

Movement Ethology: Body Kinematics Meets Relocation Data

Wayne M. Getz^{a,b,1} and Joris Komen^{c,d}

This manuscript was compiled on October 26, 2026

Movement Ethology, as we envision it, links an individual's body positions and actions to the ecological functions of its movement behavior, during both displacement and approximate stationarity. Movement Ecology, underpinned by relocation data, studies changes in the location of whole organisms, explaining path selection through internal state, motion and navigation capacities, and external factors. Computational Ethology, underpinned by body-pose sequences, infers behavior from the coordinated motion of body parts. These fields measure complementary aspects of animal movement, yet share few concepts and data standards. Fixed-length sequences of video-derived postures are clustered into an alphabet of body-pose syllables (BoPS). Fixed-length strings of BoPS and corresponding-length segments of regularly sampled relocation tracks represent two ways of generating Shannon-coded statistical movement elements (StaMEs), pose-derived P-StaMEs and location-derived L-StaMEs, temporally aligned for mutual information analysis. Pose sequences of variable duration, identified as relevant to Movement Ecology using machine learning, are the fundamental movement elements (FuMEs) posited in the Movement Ecology paradigm. Behavioral words built from fixed-length BoPS strings contain body-part movement information, allowing for a richer multi-task canonical activity mode (CAM) labeling, such as “scanning the horizon while walking,” than can be obtained from relocation data alone, while overlaid FuME sequences show how CAMs are executed. Independent behavioral observations test whether P-StaMEs distinguish behaviors that L-StaMEs alone cannot resolve. Linked P- and L-StaMEs generate testable hypotheses connecting posture, gait and action to space use, energetics and individual health, enabling a finer understanding of behavior along lifetime tracks.

movement ecology | computational ethology | pose estimation | path segmentation | information theory

Two research communities study animal movement, one at the level of the whole individual and the other at the level of the articulated body, but as yet there has been little cross-fertilization. Movement Ecology defines movement as “a change in the spatial location of the whole individual in time” (1). It asks why, how, when and where organisms move, and with what ecological and evolutionary consequences. Its primary data are relocation time series from GPS and other tags, now collected at volumes that make Movement Ecology a big-data science (2, 3).

Ethology asks Tinbergen's four questions about behavior (4): its causation, development, function and evolution. Much of what ethologists observe is the movement of body parts rather than of the whole body: a head turn, a stride, a rear, a groom. Computational Ethology now measures these movements using machine vision (5). Markerless pose estimation tracks user-defined keypoints in ordinary video (6), and unsupervised models parse pose image time series into sub-second behavioral “syllables” (7, 8). Along with Movement Ecology, Computational Ethology is becoming a big-data science (9–11).

The Movement Ecology paradigm anticipated a hierarchy extending below the resolution of whole-animal relocation data, but could not measure its lowest level. Nathan et al. (1) wrote that movement phases “are made up

Significance

Animal movement is studied by two different communities. Movement ecologists use GPS location point sequences to track where whole animals range over extended periods of time. Computational ethologists use video and machine learning to measure how body parts move over short periods of time, mostly in the laboratory. We propose Movement Ethology to link the study of body configurations and actions in relation to local surroundings with the ecological functions of animal movement over landscapes. Pose and accelerometer data reveal locomotor cycles invisible to relocation data, and information measures test, using Shannon information theory, whether body codes add behavioral information beyond path codes. Synchronized observations and independent validation test how body data improve models of space use and energetics, and evaluate movement decisions and individual health.

Author affiliations: ^aDepartment of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA; ^bSchool of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa; ^cCollege of Environmental Design, University of California, Berkeley, CA, USA; ^dColossal Biosciences, Austin, TX 78701, USA

WMG proposed the new field of Movement Ethology, wrote the first draft, and contributed to the figures. JK edited the manuscript and contributed to the figures. Many of the ideas arose through extensive discussions between WMG and JK.

The authors declare no competing interest.

¹To whom correspondence should be addressed. E-mail: wgetz@berkeley.edu

of smaller units that we call canonical activity modes ... such as runs and tumbles in bacteria and standing, walking, running, and gliding in larger animals, much as genes coding for proteins are composed of a limited number of trinucleotide codons.” Getz and Saltz (12) placed fundamental movement elements (FuMEs) at the base of that hierarchy. Yet FuMEs “are characterized by the movement of body parts while location data apply to a body as a whole” (13). They are unobservable in relocation data alone and require “video, accelerometer, or other body-parts movement data for identification” (14). Movement ecologists have therefore substituted statistical movement elements (StaMEs) computed from short runs of relocation steps (15).

Computational Ethology produces the body-part movement data that relocation records lack, but mostly in laboratory arenas, over minutes to hours, and with little landscape context. Movement Ecology has the context, the long time scales and the ecological questions, but not body-part movement itself. Body-part and movement-path data are therefore rarely recorded on the same animals, on the same clock or in a shared representation. We propose to link these two areas of research under the sobriquet **Movement Ethology**. We define this new research field as the study of body-part positions, configurations and actions relative to the local environment and other individuals, relating them to ecological functions and to patterns of displacement and approximate stationarity (i.e., stationarity within the accuracy of the relocation data) throughout an animal’s lifetime. It provides a body-part level foundation for Movement Ecology and draws on Computational Ethology methods for pose estimation and behavioral quantification. Sequences of point location (i.e., relocation) data, augmented by other kinds of data (accelerometer, contextual environmental), drive Movement Ecology research, while Movement Ethology resolves the body parts within that point and relates its actions to the surrounding world.

Movement Ecology and Computational Ethology are heavily method-oriented. Movement Ethology, on the other hand, is more defined by questions than methods. For example, it is concerned with: what body-part actions produce displacement; what individuals are doing during approximate stationarity along their movement tracks; what types of body-part actions contribute to ecologically relevant behaviors such as resource acquisition, sensing the environment, and social interactions of an ecological nature (resource competition, predator avoidance). Section 5 and Table 3 develop these questions as testable hypotheses.

Just as relocation data sets form a foundation for Movement Ecology research, so we propose sequences of BoPS to form the foundation for Movement Ethology research. Thus it becomes very natural for us to extend Movement Ecology’s path segmentation (16) and Shannon’s hierarchical path-segment coding (13, 15) paradigms to Movement Ethology’s pose sequence segments. Beyond this coding hierarchy, other methods that connect body-level behavior, ecological context, and movement history will also be encompassed by our proposed Movement Ethology rubric. In particular, when movement paths dissolve into approximate stationarity—i.e., repeated location measurements cannot be resolved beyond a single point location—Movement Ethology methods become

essential in determining the ecological relevance of the observed periods of approximate stationarity.

How do we envision Movement Ethology as distinct from the current field of Computational Ethology? One difference is the context in which the data are interpreted and through that interpretation the questions that are then asked. We accept that Computational Ethology already studies naturalistic behavior and ecological context (5, 17), and field applications are growing (Section 1), but the focus is local or inward-looking with respect to time and space and the individual as an entity in its own right. Movement Ethology draws on the same methods in organizing body observations but is more outward-looking in its view of the individual as wide-ranging, as a member of a population, and as an entity within an ecosystem.

Our implementation rests on three constructs:

1. A single hierarchy spans both fields: aligned body and path representations connect postures, body-pose syllables (BoPS) built from fixed-length posture sequences, behavioral words and FuMEs to statistical movement elements (StaMEs), activity modes and lifetime tracks (Section 2).
2. Body configurations, orientations and measured environmental and social relationships constrain models of motion, navigation and internal state within the Nathan et al. (1) framework, processes so far inferred largely from paths (Section 4).
3. Shannon coding theory, recently applied to track segmentation (14), provides the theory needed to compare codes across levels, sensors and species. Each data stream is coded separately, BoPS and path codes are aligned on shared time intervals, and their information content is compared. This supports two claims that require different tests. First, intervals assigned the same location-derived StaME (L-StaME) can receive different pose-derived StaMEs (P-StaMEs), built from the temporally aligned BoPS track, so P-StaMEs can separate intervals that L-StaMEs merge. Second, independently validated behavioral labels test whether P-StaMEs carry behavioral information beyond L-StaMEs (Section 3.4). Analysis proceeds in two stages: P- and L-StaMEs are first built, paired and strung into P- and L-CAMs without behavioral labels, and expert-labeled FuME sequences are then laid over the resulting CAM tracks (Sections 3 and 5; SI Appendix, Fig. S3).

The SI Appendix provides a glossary of acronyms and symbols (section S1, Tables S1 and S2), extended tables (section S2, Tables S3–S9), a prospective workflow for elephants (section S3), the BoPS and path codes and the construction of P-StaMEs (section S4 and Fig. S1), sensors, clocks and reference frames (section S5), the information measures computed from these codes and their estimation, including paired codes and the two stages of analysis (section S6, Figs. S2 and S3 and Table S11), and a schematic state-space model linking pose and path (section S7).

1. The Precursors

Extraordinary advances in remote sensing and computation have transformed both Movement Ecology (3) and Computational Ethology (18) over the past two decades,

but in different ways (Table 1). Movement Ecology gained in sampling frequency and spatial extent; Computational Ethology gained anatomical resolution and behavioral pattern identification. Movement Ethology connects these gains within an ecological framework.

Table 1. Movement Ecology, Movement Ethology and Computational Ethology compared, in brief. Entries are typical emphases; the fields overlap. Full version: SI Appendix, Table S3.

	Movement Ecology	Movement Ethology	Computational Ethology
Unit	Whole-body location	Body-pose in its local and social context	Body-part keypoints
Sensors	GPS, reverse GPS, accelerometers	Synchronized GPS, pose and inertial sensors	Video with pose estimation
Sampling	0.1 s to days	Aligned pose and GPS clocks	10–100s of frames per second
Record	Minutes to a lifetime	Activity bouts to lifetimes	Seconds to hours
Setting	Free-ranging, landscape scale	Captive or drone-filmed; wild via collars	Mostly laboratory arenas
Lowest unit	StaME	BoPS; FuME (one locomotor cycle)	Pose, or syllable (dynamic motif)
Questions	Space use, migration, energetics	How body actions shape ecological function	Neural control, action selection, social behavior

Since 2008, GPS and reverse-GPS systems such as ATLAS (19) have given Movement Ecology relocations at intervals of tenths of seconds to days, for periods of months to years (2, 3). Tracks are segmented into behavioral modes using step lengths and turning angles, behavioral change-point analysis (20) and hidden Markov models (21), as reviewed by Edelhoff et al. (16). Tri-axial accelerometers added a body-level auxiliary data channel (e.g., Nathan et al. (22) classified behavioral modes of griffon vultures from acceleration bursts), AcceleRater made supervised classification of such modes widely available (23), and dynamic body acceleration became a proxy for energy expenditure (24, 25). Accelerometry is Movement Ecology’s existing foothold below the whole-body level, but a tag with an accelerometer records one point on the body rather than multi-point postures.

