Using open ecology data in research and education

John-Arvid Grytnes

Friday 7th February - open symposium on DiSSCo and GBIF

UNIVERSITETET I BERGEN

Ecological data

- Ecologists collect a large amount of species data
  - (often accompanied with a variety of environmental data)
- Data are variable
  - Often area-restricted (plots)
    - With variable size
  - Measured abundance
    - With variable units
      - Number of individuals
      - Cover (percent or an ordinal scale)
      - Tailormade solutions
  - Other issues
    - Variable authorities for taxonomy between data sets
    - Variable taxonomic resolution between and within data sets
    - Additional variables (environment, treatments, etc)

Obstacles for data sharing

- Collection of data is hard work
- Little gain
- Time and resources for data preparation
  - Ecological data sets are variable
  - Ecologists are lazy…
- Want to ‘use’ the data first
- Culture!

Little gain

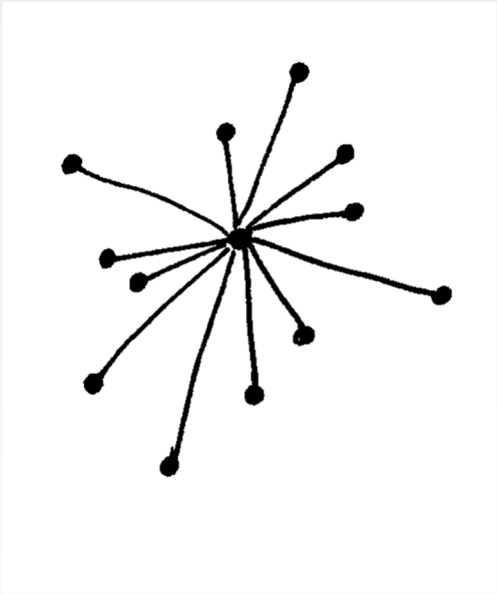
- Publish data papers
- Make data citable
- Make it countable (challenge)
  - Ask for this in evaluations

Time and resources for data preparation

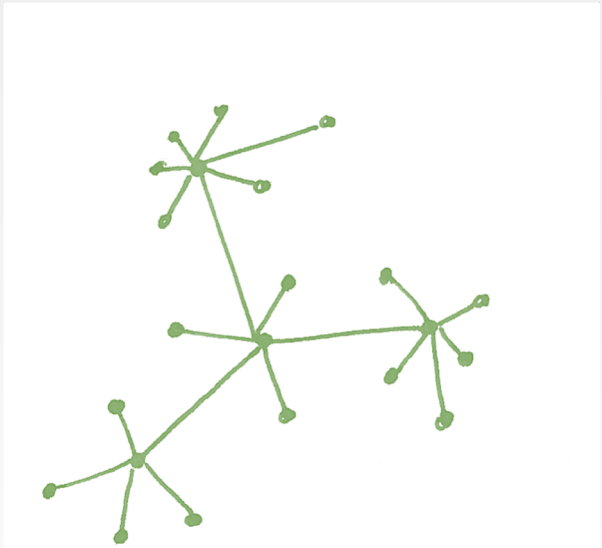
- Lower the threshold for submitting data
  - One-size-fits-all
    - Easy to get data in and out from the data custodian perspective
    - High work load for data provider
    - Variable work load for data user
- Put the work load on the data user
  - He/she has the motivation
- Many database solutions for this

How to share ecological data?

