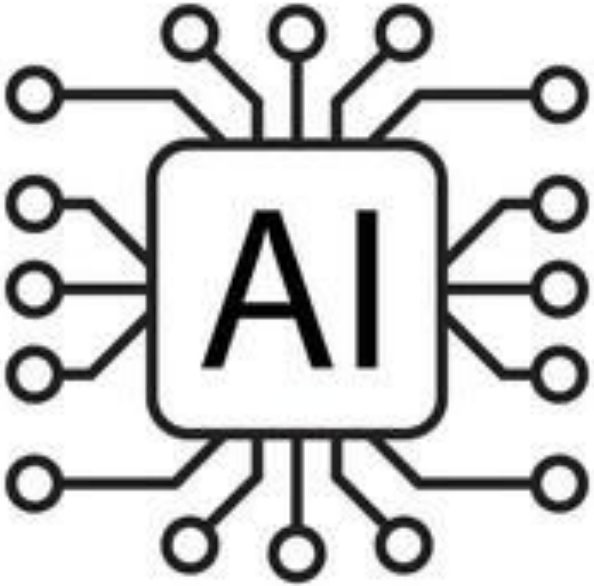
Anders G. Finstad
NTNU

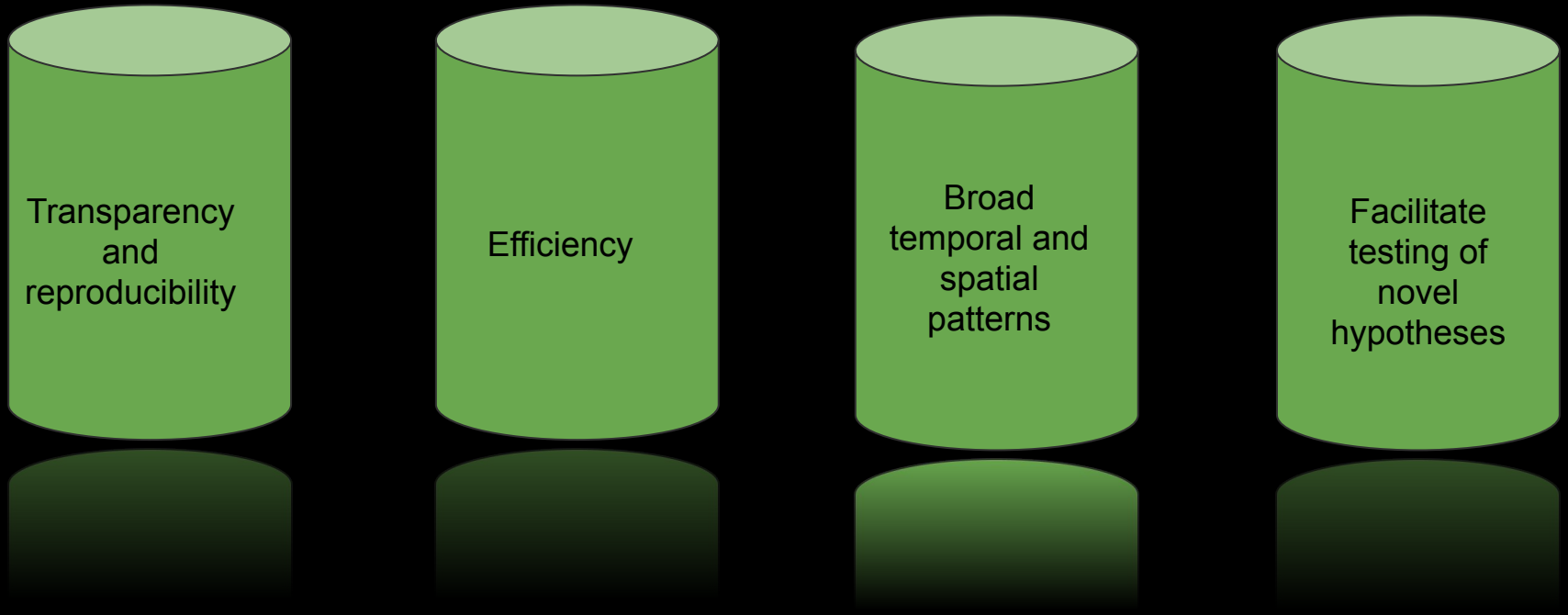
Uses of biodiversity evidence

Transform

Apply and use data




AI

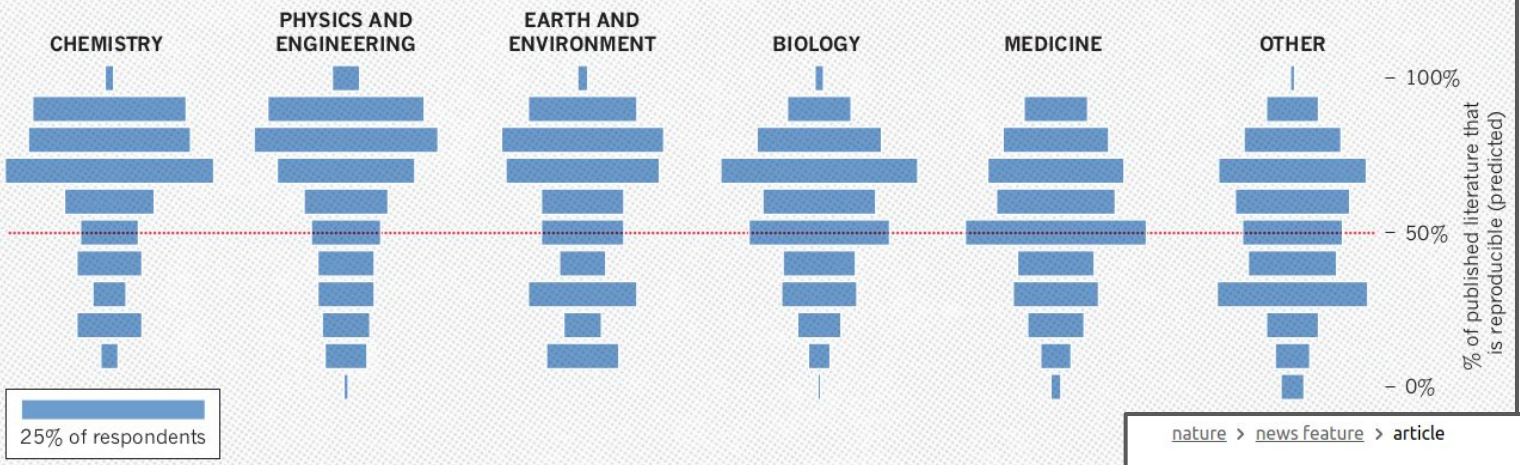

Transparency and reproducibility
Efficiency
Broad temporal and spatial patterns
Facilitate testing of novel hypotheses

A 'CRISIS' IN NUMBERS

Nature surveyed 1,576 scientists online to get their thoughts on reproducibility in their field and in science in general. See **go.nature.com/2vjr4y** for more charts and access to the full data.

