How to implement Indigenous digital sovereignty in AI for biodiversity monitoring

COP16 A&R Flash Talk

Saturday October 26th, 2024

Magali de Bruyn

DSE


A

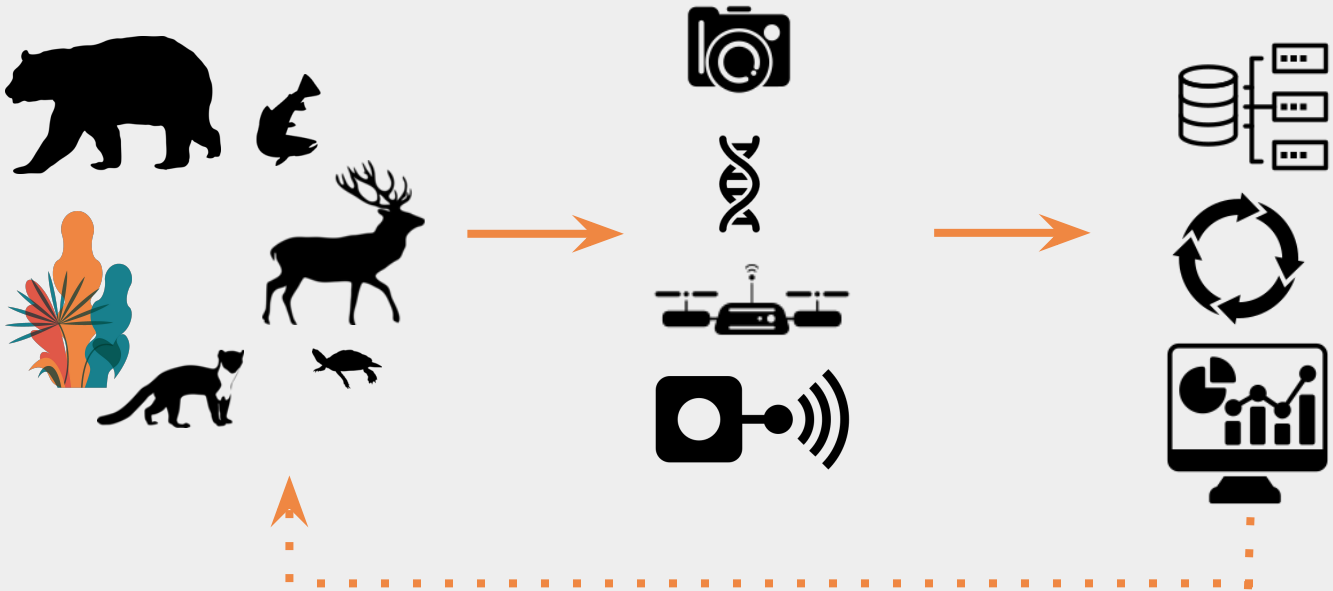

B

Real, physical things.

Related to land.

Related to people.

Biodiversity Monitoring Networks



Biodiversity monitoring networks affect people.

Real, physical things.

Related to land.