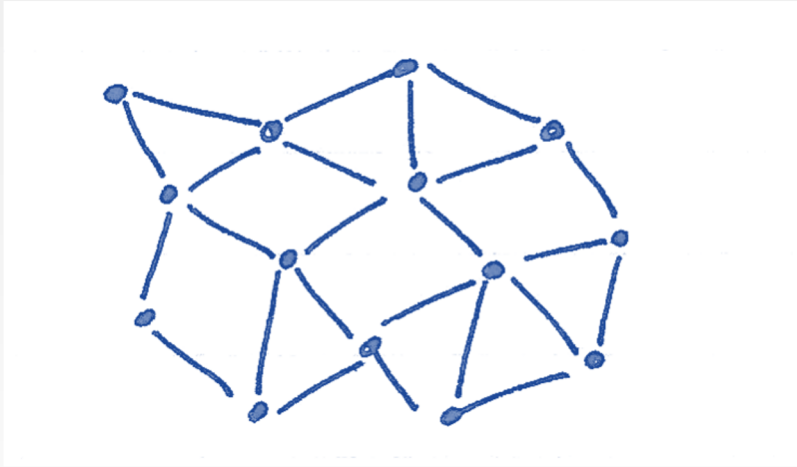
- Organization of the database



Centrally managed database
Data sent to a central point



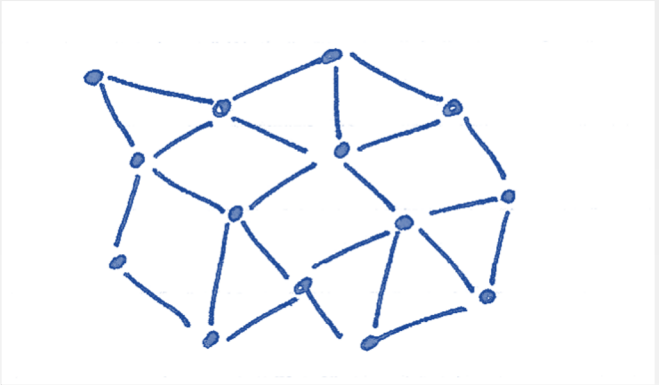
Database with several nodes
Data sent to a regional point



Data kept and managed locally
Smart contracts to run the sharing

Blockchain and smart contracts

- Data producer stays in control
- Data made available with little effort
- Building trust
  - accountability and transparency

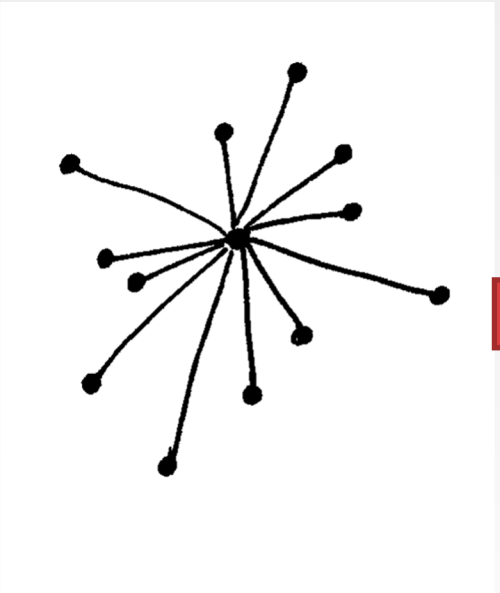


A blockchain is a peer-to-peer **distributed ledger** forged by **consensus**, combined with a system for "**smart contracts**" and other assistive technologies used to build a new generation of transactional applications that establishes **trust**, **accountability** and **transparency** at their core.

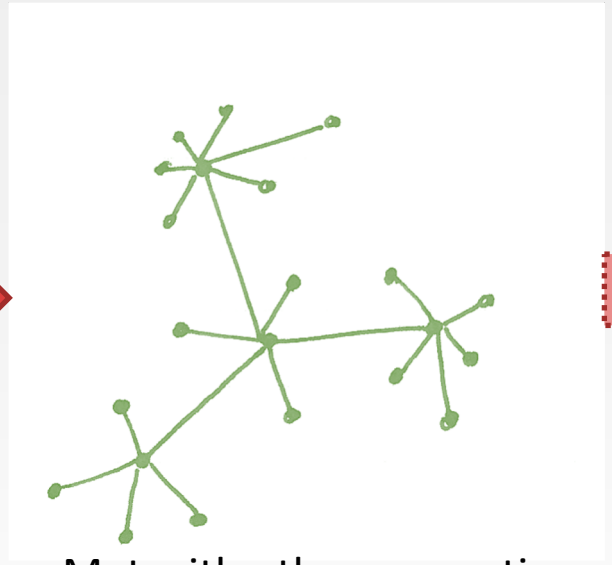
A smart contract is piece of computer code ("A what If / Else Command") that stores **rules** for negotiating the terms of a **contract**, automatically verifies the contract and then **executes** the agreed terms.

Example from an ecological consortium

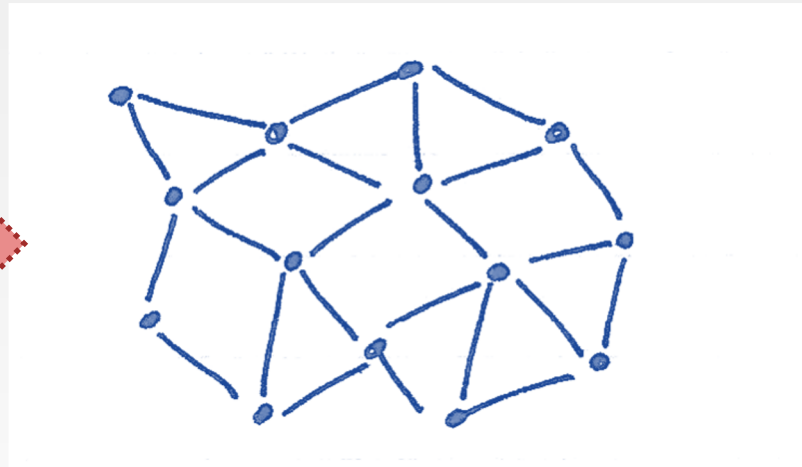
- Consortium on resampled data



Started off as one of the points
Agreed with other nodes that we trust to build a common data base



