

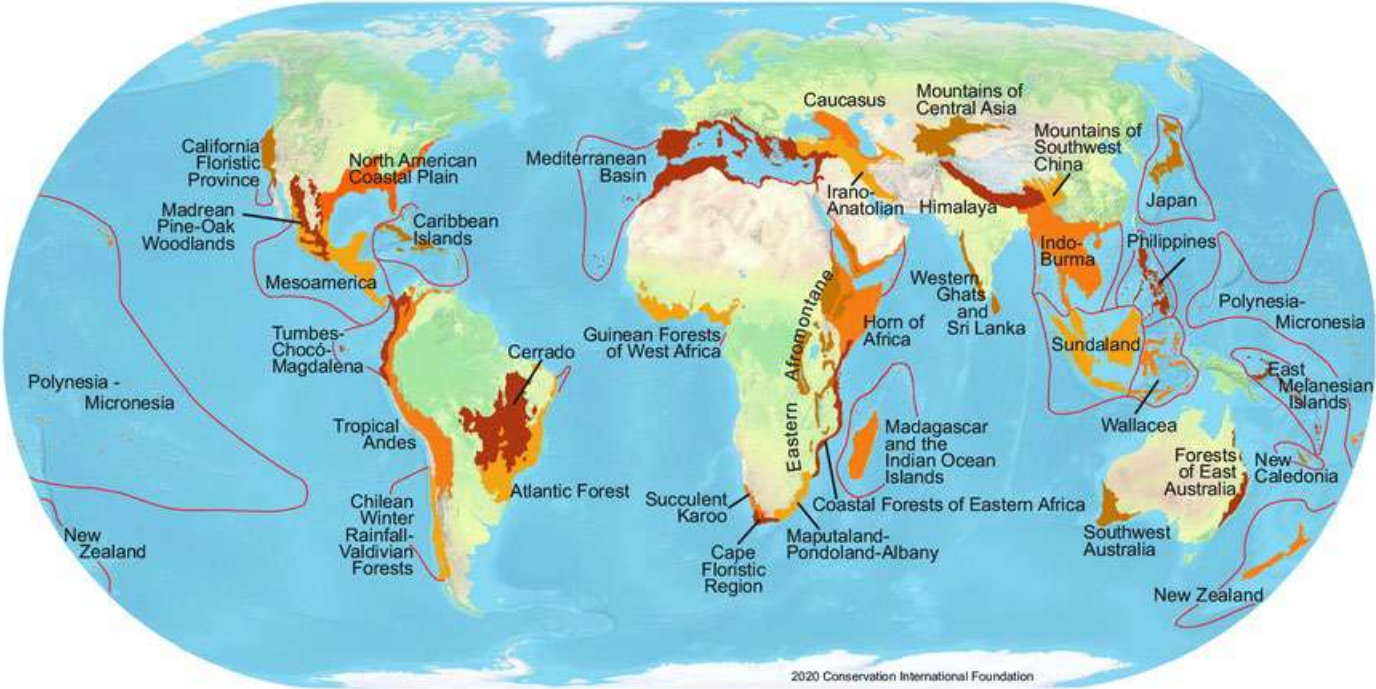
Multimodal AI for Ecological Monitoring: Images and DNA Barcodes

Joakim Bruslund Haurum

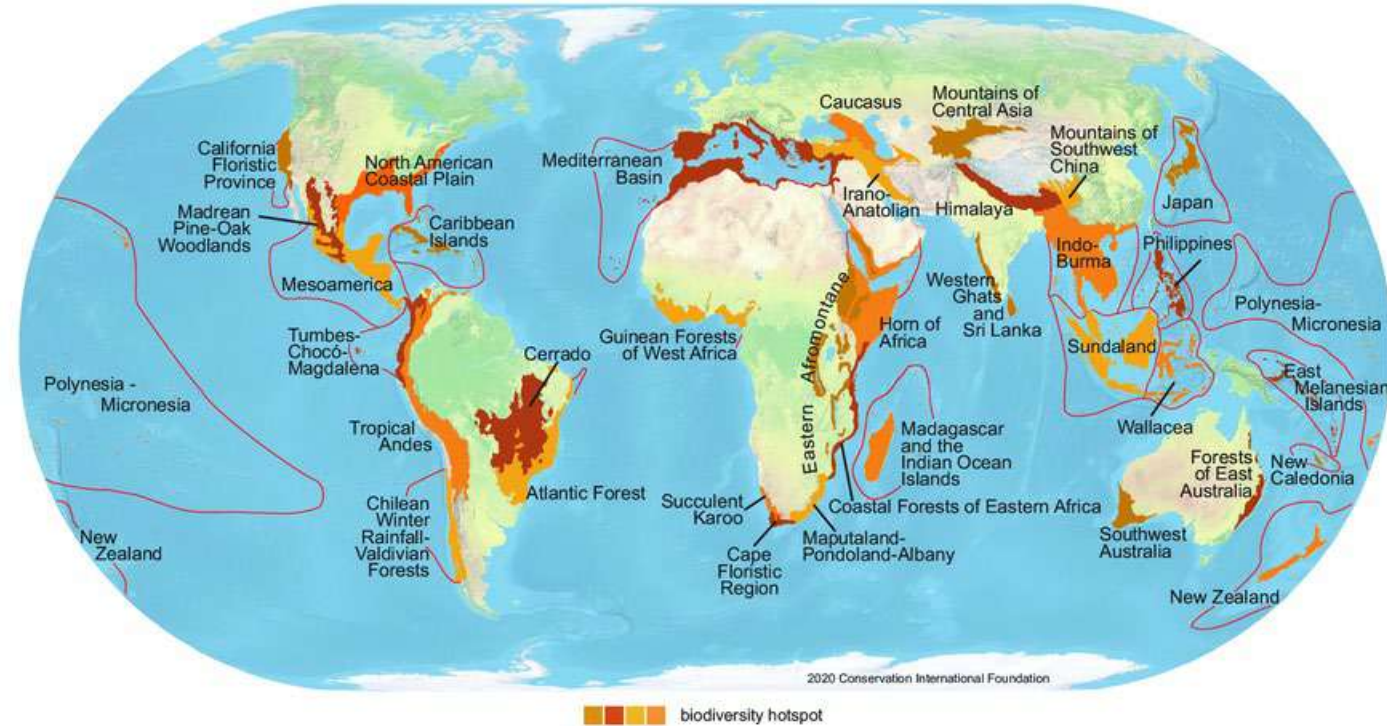
Tenure-Track Assistant Professor
SDU Center for Software Technology
Pioneer Centre for AI

jhau@mmmi.sdu.dk

Workshop on Computer Vision for Natural Heritage



Conservation International (conservation.org) defines 36 biodiversity hotspots — extraordinary places that harbor vast numbers of plant and animal species found nowhere else. All are heavily threatened by habitat loss and degradation, making their conservation crucial to protecting nature for the benefit of all life on Earth.



Conservation International (conservation.org) defines 36 biodiversity hotspots — extraordinary places that harbor vast numbers of plant and animal species found nowhere else. All are heavily threatened by habitat loss and degradation, making their conservation crucial to protecting nature for the benefit of all life on Earth.

The planet has lost **69% of its wildlife populations** since 1970 (WWF Living Planet Report).

Most biodiversity is concentrated in a handful of **hotspots** — and these are under the greatest threat.

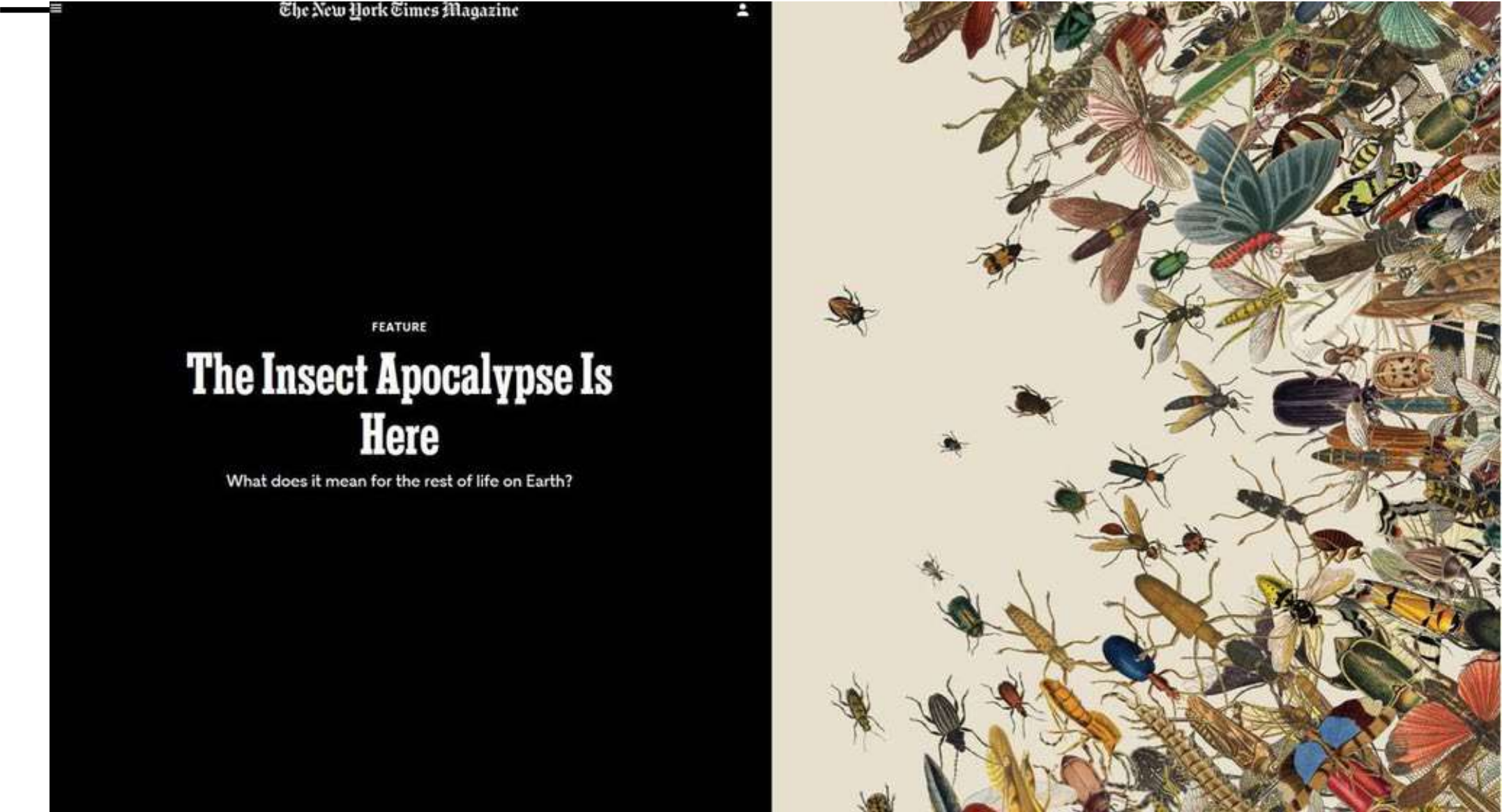
We are living through the sixth mass extinction, and most of it is invisible to us

Agenda

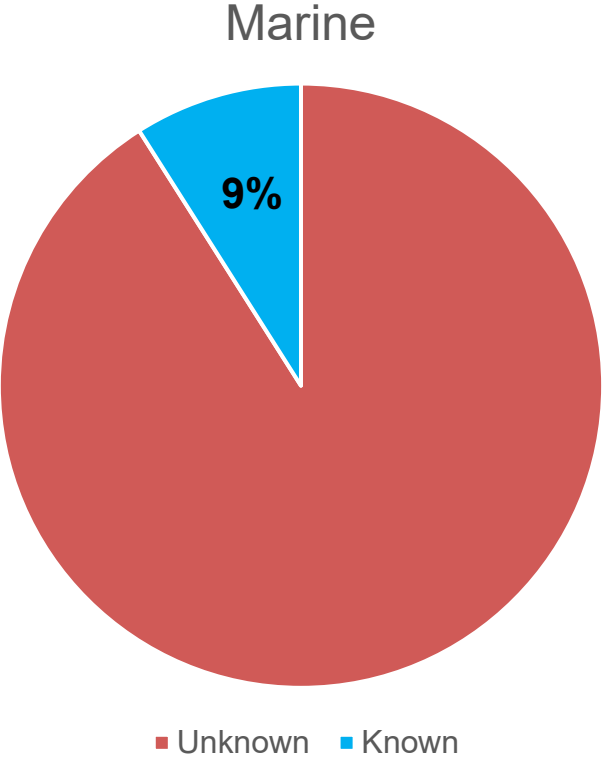
- How can AI help Global Biodiversity Monitoring?
- The BIOSCAN Project
- What's Next?