Anderson and Perona (5) articulated a program for automated, quantitative measurement of behavior, a call later renewed for Computational Neuroethology (17). Berman et al. (26) embedded the postural dynamics of fruit flies in a map of stereotyped behaviors. Wiltshko et al. (7) found that mouse behavior in depth video decomposes into sub-second syllables strung together by a probabilistic grammar. Markerless pose estimation then made keypoint tracking routine: DeepLabCut (6), LEAP (27), multi-animal DeepLabCut (28) and SLEAP (29). Unsupervised pipelines turn keypoints into motifs, notably B-SOiD (30) and keypoint-MoSeq (8). Pretrained SuperAnimal models track keypoints across more than 45 species without new labels (11).

A few studies have already begun to bridge the two fields, connecting body behavior and ecological context through drones, computer vision and on-animal sensors that now work outside the laboratory (31). Koger

et al. (32) combined drone video, keypoint tracking and 3D landscape reconstruction for Grevy’s zebras and other ungulates in Kenya and geladas in Ethiopia. Wiltshire et al. (33) applied DeepLabCut to wild chimpanzees and bonobos. Chakravarty et al. (34) used biomechanical principles to classify meerkat behavior from accelerometers. Berman et al. (35) showed that fly behavior is hierarchically organized and predictable over long time scales, an information-theoretic result that parallels the track-coding approach of Getz (14). Software is also beginning to bridge the gap. For example, the open-source *movement* package (36) processes the outputs of DeepLabCut, SLEAP and other trackers for analysis of body movement across space and time.

2. A new lowest rung for path segmentation

Relocation data are used to build the upper rungs of the Movement Ecology hierarchy (13, 14) but not its base, because FuMEs require body-part movement data for their construction. Movement Ethology, through its mobilization of body-part data, connects finer observations of the animal’s body to the organization of its lifetime track (LiT). Body-part observations are generally obtained from video, in which consecutive frames record a sequence of poses. Computational Ethology methods differ in how they segment that sequence. MoSeq and keypoint-MoSeq (7, 8) fit an autoregressive hidden Markov model to the pose time series, so their units are dynamic motifs of variable duration, typically a few hundred milliseconds; Wiltshko et al. (7) called these units “syllables” and their transition structure a “grammar.” A body-pose syllable (BoPS) is the fixed-length counterpart of such a syllable. It is obtained by clustering sequences of a fixed number of consecutive poses (video frames), never a single frame, so every BoPS spans the same short duration, just as a StaME is obtained by clustering runs of a fixed number of relocation steps (15). Strings of BoPS form behavioral words, whereas FuMEs, complete locomotor cycles of variable length, are detected directly from poses (Section 3.2); a MoSeq syllable that spans a complete locomotor cycle likewise corresponds to a FuME.

Because every BoPS has the same duration, a fixed number of consecutive BoPS can be chosen to span exactly the duration of the segment of relocation steps from which an L-StaME is built (37). StaMEs derived from BoPS, pose StaMEs (P-StaMEs), and from relocations, location StaMEs (L-StaMEs), thus provide two complementary ways of coding the same track segments. Comparing them with Shannon information theory, including how each codes canonical activity modes (CAMs: fixed strings of StaMEs clustered into movement types, called P-CAMs or L-CAMs according to the StaMEs used; Fig. 1), leads to the new insights discussed in Section 5.

3. From video to behavioral words: an alphabet of poses

The coding scheme developed for relocation data (14) can be applied to pose time series. An *alphabet* of BoPS supports a *lexicon* of behavioral words. A separate sequence records complete locomotor cycles as FuMEs, which overlay CAMs rather than StaMEs. These aligned representations preserve concurrent actions before ecological interpretation (Fig. 1); Section 3.3 defines their

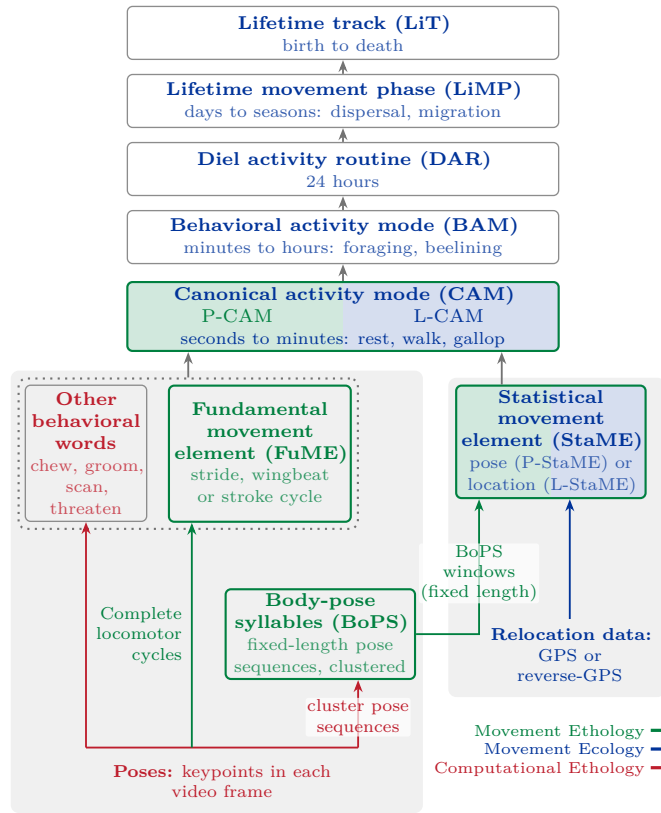


Fig. 1. Movement Ethology in a nutshell. Body-part data (keypoints in each video frame) supply postures, which feed three units directly: complete locomotor cycles of variable length form FuMEs, other actions are coded as behavioral words, and fixed-length sequences of consecutive postures are clustered into an alphabet of body-pose syllables (BoPS). Fixed-length strings of BoPS form P-StaMEs, and fixed-length strings of P-StaMEs form P-CAMs, which FuMEs and other behavioral words (dotted box) also inform. Relocation data are used to construct L-StaMEs (37). Tracks coded with size-matched L- and P-StaMEs can be compared using behavioral observations and Shannon information theory, as can those coded with L- and P-CAMs. Also, the FuMEs that overlay these segments can be used to obtain additional behavioral insights into the structure of CAMs and the movement segment hierarchy of behavioral activity modes (BAMs) and diel activity routines (DARs) (14, 38) and into the state and health of individuals. Colors identify the field each element comes from (key, lower right): green for Movement Ethology, dark blue for Movement Ecology, and red for Computational Ethology (postures and behavioral words). The CAM and StaME boxes are shaded green on the side informed by pose-derived and blue on the side informed by location-derived codes.

relationships with the path. The equations and estimation details are given in SI Appendix, sections S4 and S6, and the order in which BoPS and path codes are analyzed in SI Appendix, Fig. S3.

3.1. From frames to an alphabet of BoPS. Each video frame is first reduced to time-stamped positions of K identified keypoints, each with a confidence score, for every identified individual, using markerless pose estimation (6, 28, 29) or pretrained multi-species models (11). The keypoints are then put into the animal's own reference frame: centered on the body, rotated to the body axis, scaled by the individual's body length and, in 2D video, reflected to a standard facing direction, so that the same posture is not split by facing direction or viewing angle (SI Appendix, section S4). The result is an egocentric pose vector q_t . Whole-body location u_t and orientation R_t are kept separately, and so, where observable, is a contextual record c_t of relative positions, distances and contacts involving environmental features and neighboring animals, so that

the contributions of body shape, displacement and context can be compared.

The pose track is then cut into consecutive windows of a fixed number γ of frames, each spanning $\gamma\Delta$ seconds on a pose clock of Δ seconds per frame. Each window, a short sequence of egocentric poses, is reduced in dimension and clustered, by cluster analysis or other machine-learning methods, into a archetypal pose sequences, the body-pose syllables, which form the BoPS alphabet $\mathcal{P} = \{\pi_1, \dots, \pi_a\}$. Assigning each window to its nearest BoPS turns video of an animal into a BoPS track, with one symbol per window. A BoPS is therefore never obtained from a single frame: it describes a brief movement of fixed duration, or a maintained posture when the poses within the window barely change. The window length $\gamma\Delta$ is chosen short relative to the locomotor cycles of interest, so that a stride or wingbeat spans several BoPS. Strings of BoPS build behavioral words (Section 3.2) and windows of the BoPS track build P-StaMEs (SI Appendix, section S4); FuMEs are detected separately from the pose sequence. As with the StaME number of Getz (14), the alphabet size a and the window length γ are tunable parameters, chosen by the trade-off between coding entropy and distortion (SI Appendix, section S4). Like a syllable in speech, a BoPS carries meaning mainly in combination with others.

3.2. From strings of BoPS to behavioral words. Following Tasks 3–6 of Getz (14), the BoPS track is cut into fixed-length strings of m consecutive BoPS. Each string is described by its BoPS sequence and by egocentric kinematic features, such as keypoint velocities and the periodicity of limb motion; displacement is deliberately left out, so that it is not built into the BoPS features later compared with the path. The strings are clustered into b word types, which form a lexicon $\mathcal{W} = \{\omega_1, \dots, \omega_b\}$. Because exact BoPS sequences rarely recur, each string is reassigned to the nearest word under a string distance that tolerates jitter and phase, the analogue of the rectification rule of Getz (14). Once fitted, the lexicon is frozen and new strings are assigned to the nearest word. Runs of a repeated BoPS encode how long a posture is held, so words capture both the order and the timing of postures. Runs of mixed BoPS can encode repeated movements that no single BoPS captures, such as head nodding.

P-StaMEs and the P-CAMs built from BoPS support behavioral interpretation at different scales. L- and P-StaMEs share segment length and timing, so each StaME interval of μ relocation steps has one of each. L-StaMEs, the lowest level of the Movement Ecology hierarchy, are generally interpreted through step length, turning angle and displacement; their paired P-StaMEs can split intervals that share an L-StaME, such as walking while chewing and walking while scanning. Richer interpretations, such as searching or directed movement, emerge only at the L-CAM level, and P-CAMs aligned one to one with L-CAMs can split these in turn, for example searching with the head close to the ground from searching with the head held high (SI Appendix, section S4).