HOW MUCH PUBLISHED WORK IN YOUR FIELD IS REPRODUCIBLE?

Physicists and chemists were most confident in the literature.

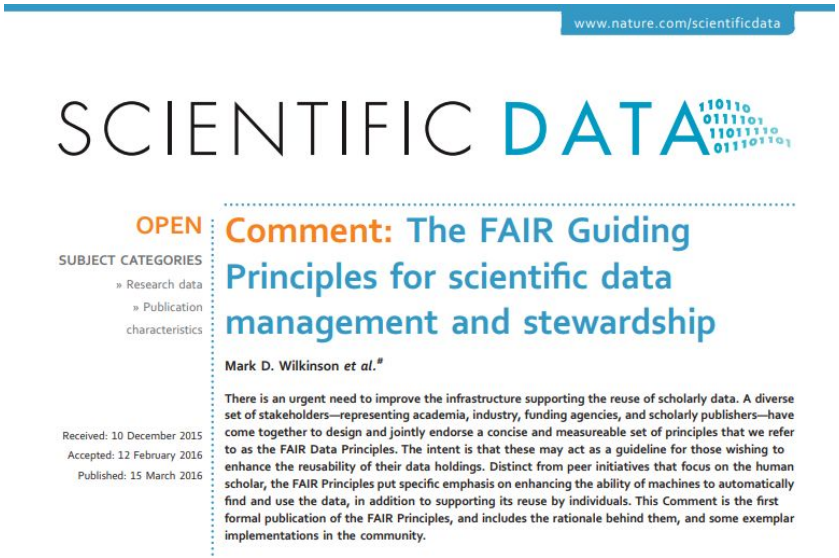


News Feature | Published: 25 May 2016

1,500 scientists lift the lid on reproducibi

Monya Baker

Nature **533**, 452–454 (2016) | Cite this article



www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

» Research data

» Publication characteristics

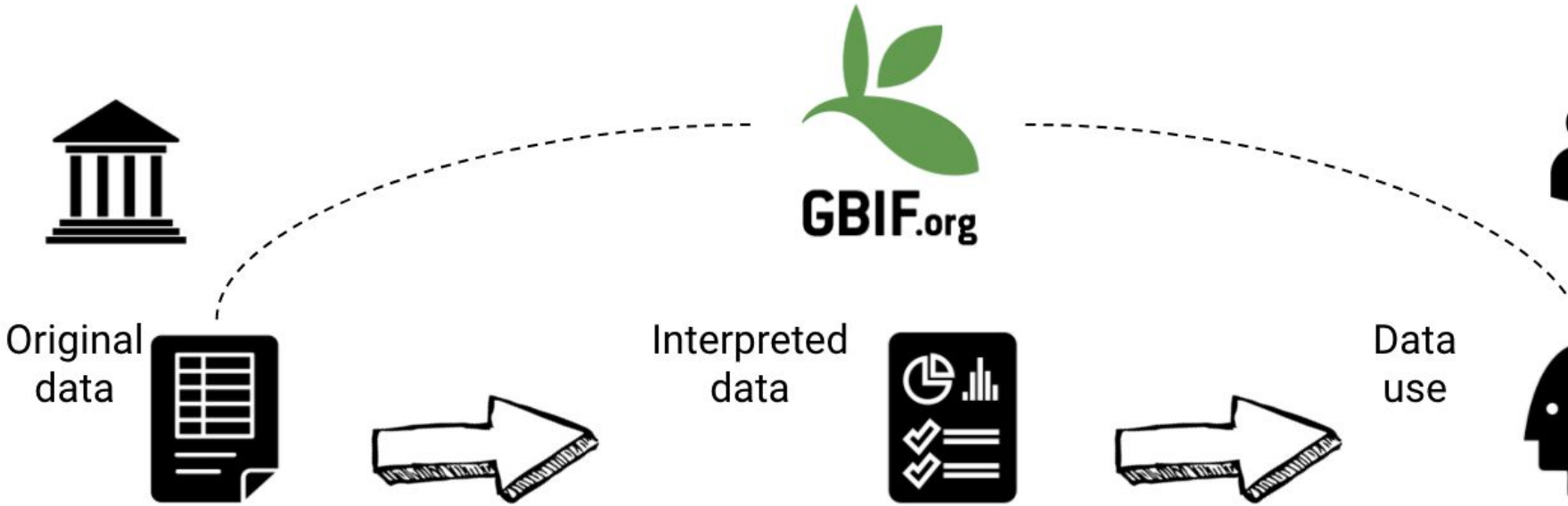
Received: 10 December 2015
Accepted: 12 February 2016
Published: 15 March 2016