Agenda

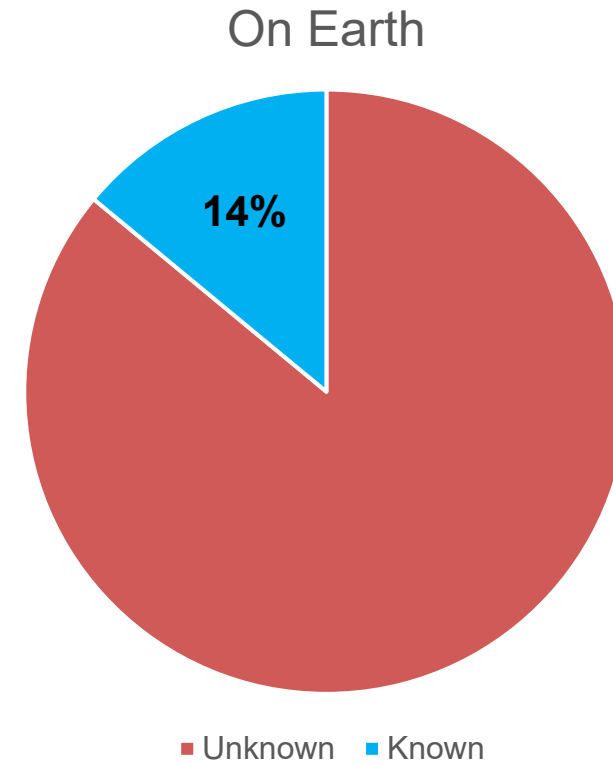
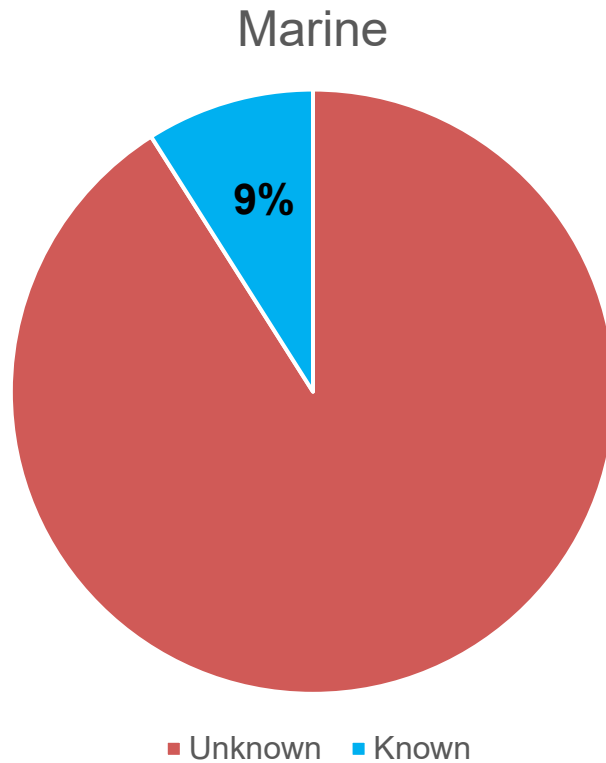
- **How can AI help Global Biodiversity Monitoring?**
- The BIOSCAN Project
- What's Next?



We don't know what we are losing!



We don't know what we are losing!



nature reviews biodiversity

<https://doi.org/10.1038/s44358-025-00022-3>

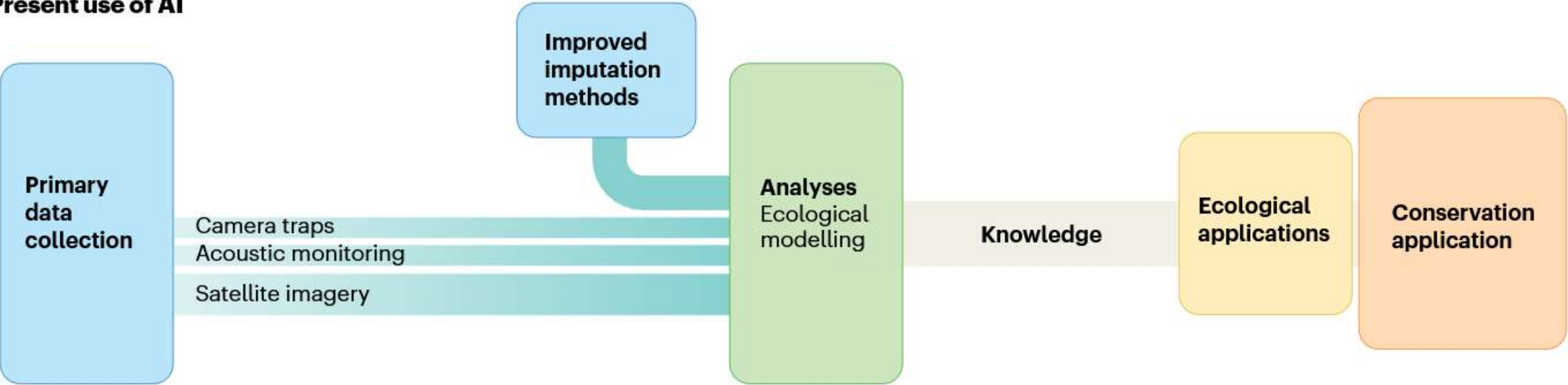
Review article

 Check for updates

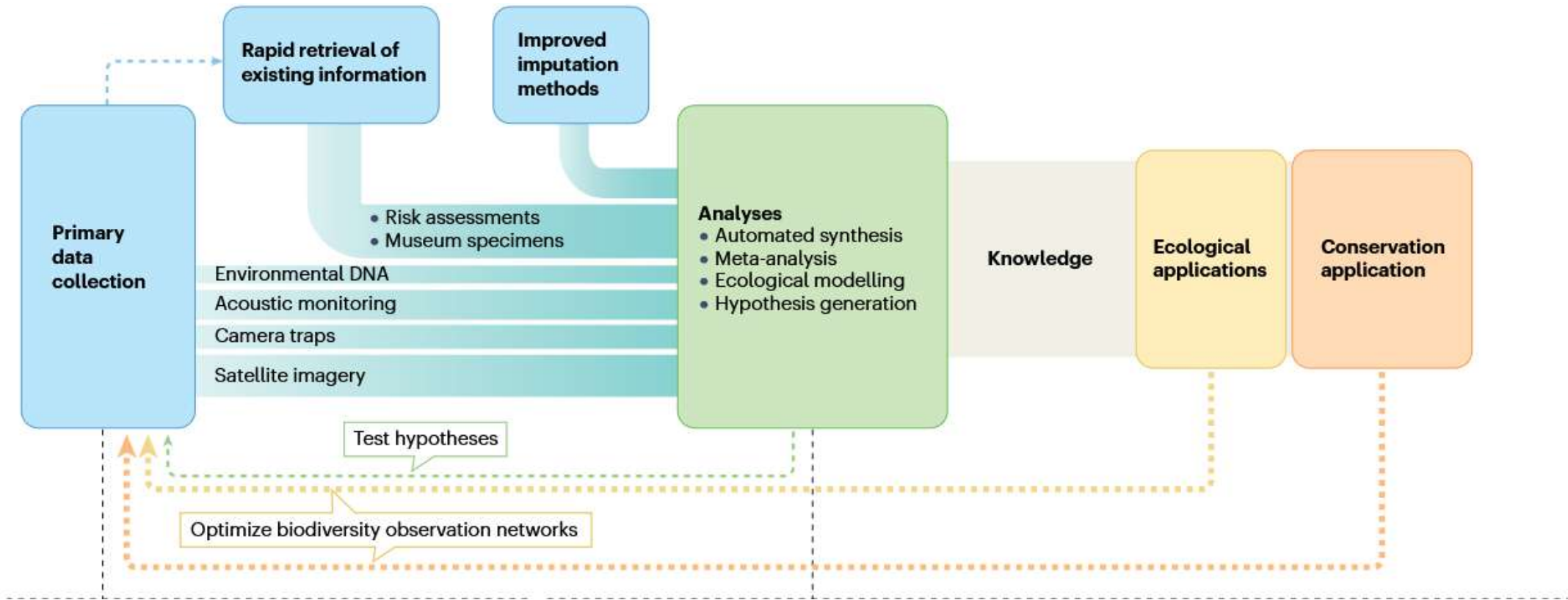
Harnessing artificial intelligence to fill global shortfalls in biodiversity knowledge

Laura J. Pollock ^{1,2,17} , Justin Kitzes^{3,17}, Sara Beery⁴, Kaitlyn M. Gaynor^{5,6}, Marta A. Jarzyna^{7,8}, Oisín Mac Aodha⁹, Bernd Meyer ¹⁰, David Rolnick^{2,11}, Graham W. Taylor^{12,13}, Devis Tuia ¹⁴ & Tanya Berger-Wolf^{7,8,15,16}