FuMEs are coded separately from BoPS, and their catalog is not constrained by fixed intervals. Each FuME is a phase-locked limb, wing or fin cycle bounded by successive occurrences of a reference phase, such as touch-down of a limb, and is typically identified in the pose sequence, or in accelerometer signals, by detectors trained

Table 2. Body actions and candidate functional categories, which may co-occur. Context and independent observations are needed to validate functional labels. Further details are given in SI Appendix, Table S5.

Category	Examples	Link to the hierarchy
Locomotion (FuMEs, $\mathcal{F}$)	Stride, wingbeat or stroke cycles; turning steps	FuME sequences overlay L- and P-CAMs (Eq. 1)
Rest and postural transitions	Lie down, rise, rest	Approximate stationarity; DAR timing
Ingestion and resource handling	Bite, chew, drink, dig	Foraging BAMs
Sensing and orientation	Scan, sniff, freeze	Navigation; BAM transitions (H5)
Maintenance and thermoregulation	Groom, wallow, pant	Internal state; DARs
Social and communicative	Greet, display, play	External factors r ; group movement
Agonistic and antipredator	Threaten, fight, flee	Risk; escape bursts
Reproductive and parental	Court, mate, nurse	Breeding LiMPs
Elimination and marking	Defecate, scent-mark	Marking routes; health
Anomalous and clinical	Stereotypies, lameness	Welfare monitoring
Unassigned	Occlusion, tracking error	Flagged as missing data

on expert-labeled examples. The cycle types form a locomotor catalog $\mathcal{F} = \{\phi_1, \dots, \phi_f\}$, distinct from the behavioral lexicon $\mathcal{W}$ (Fig. 1), and displacement is not used to identify them. Because FuMEs vary in duration, they are not paired with fixed-length StaMEs; sequences of FuMEs instead overlay CAMs (Section 3.3). Concurrent actions are retained: a stride taken while chewing remains a FuME, and the chewing remains a word in the behavioral record. Table 2 groups candidate interpretations of co-occurring body actions (SI Appendix, Table S5).

3.3. Coding the body and the path. StaMEs and CAMs can be treated as symbols of a Shannon code, whose entropy, assignment error and Jensen-Shannon divergences compare codings and clustering methods (14, 37). We write S^L and S^P , with types σ_i^L and σ_i^P , for the L- and P-StaMEs of a StaME interval, and C^L and C^P for the L- and P-CAMs of a CAM interval; SI Appendix, sections S4 and S6 give formal definitions and estimators.

FuMEs overlay CAMs, which span many StaMEs: the FuMEs whose reference phase falls within a CAM interval, whether of an L-CAM or a P-CAM, form the FuME sequence of that CAM. For any movement segment that is coded by a substantial number of CAMs, let C be the CAM (C^L or C^P) of a randomly chosen CAM interval and Ψ a prespecified summary of its overlaying FuME sequence, such as the composition, rates or order of FuME types. An empty sequence means that no complete cycle was detected or properly identified in the interval. The animal may have been stationary, passively transported or moving without completing a cycle (SI Appendix, section S6). The coupling between CAM and FuME codes is then the mutual information (39)

$$I(C; \Psi) = H(C) - H(C | \Psi), \quad [1]$$

where $H(C)$ is computed from the frequencies of CAM types across the segment's CAM intervals and $H(C | \Psi)$ from how often each CAM type co-occurs with each FuME summary; $I(C; \Psi)$ is thus the reduction in uncertainty about the CAM given its overlaying FuMEs, in bits. A value near $H(C)$ means CAMs are statistically predictable from their FuMEs, without establishing a causal mechanism. A low value means the same FuME sequences occur within many CAM types. That can reflect terrain, wind, current or social context, but measurement error, coarse codes and misaligned clocks must first be excluded. Conversely, $H(\Psi | C)$ measures how much locomotor variation a CAM type merges: a directed-travel L-CAM, for example, may contain both walking and trotting strides, which the FuME overlay separates. Because $I(C; \Psi) \leq \min\{H(C), H(\Psi)\}$, the normalized coupling $I(C; \Psi)/H(C)$ used in H1 also falls when locomotion becomes rarer or less varied; it is therefore compared among locomotor CAMs, with $H(\Psi)$ reported. SI Appendix, section S6 gives the plug-in estimator, the controls needed to compare habitats, ages or species, and the treatment of cycles that cross CAM boundaries.

Both locomotor and other actions contribute to interpreting the trajectory. During approximate stationarity, chewing, scanning or resting can distinguish intervals with similar path geometry. During displacement, the same actions can distinguish walking while feeding from walking while inspecting the surroundings. Let B be a behavioral label supported by observations independent of the codes being evaluated, and A a prespecified summary of the behavioral words of a StaME interval, such as the word types present or the fraction of the interval each occupies (SI Appendix, section S6). The additional information associated with the BoPS record is

$$I(B; A | S^L) = H(B | S^L) - H(B | S^L, A). \quad [2]$$

A large value indicates that the behavioral-word summary reduces uncertainty about the independent label beyond that resolved by the L-StaME.

3.4. Comparing L- and P-StaMEs. The two constructions yield separately fitted codebooks whose symbols are paired by time, not translated into one another, so each valid StaME interval j yields one paired observation (S_j^L, S_j^P) (SI Appendix, section S4, Figs. S1 and S2, and Table S6). The directional association between them is the fraction of L-StaME uncertainty reduced by P-StaMEs,

$$U(S^L | S^P) = \frac{I(S^L; S^P)}{H(S^L)} \in [0, 1], \quad H(S^L) > 0. \quad [3]$$

U tests whether pose predicts the path only when P-StaMEs are built from BoPS content, alone or with gait descriptors; displacement descriptors from georeferenced video or dead reckoning come from the motion that the path also records and can inflate U , which should therefore be reported for each construction (SI Appendix, section S4). Two claims about what P-StaMEs add to L-StaMEs require different tests:

- **Representation.** Intervals assigned the same L-StaME σ_i^L can receive different P-StaMEs: limited displacement may conceal rest in one interval and

scanning in another (SI Appendix, Fig. S3A). Sharing an L-StaME implies relocations that are similar at the code's resolution, not identical, so different P-StaMEs within a category show what the L-StaME code discards, not necessarily what relocations cannot resolve. The claim is measured by the entropy $H(S^P | S^L = \sigma_i^L)$ of P-StaMEs within that L-StaME category. This within-category heterogeneity is distinct from global association, because P-StaMEs can vary within every L-StaME category even when $I(S^L; S^P) = 0$, and can reflect noise or viewpoint as well as behavior.

- **Behavior.** Independently obtained labels B are more predictable from P- and L-StaMEs together than from L-StaMEs alone. Eq. 2 applies with any prespecified BoPS summary in place of A ; with the P-StaME it becomes $I(B; S^P | S^L)$, to be estimated on held-out individuals or time blocks against dependence-preserving nulls (SI Appendix, section S6 and Table S11). Conditioning on S^L tests information beyond the interval's L-StaME; a claim that the information cannot be recovered from relocations at all requires conditioning on a prespecified richer path summary, such as the L-StaME together with its L-CAM and preceding path history.

Only the behavior claim shows that P-StaMEs add behavioral information; the representation claim shows only that they separate intervals that L-StaMEs merge, which is necessary but not sufficient.

The same comparison applies to CAMs. Sethi et al. (37) built L-StaMEs from 4-s ATLAS relocations of a barn owl. Judged by their L-StaME composition, none of the owl L-CAMs was homogeneous, and the authors discuss higher relocation frequencies, shorter L-StaMEs and fewer L-StaMEs per L-CAM as possible remedies. The frame rate of video allows much shorter windows, which may help P-CAMs to be homogeneous (SI Appendix, section S4).

4. Linking Movement Ethology and Movement Ecology

4.1. Opening the black boxes of the Movement Ecology framework.

Body-configuration data do not make the motion and navigation processes featured in Nathan et al. (1) directly observable, but they supply measured kinematic time series between those processes and the path. Nathan et al. identified the difficulty of understanding movement mechanisms as an important limitation of their framework; addressing it “requires further technological and methodological advances . . . to quantify not only movement paths but also the internal and external components.”

Body configuration and its relationships to the surroundings enter as measured quantities. Let u_t denote whole-body position, q_t normalized pose, R_t orientation, c_t observed environmental and social relationships, w_t latent internal state and r_t external conditions, on an analysis clock t . A schematic state-space model (SI Appendix, section S7) is

$$\begin{aligned} q_{t+1} &\sim p_Q(\cdot | q_t, R_t, u_t, c_t, w_t, r_t, \Omega, \Phi), \\ (u_{t+1}, R_{t+1}) &\sim p_U(\cdot | u_t, R_t, q_t, q_{t+1}, c_t, w_t, r_t, \beta), \\ w_{t+1} &\sim p_W(\cdot | w_t, q_t, u_t, c_t, r_t), \end{aligned} \quad [4]$$

where β is the body length that returns egocentric poses to world coordinates. In this model, the motion and

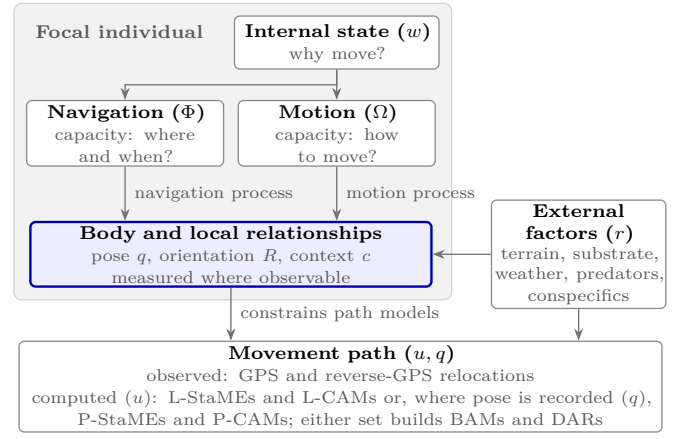


Fig. 2. Body and contextual observations within the Movement Ecology framework of Nathan et al. (1). Recorded poses, orientations and local environmental or social relationships constrain models linking internal state and movement capacities to the path. Arrows show proposed model relationships rather than identified causal effects. Environmental transport, unmeasured states and feedback require explicit treatment in particular applications (SI Appendix, section S7).

navigation capacities Ω and Φ act on location through the body (Fig. 2). They shape the next pose. Then that pose, together with propulsion or environmental transport, shapes the next location and orientation. Integrating out the body states recovers a path model of the general form $u_{t+1} = F(\Omega, \Phi, r_t, w_t, u_t)$ of Nathan et al. (1). The model is a schematic factorization: it organizes what can be measured, and does not identify mechanisms from poses alone. Discrete BoPS and behavioral-word models provide coarser alternatives, and latent-state models can represent unmeasured internal conditions (21, 40).