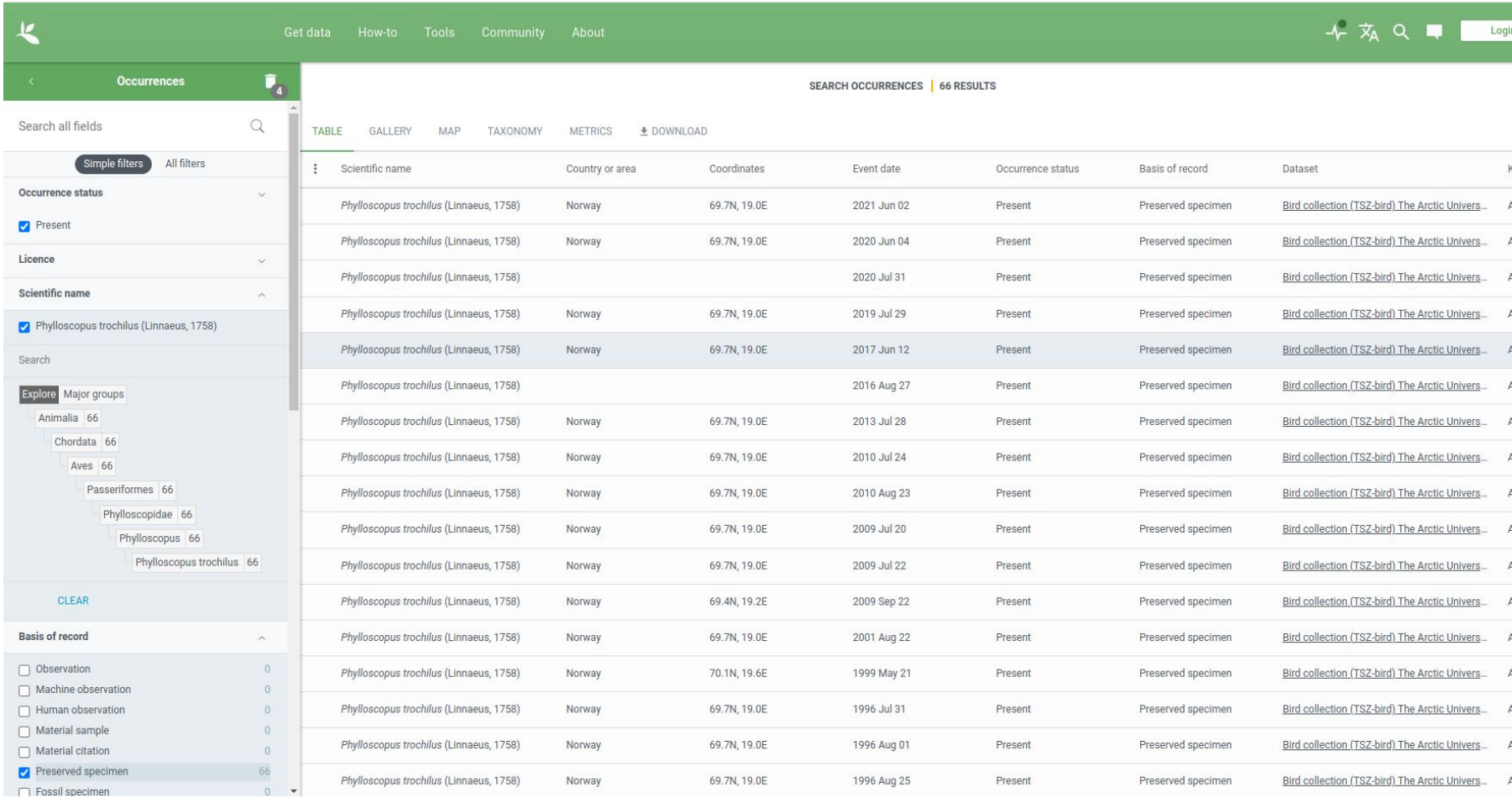
Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*#

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

- **F**indable
  - Rich metadata
  - Registered and indexed in searchable resource
- **A**ccessible
  - Data and metadata are retrievable using free and open protocols
  - Allows for authentication and authorization when needed
- **I**nteroperable
  - Shared in broadly accessible digital language
  - Use vocabularies that follow FAIR principles
- **R**eusable
  - Clear and accessible data licence
  - Use relevant and dominant community standards for data


GBIF.org
Original data
Interpreted data
Data use


Get data
How-to
Tools
Community
About
Occurrences
Search all fields
Simple filters
All filters
Occurrence status
Present
Licence
Scientific name
Phylloscopus trochilus (Linnaeus, 1758)
Search
Explore
Major groups
Animalia 66
Chordata 66
Aves 66
Passeriformes 66
Phylloscopidae 66
Phylloscopus 66
Phylloscopus trochilus 66
CLEAR
Basis of record
Observation 0
Machine observation 0
Human observation 0
Material sample 0
Material citation 0
Preserved specimen 66
Fossil specimen 0
SEARCH OCCURRENCES | 66 RESULTS
TABLE
GALLERY
MAP
TAXONOMY
METRICS
DOWNLOAD
Scientific name
Country or area
Coordinates
Event date
Occurrence status
Basis of record
Dataset
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2021 Jun 02 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2020 Jun 04 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) 2020 Jul 31 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2019 Jul 29 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2017 Jun 12 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) 2016 Aug 27 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2013 Jul 28 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2010 Jul 24 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2010 Aug 23 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2009 Jul 20 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2009 Jul 22 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.4N, 19.2E 2009 Sep 22 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 2001 Aug 22 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 70.1N, 19.6E 1999 May 21 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 1996 Jul 31 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 1996 Aug 01 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...
Phylloscopus trochilus (Linnaeus, 1758) Norway 69.7N, 19.0E 1996 Aug 25 Present Preserved specimen Bird collection (TSZ-bird) The Arctic Univers...

https://api.gbif.org/v1/species/match?name=Passer%20domesticus

You should see a **wall of text** known as JSON.