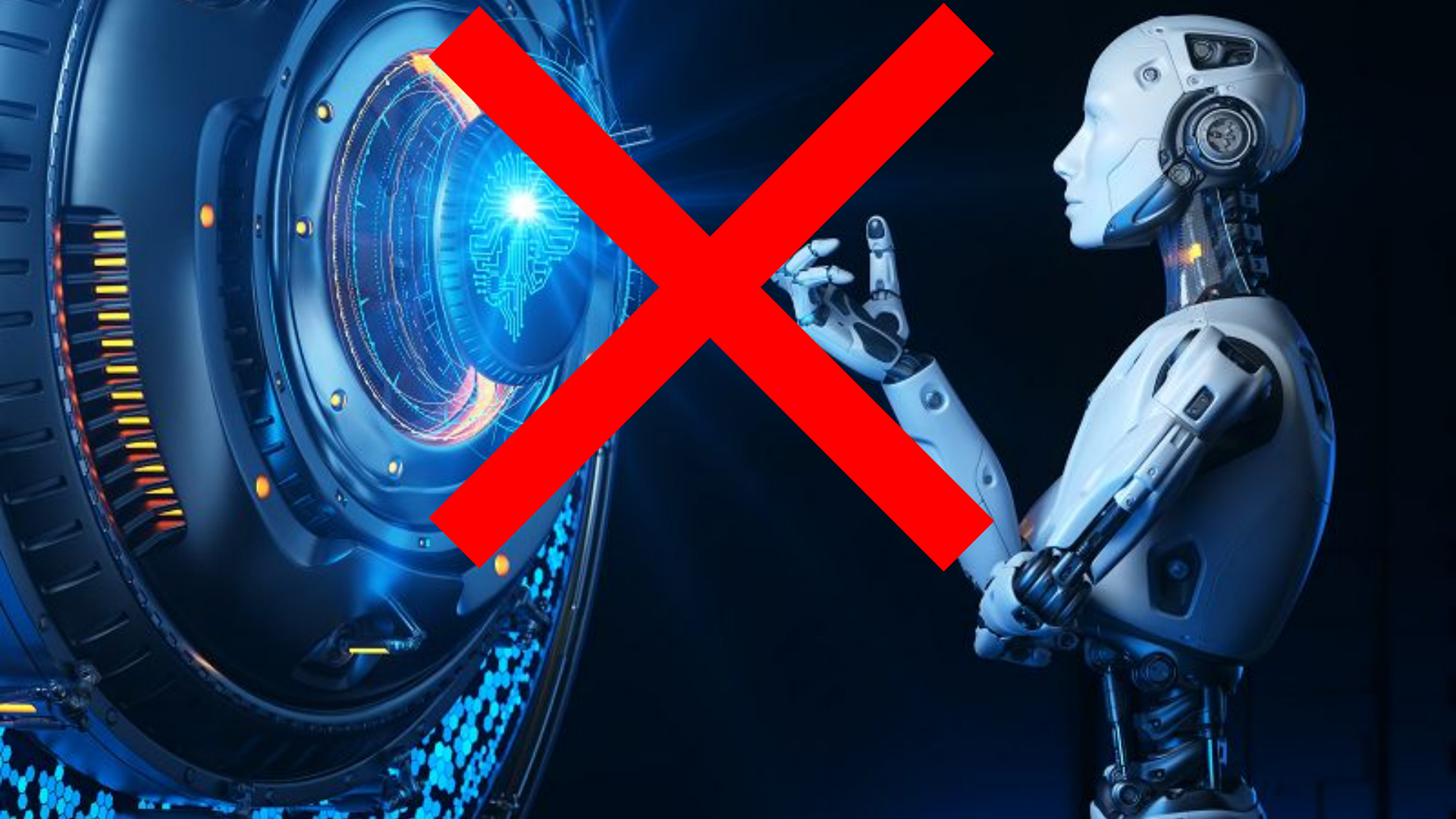
Related to people.

What the hell is AI?

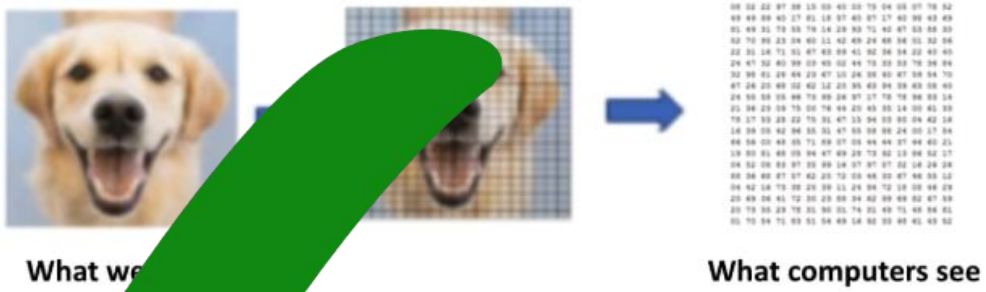
Imagine a world where there is only one word for transportation—only the collective noun “vehicle.” All vehicles from cars, buses, bikes, spacecraft are simply called “vehicles. **Conversations in this world are confusing.** There are furious debates about whether or not vehicles are environmentally friendly, even though no one realizes that one side of the debate is talking about bikes or trucks….Meanwhile, **fraudsters have capitalized on the fact that consumers don’t know** what to believe when it comes to vehicle technology, so **scams are rampant in the vehicle sector.**

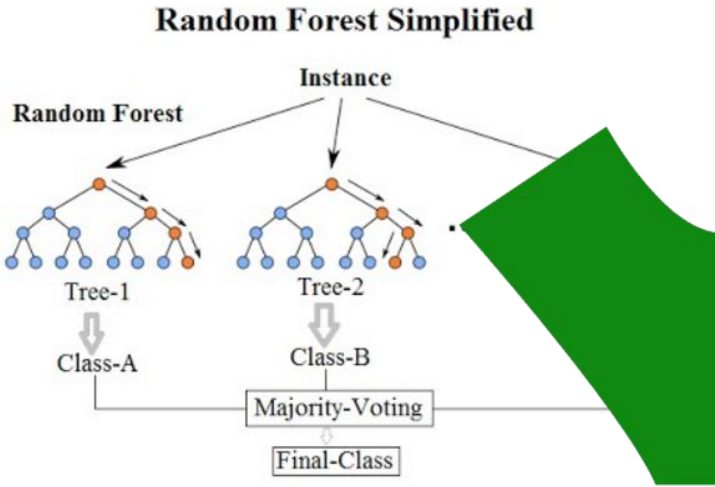
Now **replace the word “vehicle” with “artificial intelligence,”** and we have a pretty good description of the world we live in.