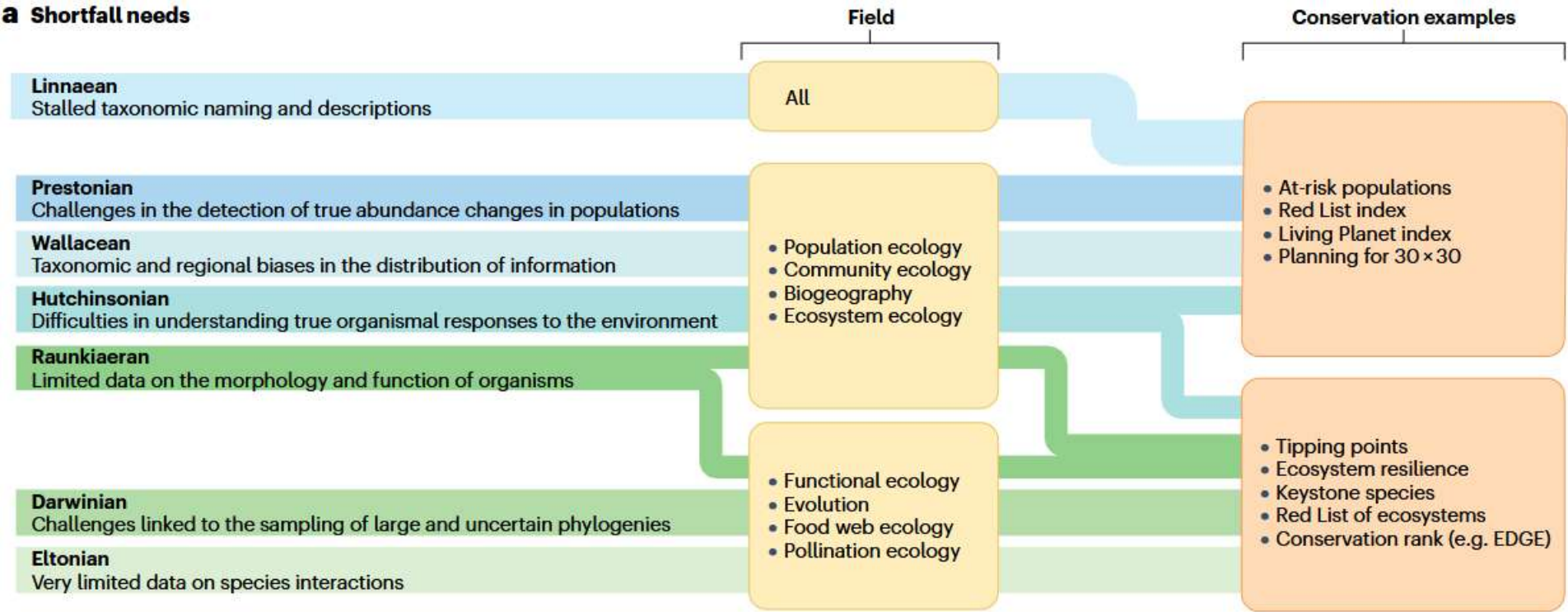
a Present use of AI



Future use of AI

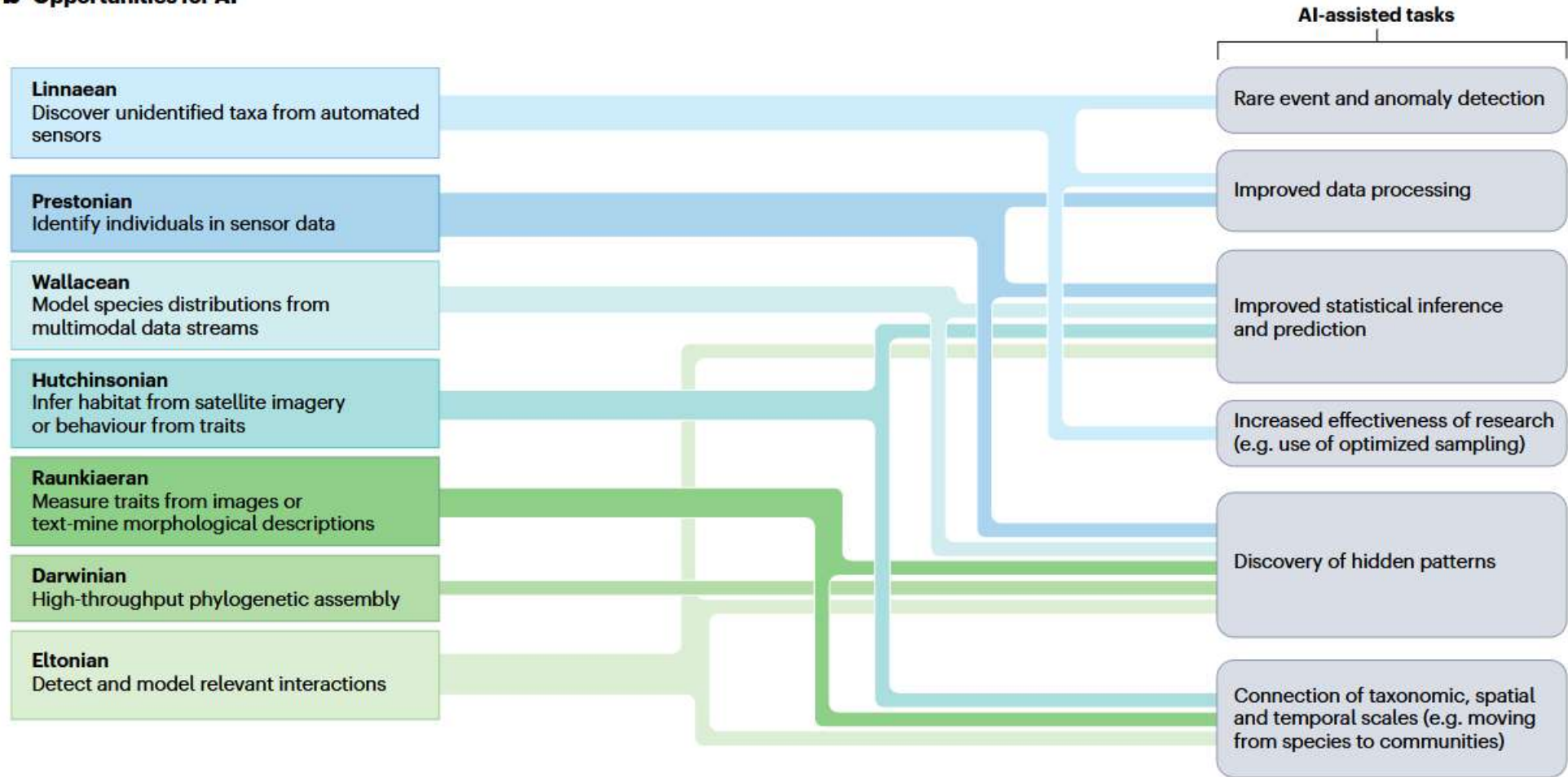


7 Global Shortfalls



Addressing shortfalls with AI

b Opportunities for AI



Agenda

- How can AI help Global Biodiversity Monitoring?
- **The BIOSCAN Project**
- What's Next?



BIOSCAN: A Global Scientific Mission

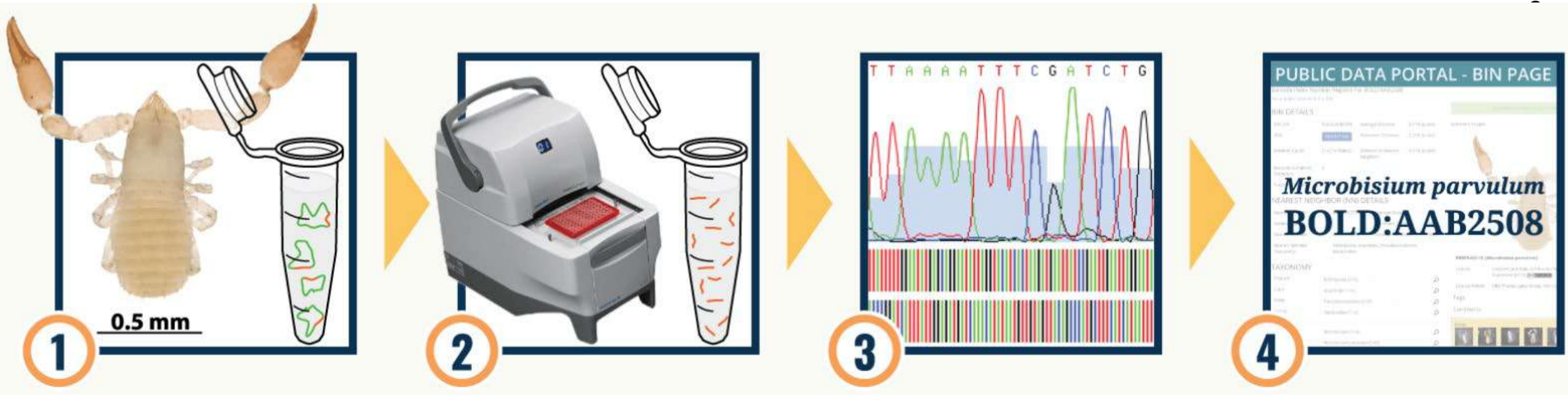
→ *To illuminate biodiversity by developing globally accessible **DNA-based systems** to speed species discovery, probe species interactions, and establish biodiversity baselines*



How do you count what you can't see?

- **Malaise traps** are tent-like structures that passively collect flying insects 24/7.
- A single trap can capture **thousands of specimens per week**.
- The problem? Someone has to sort and identify all of them. That's where **DNA and AI** comes in.

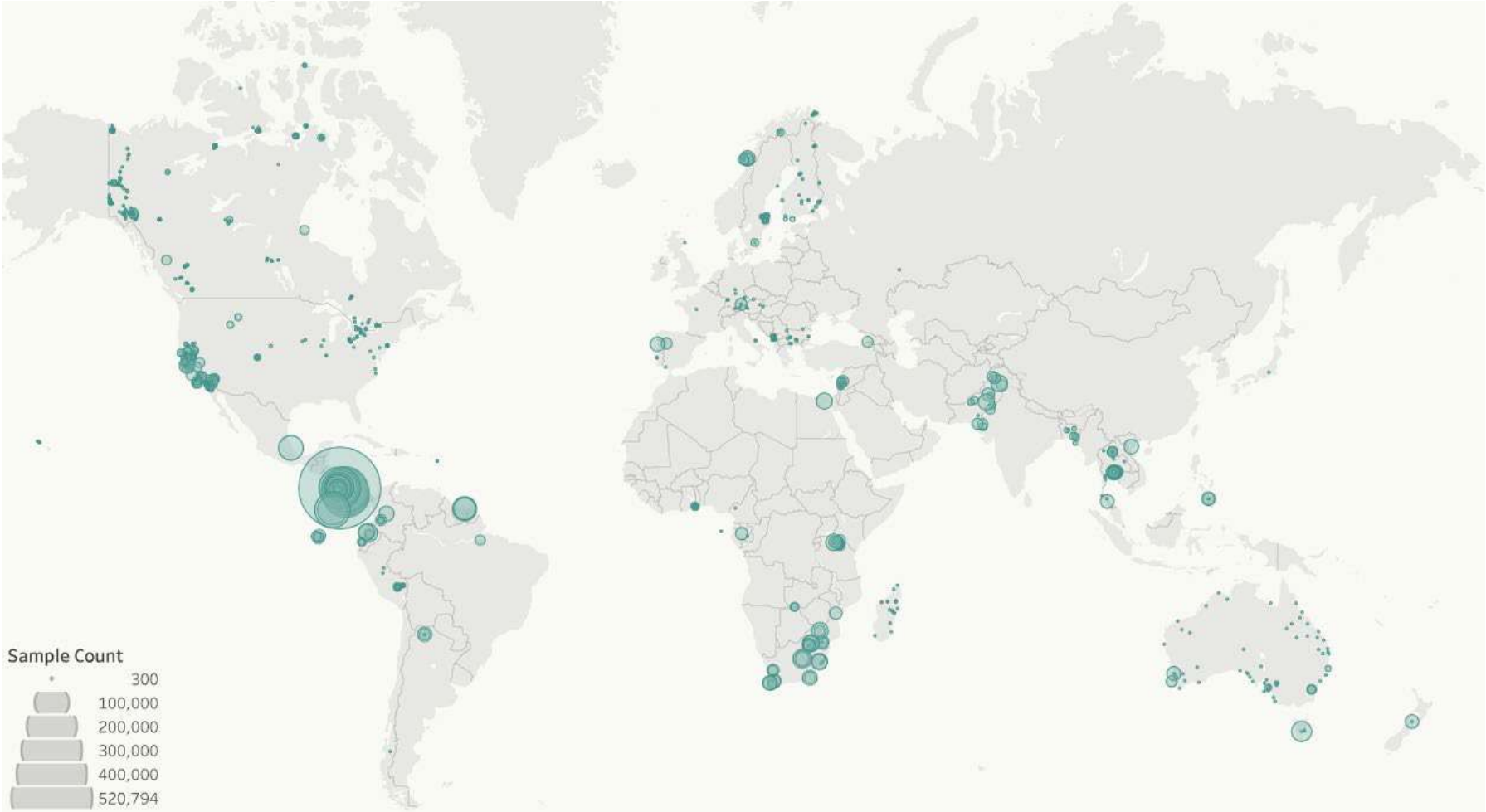
DNA Barcoding



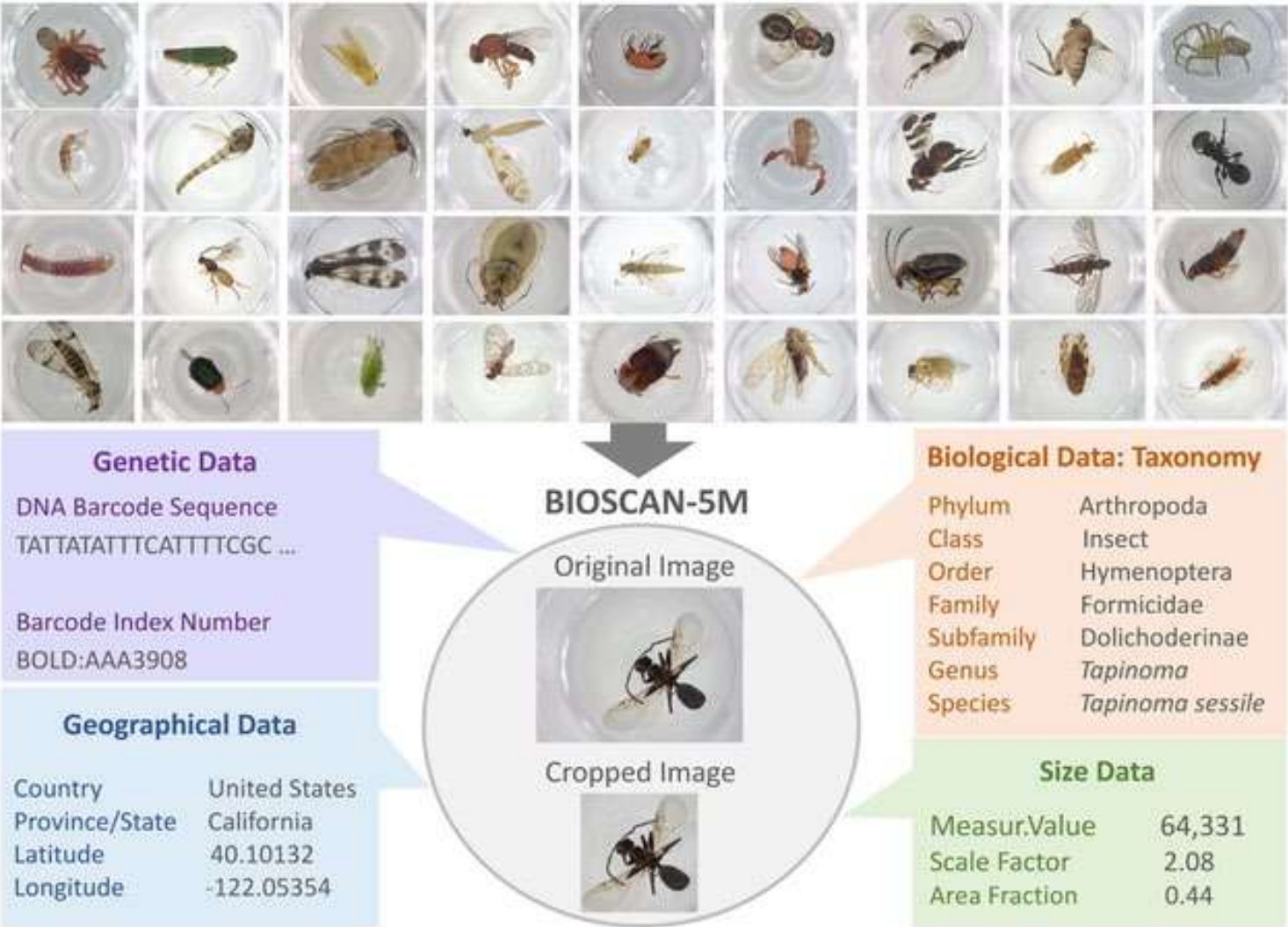
#sdruk



BIOSCAN – 47 Countries, 5 Million specimen



BIOSCAN-5M: A Multimodal Dataset for Insect Biodiversity (NeurIPS'24)