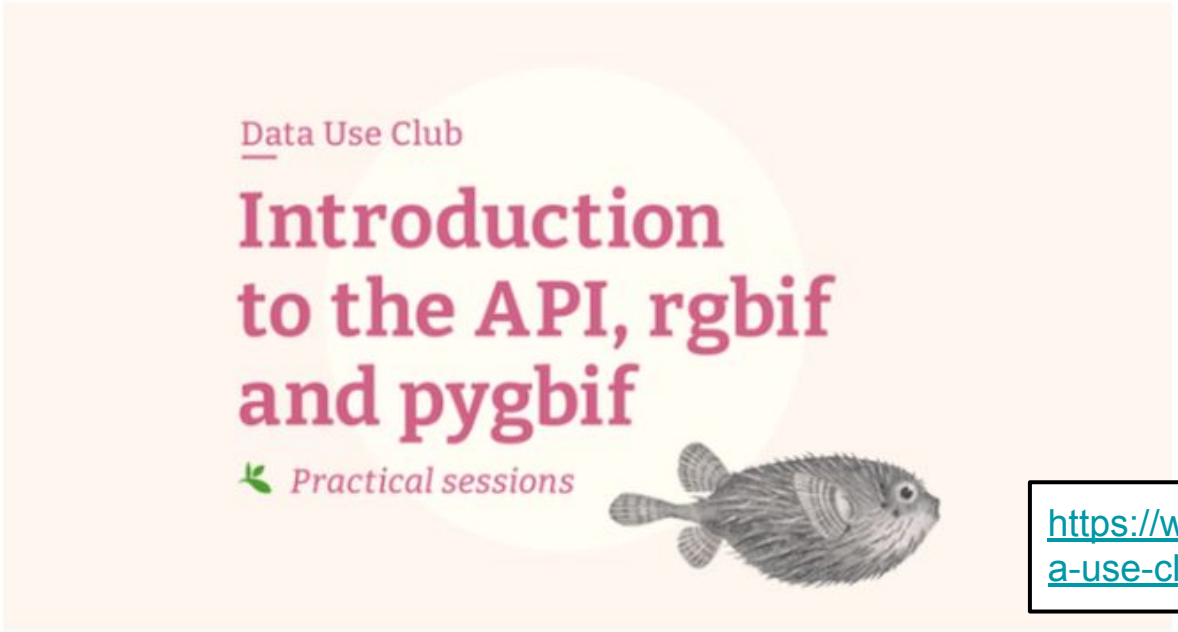
Code


```
{
  "usageKey": 5231190,
  "scientificName": "Passer domesticus (Linnaeus, 1758)",
  "canonicalName": "Passer domesticus",
  "rank": "SPECIES",
  "status": "ACCEPTED",
  "confidence": 98,
  "matchType": "EXACT",
  "kingdom": "Animalia",
  "phylum": "Chordata",
  "order": "Passeriformes",
  "family": "Passeridae",
  "genus": "Passer",
  "species": "Passer domesticus",
  "kingdomKey": 1,
  "phylumKey": 44,
  "classKey": 212,
  "orderKey": 729,
  "familyKey": 5264,
  "genusKey": 2492321,
  "speciesKey": 5231190,
  "synonym": false,
  "class": "Aves"
}
```

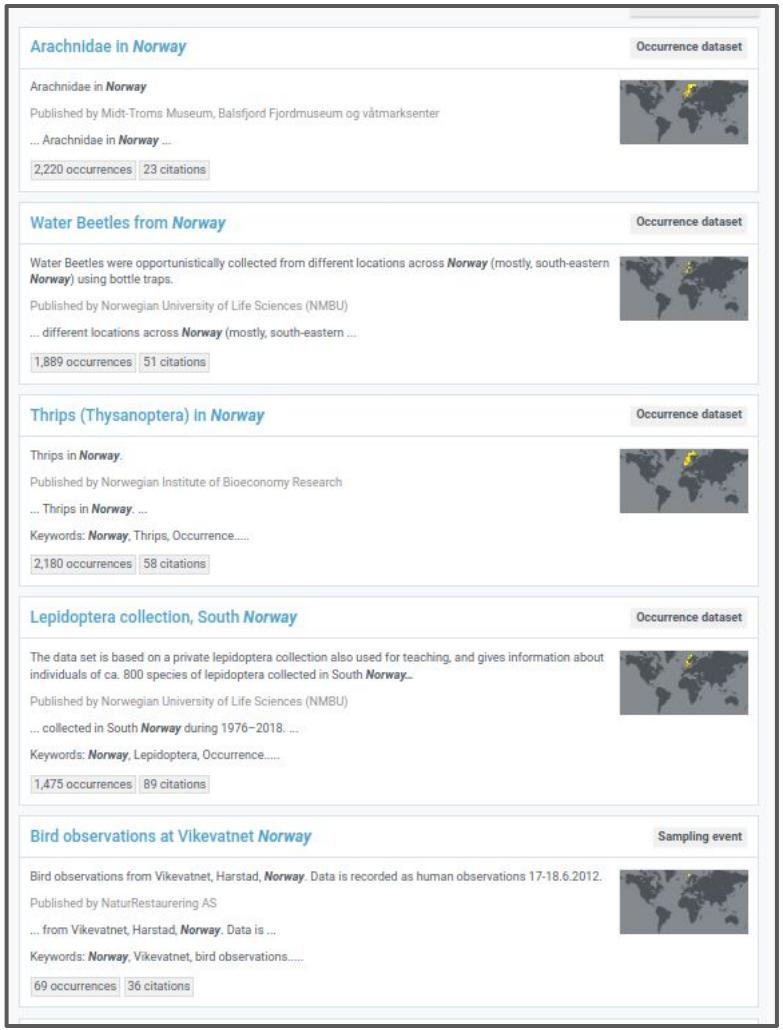

Data use club, API, rgbif, pygbif

The GBIF Data Use Club Practical Sessions: The API and introduction to rgbif and pygbif. Online on 8th February 2023 (recording is available).

Time and place: Feb. 8, 2023 3:00 PM – 4:00 PM, online

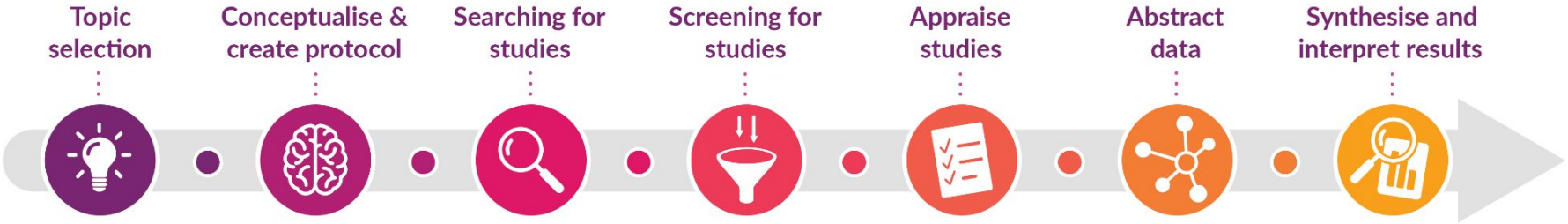


https://www.gbif.no/events/2023/data-use-club-api-rgbif-pygbif.html



Remember: GBIF is not a dataset - it’s an data aggregator

Steps in a systematic review



Example use case: Biodiversity hotspots of Norway

https://github.com/gjearevoll/BioDivMapping/

Modelling of comprehensive distribution maps for species: Background, concept and workflow