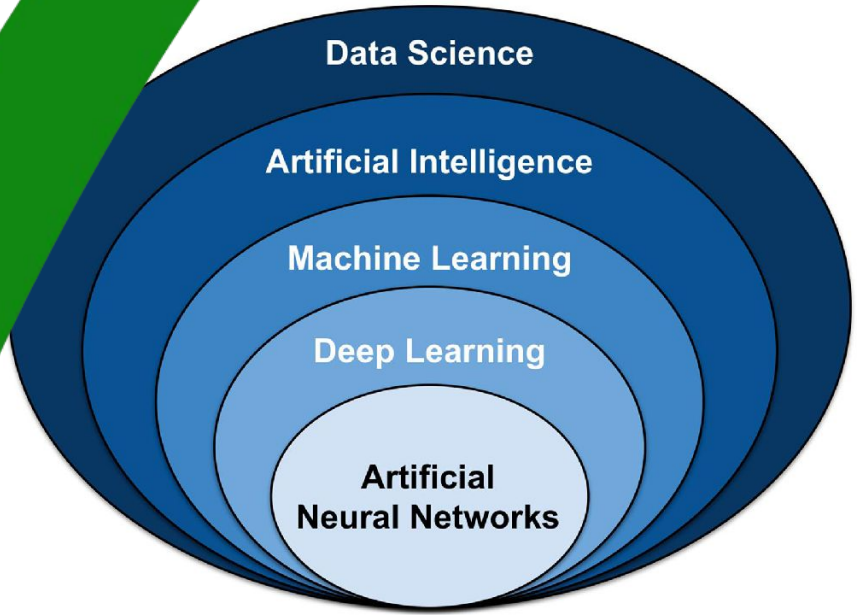
- AI Snake Oil. Arvind Narayanan (*modified version)*



$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$







How do you implement Indigenous digital sovereignty in AI for biodiversity monitoring?

1.

Work with the people on the land.

Credits: Karuk Tribe, Emilio Tripp, McKalee Stein et. al., and many others

Ask ourselves

Output

- Who is this for?
- Who does this benefit?
  - *Did they ask for it?*
- Who does this leave out?

