

INTELLIGENCE MONOGRAPH · 2026

Animal Language Processing *in 2026*

*Computational Foundations, Robotic Epistemology,
and the Horizon of Interspecies Translation*

Marie-Soleil Seshat Landry

CEO & FOUNDER · LANDRY INDUSTRIES INC. · MONCTON, NB, CANADA

AI-assisted research: Google Gemini & Anthropic Claude

Table of Contents

i	Introduction & Computational Foundations
ii	Taxonomic Initiatives & Key Species
iii	The Robotic & Sensor Revolution
iv	Information-Theoretic Limits of Unsupervised Translation
v	Consumer Commercialization & the Anthropomorphic Illusion
vi	Epistemic, Legal & Ethical Frontiers
vii	Diagnostic Synthesis: The Horizon of Interspecies Translation
—	Works Cited

Introduction & Computational Foundations

The year 2026 represents a critical watershed in our scientific understanding of nonhuman communication — where deep learning, bioacoustic telemetry, and computational ethology converge.

The convergence of deep learning foundation models, high-density bioacoustic telemetry, and computational ethology has established the field of **Nonhuman Animal Language Processing (ALP)**. Moving far beyond historical methods of manual audio classification, modern ALP researchers are mapping the acoustic, gestural, and contextual worlds of nonhuman species using Unsupervised Machine Translation (UMT).

Despite these computational breakthroughs, a profound tension remains between semantic **classification** and genuine **translation**. Deciphering the vocal signatures of aggression, naming, or alarm does not equate to achieving open-ended, two-way dialogic exchange.

Foundation Models for Bioacoustics

The shift from classic bioacoustics to ALP has been accelerated by self-supervised foundation models designed specifically for nonhuman vocalizations. The **Earth Species Project (ESP)** has pioneered this paradigm, releasing:

***NatureLM-audio** — the first large-scale audio-language model tailored for bioacoustics, alongside **AVES** (and its avian counterpart **BirdAVES**), which functions as a self-supervised transformer-based acoustic representation encoder — effectively a "BERT for animals."*

To standardize performance metrics across computational ethology tasks, the academic community relies on **BEANS** (Benchmark of Animal Sounds), the nonhuman equivalent to human natural language benchmarks.

The SETI Parallel

The computational challenges of ALP have established deep parallels with the **Search for Extraterrestrial Intelligence (SETI)**. Because both fields involve decoding completely unknown, nonhuman communication systems with zero access to bilingual parallel corpora, researchers are sharing methodologies. This theoretical overlap was demonstrated in a collaborative study where scientists recorded a structured, 20-minute acoustic exchange with a humpback whale named "Twain," establishing a primitive communication loop as a terrestrial analog for exobiological contact scenarios.

"Deciphering the vocal signatures of aggression, naming, or alarm does not equate to achieving open-ended, two-way dialogic exchange."

Taxonomic Initiatives & Key Species

From sperm whale phonology to elephant naming — the species most illuminated by ALP research reveal startling parallels to human linguistic capacity.

TABLE 1 — KEY TAXONOMIC INITIATIVES IN ANIMAL LANGUAGE PROCESSING (2024-2026)

TARGET SPECIES	LEAD ORGANIZATION	ARCHITECTURE	KEY FEATURE	MILESTONE (2025-2026)
Sperm Whale <i>Physeter macrocephalus</i>	Project CETI	Transformer, Graph Clustering, HMMs	Structured click sequences ("codas")	Phonetic alphabet & human-like coda vowel phonology demonstrated
African Elephant <i>Loxodonta africana</i>	ElephantVoices / Save the Elephants	Random Forest, Variational Autoencoders	Low-frequency rumbles	Verification of non-imitative "vocal labels" (individual names)
Egyptian Fruit Bat <i>Rousettus aegyptiacus</i>	Tel Aviv University (Yovel Lab)	CNN, Multi-class LSTMs	High-frequency aggressive squeaks	Real-time classification into social conflict categories — 88.5% accuracy
Marmoset Monkey <i>Callithrix jacchus</i>	Hebrew University / Partners	Random Forest, Leave-one-session-out	Spontaneous "phee-calls"	Discovery of vocal labeling of conspecifics & social vocal learning
Bottlenose Dolphin <i>Tursiops truncatus</i>	Google DeepMind / Georgia Tech / Wild Dolphin Project	DolphinGemma LLM	Signature whistles & click trains	Unsupervised segmentation of complex aquatic vocalizations at continental scale

Sperm Whales & the Steganographic Hypothesis

Under the leadership of **David Gruber**, Project CETI has focused on decoding the communication structures of sperm whales off the coast of Dominica. A study accepted to **ICLR 2026** by Michael Bronstein and Antonio Norelli reimaged cetacean communication through the lens of steganography — the art of concealing a message within a seemingly benign signal.