- Operationalization of SDMs and community models at management relevant scales - where is threatened biodiversity located?
- Financed by the Norwegian Environment Agency (2023-24)
- **We need models that are reproducible and verifiable - this require open and FAIR data**
- **Challenge:** Quite a lot of data on Norwegian biodiversity from public or private impact assessments etc remain unpublished or published in a non accessible form

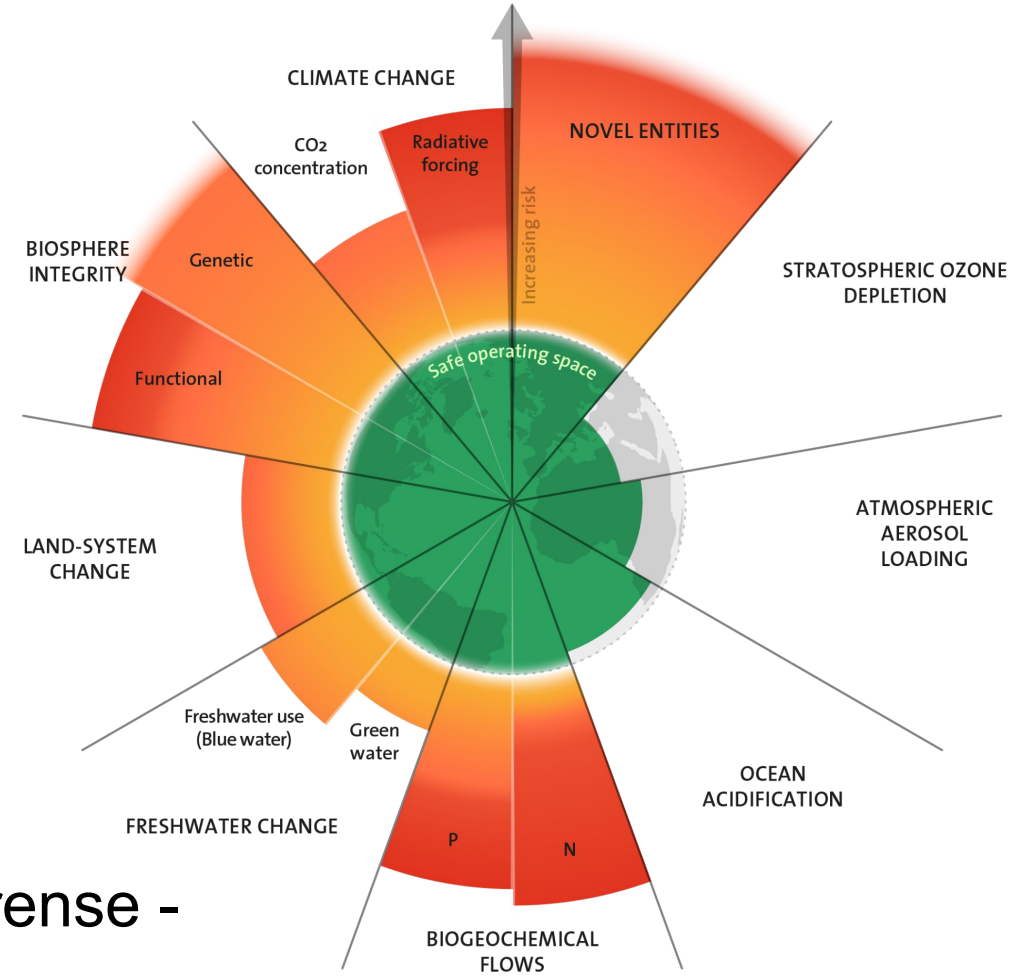
Finstad AG, Herfindal I, Perrin S, O'Hara B, Chipperfield T, Töpper J. (2023). Modellering av heildekkande utbreiingskart for arter: Bakgrunn, konsept og arbeidsflyt. NTNU Gjærevollsenteret, rapport 1/2023. Norges Teknisk Naturvitenskapelige Universitet