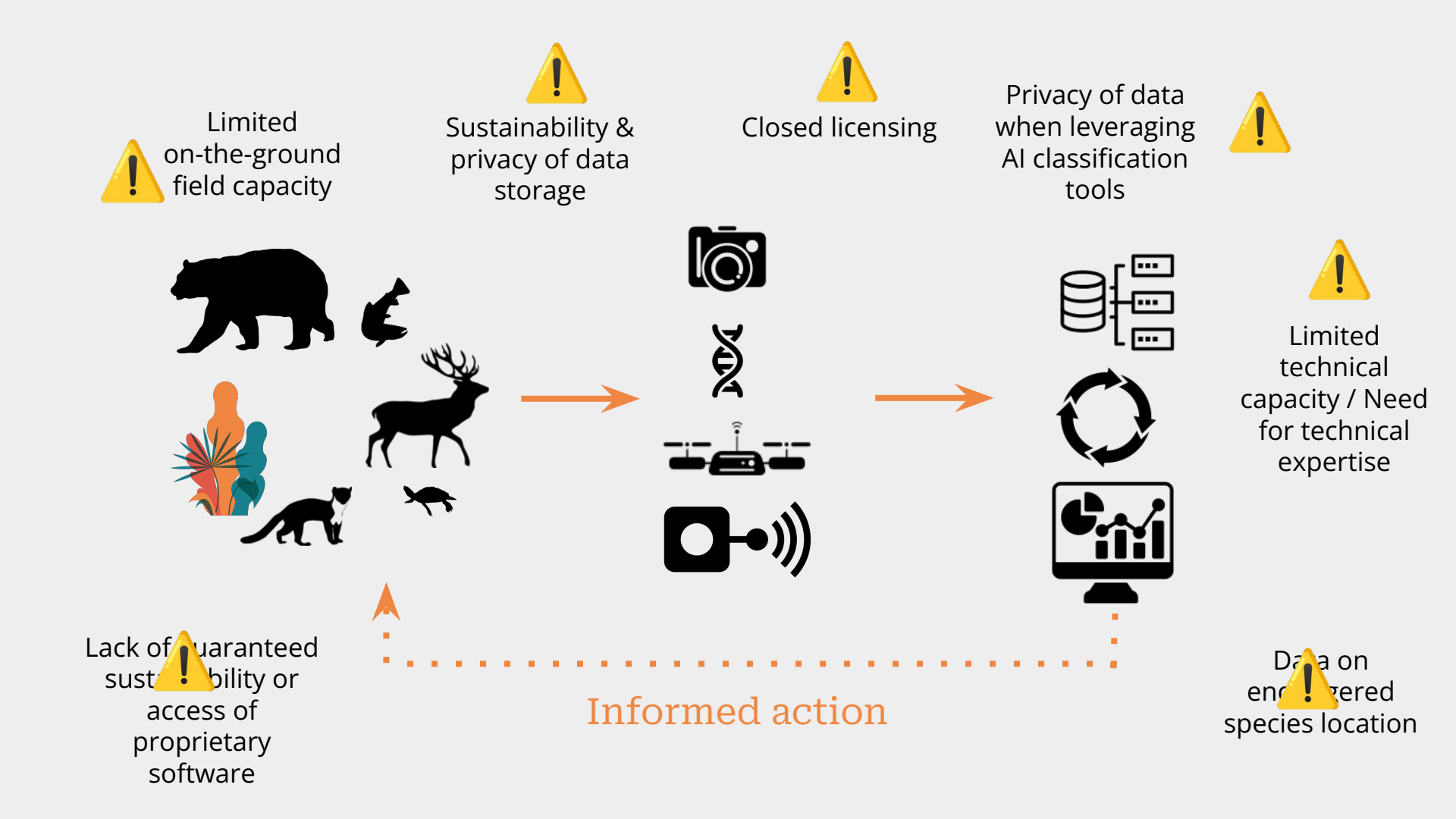
Input

- Who is excluded from this data (collection)?
- How might we incorporate those perspectives?
- Which labor was involved in this project? Which labor might be not recognized?

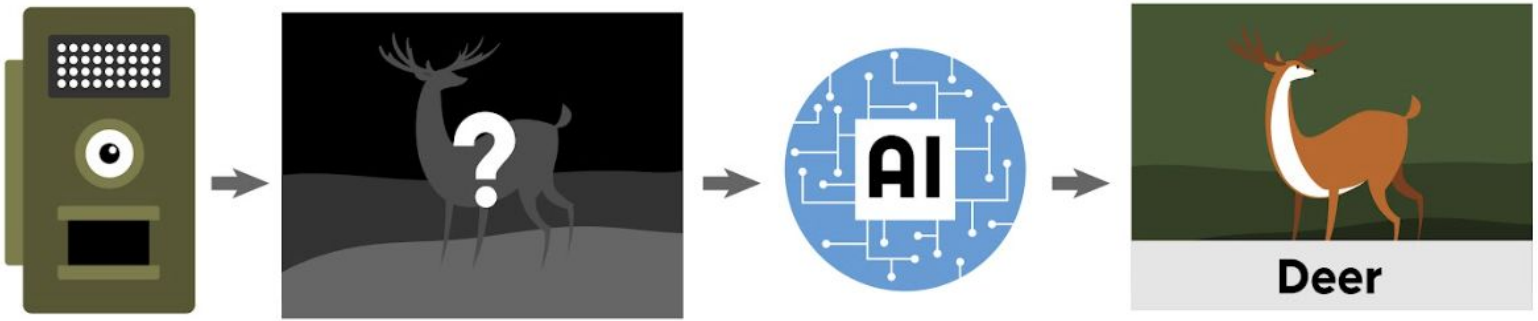
Credits: Catherine D'Ignazio & Lauren Klein

Credits: Karuk Tribe, Emilio Tripp, and many other Indigenous partners

- Listening
- On-the-ground visits
- Weekly meetings with the land inhabitants & stewards
- Talking to (tribal) elders, leaders
- Going to tribal council & community meetings
- Community workshops
- Written dissemination: community newsletter