Body-configuration data also supply candidate indicators for each organism-centered component. Each needs validation against independent evidence before it is read as a measure of that component:

- **Motion process.** Stride frequency, duty factor, limb phase and the BoPS composition of each stride describe how motion capacity is realized under current load, injury, terrain and fatigue. Calibrated against independent measures of energy expenditure, they can refine energy landscapes (24, 41). For example, they can help assess whether an individual is traversing flat, hard terrain, marshy ground, or a sandy or stony landscape.
- **Navigation process.** Head-orientation and vigilance BoPS, and the scanning and sniffing words they spell, are candidate indicators of the sensing actions by which navigation capacity acquires information, and movement in turn shapes what information an animal can acquire (42). A head turn may correlate with an upcoming heading change; interpreting it as a navigation decision requires independent evidence. Head and gaze direction can help test responses to navigational beacons (14).
- **Internal state.** Hunched or guarded postures, and shifts in how often BoPS occur, may carry signals of pain, sickness, hunger, fatigue and reproductive state, to be validated against independent measures of those states. Facial and postural pain scales in laboratory mice are an early example (43). Whether

illness or parasitism reduces the information content of movement (14) is testable on posture and gait.

4.2. Sensors, clocks and frames. No single sensor yet records postures, contextual relationships and relocations across a lifetime (44), so Movement Ethology relies on fusing sensors and on using short, rich recordings to train detectors that run on long, sparse ones. SI Appendix, section S5 and Table S7 set out the ways of combining body and path data, the clocks they require and the reference frames they use. Pairing P- and L-StaMEs on shared intervals further requires timestamps corrected for clock offset and drift, a stated interval convention, the coverage of each stream within each interval, and missing intervals flagged rather than coded. The most practical way to obtain FuMEs in the wild is pose-to-sensor transfer. Classifiers trained on synchronized video and tag data carry FuME detection into months of collar data, where no camera can follow. In addition, classifiers that recognize behavior in collar accelerometer data have long been trained on behavior labeled by human observers, often from synchronized video (23, 34, 45). The new element provided by Movement Ethology is that the labels derived from pose are much more informative than those from a tag that detects movement at its single point of placement. A neck collar registers grazing or browsing, but not whether an elephant's trunk plucks a leaf or moves an object.

5. Research agenda and challenges

Movement Ethology is justified as a field only if it answers questions that none of its progenitors—Computational Ethology, Movement Ecology, Behavioral Ecology, Ethology and Biomechanics—answers alone. As they now stand, the parent fields of Computational Ethology and Movement Ecology leave a gap not covered by any other field of research. Movement Ecology follows identified individuals across landscapes and lifetimes, but as points. A movement track shows that an animal moved or stopped but not which body actions produced the displacement or what the animal did while stationary. Computational Ethology resolves those actions, increasingly in the field (32, 33), but mostly over minutes to hours and without the individual's movement history, and without providing ecological context. Questions that pair a body-level measurement with a path-level outcome on the same individual fill this gap (Table 3).

How do terrain, flow and load change the conversion of locomotor cycles into displacement, and its energy cost (H1, H4)? Collar studies of hunting predators have begun to answer this (46, 47), but not across the landscapes and seasons of a lifetime track. Does what an animal does at a stop predict when and where it goes next (H9)? Patch-departure theory turns on intake rate (48), which tracks do not record and arenas record without patches. Do sensing actions and body cues anticipate individual and collective movement decisions (H5, H10)? Head scanning at choice points has been studied in mazes (49), and GPS tracking has shown how a wild baboon troop decides where to go (50), but not which body cues precede departure. Does sickness show in posture and gait before it shows in the track (H3) (51)? Does learning a landscape reshape locomotor sequences, as reinforcement reshapes

syllable use in mice (H2) (52)? Each hypothesis in Table 3 links at least one body-level and one path-level rung of the hierarchy and is testable with sensor combinations available today. SI Appendix, Table S8 sets out, for each, what Movement Ecology and Computational Ethology can test as they now stand, and the test that requires both.

Table 3. Questions (bold) and testable hypotheses (italics) that need Movement Ethology. Each links a body-level and a path-level rung on the same individuals. What Movement Ecology and Computational Ethology can test alone, and full tests, are given in SI Appendix, Table S8.

Question and hypothesis	Levels linked
H1. Do landscapes decouple body from path? <i>Among locomotor CAMs, $I(C; \Psi)/H(C)$ falls as terrain, wind or current grow more complex</i>	FuME → CAM
H2. Does learning change action predictability? <i>The FuME sequence becomes more predictable with experience, after matching activity and context</i>	FuME → BAM
H3. Does sickness show in the body first? <i>Posture and gait change before displacement or DARs</i>	BoPS, FuME → DAR
H4. Does gait choice set the energy budget? <i>The FuME mix predicts independently measured energy use better than path speed</i>	FuME → CAM
H5. Does sensing precede behavioral transitions? <i>Head-scanning words peak before independently detected BAM transitions</i>	Sensing words → BAM
H6. Is the hierarchy conserved? <i>Transition structure is hierarchical at every level, beyond non-hierarchical null models (53)</i>	All levels
H7. Do syndromes have a body signature? <i>FuME repertoires predict movement syndromes (54–57)</i>	FuME → LiT
H8. Do similar body plans share syllables but not words? <i>Related species that share BoPS differ in their FuMEs</i>	BoPS → CAM
H9. Do actions at stops set departures? <i>Declining intake actions during approximate stationarity predict patch departure and the next L-StaME</i>	Ingestion words → L-StaME, BAM
H10. Do body cues lead collective movement? <i>Orientation and posture of initiators predict when and where a group departs</i>	Social words → BAM
H11. Does the same L-StaME hide different behavior? <i>Independently validated behavioral labels are better predicted by P- and L-StaMEs together than by L-StaMEs alone</i>	P-StaME → L-StaME

Testing these hypotheses runs into the hardest problems of the new field, which sit where the two parent fields meet (Table 4). The first set concerns data. Most of the hypotheses require pose from free-ranging animals, and occlusion, vegetation, lighting, distance and viewpoint all degrade pose models trained in laboratory arenas. Pretrained multi-species models (11) and drone video combined with terrain models (32) reduce the problem for H1 and H8, which compare habitats and species. The long-term hypotheses, H2, H3 and H7, also need body data tied to a known, tagged individual across occlusions and over days to years, which multi-animal tracking with identity models (28, 29) and matching of tags to video by their synchronized motion can provide. H9 and H10 need video of identified, collared animals at resources and at group departures, which fixed cameras and drones can supply (32). Where cameras cannot follow, pose-to-sensor transfer (Section 4.2) carries FuMEs identified

Table 4. Challenges for Movement Ethology and the main remedies. Why each matters, and further remedies, are given in SI Appendix, Table S9.

Challenge	Main remedies
Field imaging	Pretrained multi-species pose models; drones; pose-to-sensor transfer
Identity	Multi-animal tracking with identity models; tag-to-video matching
Arbitrary vocabularies	Report parameters; select them by entropy, assignment error and divergence
Pattern versus process	Name units by observed form; validate functions with independent evidence and context; retain concurrent actions
Ground truth	Synchronized video on tagged subsets; held-out validation; shared benchmarks
Data volume and welfare	Keypoint extraction on the device; welfare limits on tags and drones

on video into months of collar accelerometry, which also lets H4 relate gait to independently measured energy expenditure. Keypoints can be extracted on the device to limit storage, and tag mass and drone disturbance must be kept within welfare limits, a constraint that weighs most on the repeated sampling that H3 and H7 require.

The second set concerns inference. BoPS, behavioral words and StaMEs all depend on window length, cluster number and clustering method, so the units differ from study to study. This matters most for H1, H2, H6 and H8, which compare entropies or mutual information across habitats, ages, levels and species. Such comparisons require fixed or explicitly aligned codebooks, matched sampling, appropriate null models and uncertainty estimates. The parameters of every code, such as alphabet sizes, word and segment lengths and cluster numbers (SI Appendix, Tables S1 and S2), should be reported explicitly, and codings compared by entropy, assignment error and Jensen–Shannon divergence (14), so that a result can be shown to hold across a range of settings. Ground truth is a related problem. Movement behaviors of sheep derived from relocation data agreed with field observation only about 81% of the time (58). Independently annotated synchronized pose can validate the lower levels of relocation-based segmentation, and measuring how much of that segmentation P-StaMEs recover (Eq. 3) is itself part of the agenda. Lastly, a kinematic unit says what the body did, not which goal it served, the confusion between pattern and process that Nathan et al. (1) warned against. Hypotheses that read BoPS codes as signs of learning, decision or sickness, such as H2, H3, H5 and H9, require context, a stated temporal scale, and validation labels independent of the codes being evaluated.

6. Conclusion

Nathan et al. (1) predicted that tracking technology would “do for movement ecology what genetic sequencing did for molecular genetics: to provide an elemental view of a movement track.” Relocation data alone could not deliver the elements. The fundamental movement elements at the base of the Movement Ecology hierarchy (12, 13) are movements of body parts, and they have so far been replaced by statistical proxies computed from short runs of relocations (14, 15). Pose estimation, multi-site inertial

sensors and their fusion with GPS now make the elements themselves observable.

Movement Ethology provides a body-level foundation for interpreting the animal represented by a point in Movement Ecology. It studies body positions, configurations and actions relative to local surroundings and other individuals, connecting them to ecological functions and the lifetime trajectory. An animal continues to behave during approximate stationarity, and feeding, sensing or social interaction may also accompany locomotion. The reference frame, spatial tolerance and observation interval make this distinction operational.

The implementation developed here aligns body and path observations across scales. BoPS describe brief, fixed-length sequences of postures; behavioral words preserve maintained configurations and concurrent actions; complete gait cycles form a separate FuME sequence that overlays the CAM track. Pairing P- with L-StaMEs, and laying FuMEs over CAMs, supplies testable statistical relationships rather than a deterministic reconstruction of the path from pose. Two claims can be tested separately: that intervals sharing an L-StaME differ in their P-StaMEs, and that these differences carry behavioral information, validated independently, beyond L-StaMEs. Importantly, intervals of approximate stationarity become sequences of resting, feeding, grooming, scanning and social behavior. StaMEs and CAMs can be compared with independently observed body actions. Behavioral syllables measured in arenas over minutes form a basis for understanding movement over landscapes supported by environmental and social observations.