GJÆREVOLLSENTERET
Framtidsanalyser av naturmangfold

NTNU

NINA
Norsk institutt for naturforskning

UNIVERSITAS BERGENSIS

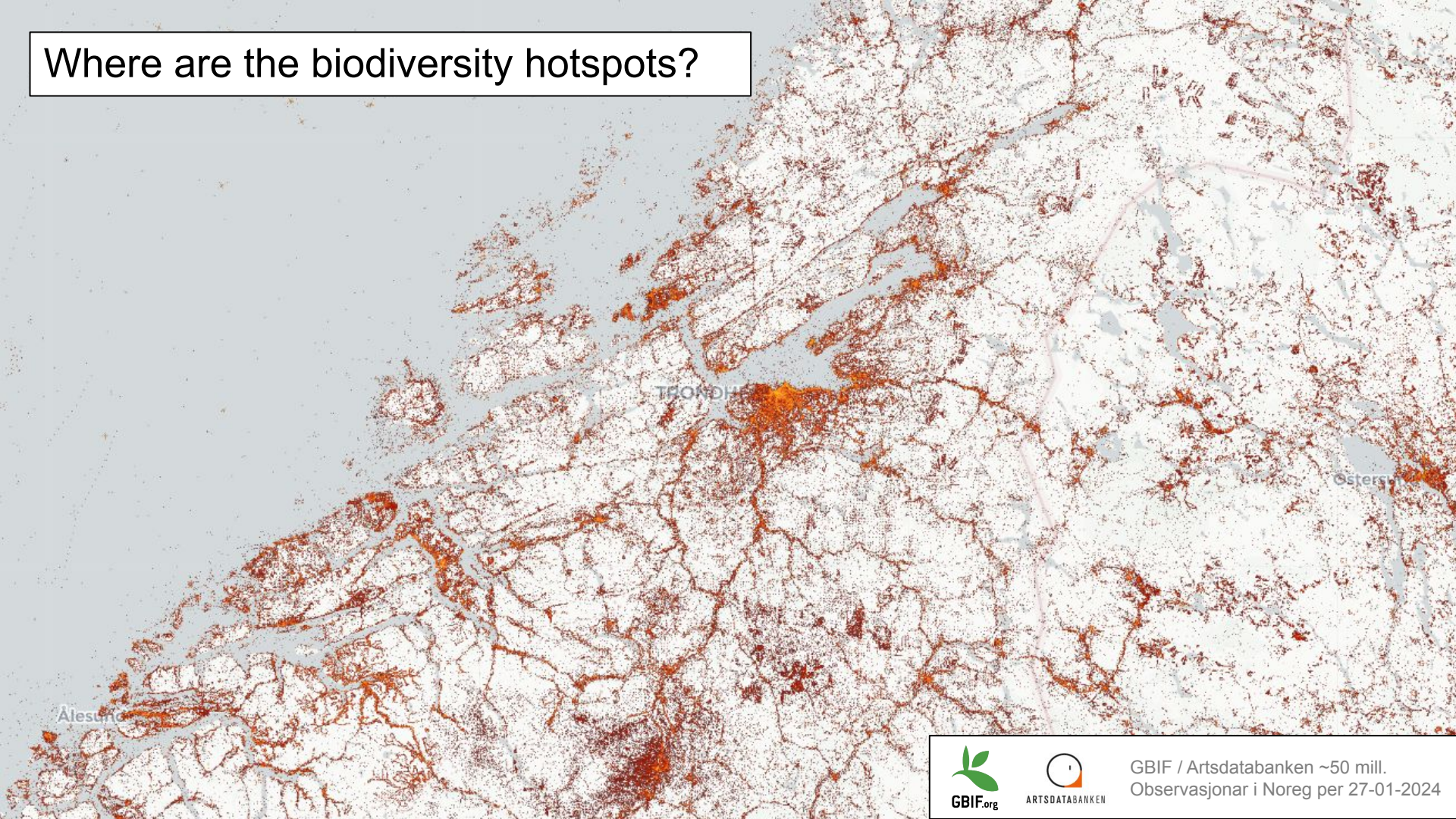


Vi trør over naturens tolegrense - utan at vi veit kor den er

2020 UN BIODIVERSITY CONFERENCE
COP 15 - CP/MOP10 - NP/MOP4
Ecological Civilization-Building a Shared Future for All Life on Earth
KUNMING – MONTREAL

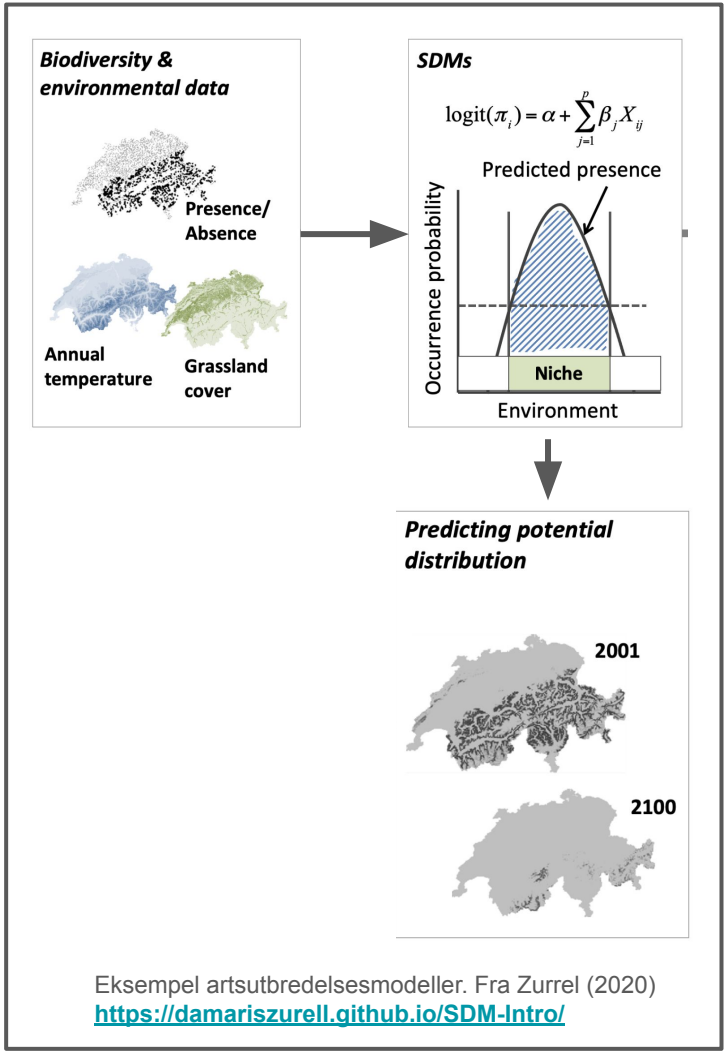
KUNMING MONTREAL GLOBAL BIODIVERSITY FRAMEWORK

Where are the biodiversity hotspots?



Comprehensive maps based on fragmented information must be based on models (in one form or another)

- **Environmental variables** can be (i) directly observed or (ii) modeled
- **Modeled environmental variables** can introduce "hidden" uncertainties
- **Skewness in data collection** leads to biases in results if not corrected for



Eksempel artsutbredelsesmodeller. Fra Zurrel (2020)
https://damariszurell.github.io/SDM-Intro/