Attributes	BIOSCAN-5M (Ours)			
	IR	Categories	Labelled	Labelled (%)
phylum	1	1	5,150,850	100.0
class	719,831	10	5,146,837	99.9
order	3,675,317	55	5,134,987	99.7
family	938,928	934	4,932,774	95.8
subfamily	323,146	1,542	1,472,548	28.6
genus	200,268	7,605	1,226,765	23.8
species	7,694	22,622	473,094	9.2
dna_bin	35,458	324,411	5,137,441	99.7
dna_barcode	3,743	2,486,492	5,150,850	100.0

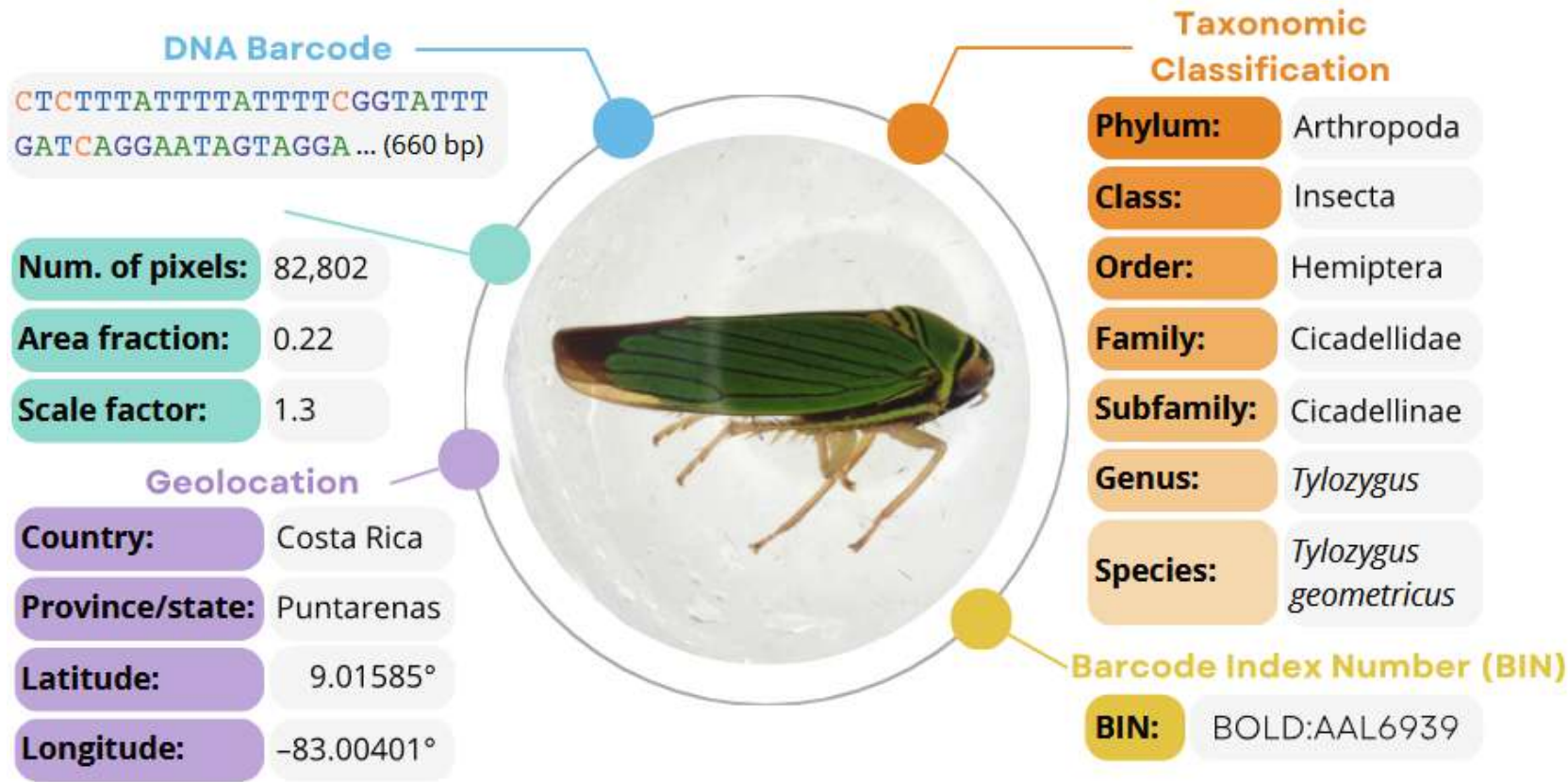
BIOSCAN-5M (Ours)				
Attributes	IR	Categories	Labelled	Labelled (%)
phylum	1	1	5,150,850	100.0
class	719,831	10	5,146,837	99.9
order	3,675,317	55	5,134,987	99.7
family	938,928	934	4,932,774	95.8
subfamily	323,146	1,542	1,472,548	28.6
genus	200,268	7,605	1,226,765	23.8
species	7,694	22,622	473,094	9.2
dna_bin	35,458	324,411	5,137,441	99.7
dna_barcode	3,743	2,486,492	5,150,850	100.0

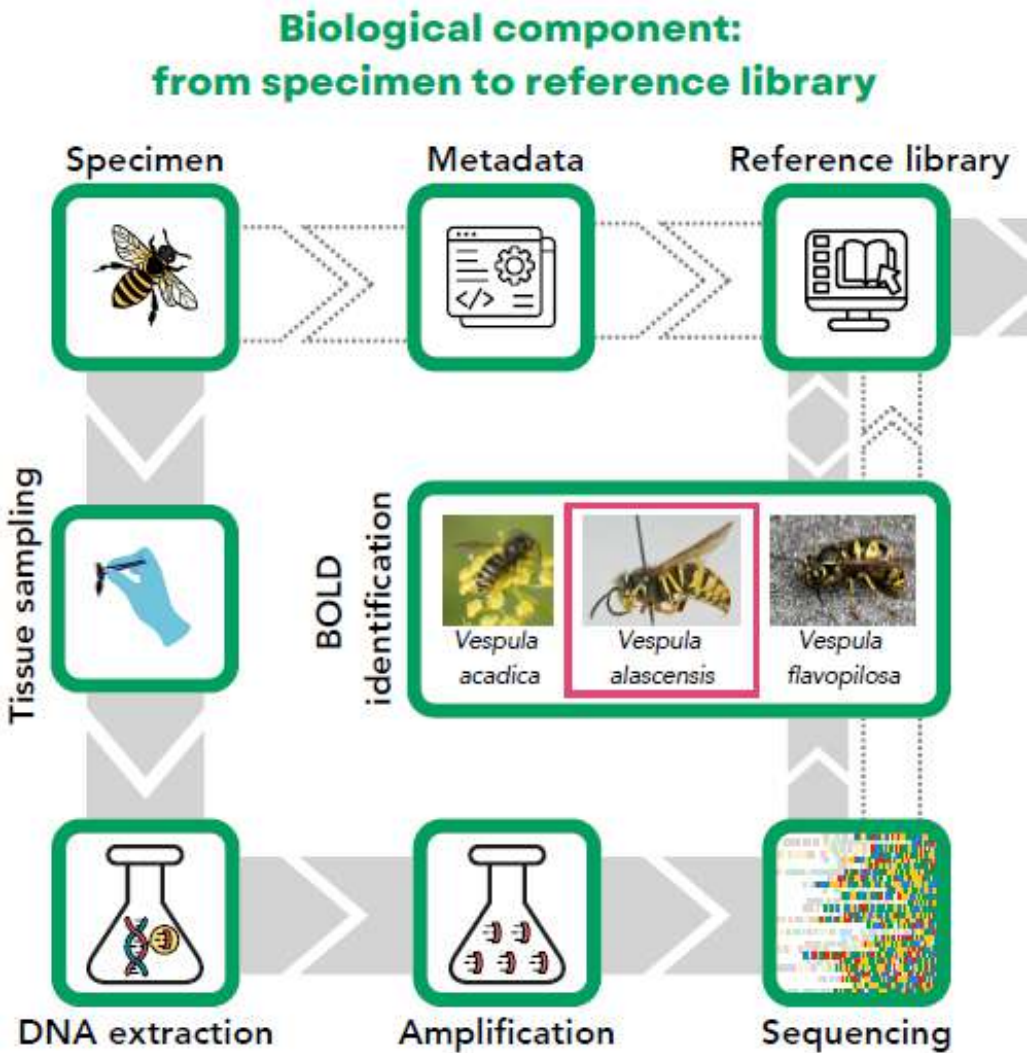
BIOSCAN-5M (Ours)				
Attributes	IR	Categories	Labelled	Labelled (%)
phylum	1	1	5,150,850	100.0
class	719,831	10	5,146,837	99.9
order	3,675,317	55	5,134,987	99.7
family	938,928	934	4,932,774	95.8
subfamily	323,146	1,542	1,472,548	28.6
genus	200,268	7,605	1,226,765	23.8
species	7,694	22,622	473,094	9.2
dna_bin	3,743	324,411	1	99.7
dna_barcode	3,743	2,486,492	5,150,850	100.0