Using their **Calgacus protocol**, the team demonstrated that modest 8-billion-parameter open-source LLMs can successfully hide and extract highly coherent messages within text blocks. In April 2026, CETI published findings in *Proceedings B* titled "*The Phonology of Sperm Whale Coda Vowels*," demonstrating that sperm whale codas possess a vocalic architecture representing one of the closest evolutionary parallels to human phonology ever analyzed.

Elephants & Non-Imitative Naming

Researchers led by **Mickey Pardo** and **George Wittemyer**, in collaboration with **Joyce Poole** (ElephantVoices), uncovered that elephants address one another using **non-imitative vocal labels**. Unlike bottlenose dolphins or certain parrots, which address individuals by imitating the receiver's own signature whistle, elephants construct vocal labels that do not sound like the target — indicating a capacity for **abstract, symbolic thought**.

*At the SMU Hackathon (April 2026), students developed **Project Tusk** — an algorithm that isolates low-frequency elephant rumbles while filtering overlapping mechanical interference from aircraft, generators, and land vehicles, significantly expanding training datasets.*

Prairie Dogs, Bats & Primate Dialogues

Prior research by **Con Slobodchikoff** revealed that prairie dog alarm calls vary systematically based on the size, shape, speed, and color of approaching predators. At Tel Aviv University, **Yossi Yovel's** continuous monitoring of Egyptian fruit bat colonies — aligning 91,080 annotated vocal events with continuous video — proved that bat vocalizations contain distinct features identifying both caller and addressee.

Marmoset monkeys similarly use spontaneous "phee-call" dialogues to vocally label their conspecifics, with family members utilizing shared acoustic features — positioning marmosets as a critical evolutionary model for understanding human social language transmission.

The Robotic & Sensor Revolution

The massive datasets required to train ALP foundation models are generated by a parallel revolution in soft robotics, autonomous vehicles, and non-invasive telemetry.

TABLE 2 — ROBOTIC & HARDWARE INNOVATIONS IN BIOACOUSTIC DATA GATHERING (2024-2026)

TECHNOLOGY	DEVELOPER	MECHANISM	DATA / FUNCTION	REFERENCE
"Tap-and-Go" Drone Delivery	Project CETI / Harvard	FPV racing drones deploy suction-cup sensors onto dorsal skin	Minimizes tag-deployment disturbance	PLOS One (Aug 2025)
Passive Vacuum Regeneration	Harvard Robotics / CETI	Bioinspired suction cups powered by pressure changes during deep dives	Extends sensor attachment on marine mammals	IEEE Oceans (Sep 2025)
Multimodal Bio-Logger	Harvard Robotics / CETI	Suction-mounted array: 3 synchronized hydrophones + motion + environmental sensors	High-bandwidth audio; maps angle-of-arrival to isolate individual voices	Harvard SEAS Report (2025)
"Backseat Driver" Glider	Project CETI	SeaExplorer glider with real-time acoustic source separation on 4-element hydrophone array	Passively tracks individual sperm whales via real-time vocalizations	Scientific Reports (Apr 2026)
VAE Anomalous Filter	Save the Elephants	Variational Autoencoders on ultra-low-power edge monitoring devices	Filters redundant acoustic data; retains rare vocalizations — 80-85% accuracy	Save the Elephants Tech Paper (Feb 2026)

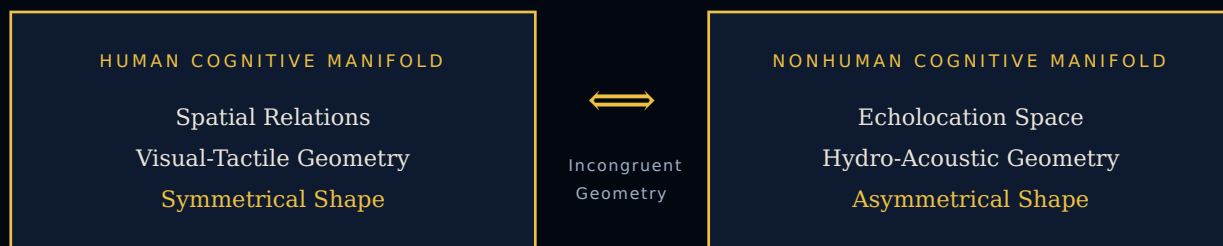
This hardware ecosystem has resolved several historical bottlenecks. The **"cocktail party problem"** — isolating individual voices within densely overlapping animal groups — has been mitigated by both the Harvard tri-hydrophone bio-logger and Earth Species Project's source-separation algorithms.