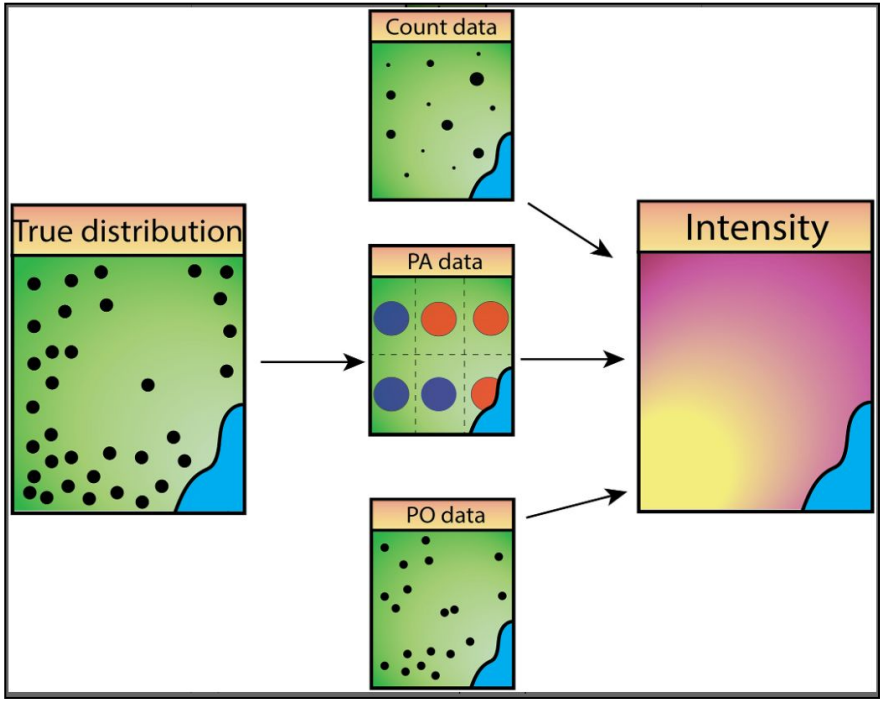
Comprehensive distribution

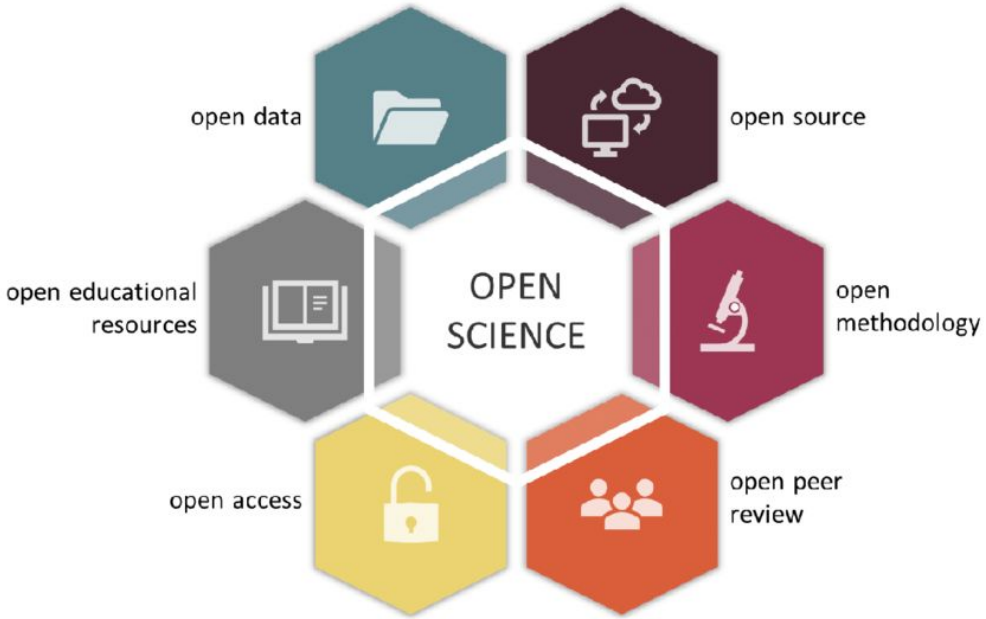
Methods in Ecology and Evolution
BRITISH ECOLOGICAL SOCIETY
APPLICATION | Open Access

PointedSDMs: An R package to help facilitate the construction of integrated species distribution models