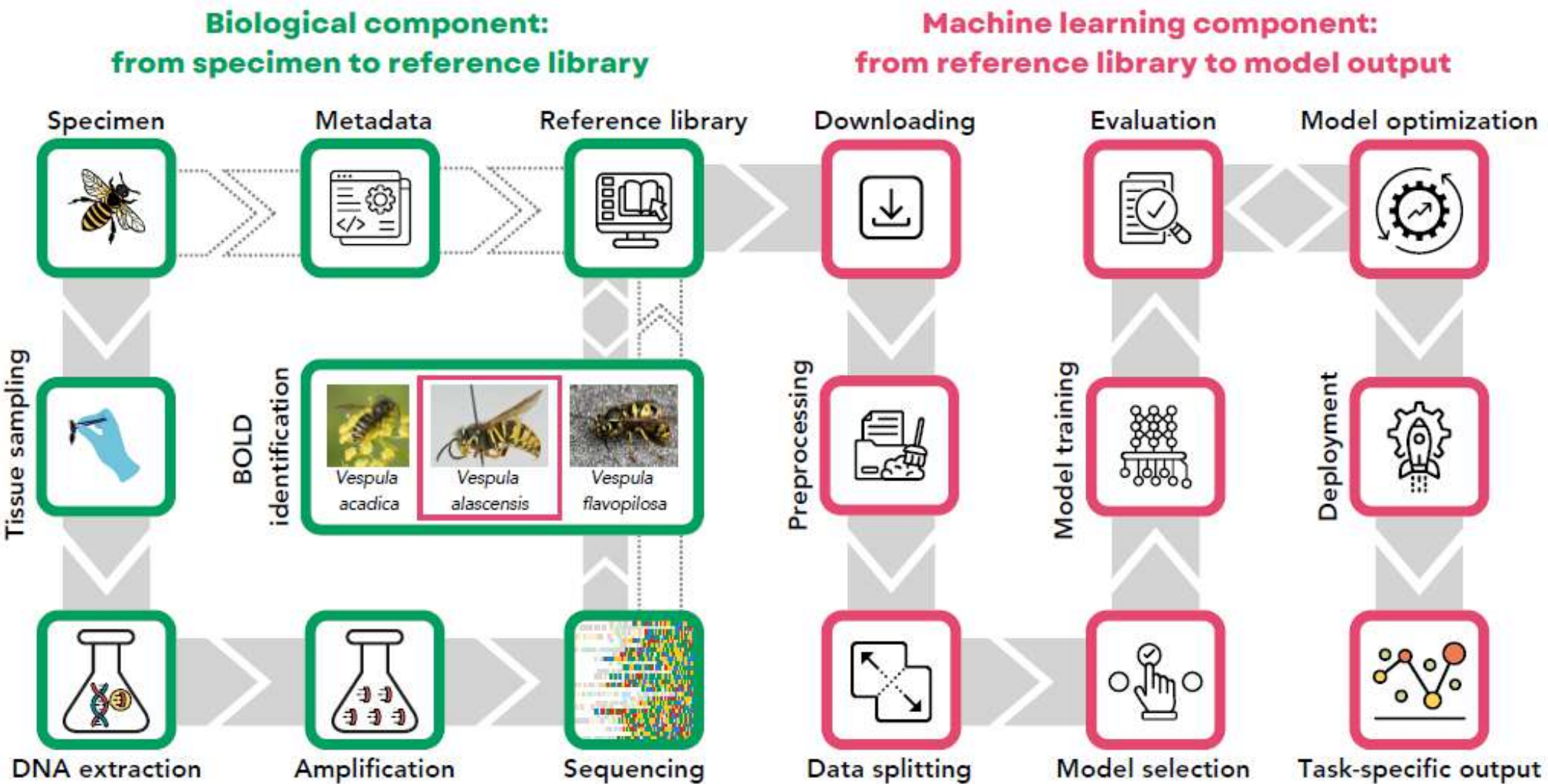
Very *Fine-Grained* task



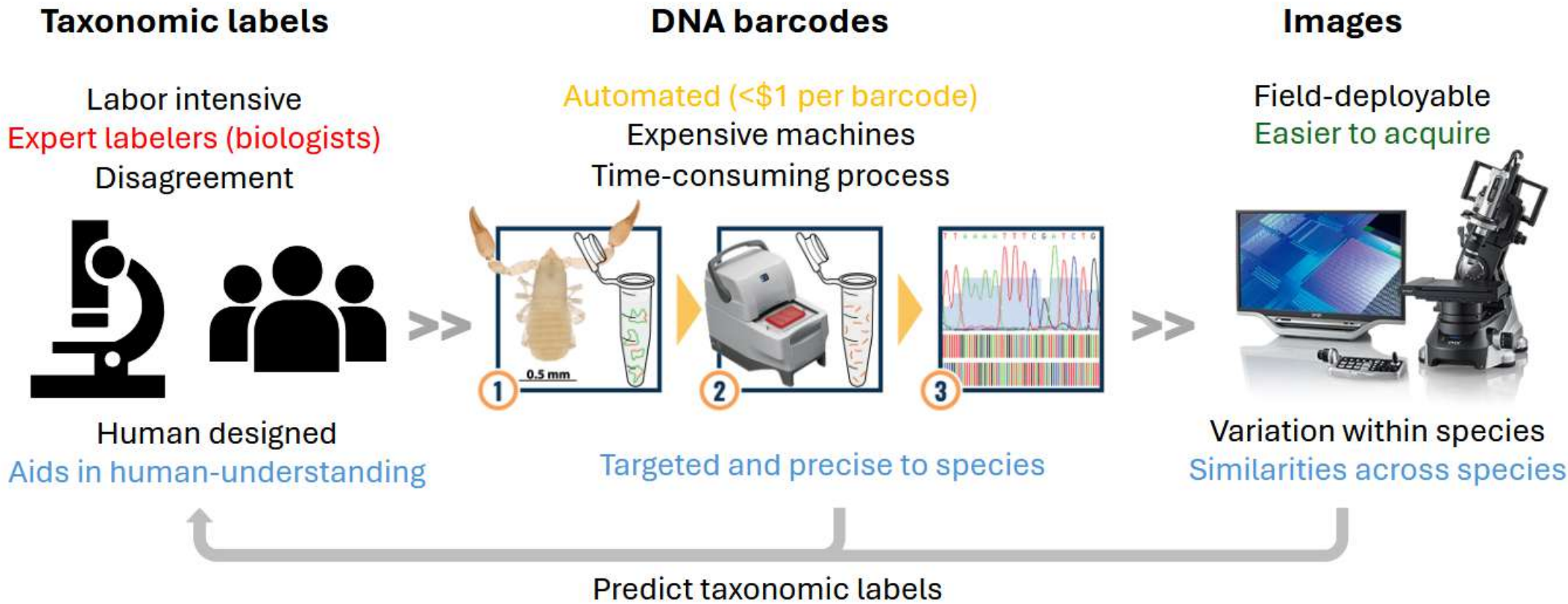
But also, *very data rich*



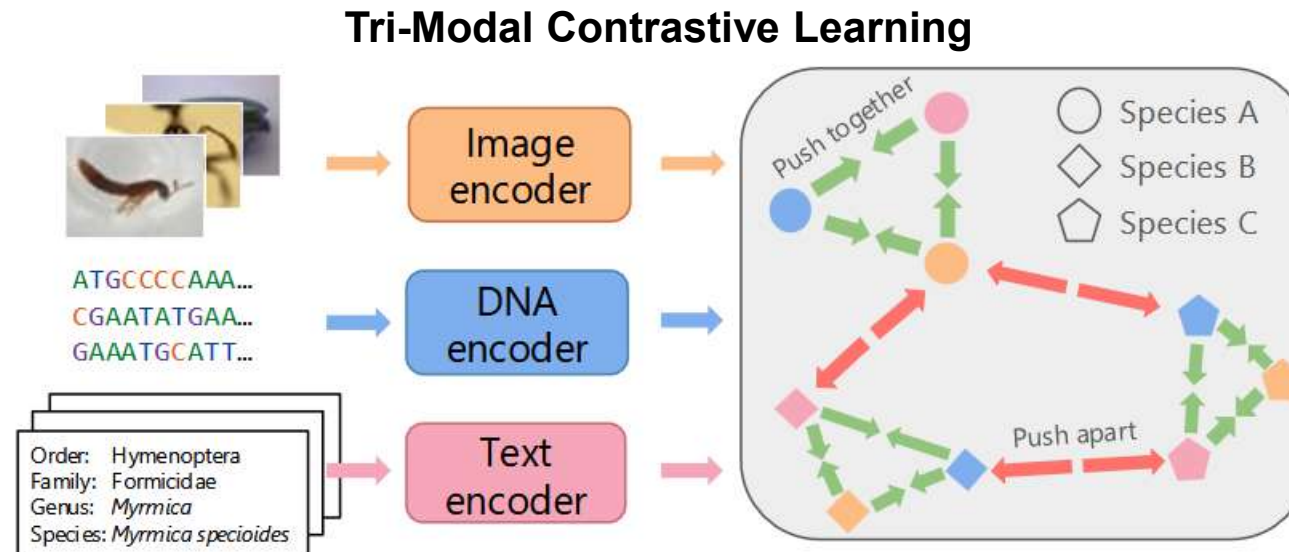




**But how do we use all these
modalities together?**



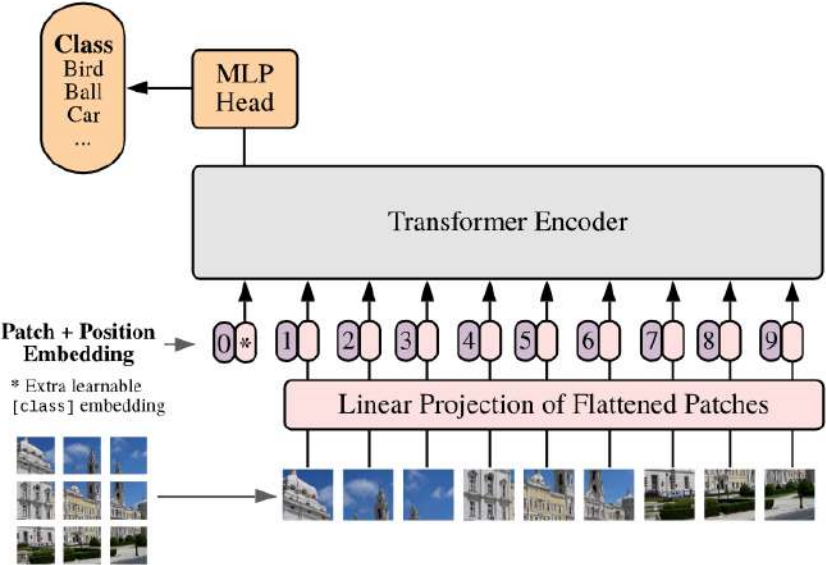
CLIBD: Bridging Vision and Genomics for Biodiversity Monitoring at Scale (ICLR'25)



Encoder Overview

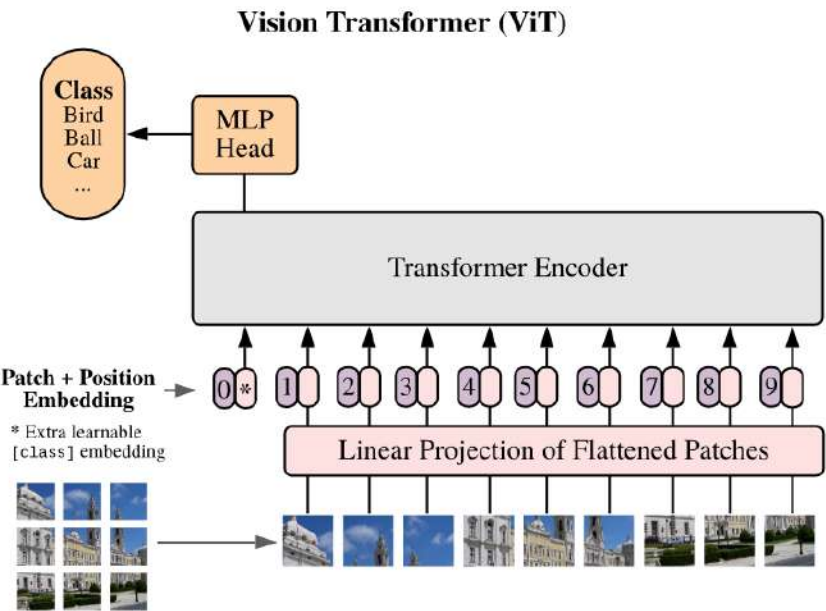
ViT-B

Vision Transformer (ViT)