On terrestrial landscapes, deployment of low-power, unsupervised edge devices running compressed VAE architectures allows researchers to monitor wildlife across vast geographies without exhausting local memory storage, capturing critical behavioral changes in real time.

Information-Theoretic Limits of Unsupervised Translation

The theoretical pathway to true interspecies translation relies on Unsupervised Machine Translation — a field whose mathematical constraints are as revealing as its possibilities.

In human UMT, two distinct languages are transformed into high-dimensional geometric shapes (manifolds) where semantically similar words occupy adjacent spatial coordinates. Because human experiences of gravity, spatial relations, and emotional states are universal, these geometric shapes align, permitting direct coordinate-to-coordinate translation.



The Goldwasser-Gruber-Kalai-Paradise Framework

To address the "domain gap" arising from divergent cognitive worlds (*Umwelten*), Shafi Goldwasser, David Gruber, Adam Tauman Kalai, and Orr Paradise formulated an information-theoretic framework for analyzing UMT between languages with highly disparate structural properties and no parallel bilingual data. The framework establishes that unsupervised translation is mathematically possible under two conditions:

Prior Access — The system has access to an informational "prior" distribution over the target language that can estimate the statistical likelihood that a candidate sentence is a valid translation.

High-Probability Rejection — Most incorrect candidate translation functions output sentences deemed highly implausible by the prior, allowing systematic elimination of invalid translation pathways.

SAMPLE COMPLEXITY FORMULA

$$N \geq (d \cdot \log(1/\delta)) / \kappa^2$$

Where:

- N = required sample volume (unlabeled data)
- d = dimensional complexity of the latent language space
- δ = target error bound
- κ = shared common ground parameter
(overlap in semantic states / physical interactions)

Critical implication:

As $\kappa \rightarrow 0$ (e.g., mapping human symbols onto the hydro-acoustic sensory world of a deep-diving sperm whale), $N \rightarrow \infty$

This inequality demonstrates that the mathematical feasibility of UMT in animal communication depends not on computing power, but on the **physical, ecological, and cognitive overlap** shared between human and nonhuman subjects.

"If κ approaches zero, the required sample complexity approaches infinity — regardless of computational power."

Consumer Commercialization & the Anthropomorphic Illusion

The rapid acceleration of academic ALP has generated a parallel consumer market — one that blurs the boundary between genuine translation and emotionally satisfying projection.

By 2035, the AI in Pet Care Market is projected to reach **\$15.34 billion**, driven by pet owners' desire to decode the emotional and physical states of companion animals.

Case Study: PettiChat

A prominent example is **PettiChat**, a \$118, 27-gram AI-powered collar launched on Kickstarter in April 2026 by the Chinese startup *Meng Xiaoyi* ("cute little translator"). Backed by \$1 million in angel funding and securing over 10,000 pre-orders within weeks, the device claims to translate dog and cat vocalizations into complete human sentences with **94.6% real-time translation accuracy**.

Developed in partnership with Zhejiang University's College of Animal Sciences, PettiChat captures raw audio and motion data locally, then transmits feature coordinates to Alibaba Cloud's **Qwen LLM** platform, processed through a proprietary classifier trained on five million voiceprint samples from over 1,000 animals.

*The system does not translate syntax. It **classifies emotional states** across 20+ distinct categories (excitement, frustration, fear, hunger, etc.) and uses a generative LLM to reconstruct these into playful, human-readable sentences such as "Feed me, I am hungry!"*

The Accuracy Problem

The claimed 94.6% accuracy is a managed corporate metric. Independent bioacoustic validation indicates the physical ceiling for classifying dog vocalizations from acoustic

signals alone under laboratory conditions is approximately **57.3%**. Even integrating multi-modal telemetry, the theoretical ceiling remains **capped at 89%**.

The discrepancy is resolved by **anthropomorphic projection**: pet owners naturally project human-like intentions onto minor behavioral variations. When a generative AI collar outputs a sentence aligned with an owner's pre-existing emotional narrative (e.g., "I don't like when you work so late!"), the output is experienced as highly accurate — despite the absence of semantic proof.

While these tools are valuable for identifying broad health alerts, stress levels, or veterinary pain indicators, they represent **creative, generative reconstructions** of emotional states rather than literal language translation.

Epistemic, Legal & Ethical Frontiers

As ALP technologies transition from experimental laboratories to the public sphere, they are challenging long-standing anthropocentric paradigms in environmental law, animal ethics, and conservation biology.