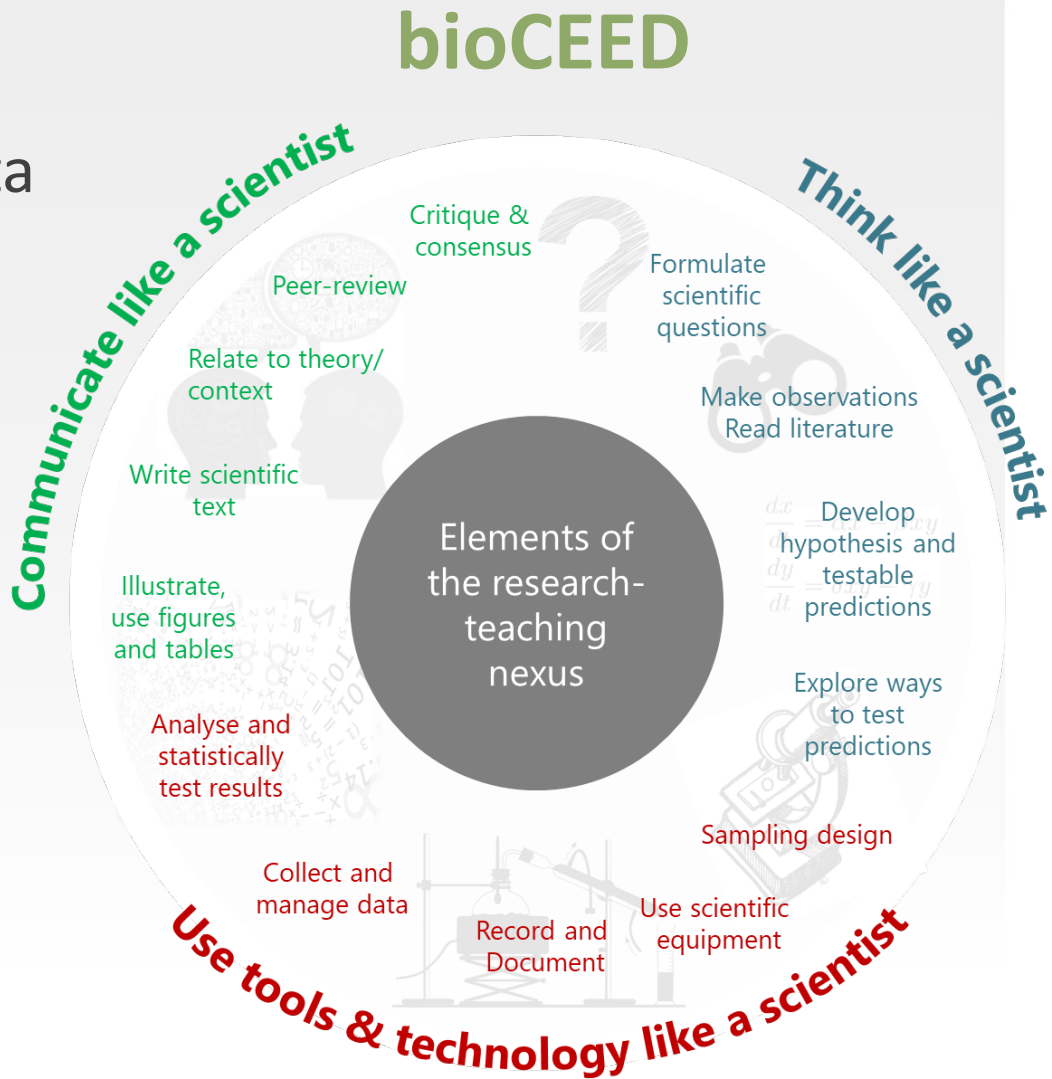
Met with other consortia
Became one of the nodes