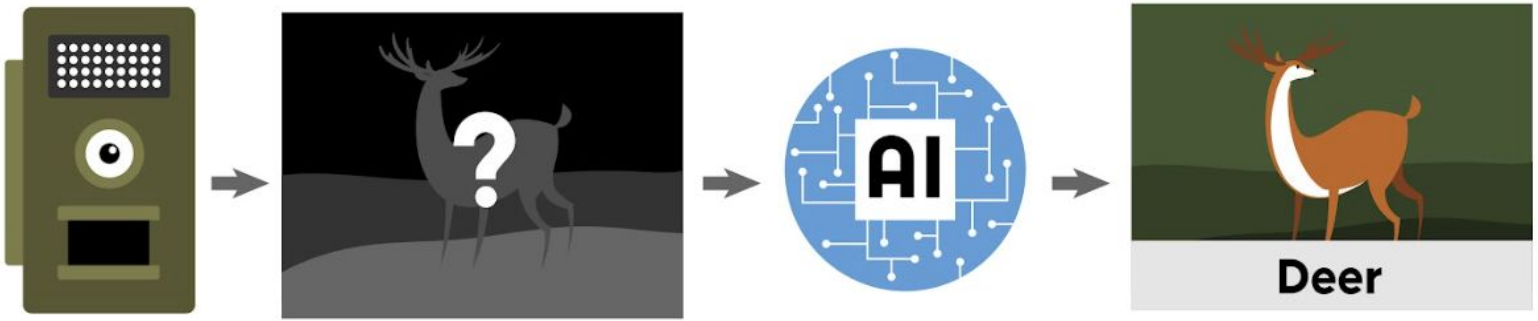

Limited on-the-ground field capacity
Sustainability & privacy of data storage
Closed licensing
Privacy of data when leveraging AI classification tools
Limited technical capacity / Need for technical expertise
Lack of guaranteed sustainability or access of proprietary software
Informed action
Data on endangered species location

For example: Species classification from audio and image data



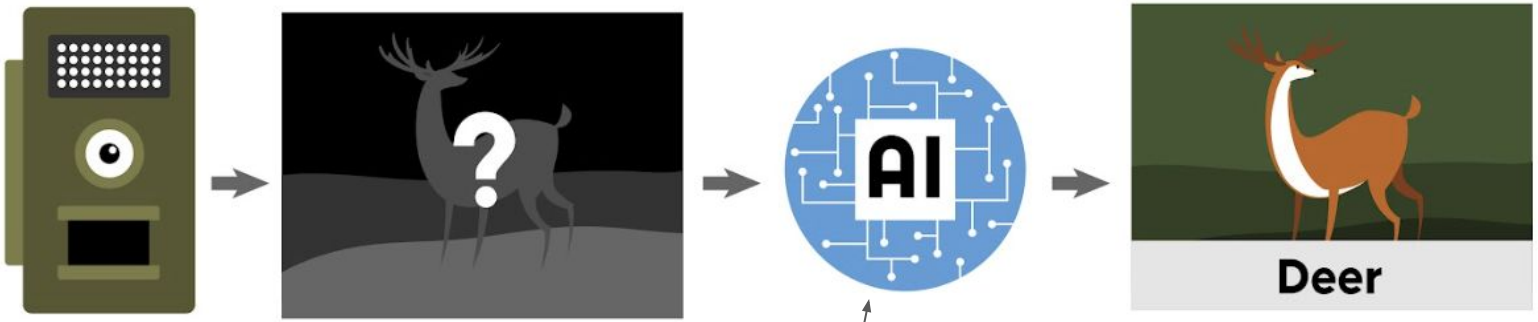
Credits: Karuk Tribe, Emilio Tripp

For example: Species classification from audio and image data



Benefit: reduce time and effort needed to label all of the data.
→ more bandwidth for other activities.

For example: Species classification from audio and image data



⚠️ **Ethical Consideration**: AI models are always imperfect. They will always make some level of mistakes. What is at stake when the model is wrong?

Wildlife Insights

← Stories

Wildlife Insights helps capture the beauty of biodiversity, as well as its fragility

March 2021 • Share

Featured technology

Motion-triggered cameras

Google Cloud

AI models

Who we're helping

Conservation experts

Local communities

Our role

Developed the Wildlife Insights tool by using Google Cloud AI Platform Predictions, a custom AI model and images from conservation partners Wildlife Conservation Society, Smithsonian Conservation Biology Institute, North Carolina Museum of Natural Sciences, WWF, Zoological Society of London, and Conservation International

License to Wildlife Insights

By agreeing to the Terms of Use, you agree to license the use of your data to Friends of the North Carolina Museum of Natural Sciences ("Friends"), on behalf of the Wildlife Insights partners, for the purposes described below. The license you grant to Friends will be royalty free (Friends will not have pay any royalties to you for the uses described above), non-exclusive (you retain the right to grant a license to your data to any other entity), worldwide (it will apply anywhere in the world), sub-licensable (Friends can sub-license your data to others as described in the Terms of Use) and perpetual (it does not expire).