Taxonomic Bias & Selective Silencing

A 2025 systemic analysis of 14,600 conservation projects over 25 years revealed profound funding bias: **29%** of resources were allocated to species of "least concern," whereas only **6%** directly supported threatened species. This is mirrored in ALP, which concentrates on "charismatic megafauna" while systematically ignoring highly exploited agricultural species (pigs, cows, chickens).

Ethicists argue this dynamic is maintained because acknowledging the linguistically legible preferences, pain states, or active resistance of agricultural animals would threaten the global food industry — a phenomenon called **selective silencing**.

The Interspecies Epistemic Tiers

TIER 1

Acute Affective States

Immediate emotional expressions such as terror, excitement, or acute pain. Highly correlated with physiological behaviors and classifiable by modern ML — but does not offer insight into the animal's unique conceptual world.

TIER 2

Continuous Aversions & Preferences

Expressed states of sustained desire, such as a captive animal's aversion to its cage. Technically viable but disruptive — directly challenging the legal justification for zoological confinement, shifting "ownership" toward legal "guardianship."

Narrative & Relational Inner States

Complex cognitive processes: long-term grief, existential reflection, abstract familial memories. These lack immediate behavioral markers and remain opaque behind the interspecies veil — beyond current model capacity.

The MOTH Paradigm: Jurisprudential Transformation

In October 2025, researchers from Project CETI and the **NYU More-Than-Human Life (MOTH) Program** published "What if We Understood What Animals Are Saying?: The Legal Impact of AI-Assisted Studies of Animal Communication" in *Ecology Law Quarterly*. The authors argue that empirical proof of phonetic alphabets and non-imitative naming undermines the legal justification for treating animals as property, strengthening arguments for:

Legal Personhood — Building on Ecuador's 2008 constitutional recognition of Rights of Nature, the Los Cedros ruling, and Argentina's habeas corpus case releasing Cecilia the chimp.

Right to Culture — Legally protecting shared social structures, parenting protocols, and distinct communication dialects from human disruption.

Acoustic Sanctuary — Regulating undersea noise pollution, shipping lanes, and military sonar as violations of marine mammals' right to communicate.

The PEPP Framework

To address the regulatory vacuum surrounding Nonhuman Animal Communication Technologies (NACTs), the NYU MOTH Program released the **PEPP Framework** (November 2025) — twelve principles across four pillars for voluntary adoption by researchers, engineers, and funders:

TABLE 3 — THE PEPP FRAMEWORK FOR NACT GOVERNANCE (NOVEMBER 2025)

PILLAR	PRINCIPLE	CORE OBLIGATION	OPERATIONAL IMPACT
	Research Design & 3Rs		Forces institutional review before NACT deployment

PILLAR	PRINCIPLE	CORE OBLIGATION	OPERATIONAL IMPACT
Prepare		Adhere to rigorous experimental design; prioritize reduction, replacement, and refinement of animal subjects	
	Transparency & Governance	Openly publish ethics protocols, share data risk assessments, and report all unintended harms	Minimizes proprietary, closed-source animal testing
Engage	Stakeholder Inclusion	Consult independent scientific experts and Indigenous knowledge-holders sharing lands/waters with subjects	Integrates Western ethology with ancestral ecological knowledge
	Non-Coercion	Respect animal autonomy; facilitate only non-coerced participation and disengage if avoidance appears	Eliminates invasive restraint during acoustic recording
Prevent	Precautionary Principle	Cease any planned NACT activity if threats of serious, irreversible harm to animals or ecosystems arise	Shifts burden of proof of safety onto technology developers
	Misuse Mitigation	Anticipate and block potential commercial, military, or recreational exploitation of synthesized animal signals	Limits development of weaponized or poaching-related bioacoustics
Protect	Best Interests of the Animal	Let physical, mental, relational, and environmental best interests guide every computational design decision	Centers animal welfare in engineering design choices
	Remediation of Harms	Urgently address and remediate any direct or indirect ecological harms caused by NACT deployment	Establishes legal accountability for computational ethologists

Ethicists warn that generative AI capable of "talking back" to animals represents a profound ecological danger. Emitting synthetic AI-generated calls into the wild could cause severe cognitive distress, disrupt mating patterns, or fracture social bonds. Poachers could exploit synthesized contact calls to lure endangered species; commercial fleets could deploy simulated dolphin signals to herd marine life into nets.

Diagnostic Synthesis: The Horizon of Interspecies Translation

A clear bipartite distinction is required to evaluate how close humanity truly is to meaningful "discussion" with animals in 2026.