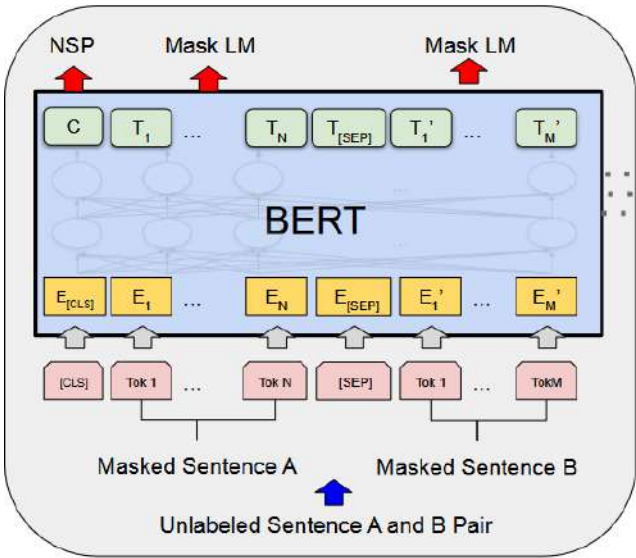


Encoder Overview

ViT-B

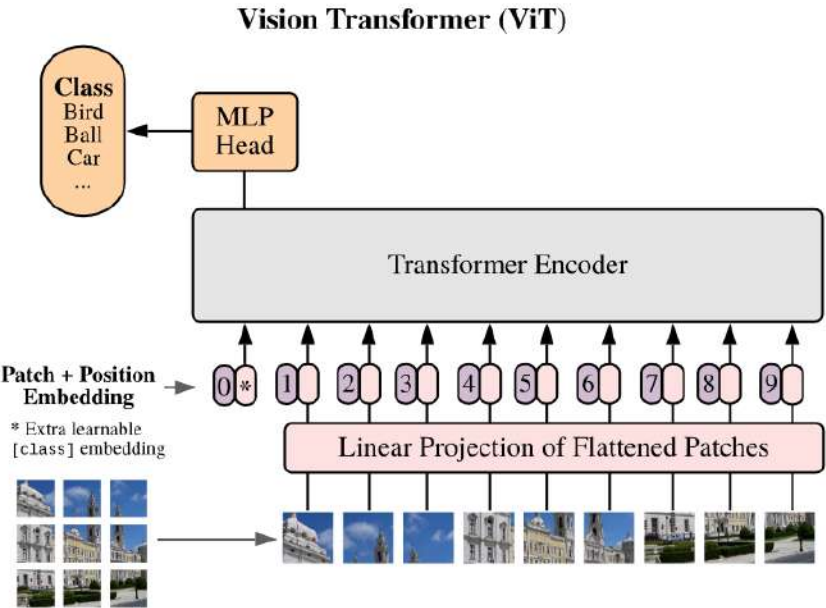


BERT-Small

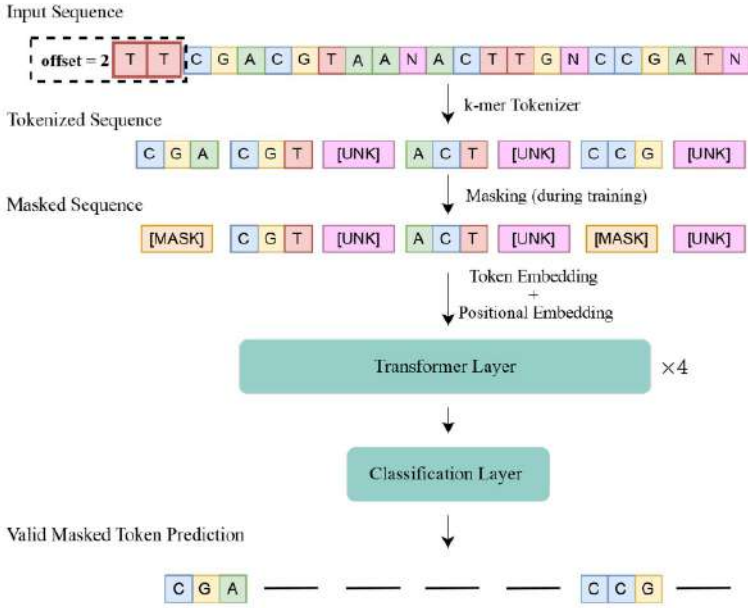


Encoder Overview

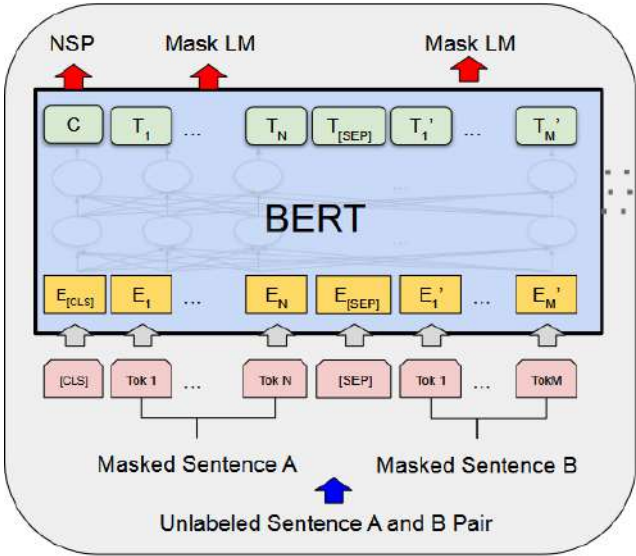
ViT-B



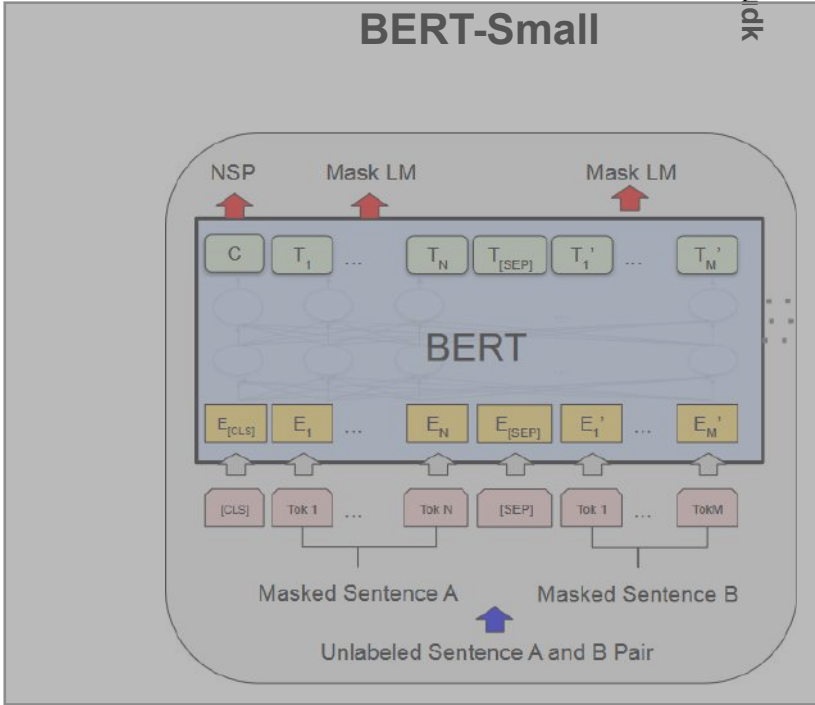
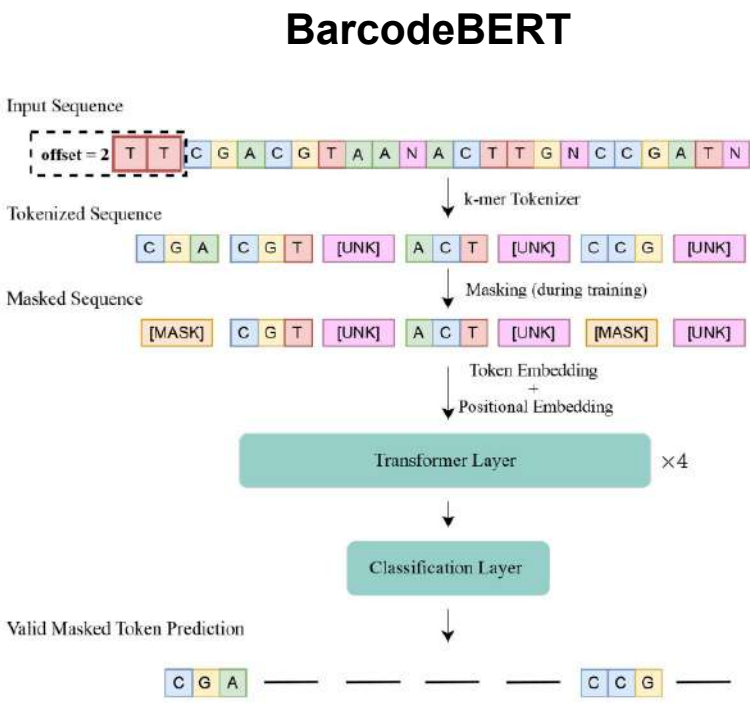
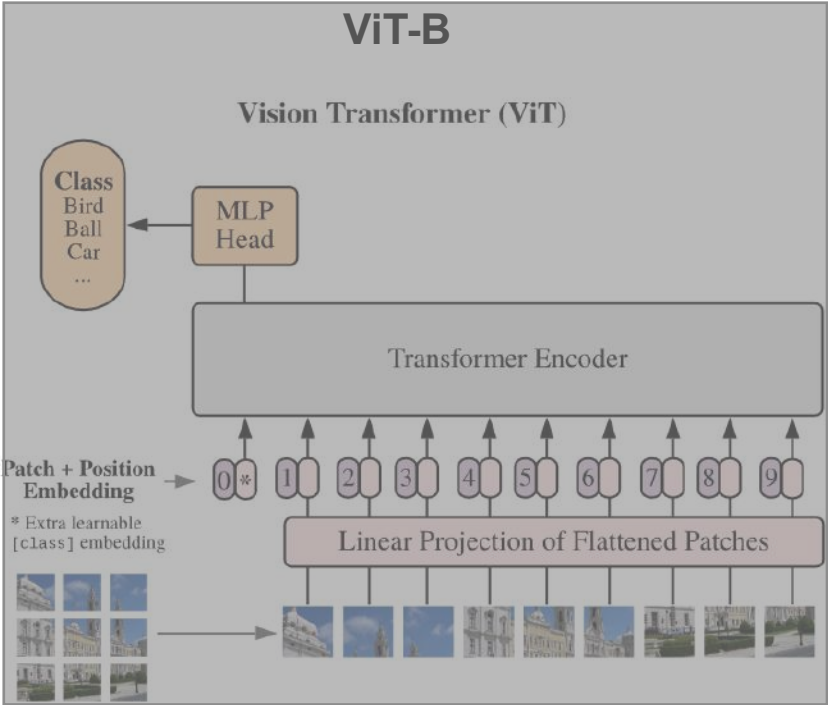
BarcodeBERT



BERT-Small



Encoder Overview



Tri-Modal Contrastive Learning

Tri-Modal Contrastive Learning

$$L_i^{(V \rightarrow D)} = -\log \frac{\exp(V_i^T D_i / \tau)}{\sum_{j=1}^n \exp(V_i^T D_j / \tau)}, \quad L_i^{(D \rightarrow V)} = -\log \frac{\exp(D_i^T V_i / \tau)}{\sum_{j=1}^n \exp(D_i^T V_j / \tau)},$$

- V, D are modality-embeddings (e.g. image and DNA)

Tri-Modal Contrastive Learning

$$L_i^{(V \rightarrow D)} = -\log \frac{\exp(V_i^T D_i / \tau)}{\sum_{j=1}^n \exp(V_i^T D_j / \tau)}, \quad L_i^{(D \rightarrow V)} = -\log \frac{\exp(D_i^T V_i / \tau)}{\sum_{j=1}^n \exp(D_i^T V_j / \tau)},$$

$$L_{VD} = \sum_{i=1}^n \left(L_i^{(V \rightarrow D)} + L_i^{(D \rightarrow V)} \right)$$

- V, D are modality-embeddings (e.g. image and DNA)

Tri-Modal Contrastive Learning

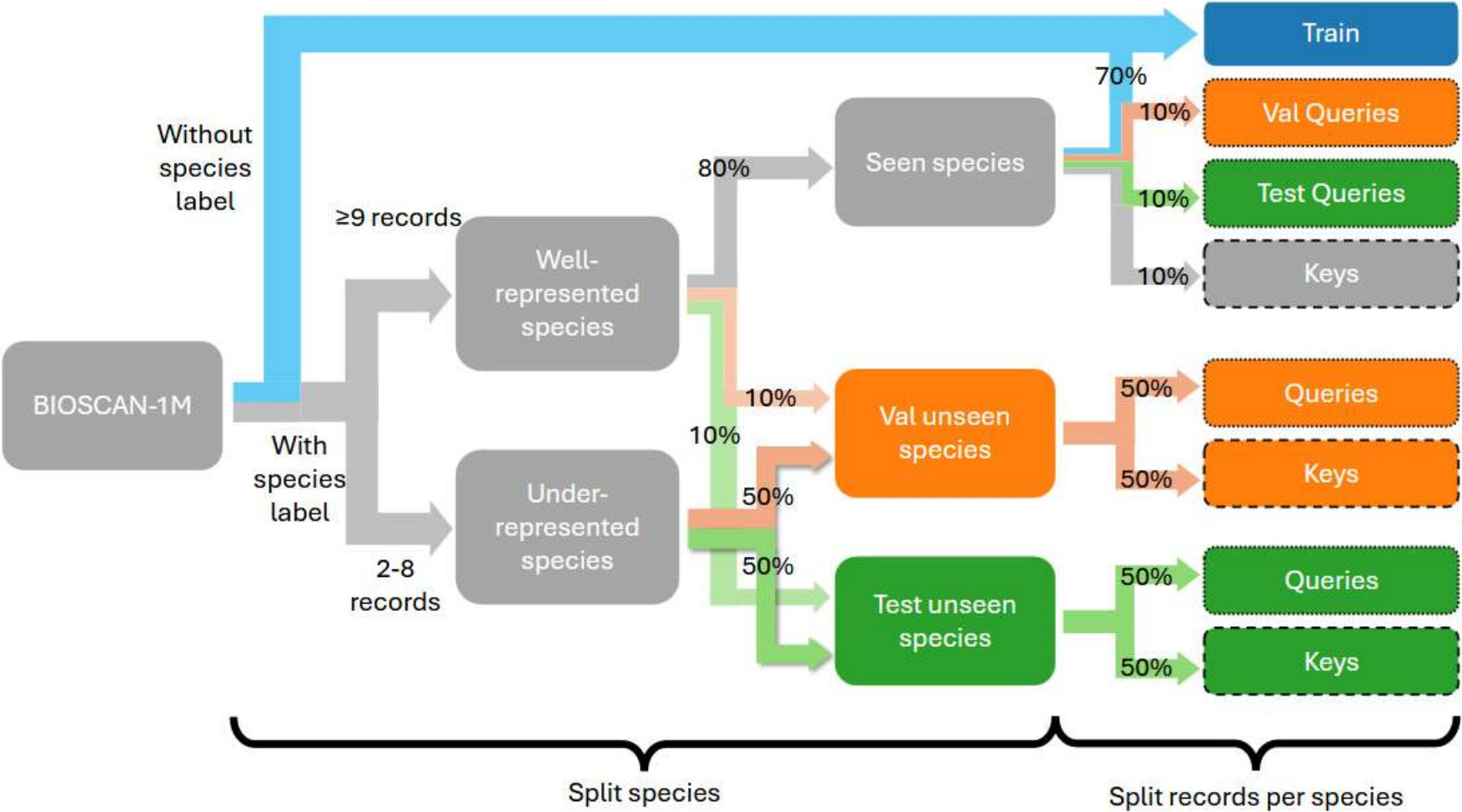
$$L_i^{(V \rightarrow D)} = -\log \frac{\exp(V_i^T D_i / \tau)}{\sum_{j=1}^n \exp(V_i^T D_j / \tau)}, \quad L_i^{(D \rightarrow V)} = -\log \frac{\exp(D_i^T V_i / \tau)}{\sum_{j=1}^n \exp(D_i^T V_j / \tau)},$$

$$L_{VD} = \sum_{i=1}^n \left(L_i^{(V \rightarrow D)} + L_i^{(D \rightarrow V)} \right)$$