All uses of your data described below will be with attribution to you and/or your Organization. Friends may:

1. Use your data to **develop or improve computer vision ("CV") models for the purpose of advancing technology related to conservation**. Friends may sub-license data to third parties, including Google, only for this purpose.

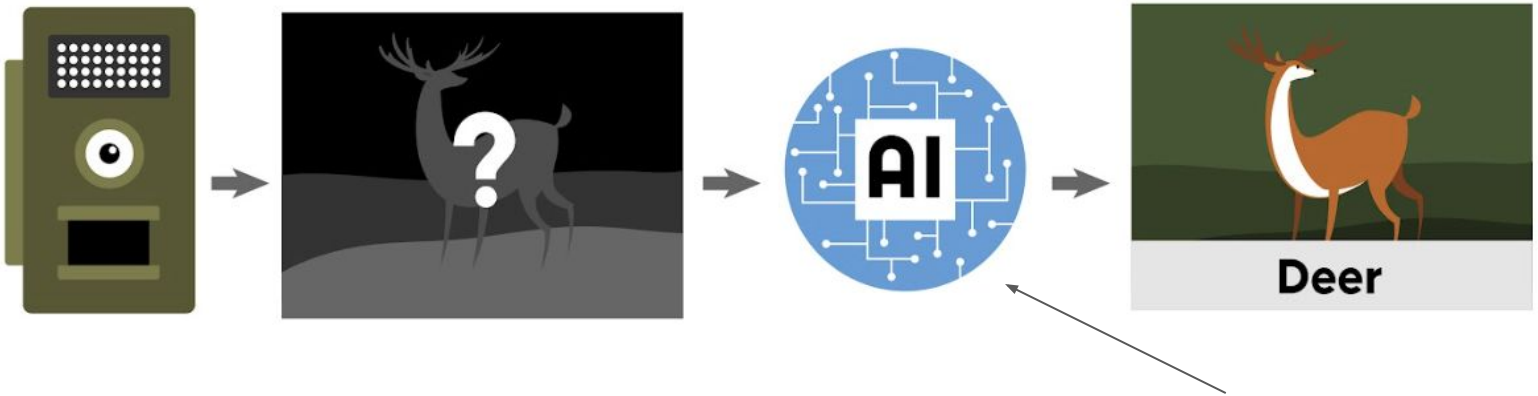
2. Use your data to **develop derived regional and global analytical products**. These products will not identify individual projects or locations, but will be aggregated at a suitable spatial scale (e.g. global, regional or national level). These products may be publicly available on the WI platform, used in online forums, in presentations and in peer-reviewed or other publications.

3. **Sub-license your data to WI Core Members to develop derived regional and global analytical products** as described in section 2 above.