Structural Decoding & Intentional Classification: Extremely Close

If "discussion" is defined as the automated, high-fidelity classification of emotional states, context-specific intentions, and individual identities — such as detecting whether a bat is arguing over food, or recognizing which elephant is being addressed by name — humanity is **remarkably close**. State-of-the-art models like NatureLM-audio and DolphinGemma have successfully mapped the baseline structural units and acoustic alphabets of several species, transforming bioacoustics from passive listening into systematic decoding.

Reciprocal, Open-Ended Semantic Dialogue: Profoundly Distant

If "discussion" means a mutual, open-ended exchange of abstract ideas — a literal "Google Translate" converting a whale's inner thoughts into English sentences — humanity remains **far from this goal**. The primary bottlenecks are not merely computational:

The Data Asymmetry — Modern human LLMs are trained on trillions of words of written text. Nonhuman animal datasets, even the most extensive, are many orders of magnitude smaller. Because animal communication is continuous, lacks a written form, and is deeply embedded in physical behavior, gathering the necessary scale of multimodal data cannot match the scaling velocity of human AI.

The Geometry of Subjective Experience — Unsupervised machine translation requires that the semantic structures of two languages share a congruent geometric shape. A sperm whale's or bat's sensory experience — built on echolocation or ultrasonic tracking — may share very little "common ground" with human language. If the geometric shapes of our respective cognitive worlds do not align, no computational power will find a coherent mathematical mapping function.

"The future of interspecies communication lies not in forcing animals to speak human languages through creative AI generation, but in using computational tools to appreciate their unique cognitive structures."

The future lies in establishing rigorous ethical standards to protect animal habitats and allowing their decoded voices to reshape environmental law and human ecological responsibility. The question is not whether we can make animals speak our language, but whether we are willing to learn theirs.

WORKS CITED

- [1] Parley — The Organization Using AI to Unlock Communication with the Animal Kingdom. parley.tv/journal/earth-species
- [2] Earth Species Project — earthspecies.org
- [3] A Theory of Unsupervised Translation Motivated by Understanding Animal Communication — NeurIPS 2023. proceedings.neurips.cc
- [4] AI-assisted animal language translation — Wild Animal Initiative. wildanimalinitiative.org/blog/ai-animal-translation
- [5] Are We Ready to Hear a Non-Human 'No'? — EA Forum. forum.effectivealtruism.org
- [6] PettiChat AI collar launch — Make Tech Easier. maketechesasier.com
- [7] Bridging Species: Legal and Ethical Implications of AI-Assisted Animal Communication — American Bar Association, Winter 2025. americanbar.org
- [8] AI-enabled Decoding of Whale Communication — NYU Law. law.nyu.edu/news/ideas/cesar-rodriguez-garavito-animal-communication-ai
- [9] Project CETI — Research. projectceti.org/research/index
- [10] Like Us, Elephants Have Names Too! — ElephantVoices. elephantvoices.org
- [11] iMasons / SMU Hackathon — Project Tusk. imasons.org/blog
- [12] Vocal Labeling of Others by Nonhuman Primates — ResearchGate. researchgate.net
- [13] Researchers translate bat language — ZME Science. zmescience.com
- [14] AI-Driven Communication Analysis for Biodiversity Conservation: A Fruit Bat Prototype — ResearchGate. researchgate.net
- [15] Animal-Borne Adaptive Acoustic Monitoring — Save the Elephants Technical Paper (Feb 2026). savetheelephants.org
- [16] China's AI pet translator sells a dream that animal science cannot yet prove. webiano.digital
- [17] Legal and Ethical Principles for NACTs — NYU MOTH Program (Nov 2025). mothlife.org
- [18] The PEPP Framework: Ethical Standards for NACTs. ouranimalkin.com/pepp-framework
- [19] Listening Is Not Neutral: New Ethical Guardrails for Animal Communication Technologies — Bioneers. bioneers.org
- [20] AI for Decoding Non-Human Languages — Gemma's Open Research. open-research.gemmadanks.com



Landry Industries Inc.

This intelligence monograph was produced by **Marie-Soleil Seshat Landry**, CEO & Founder of Landry Industries Inc., Moncton, New Brunswick, Canada. Research and typesetting was AI-assisted using Google Gemini and Anthropic Claude.

Published 2026 under the Landry Industries Intelligence Division.

MARIELANDRYSPYSHOP.COM

LANDRYINDUSTRIES.CA

[ZENODO ARCHIVE](#)

[ORCID 0009-0008-5027-3337](#)