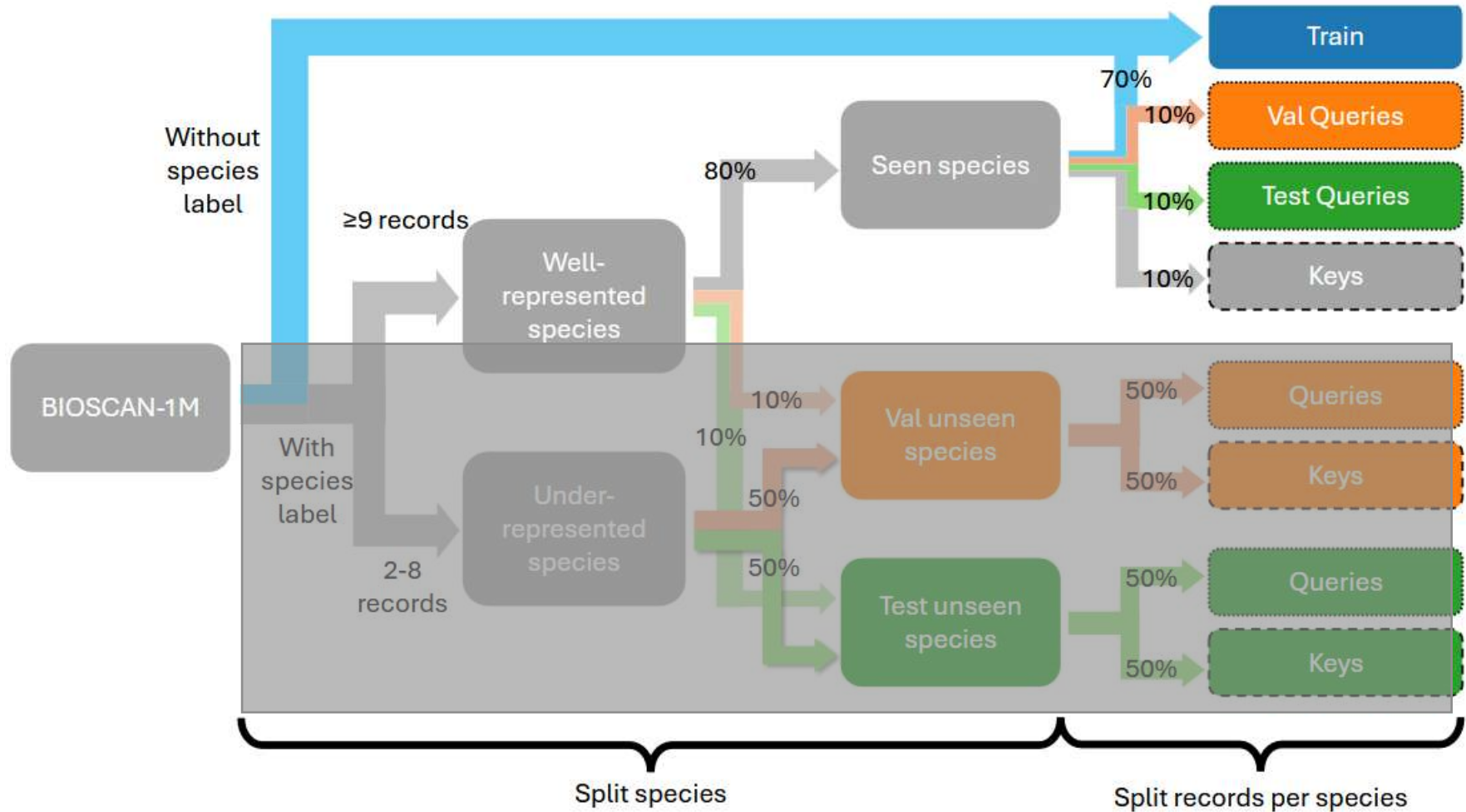
$$L = L_{VD} + L_{DT} + L_{VT}$$

CLIBD Datasplit

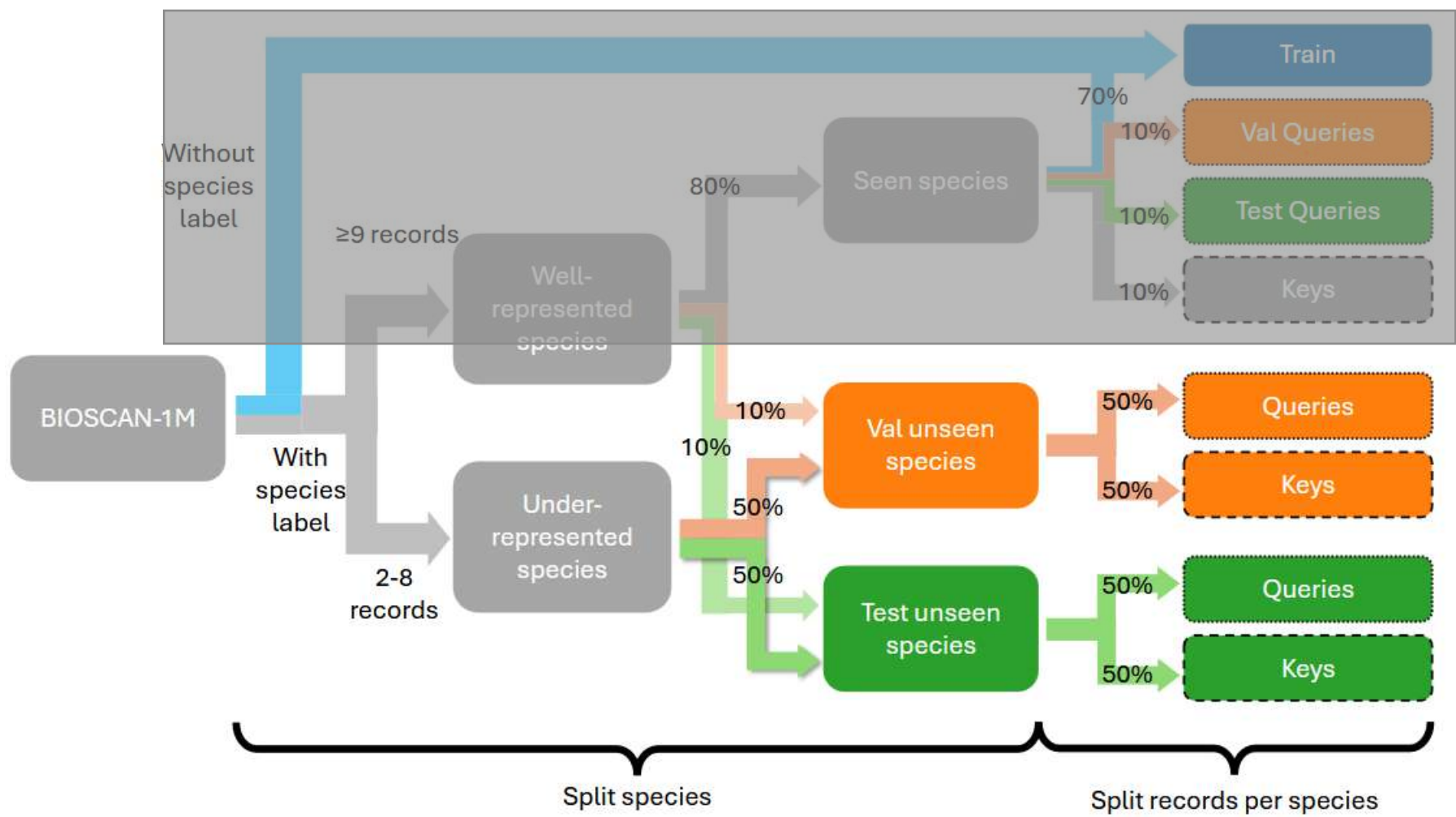
CLIBD Datasplit



CLIBD Datasplit – “Known” Species

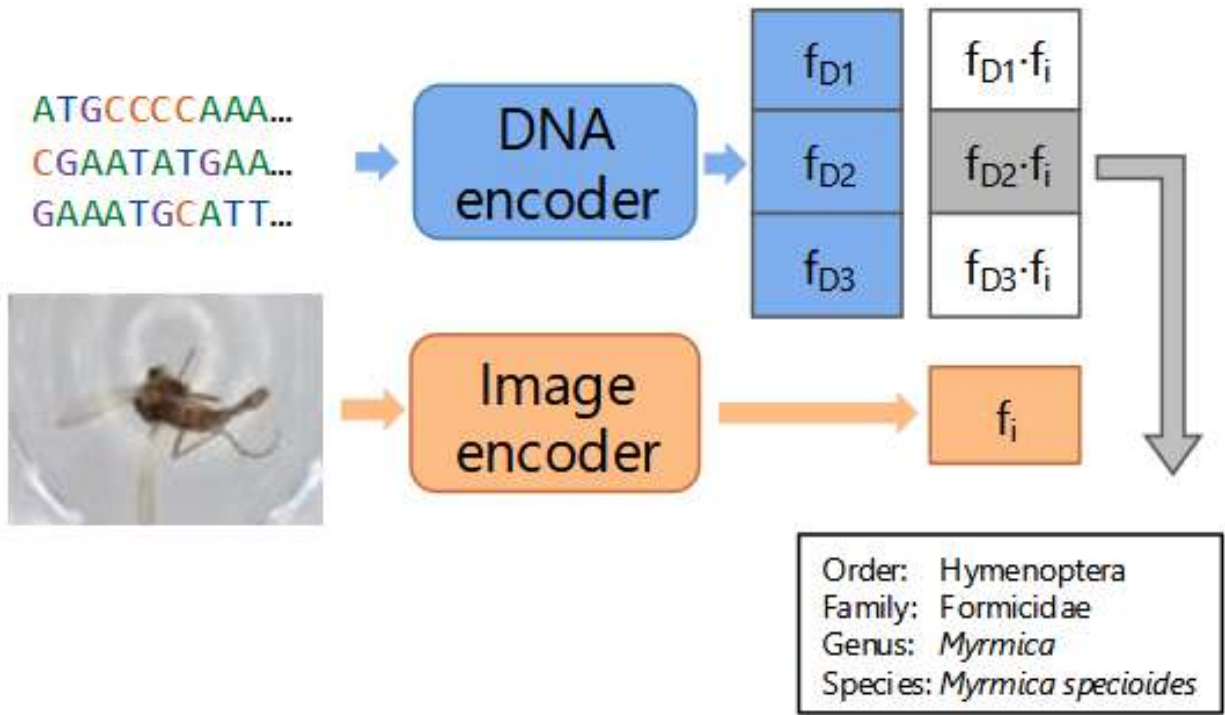


CLIBD Datasplit – “Unknown” Species



Inference of CLIBD

Inference: Reference DNA/Image database query



Retrieval Performance

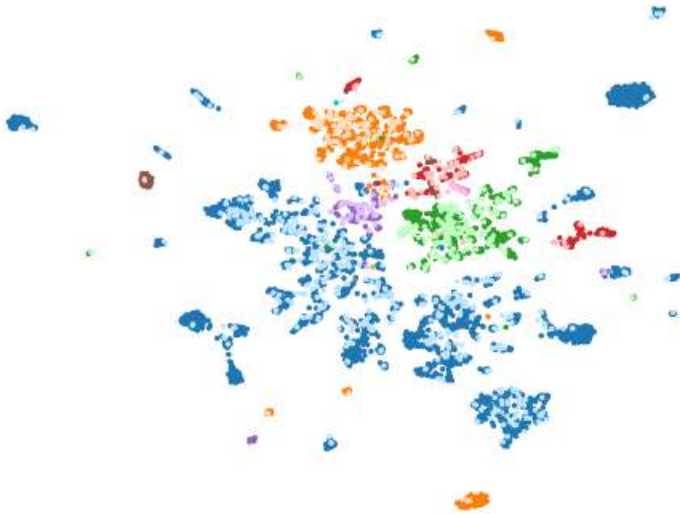
Model	Aligned embeddings			Image-to-Image			Image-to-DNA			Image-to-Text		
	Img	DNA	Txt	Seen	Unseen	H.M.	Seen	Unseen	H.M.	Seen	Unseen	H.M.
BioCLIP	✓	✗	✓	20.4	14.8	17.1	—	—	—	4.2	3.1	3.6
CLIBD	✓	✗	✓	54.2	33.6	41.5	—	—	—	57.6	4.6	8.5
CLIBD	✓	✓	✓	59.3	45.0	51.2	51.6	8.6	14.7	56.0	4.8	8.9

	Original	DNA to DNA			Image to Image			Image to DNA		
Caenocholax sp. 10 DPM-2010										
Micromus subanticus										
Phyllonorycter anderidae										
Glyptotermes barbouri										

#sdu.dk



(a) No alignment



(b) Align image and text



(c) Align image and DNA



(d) Align all three modalities

You can try it yourself!