The hypotheses in Table 3 pose tests that neither parent field can carry out alone. They require synchronized records of identified individuals, first from captive, habituated or drone-filmed animals and then, through collar sensors, from free-ranging ones. Records that keep location, pose, orientation, environmental and social context, concurrent actions and code parameters together, building on open tools such as *movement* (36), will make the links between body actions and space use, energetics and risk testable.

Aristotle wrote two treatises on movement: *De Motu Animalium* on why and how animals move, and *De Incessu Animalium* on their limbs and gaits. Their two questions have since grown into largely separate sciences, and over the past two decades into two research communities with different data and tools. The tools now exist to study both questions in the same animal, on the same clock and in the same information-theoretic currency, from a single posture to a lifetime track.

Data, Materials, and Software Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. We thank Ran Nathan for comments on an earlier version that helped us redirect the focus of the paper.

1. R Nathan, et al., A movement ecology paradigm for unifying organismal movement research. *Proc. Natl. Acad. Sci. USA* **105**, 19052–19059 (2008).
2. R Kays, MC Crofoot, W Jetz, M Wikelski, Terrestrial animal tracking as an eye on life and planet. *Science* **348**, aaa2478 (2015).
3. R Nathan, et al., Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science* **375**, eabg1780 (2022).
4. N Tinbergen, On aims and methods of ethology. *Zeitschrift für Tierpsychologie* **20**, 410–433 (1963).
5. DJ Anderson, P Perona, Toward a science of computational ethology. *Neuron* **84**, 18–31 (2014).

1019	6. A Mathis, et al., DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. <i>Nat. Neurosci.</i> 21 , 1281–1289 (2018).	47. TM Williams, et al., Instantaneous energetics of puma kills reveal advantage of felid sneak attacks. <i>Science</i> 346 , 81–85 (2014).	1083
1020	7. AB Wiltischko, et al., Mapping sub-second structure in mouse behavior. <i>Neuron</i> 88 , 1121–1135 (2015).	48. EL Charnov, Optimal foraging, the marginal value theorem. <i>Theor. Popul. Biol.</i> 9 , 129–136 (1976).	1084
1021			1085
1022	8. C Weinreb, et al., Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics. <i>Nat. Methods</i> 21 , 1329–1339 (2024).	49. AD Redish, Vicarious trial and error. <i>Nat. Rev. Neurosci.</i> 17 , 147–159 (2016).	1086
1023	9. L von Ziegler, O Sturman, J Bohacek, Big behavior: challenges and opportunities in a new era of deep behavior profiling. <i>Neuropsychopharmacology</i> 46 , 33–44 (2021).	50. A Strandburg-Peshkin, DR Farine, ID Couzin, MC Crofoot, Shared decision-making drives collective movement in wild baboons. <i>Science</i> 348 , 1358–1361 (2015).	1087
1024	10. TD Pereira, JW Shaevitz, M Murthy, Quantifying behavior to understand the brain. <i>Nat. Neurosci.</i> 23 , 1537–1549 (2020).	51. BL Hart, Biological basis of the behavior of sick animals. <i>Neurosci. & Biobehav. Rev.</i> 12 , 123–137 (1988).	1088
1025	11. S Ye, et al., SuperAnimal pretrained pose estimation models for behavioral analysis. <i>Nat. Commun.</i> 15 , 5165 (2024).	52. JE Markowitz, et al., Spontaneous behaviour is structured by reinforcement without explicit reward. <i>Nature</i> 614 , 108–117 (2023).	1089
1026	12. WM Getz, D Saltz, A framework for generating and analyzing movement paths on ecological landscapes. <i>Proc. Natl. Acad. Sci. USA</i> 105 , 19066–19071 (2008).	53. S Cain, et al., Movement predictability of individual barn owls facilitates estimation of home range size and survival. <i>Mov. Ecol.</i> 11 , 10 (2023).	1090
1027	13. WM Getz, A hierarchical path-segmentation movement ecology framework. <i>Ecol. Process.</i> 11 , 56 (2022).	54. O Spiegel, ST Leu, CM Bull, A Sih, What's your move? Movement as a link between personality and spatial dynamics in animal populations. <i>Ecol. Lett.</i> 20 , 3–18 (2017).	1091
1028	14. WM Getz, An information theory treatment of animal movement tracks in <i>The Mathematics of Movement</i> , eds. L Giuggioli, PK Maini. (Springer, Cham), pp. 175–193 (2026).	55. B Abrahms, et al., Suite of simple metrics reveals common movement syndromes across vertebrate taxa. <i>Mov. Ecol.</i> 5 , 12 (2017).	1092
1029	15. WM Getz, et al., The statistical building blocks of animal movement simulations. <i>Mov. Ecol.</i> 12 , 67 (2024).	56. O Spiegel, ST Leu, A Sih, SS Godfrey, CM Bull, When the going gets tough: behavioural type-dependent space use in the sleepy lizard changes as the season dries. <i>Proc. Royal Soc. B</i> 282 , 20151768 (2015).	1093
1030	16. H Edelhoff, J Signer, N Balkenhol, Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. <i>Mov. Ecol.</i> 4 , 21 (2016).	57. AG Hertel, PT Niemelä, NJ Dingemanse, T Mueller, A guide for studying among-individual behavioral variation from movement data in the wild. <i>Mov. Ecol.</i> 8 , 30 (2020).	1094
1031	17. SR Datta, DJ Anderson, K Branson, P Perona, A Leifer, Computational neuroethology: a call to action. <i>Neuron</i> 104 , 11–24 (2019).	58. M Teimouri, UG Indahl, H Sickel, H Tveite, Deriving animal movement behaviors using movement parameters extracted from location data. <i>ISPRS Int. J. Geo-Information</i> 7 , 78 (2018).	1095
1032	18. TIR Authors, Technological advances in computational neuroethology. <i>The Innov. Rev.</i> 1 , 100012 (2026).		1096
1033	19. S Toledo, et al., Cognitive map-based navigation in wild bats revealed by a new high-throughput tracking system. <i>Science</i> 369 , 188–193 (2020).		1097
1034	20. E Gurarie, RD Andrews, KL Laidre, A novel method for identifying behavioural changes in animal movement data. <i>Ecol. Lett.</i> 12 , 395–408 (2009).		1098
1035	21. R Langrock, et al., Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. <i>Ecology</i> 93 , 2336–2342 (2012).		1099
1036	22. R Nathan, et al., Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. <i>J. Exp. Biol.</i> 215 , 986–996 (2012).		1100
1037	23. YS Resheff, S Rotics, R Harel, O Spiegel, R Nathan, AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. <i>Mov. Ecol.</i> 2 , 27 (2014).		1101
1038	24. RP Wilson, et al., Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. <i>J. Animal Ecol.</i> 75 , 1081–1090 (2006).		1102
1039	25. DD Brown, R Kays, M Wikelski, R Wilson, AP Klimley, Observing the unwatchable through acceleration logging of animal behavior. <i>Animal Biotelemetry</i> 1 , 20 (2013).		1103
1040	26. GJ Berman, DM Choi, W Bialek, JW Shaevitz, Mapping the stereotyped behaviour of freely moving fruit flies. <i>J. Royal Soc. Interface</i> 11 , 20140672 (2014).		1104
1041	27. TD Pereira, et al., Fast animal pose estimation using deep neural networks. <i>Nat. Methods</i> 16 , 117–125 (2019).		1105
1042	28. J Lauer, et al., Multi-animal pose estimation, identification and tracking with DeepLabCut. <i>Nat. Methods</i> 19 , 496–504 (2022).		1106
1043	29. TD Pereira, et al., SLEAP: a deep learning system for multi-animal pose tracking. <i>Nat. Methods</i> 19 , 486–495 (2022).		1107
1044	30. AI Hsu, EA Yttri, B-SOId, an open-source unsupervised algorithm for identification and fast prediction of behaviors. <i>Nat. Commun.</i> 12 , 5188 (2021).		1108
1045	31. ID Couzin, C Heins, Emerging technologies for behavioral research in changing environments. <i>Trends Ecol. & Evol.</i> 38 , 346–354 (2023).		1109
1046	32. BR Koger, et al., Quantifying the movement, behaviour and environmental context of group-living animals using drones and computer vision. <i>J. Animal Ecol.</i> 92 , 1357–1371 (2023).		1110
1047	33. C Wiltshire, et al., DeepWild: application of the pose estimation tool DeepLabCut for behaviour tracking in wild chimpanzees and bonobos. <i>J. Animal Ecol.</i> 92 , 1560–1574 (2023).		1111
1048	34. P Chakravarty, G Cozzi, A Ozgul, K Aminian, A novel biomechanical approach for animal behaviour recognition using accelerometers. <i>Methods Ecol. Evol.</i> 10 , 802–814 (2019).		1112
1049	35. GJ Berman, W Bialek, JW Shaevitz, Predictability and hierarchy in <i>Drosophila</i> behavior. <i>Proc. Natl. Acad. Sci. USA</i> 113 , 11943–11948 (2016).		1113
1050	36. N Sirmpllatze, et al., movement: a Python toolbox for analysing animal body movements across space and time (Zenodo) (2024).		1114
1051	37. V Sethi, et al., An information theory framework for hierarchical path segmentation and analysis of animal movement. <i>Ecol. Informatics</i> 92 , 103401 (2025).		1115
1052	38. L Luisa Vissat, S Cain, S Toledo, O Spiegel, WM Getz, Categorizing the geometry of animal diel movement patterns with examples from high-resolution barn owl tracking. <i>Mov. Ecol.</i> 11 , 15 (2023).		1116
1053	39. CE Shannon, A mathematical theory of communication. <i>Bell Syst. Tech. J.</i> 27 , 379–423 (1948).		1117
1054	40. W Zucchini, IL MacDonald, R Langrock, <i>Hidden Markov Models for Time Series: An Introduction Using R</i> . (Chapman and Hall/CRC, Boca Raton, FL), Second edition, (2016).		1118
1055	41. ELC Shepard, et al., Energy landscapes shape animal movement ecology. <i>Am. Nat.</i> 182 , 298–312 (2013).		1119
1056	42. O Spiegel, MC Crofoot, The feedback between where we go and what we know: information shapes movement, but movement also impacts information acquisition. <i>Curr. Opin. Behav. Sci.</i> 12 , 90–96 (2016).		1120
1057	43. DJ Langford, et al., Coding of facial expressions of pain in the laboratory mouse. <i>Nat. Methods</i> 7 , 447–449 (2010).		1121
1058	44. HJ Williams, et al., Optimizing the use of biologgers for movement ecology research. <i>J. Animal Ecol.</i> 89 , 186–206 (2020).		1122
1059	45. J Soltis, et al., Accelerometers in collars identify behavioral states in captive African elephants <i>Loxodonta africana</i> . <i>Endangered Species Res.</i> 18 , 255–263 (2012).		1123
1060	46. AM Wilson, et al., Locomotion dynamics of hunting in wild cheetahs. <i>Nature</i> 498 , 185–189 (2013).		1124
1061			1125
1062			1126
1063			1127
1064			1128
1065			1129
1066			1130
1067			1131
1068			1132
1069			1133
1070			1134
1071			1135
1072			1136
1073			1137
1074			1138
1075			1139
1076			1140
1077			1141
1078			1142
1079			1143
1080			1144
1081			1145
1082			1146

Supplementary Information for

Movement Ethology: Body Kinematics Meets Relocation Data

Wayne M. Getz and Joris Komen

Corresponding author: Wayne M. Getz.