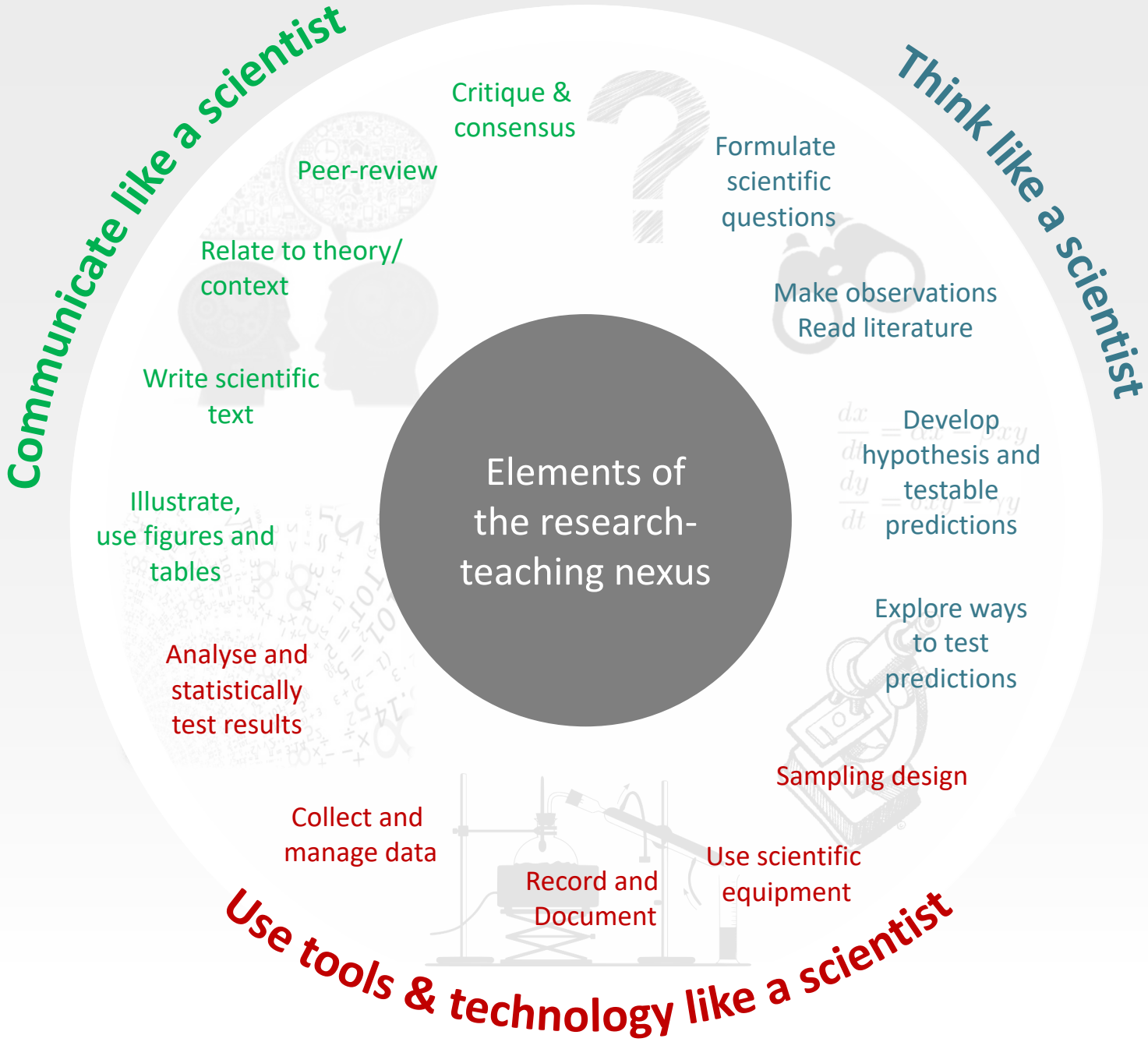
Aiming for making this

Culture

- Educating the new generation
- Get the students to see the value of shared data
  - Use real data in teaching
    - ArtsApp
    - Artskart/Artsobservasjoner
  - Produce real data in teaching
    - Quality
    - Focus on building knowledge
    - And interest for species



bioCEED


Communicate like a scientist
Think like a scientist
Use tools & technology like a scientist
Elements of the research-teaching nexus
Critique & consensus
Peer-review
Formulate scientific questions
Relate to theory/ context
Make observations
Read literature
Write scientific text
Develop hypothesis and testable predictions
Illustrate, use figures and tables
Explore ways to test predictions
Analyse and statistically test results
Sampling design
Collect and manage data
Use scientific equipment
Record and Document