BIOSCAN Browser

Visualizing a Multimodal Dataset for Insect Biodiversity

BIOSCAN-5M is a large-scale multimodal dataset of over 5 million insect specimens. Each record links high-resolution images with taxonomic labels, raw DNA barcode sequences, Barcode Index Numbers (BINs), and geographic information.

 Browse records

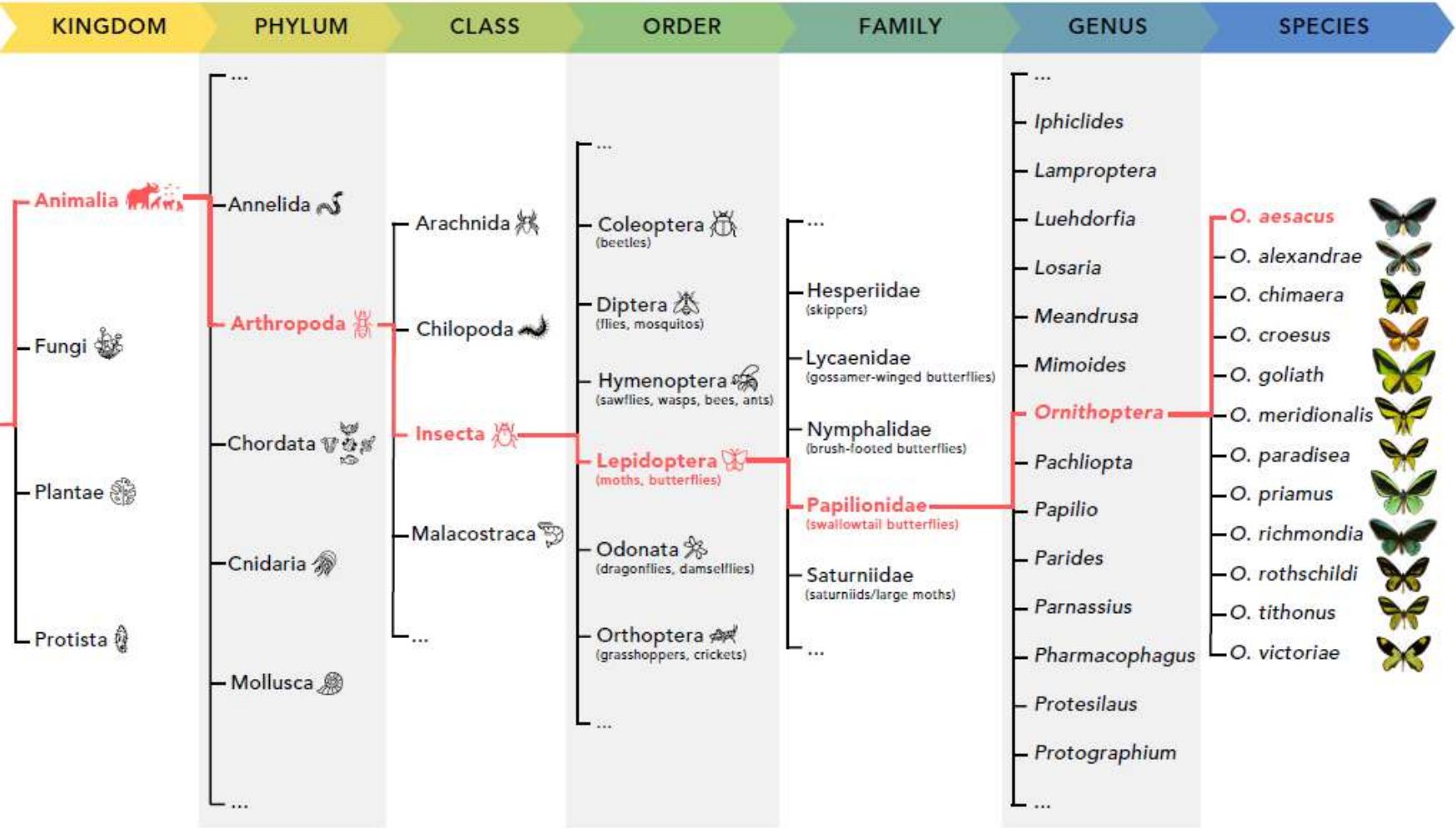


<https://browser.bioscan-ml.org/>

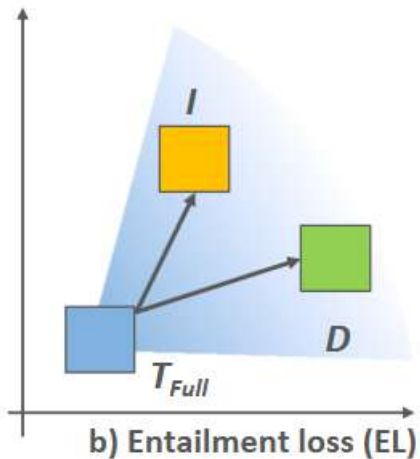
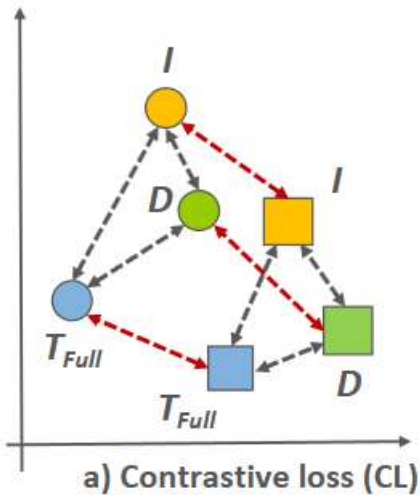
Agenda

- How can AI help Global Biodiversity Monitoring?
- The BIOSCAN Project
- **What's Next?**

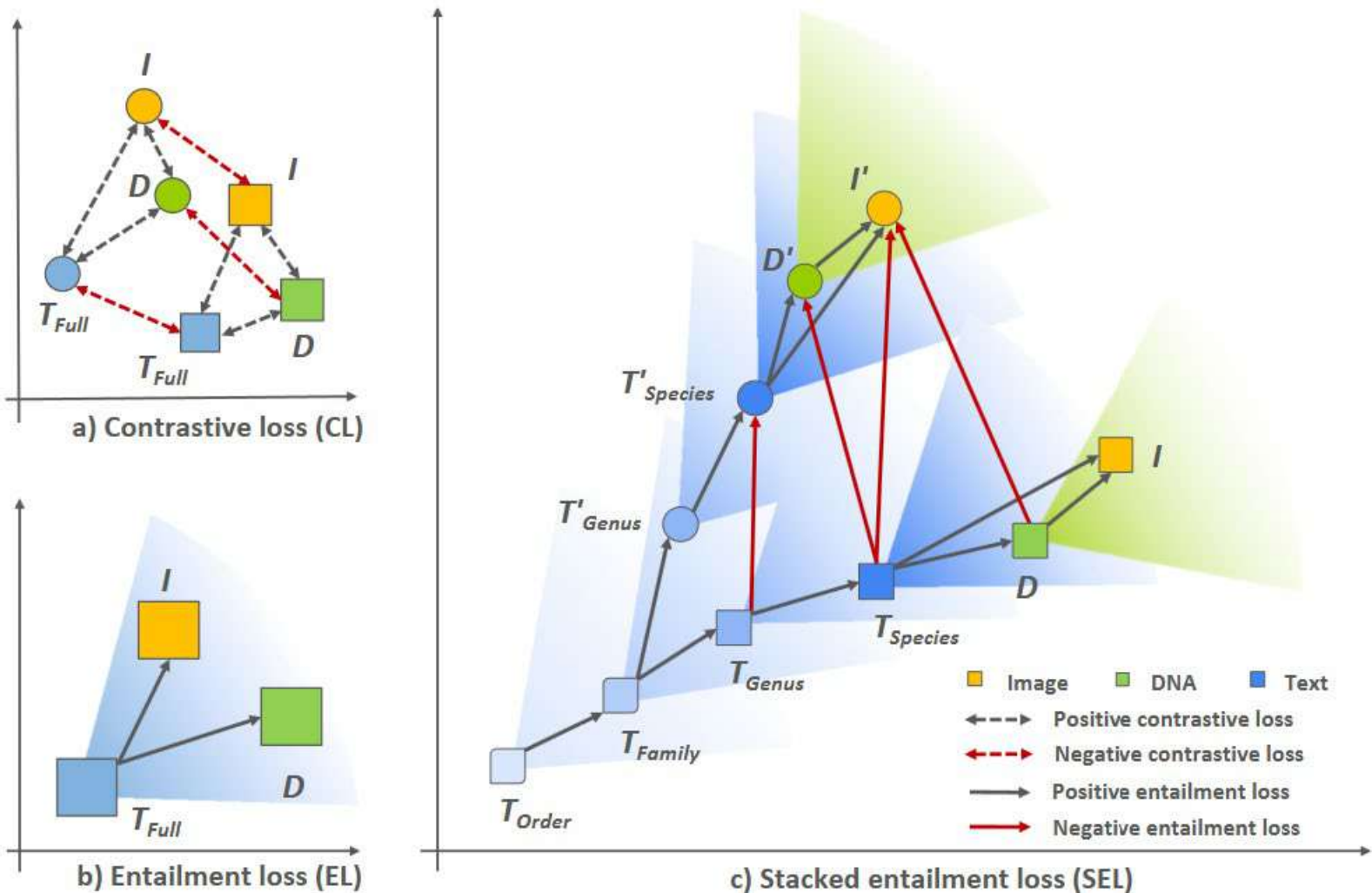
Can we use the taxonomic information?



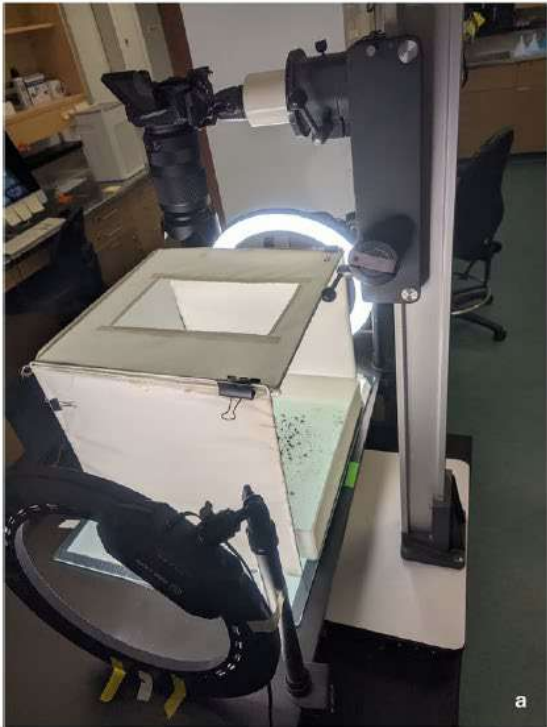
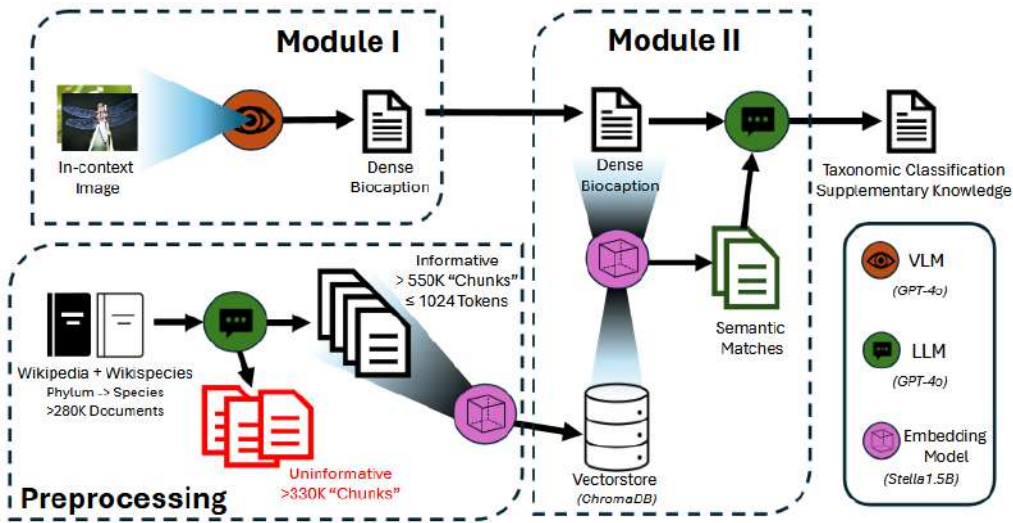
Stacked Entailment Loss



Stacked Entailment Loss



Agents and MegaBarcoding



Collaborators – The BIOSCAN-ML Group



Acknowledgements



Government of Canada
Gouvernement du Canada



New Frontiers in Research Fund
Fonds Nouvelles frontières en recherche



AI AND BIODIVERSITY
CHANGE
GLOBAL CENTER



Canada Research Chairs
Chaires de recherche du Canada



CIFAR



NSERC
CRSNG