4. Use parts of your data **for WI publicity materials or on WI's social media accounts.**

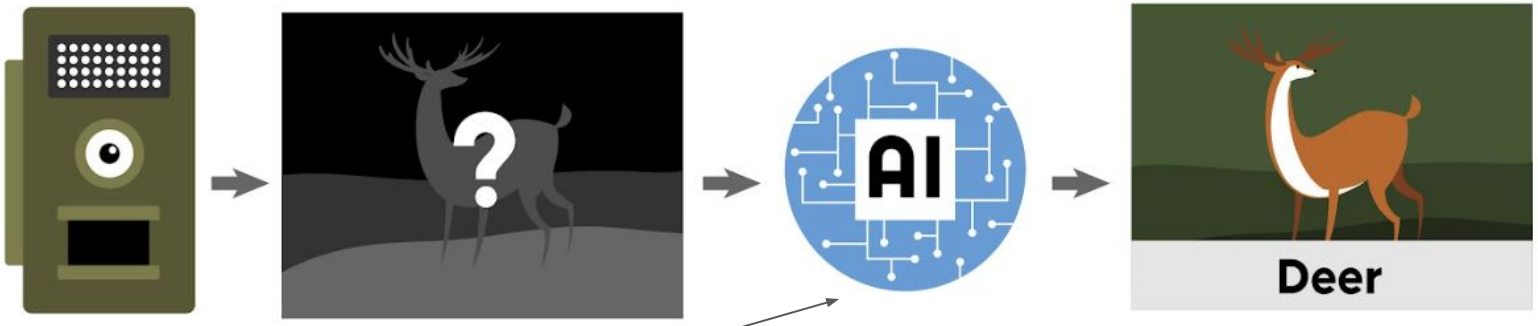
Wildlife Insights

For example: Species classification from audio and image data



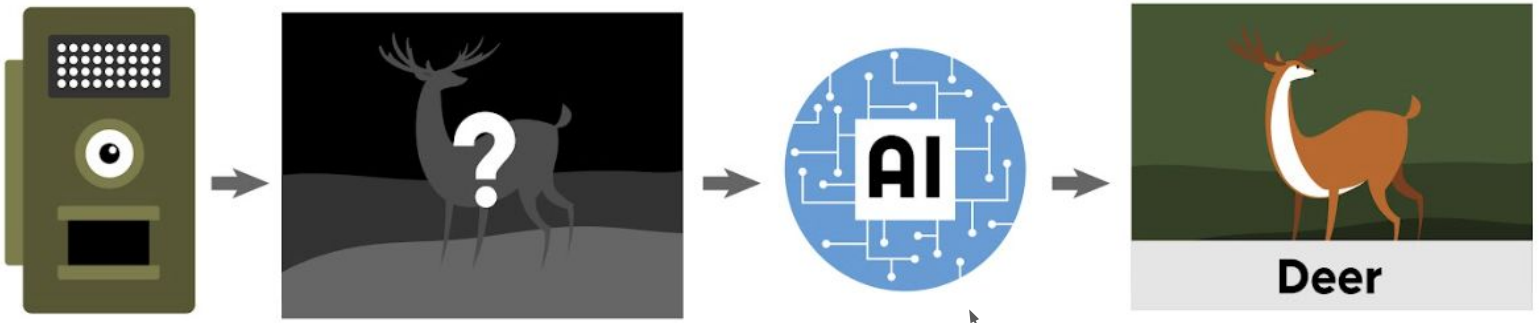
⚠️ **Ethical Consideration**: Who owns the model? Image classifiers for wildlife data are often trained on large open datasets, but if the model is further refined by the Tribe, where does ownership lie?

For example: Species classification from audio and image data



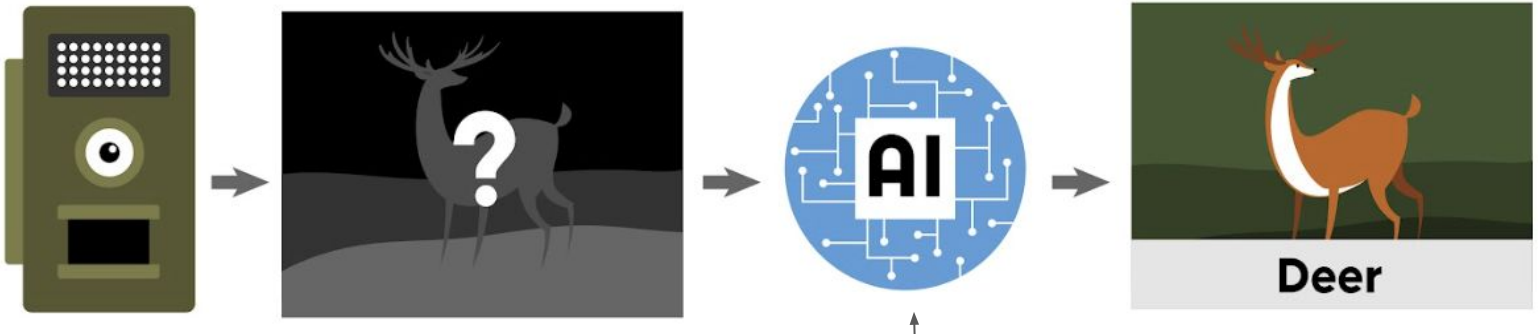
⚠️ **Ethical Consideration**: Training an AI model requires lots of data. Collecting this data is often labor intensive. Who is collecting this data? How are they being compensated?

For example: Species classification from audio and image data



⚠️ **Ethical Consideration**: Considering IDS, how is the workflow for using the AI implemented? Locally? In the cloud?

For example: Species classification from audio and image data



⚠️ **Ethical Consideration**: Human out of the loop: Going through the images can be a joyful aspect of this work which connects people to the ecosystem - this could be lost.