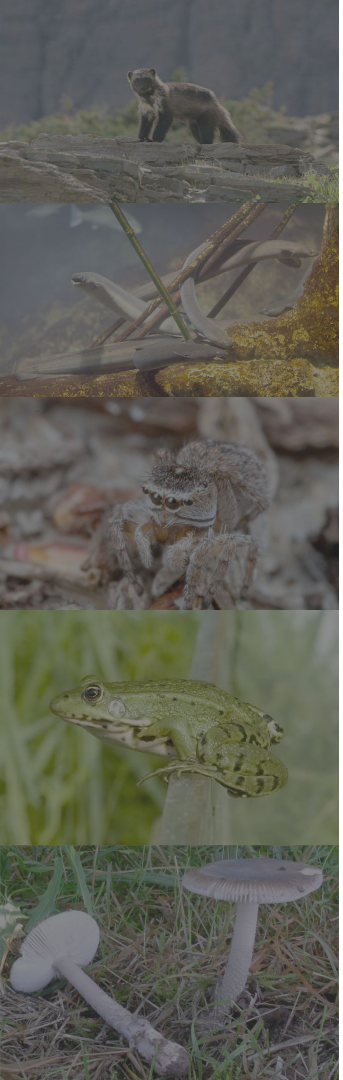
Philip S. Mostert | Robert B. O'Hara

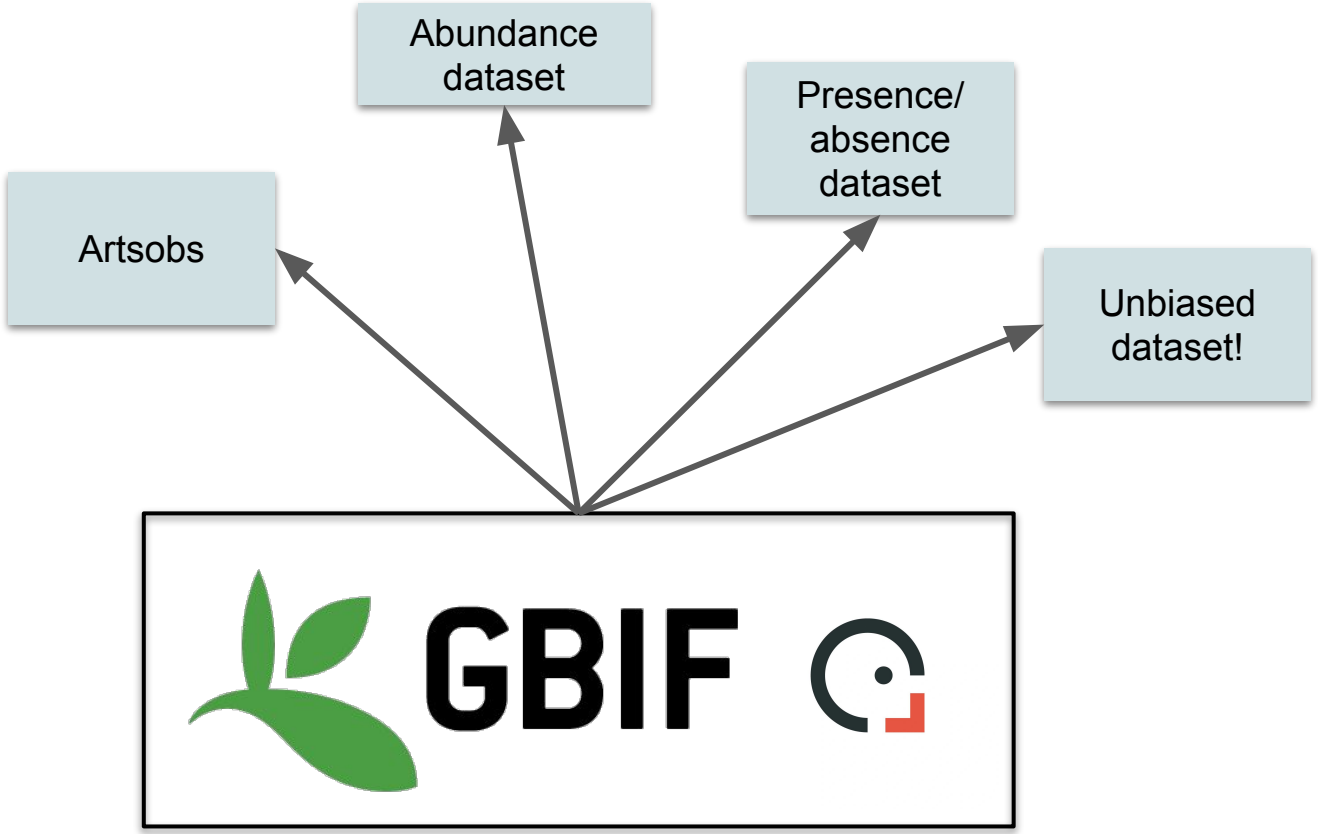
Must rely on models that estimate species distribution from a combination of opportunistic observations and observations that record absence of findings/quantities of organisms (effort data)



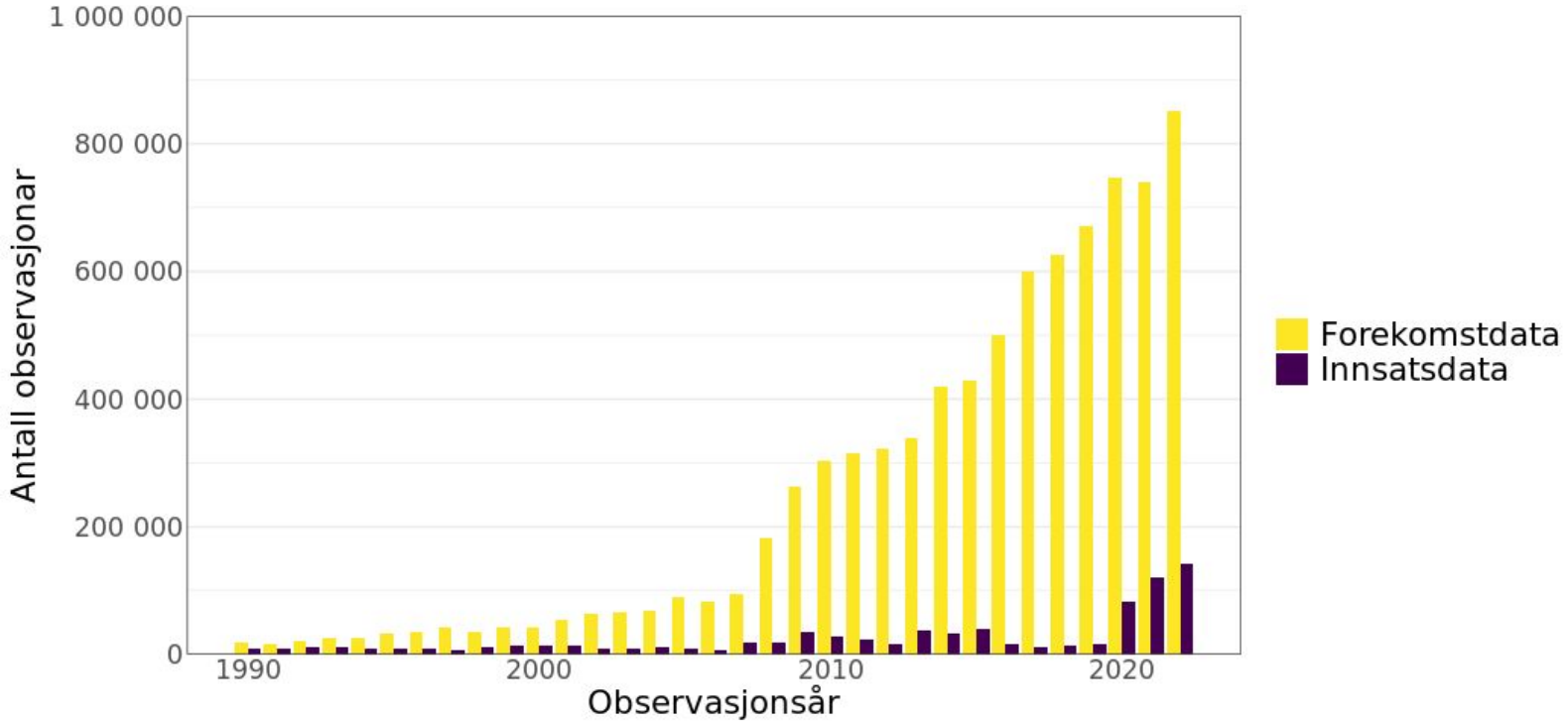


Modelling approaches puts extra demands on the reproducibility aspect





In practice it's only GBIF / Artskart that are available as large scale open source for information on Norwegian biodiversity



Large increase in number of occurrences - but mostly from opportunistically collected sources

https://www.gbif.org/data-use