E-mail: wgetz@berkeley.edu

This PDF file includes:

- Supplementary text (sections S1–S7)
- Figs. S1 to S3
- Tables S1 to S11
- SI References

Supporting Information Text

Section [S1](#) provides a glossary of acronyms and symbols, section [S2](#) displays extended tables, section [S3](#) illustrates a prospective workflow for elephants, section [S4](#) defines the BoPS and path codes summarized in Sections [2](#) and [3](#), including the construction of P-StaMEs (Fig. [S1](#)), section [S5](#) gives notes on sensors, clocks and reference frames, section [S6](#) gives the information measures computed from these codes and their estimation and interpretation, including paired L- and P-StaMEs (Fig. [S2](#) and Table [S11](#)) and the two stages of analysis (Fig. [S3](#)), and section [S7](#) gives a schematic state-space model linking body and path. Equations, figures, sections and Tables 1–4 numbered without an S refer to the main text.

S1 Glossary of acronyms and symbols

Table S1 lists the acronyms and symbols used in the main text, and Table S2 the additional ones used only in this SI Appendix, each grouped by the part of the framework in which they appear. Terms that appear only in the prospective workflow of section S3 are not listed. Calligraphic letters denote alphabets or other sets of symbols, and the superscript π marks quantities computed on the BoPS code. Logarithms are to base 2, so entropies and mutual information are in bits.

Table S1: Acronyms and symbols used in the main text.

Term	Meaning	Defined
<i>Acronyms: aligned body and path representations</i>		
BoPS	Body-pose syllable: one of a finite set of archetypal short pose sequences, obtained by clustering windows of a fixed number γ of consecutive video frames (never a single frame)	Sec. 2
FuME	Fundamental movement element: one complete cycle of a locomotor gait, such as a stride or wingbeat, identified from limb, wing or fin kinematics in pose sequences or accelerometer signals; FuME sequences overlay CAMs	Sec. 2
StaME	Statistical movement element: a cluster of fixed-length segments of μ relocation steps (L-StaME), or of BoPS windows of the same duration and timing (P-StaME)	Sec. 2
L-StaME, P-StaME	Location StaME, built from relocations; pose StaME, built from fixed-length windows of the BoPS track	Sec. 2
CAM	Canonical activity mode: a cluster of words of StaMEs, such as resting or directed stepping	Sec. 2
L-CAM, P-CAM	CAMs built from words of L-StaMEs and of P-StaMEs, respectively	Sec. 2
BAM	Behavioral activity mode: a variable-length, sub-diel string of CAMs, such as a foraging bout	Sec. 2
DAR	Diel activity routine: the 24-hour pattern of BAMs	Sec. 2
LiMP	Lifetime movement phase: a multi-day to seasonal phase such as dispersal or migration	Sec. 2
LiT	Lifetime track: an individual's movement track from birth to death	Sec. 2
<i>Other acronyms</i>		
2D, 3D	Two- and three-dimensional (video and keypoint coordinates)	Sec. 1
ATLAS	Advanced Tracking and Localization of Animals in real-life Systems, a reverse-GPS tracking system	Sec. 1
B-SOiD	Behavioral segmentation of open field in DeepLabCut, an unsupervised motif-discovery pipeline (1)	Sec. 1
GPS	Global Positioning System	Sec. 1
LEAP	Deep-learning pose-estimation software (2), forerunner of SLEAP	Sec. 1
MoSeq	Motion sequencing: unsupervised parsing of pose dynamics into syllables; keypoint-MoSeq applies it to keypoints (3)	Sec. 1
SLEAP	Social LEAP Estimates Animal Poses, multi-animal pose-tracking software (4)	Sec. 1
<i>Poses and the BoPS track</i>		
K	Number of keypoints tracked per individual	Sec. 3

Table S1, continued

Term	Meaning	Defined
t	Frame index in pose coding; analysis time step in the schematic state model	Sec. 3
Δ	Pose clock: duration of one video frame, in seconds	Eq. S3
γ	BoPS window length: number of consecutive frames clustered into one BoPS, which therefore spans $\gamma\Delta$ seconds; $\gamma \geq 2$	Eq. S1
$q_t \in \mathbb{R}^{dK}$	Egocentric pose vector of frame t : keypoints centered on the body centroid, aligned with the body axis and scaled by body length; also the body configuration of Eq. S16	Sec. 3
a	BoPS alphabet size	Eq. S2
$\mathcal{P} = \{\pi_1, \dots, \pi_a\}$	BoPS alphabet	Eq. S2
π_i	The i th BoPS: an archetypal sequence of γ poses, represented by its cluster centroid	Eq. S2
<i>Behavioral words and the lexicon</i>		
m	Word length: number of consecutive BoPS in a string, which spans $m\gamma\Delta$ seconds	Sec. 3.2
b	Number of behavioral word types	Eq. S7
$\mathcal{W} = \{\omega_1, \dots, \omega_b\}$	Lexicon of behavioral words	Eq. S7
ω_ν	Behavioral word type	Eq. S7
$\mathcal{F} = \{\phi_1, \dots, \phi_f\}$	Catalog of complete locomotor-cycle types, separate from fixed-duration words	Eq. S8
f	Number of FuME types	Eq. S8
A	Prespecified interval summary of behavioral words, retaining concurrent body-part actions and timing where relevant	Eq. 2
<i>L-StaMEs, P-StaMEs and CAMs</i>		
μ	Number of relocation steps in a StaME segment	Eq. S9
$\mathcal{S}^L = \{\sigma_1^L, \dots, \sigma_{n^L}^L\}$	L-StaME alphabet; σ_i^L is the centroid of cluster $\mathcal{C}_i^L$	Eq. S9
$\mathcal{S}^P = \{\sigma_1^P, \dots, \sigma_{n^P}^P\}$	P-StaMEs: StaMEs built from windows of the BoPS track; σ_i^P is the centroid (or medoid) of the i th cluster $\mathcal{C}_i^P$, and n^P is the number of P-StaME types	Eq. S12
$\mathcal{S}^L, \mathcal{S}^P$	L- and P-StaME codes of a randomly sampled valid StaME interval, which they share in length and timing	Sec. 3.3
$\mathcal{C}^L, \mathcal{C}^P, C$	L- and P-CAM codes of a randomly chosen CAM interval; C denotes either, with types κ_i	Eq. 1
<i>Information measures</i>		
$H(\cdot), H(\cdot \cdot)$	Shannon entropy and conditional entropy, in bits	Eq. 1
$I(\cdot; \cdot), I(\cdot; \cdot \cdot)$	Mutual information and conditional mutual information, in bits	Eq. 1
Ψ	Prespecified summary of the FuME sequence overlaying a randomly chosen CAM interval, such as the composition, rates or order of FuME types; empty, $\emptyset$, when no complete, detected cycle has its reference phase in the interval	Eq. 1
$I(C; \Psi)$	Coupling between a CAM track and its FuME overlay; bounded by $\min\{H(C), H(\Psi)\}$; compared across groups as $I(C; \Psi)/H(C)$ among locomotor CAM intervals, at matched sample sizes, with $H(\Psi)$ reported	Eq. 1
$H(\Psi C = \kappa_i)$	Locomotor variation within CAM type κ_i , which the FuME overlay resolves	Eq. S15

Table S1, continued

Term	Meaning	Defined
B	Independently validated behavioral labels; provenance and temporal support must be reported	Eq. 2
$I(B; A \mid S^L)$	Information associated with the concurrent behavioral-word summary about independent behavioral labels beyond the L-StaME code	Eq. 2
$U(S^L \mid S^P)$	Directional fraction of L-StaME uncertainty reduced by the P-StaME code; defined for $H(S^L) > 0$	Eq. 3
(S_j^L, S_j^P)	Paired L- and P-StaME codes of valid interval j , assigned by separately frozen codebooks on a common interval convention	Sec. 3.4
$H(S^P \mid S^L = \sigma_i^L)$	Entropy of P-StaME codes among intervals sharing the L-StaME σ_i^L : within-category heterogeneity, distinct from $I(S^L; S^P)$	Sec. 3.4
$I(B; S^P \mid S^L)$	Eq. 2 with the P-StaME as the prespecified BoPS summary: behavioral information added by the P-StaME code beyond the L-StaME code	Sec. 3.4
<i>The Movement Ecology framework (5) and its extension</i>		
u_t	Whole-body location at analysis time t , observed on the relocation sampling schedule	Eq. S16
R_t	Body orientation: a rotation in 3D, or a heading angle in a planar model	Eq. S16
β	Body length of an individual, estimated once rather than per frame, used to scale egocentric poses and to return them to world coordinates	Eq. S16
w_t	Internal state	Eq. S16
r_t	External factors	Eq. S16
Ω, Φ	Motion capacity and navigation capacity	Eq. S16
c_t	Observed local environmental and social relationships: relative positions, orientations, distances and contacts	Eq. S16
p_Q, p_U, p_W	Schematic conditional models for pose, location and orientation, and internal state; histories or additional states may be needed	Eq. S16
F	Path model of Nathan et al. (5), $u_{t+1} = F(\Omega, \Phi, r_t, w_t, u_t)$, recovered by integrating out the body states	Sec. 4.1

Table S2: Additional acronyms and symbols used only in the SI Appendix.