Indigenous Data Sovereignty

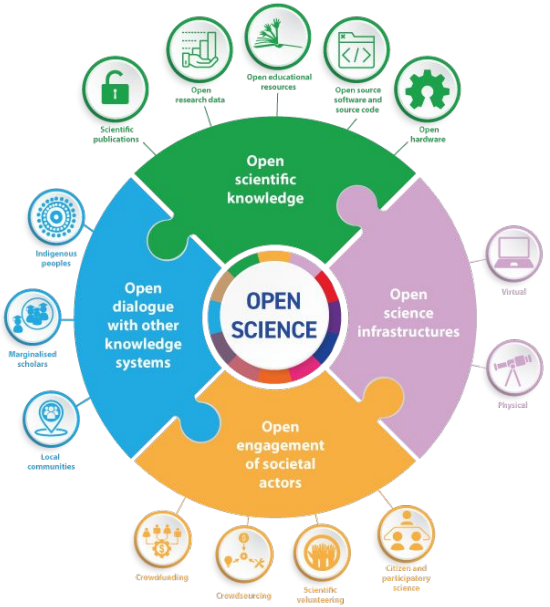
Indigenous communities require *full ownership* over *every part* of the data lifecycle for data that is generated by or pertains to them, their lands, and their resources.

2.

Operationalize CARE & FAIR principles.

Credits: McKalee Stein et. al (2024), Carroll et al. (2020), Wilkinson et al. (2016), and many others

- “Research should support community-led initiatives and secure funding for long-term investments in community”
- Disaggregate data
- Integrate TEK
- Use Indigenous language



Local Contexts

TK

TK Secret/Sacred (TK SS)
TK Women Restricted (TK WR)
TK Community Use Only (TK CO)
TK Attribution (TK A)
TK Men Restricted (TK MR)
TK Commercial (TK C)

Credits: Local Contexts, Wilkinson et al.

3.

Acknowledge and credit Indigenous Knowledge Systems

Credits: McKalee Stein et. al (2024), Carroll (2020), and many others

Some best practices

- ❏ Listen
- ❏ Ensure authorship
- ❏ Compensate community experts

4.

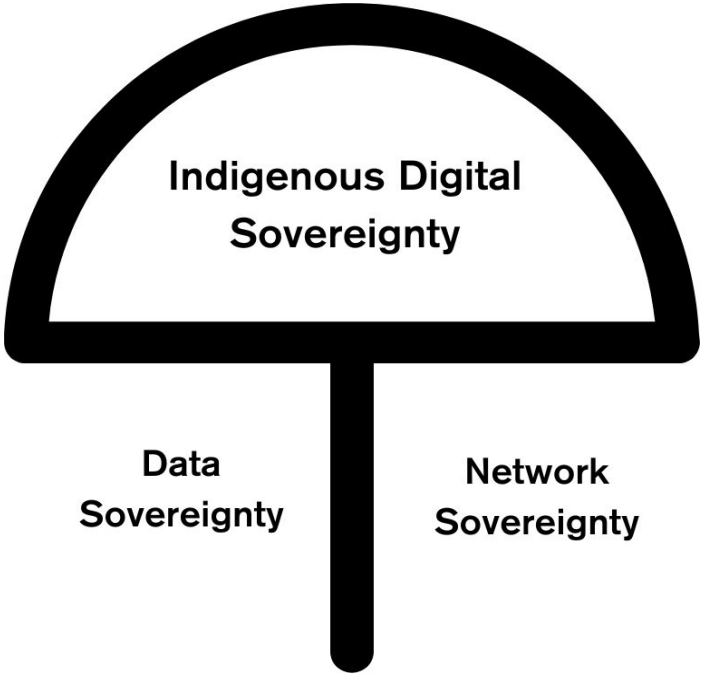
Continue engaging.

Make sustainability plan.

Credits: Karuk Tribe, Emilio Tripp, Dan Sarna, McKalee Stein et. al (2024), Carroll (2020), and many others

Indigenous digital sovereignty

- Ownership
- Privacy
- Infrastructure
- Accessibility
- Transparency



Credits: Traci Morris

Ultimately, it is up to the tribe to decide what they want to implement and how: our job is to engage with and support that process.

We need:

- More co-design of projects & tooling.
- More understanding & implementation of how to utilize TEK into technology and methodology while protecting privacy.
- Be genuine allies: Approach Indigenous communities with respect and humility. Listen to their perspectives and priorities, and act in ways that align with their values and desires. Build trust through long-term relationships.

Interested in more? Come to our side event!

Tech for TEK

Implementing Biodiversity Conservation with AI, Traditional Ecological Knowledge, and Indigenous Data Sovereignty.

When: Monday, October 28th, 11:40am Colombia time

Where: Place Quebec, Academic & Research Organizations Booth