Term	Meaning	Defined
<i>Acronyms</i>		
HMM	Hidden Markov model	Table S3
SD	Standard deviation	Table S6
<i>Poses and the BoPS track</i>		
d	Spatial dimension of the keypoints: 2 for single-camera and 3 for multi-camera or depth video	Sec. S4
T	Number of frames in a recording, which therefore lasts $T\Delta$ seconds	Eq. S3
y, N^π	Index of a BoPS window, and the number of windows in a recording, $\lfloor T/\gamma \rfloor$	Eq. S1
$Q_y \in \mathbb{R}^{\gamma dK}$	Pose sequence of window y : its γ consecutive egocentric poses	Eq. S1

Table S2, continued

Term	Meaning	Defined
i_y	Index of the BoPS assigned to window y	Eq. S3
$\mathcal{T}^\pi$	BoPS track: the sequence of BoPS assigned to windows $1, \dots, N^\pi$, one symbol per γ frames	Eq. S3
p_i^π	Proportion of BoPS windows coded by π_i	Eq. S4
H_0^π	Marginal BoPS entropy; an entropy rate only for independent symbols; at most $\log_2 a$ bits per BoPS, or $\log_2 a/(\gamma\Delta)$ bits per second	Eq. S5
$\bar{Q}_i$	Mean pose sequence of the windows coded by π_i	Eq. S6
$D(a, \gamma)$	Distortion: mean squared distance per frame between each pose sequence and the mean pose sequence of its BoPS, in squared body lengths	Eq. S6
<i>Behavioral words and the lexicon</i>		
s_j	The j th observed string of m BoPS	Eq. S7
N^{wd}	Number of strings in the BoPS track	Eq. S7
J_ν	Index set of the observed occurrences of word type ω_ν , retaining repeated strings	Eq. S7
E^ω	Assignment error of strings to words, defined as E^κ is for CAMs	Eq. S7
m_F	Resampling size for a complete cycle; original event duration is retained	Eq. S8
<i>L-StaMEs, P-StaMEs and CAMs</i>		
z	Index of a StaME segment (interval)	Eq. S9
$\mathbf{v}_z^L$	Vector of summary statistics of segment z , such as the means and SDs of speed and turning angle	Eq. S9
n^L	Number of L-StaME types	Eq. S9
$\mathcal{C}_i^L$	Set of segments in the i th L-StaME cluster	Eq. S9
η	Number of StaMEs in a StaME word, the unit clustered into L- or P-CAMs; distinct from a behavioral word of BoPS (6)	Eq. S13
k^L, k^P	Numbers of L- and P-CAM types (6)	Eq. S13
$\mathcal{K}^L, \mathcal{K}^P$	L- and P-CAM alphabets, $\{\kappa_1^L, \dots, \kappa_{k^L}^L\}$ and $\{\kappa_1^P, \dots, \kappa_{k^P}^P\}$	Eq. S13
E^κ	Error rate for assigning words of StaMEs to rectified CAMs (6)	Sec. S6
$\mathbf{v}_z^P$	Feature vector of window z : $\mathbf{h}_z$ (pose-only), $(\mathbf{h}_z, \mathbf{g}_z)$ (pose-plus-gait) or $(\mathbf{h}_z, \mathbf{g}_z, \mathbf{x}_z)$ (displacement-augmented)	Eq. S11
$h_{iz}, \mathbf{h}_z$	Number of BoPS in StaME window z coded by π_i , and the vector of these BoPS counts	Eq. S11
$\mathbf{g}_z, \mathbf{x}_z$	Gait descriptors and displacement descriptors of window z	Eq. S11
Λ	Window length, in BoPS, of a P-StaME when no relocations are recorded	Sec. S4
<i>BoPS strings of StaME intervals</i>		
ρ	Relocation interval as an integer number of frames, with $\mu\rho$ a multiple of γ (if not a multiple, see section S4)	Eq. S10
M	Number of BoPS in a StaME interval, $\mu\rho/\gamma$	Eq. S10
N^μ	Number of StaME intervals covered by the BoPS track, $\lfloor T/(\mu\rho) \rfloor$	Eq. S10
τ_z, Y_z	Sets of frames and of BoPS windows that fall within the z th StaME interval	Eq. S10
$\pi^{(z)}$	BoPS content of the z th StaME interval: the string of M BoPS assigned to its windows	Eq. S10

Table S2, continued

Term	Meaning	Defined
$\mathcal{P}^{(\mu)}$	Set of distinct BoPS strings that occur across StaME intervals	Eq. S10
Π	BoPS string of a randomly chosen StaME interval, with values in $\mathcal{P}^{(\mu)}$	Eq. S10
<i>Information measures</i>		
ψ_ℓ, N^ψ	Distinct values of Ψ , and their number	Eq. S15
p_i^κ	Proportion of CAM intervals coded as κ_i	Eq. S15
$p(\kappa_i, \psi_\ell)$	Proportion of CAM intervals coded as κ_i whose FuME summary is ψ_ℓ ; $p(\kappa_i \psi_\ell)$ is the corresponding conditional proportion	Eq. S15
$I(S^L; \Pi)$	Information about the L-StaME carried by the BoPS string; bounds that of any code assigned from Π alone, such as the pose-only P-StaME	Eq. S14

S2 Extended tables

Tables S3, S5, S8 and S9 are full versions of main-text Tables 1, 2, 3 and 4. Table S8 also sets out, for each question, what Movement Ecology and Computational Ethology can test as they now stand. Tables S4, S6 and S7 summarize the levels of the hierarchy, the two constructions of StaMEs (section S4) and the ways of combining body and path data (section S5).

Table S3: Movement Ecology, Movement Ethology and Computational Ethology compared. Entries are typical emphases, and the fields overlap substantially. Full version of Table 1.

	Movement Ecology	Movement Ethology	Computational Ethology
Unit of movement	Whole-animal location, represented by a point in geographic space	Articulated body in relation to local environmental features and other individuals	Configuration of body parts (keypoints)
Primary sensors	GPS, reverse-GPS arrays, satellite tags; collar accelerometers	Synchronized relocations, pose (video, drones), body-mounted inertial sensors and environmental or social measurements	Video (2D, multi-camera 3D, depth) with markerless pose estimation
Sampling interval	Tenths of seconds to days	Pose and relocation clocks aligned, ideally with the relocation interval an integer multiple of the pose interval	Tens to hundreds of frames per second
Record length	Minutes to a lifetime	Body observations related to activity bouts, daily routines and longer movement histories	Seconds to hours; days in home cages
Setting	Free-ranging, landscape scale	Captive, habituated or drone-filmed animals, extended to free-ranging animals through collar sensors	Mostly laboratory arenas; field use emerging
Segmentation tools	Step length and turning angle, change-point analysis, hidden Markov models, StaME and CAM clustering	Alignment of BoPS and path codes, concurrent-action coding and contextual validation	Postural embeddings, autoregressive HMM syllables (short dynamic motifs), clustering of kinematic features
Lowest unit analyzed	Statistical movement element (StaME)	BoPS, a fixed-length sequence of poses; FuME as one complete locomotor cycle	Single-frame pose, or syllable (a dynamic motif of a few hundred milliseconds)
Typical questions	Space use, migration, dispersal, energetics, movement mechanisms	How body actions and local interactions contribute to ecological functions and the lifetime trajectory	Neural control, action selection, behavioral organization, drug and genetic effects, social behavior

Table S4: Aligned body and path levels (Fig. 1). Behavioral words and maintained postures contribute alongside locomotor cycles. Corresponding terms in Getz (6), Nathan et al. (5) and Computational Ethology.

Level	Getz (2026) unit	Nathan et al. (2008) term	Computational Ethology analogue	Measured with	Time scale
Posture	(none)	(none)	Pose: keypoint configuration	Video pose estimation	One frame
BoPS	(none)	(none)	Fixed-length pose motif; a MoSeq syllable is the variable-length analogue	Clustered windows of γ consecutive video poses	γ frames ($\gamma\Delta$ seconds), never one frame
Word	(none)	(none)	Action motif or maintained posture	Whole-body or factorized BoPS strings	$m\gamma\Delta$ seconds
FuME	Fundamental movement element	Below the step	Gait-cycle motif; a syllable spanning one locomotor cycle	Pose sequences, multi-site accelerometers, audio	One locomotor cycle: sub-second to several seconds
StaME	Statistical movement element	Movement step	(none)	Relocations at 0.1 Hz or faster (L-StaMEs); or BoPS windows (P-StaMEs, section S4)	μ relocation intervals
CAM	Canonical activity mode	Canonical activity mode	Bout of a single motif class	Words of L- or P-StaMEs (L- or P-CAMs)	Seconds to minutes
BAM	Behavioral activity mode	Movement phase	Ethogram behavior, e.g., a foraging bout	Path, body and contextual observations	Minutes to hours
DAR	Diel activity routine	(none)	Circadian activity profile	Paths aligned with available body and context data	24 hours
LiMP	Lifetime movement phase	Movement phenomenon, e.g., migration	Life-stage context	Paths with available body and context data	Days to seasons
LiT	Lifetime track	Lifetime track	(none)	Relocations aligned with available body and context observations	Lifespan

Table S5: Body actions and candidate functional categories. Locomotor cycles and other actions can co-occur during displacement or relative stationarity, and both inform the movement hierarchy. Functional interpretations require context and validation. Full version of Table 2.

Category	What the words do	Keypoint signature	Link to the movement hierarchy
Locomotion (FuMEs, $\mathcal{F}$)	Complete limb, wing or fin cycles of a locomotor gait; turning steps; backing up	Phase-locked limb, wing or fin cycle with a gait support pattern (displacement not used)	FuME sequences overlay L- and P-CAMs; coupling $I(C; \Psi)$ (Eq. 1)
Rest and postural transitions	Lie down, rise, stand at rest	Changes in body height or support, followed by maintained posture; sleep requires separate validation	Relative stationarity; resting CAMs; timing of DARs
Ingestion and resource handling	Bite, pluck, chew, drink, dig, carry or cache food	Head, mouth or forelimb actions, including during concurrent stepping	Distinguish among similar L-CAMs; resource selection
Sensing and orientation	Scan, sniff, raise the head, turn the head, peck to probe, orient ears or eyes, freeze	Reorientation of head and sense organs during displacement or relative stationarity	Candidate indicator of the navigation process; precede BAM transitions (H5)
Maintenance and thermoregulation	Groom, scratch, dust, wallow, pant, seek shade, fan ears or wings	Self-directed limb or head movement; repetitive appendage oscillation	Candidate indicator of internal state; weather-dependent DARs
Social and communicative	Greet, touch, display, play, coordinate departures	Relative pose, distance, orientation or contact among identified individuals	Conspecific part of external factors r ; collective movement decisions
Agonistic and antipredator	Threaten, fight, submit, mob, flee	Rapid postural change; may grade into locomotion	Risk, landscape of fear; escape bursts in the track
Reproductive and parental	Court, mate, nurse, guard or carry young	Dyadic configurations; sex- or state-specific postures	Breeding LiMPs; parental constraints on movement
Elimination and marking	Defecate, urinate, scent-mark	Stereotyped hindquarter and tail postures	Marking routes; health indicators
Anomalous and clinical	Stereotypies, lameness, sickness postures	Invariant repetition; left-right asymmetry	Welfare and health monitoring (H3)
Unassigned	Occlusion or tracking error	Low keypoint confidence	Missing-data flag; an observed transition is not automatically missing

Table S6: Two constructions of StaMEs: the relocation construction implemented by Sethi et al. (7) and the BoPS construction proposed here (Section 3.4 and section S4).

	Relocation construction: L-StaMEs (7)	BoPS construction: P-StaMEs (Movement Ethology)
Input time series	Relocations, e.g., 4-s ATLAS fixes of a barn owl	BoPS track, one BoPS per window of γ frames of video pose recorded at tens of frames per second
Segment	μ consecutive relocation steps	Window of the same duration, or of Λ BoPS when no relocations are recorded
Segment features	Mean and SD of speed; mean and SD of absolute turning angle; net displacement	Three nested sets: BoPS counts or string (pose-only); plus stride frequency, duty factor and gait regularity (pose-plus-gait); plus speed and turning from georeferenced video or dead reckoning (displacement-augmented)
Clustering	Ward and spectral (vector methods); dynamic time warping (shape methods)	The same methods; string distances on BoPS sequences
What separates StaMEs	How far and how straight the animal moved	How the body was configured and cycled, moving or not
Relative stationarity	May share an L-StaME despite different actions; reference frame, spatial tolerance and interval must be stated	Can distinguish configurations and actions, interpreted with context
Main error sources	Positional error; sampling interval	Keypoint error; occlusion; viewpoint
Record length	Weeks to years	Minutes to hours; longer via pose-to-sensor transfer
Context	Geographic location with environmental covariates	Local features and social relationships; geographic placement requires georeferencing
Main limitation	CAMs may not be homogeneous in StaME composition; higher relocation frequency, shorter StaMEs and fewer StaMEs per CAM discussed as remedies	Video coverage and identity; drift in dead-reckoned turning; homogeneity still to be tested against independent labels

Table S7: Ways of combining body and path data (section S5).

Approach	Body data	Path data	Typical record	Example
Laboratory arena	2D or 3D video pose	Arena coordinates	Minutes to days	Weinreb et al. (3)
Drone survey	Video pose of each group member	Georeferenced trajectories on a 3D terrain model	Minutes per flight	Koger et al. (8)
Fixed field cameras	Video pose	Short local trajectories	Seconds to minutes per visit	Wiltshire et al. (9)
Biologging	Accelerometers, magnetometers, audio	GPS or reverse-GPS relocations	Weeks to years	Nathan et al. (10)
Pose-to-sensor transfer	Video pose labels accelerometer FuMEs, then accelerometers alone	GPS relocations	Weeks to years	Video-annotated training: Soltis et al. (11); pose-derived labels proposed here

Table S8: Questions and testable hypotheses that need Movement Ethology: what Movement Ecology and Computational Ethology can test as they now stand, and the test that requires both. Enlarged version of Table 3.

Question & <i>hypothesis</i>	Levels linked	Movement Ecology as it stands	Computational Ethology as it stands	Movement Ethology test
H1. Do landscapes decouple body from path? <i>Mutual information between CAMs and their FuME overlay falls as terrain, wind or current grow more complex</i>	FuME → CAM	Relates path metrics and StaMEs to terrain, wind or current, and accelerometry adds effort at one point on the body, but not which locomotor cycles produced each step	Resolves gait cycles and their kinematics, mostly on treadmills or flat arenas, without natural terrain or the displacement that results	Synchronized tags and drone pose across habitats; compare the normalized coupling $I(C^L; \Psi)/H(C^L)$ (Eq. 1) among locomotor CAM intervals ($\Psi \neq \emptyset$) or at matched locomotor fractions, reporting $H(\Psi)$, by prespecified terrain or flow covariates, with frozen codebooks, matched sampling and sensor quality, and $H(C^L) > 0$. Use held-out individuals or time blocks and nulls preserving within-stream dependence. FuME identification does not use displacement, although cycle frequencies may vary with habitat
H2. Does learning change action predictability? <i>FuME sequences within commuting bouts become more predictable with experience after matching activity and context</i>	FuME → BAM	Compares route efficiency, site fidelity or path predictability between naive and experienced animals, without the actions that produce the path	Shows that reinforcement shapes how often syllables are used and how they are sequenced (12), in arenas rather than landscapes	Separate the components of learning: familiarity (translocated versus resident animals of the same age), development (juveniles versus adults in the same landscape) and experience (the same individuals followed over time); prespecify conditional entropy or held-out next-cycle log loss, use a common codebook and equal-duration or matched-cycle samples, and distinguish sequence predictability from repertoire entropy. Addresses the question of Getz (6) on learning and information content
H3. Does sickness show in the body first? <i>Infection shifts BoPS frequencies toward guarded or hunched postures and reduces gait regularity before it changes daily displacement or DAR geometry</i>	BoPS, FuME → DAR	Detects reduced daily displacement, activity or range from tracks and accelerometry, usually once sickness behavior (13) is well developed	Detects pain expressions, lameness and abnormal gait in laboratory or farm settings (14), without their consequences for space use	Repeated pose sampling of tagged animals in disease-monitored populations; tests the parasite question of Getz (6) at the FuME level

Table S8, continued

Question & hypothesis	Levels linked	Movement Ecology as it stands	Computational Ethology as it stands	Movement Ethology test
H4. Does gait choice set the energy budget? <i>Within one CAM class, the FuME mix predicts energy expenditure better than path speed does</i>	FuME → CAM	Uses dynamic body acceleration and path speed as energy proxies (15, 16); collar studies of hunting predators have measured locomotion and its energetic cost (17, 18)	Measures gait kinematics precisely in the laboratory, usually without an energy measurement or an ecological setting	FuME mix from pose or pose-labeled accelerometry and GPS speed, compared as predictors of energy expenditure measured independently of the accelerometer: heart rate, doubly labeled water, or respirometry in captivity. Dynamic body acceleration from the same accelerometer would make the test circular. Relate to energy landscapes (16)
H5. Does sensing precede behavioral transitions? <i>Head-scanning words (sensing category) concentrate just before BAM transitions and increase with uncertainty or predation risk</i>	Sensing words → BAM	Detects transitions in tracks by change-point or hidden Markov models, but not the sensing actions that precede them	Studies head scanning at choice points in mazes (19), not in landscapes with real uncertainty or risk	Define scanning windows before independently detected transitions; compare with matched non-transition intervals and time-shifted controls. Use head orientation and contextual evidence without defining the transition from the same scanning features
H6. Is the hierarchy conserved? <i>Transition structure is hierarchically organized at every level, from syllables in the lab (20) to CAMs and BAMs in the field</i>	All levels	Finds predictability and nested structure in tracks from StaMEs upward (6, 21)	Finds hierarchy in fly and mouse behavior from syllables to minutes (20); neither field spans syllables to lifetime tracks in the same individuals	Fit the same autoregressive and information measures at each level and compare predictability across time scales (21) against non-hierarchical alternatives, such as single-level Markov or semi-Markov models fitted to the same data, since an analysis that imposes levels can find a hierarchy that is not there
H7. Do syndromes have a body signature? <i>Consistent individual differences in FuME repertoires predict path-level movement syndromes</i>	FuME → LiT	Defines movement syndromes from path metrics (22, 23)	Measures individual differences in action repertoires over minutes to days, without the individual's track	Repeat pose sampling of individuals with long tracks; relate to syndrome metrics (22–25)

Table S8, continued

Question & hypothesis	Levels linked	Movement Ecology as it stands	Computational Ethology as it stands	Movement Ethology test
H8. Do similar body plans share syllables but not words? <i>Species with similar body plans share most BoPS but differ in their FuME catalogs, and these differences predict their CAM repertoires</i>	BoPS → CAM	Compares path statistics and activity modes across species, not their gaits	Supplies keypoints across many species from pretrained models (26), without linking gait repertoires to space use	Map keypoints onto homologous anatomical landmarks across species; compare BoPS and FuME distributions by Jensen–Shannon divergence (6) at matched sample sizes, in comparative designs that separate phylogeny, body mechanics and environment. Pretrained multi-species pose models (26) supply keypoints but do not by themselves establish homologous BoPS
H9. Do actions at stops set departures? <i>Declining rates of intake actions, such as bites and chews, during relative stationarity predict patch departure and the next L-StaME</i>	Ingestion words → L-StaME, BAM	Infers patch use from residence times and revisits; intake rate, on which patch-departure theory turns (27), is not observed	Counts bites and chews in arenas or at feeders, without patches spread over a landscape or the path that follows	Video or pose-labeled accelerometry at feeding sites, synchronized with collars; estimate intake rate from ingestion words, validated against independent measures such as observer counts or offtake; test whether its decline predicts departure time and the next StaME better than residence time alone, using held-out patches or individuals
H10. Do body cues lead collective movement? <i>Orientation, posture and initiating actions of group members predict when and in which direction a group departs</i>	Social words → BAM	GPS tracking of whole troops shows who initiates and who follows (28), but not the body cues exchanged before departure	Multi-animal pose tracking resolves social postures and interactions (4, 29), mostly in arenas and without group travel across a landscape	Drone video of collared, identified group members (8); code orientation, posture and initiating actions before departures; test whether they predict departure timing and direction beyond the positions and movements recorded by GPS, against time-shifted controls

Table S8, continued

Question & hypothesis	Levels linked	Movement Ecology as it stands	Computational Ethology as it stands	Movement Ethology test
H11. Does the same L-StaME hide different behavior? <i>Among intervals sharing an L-StaME, P-StaMEs vary more than assignment noise alone would produce (representation) and carry independently validated behavioral information beyond the L-StaME (behavior)</i>	P-StaME → L-StaME	Assigns intervals of limited or similar displacement to one L-StaME; what the body did within them is not recorded	Clusters pose windows into behavioral categories, without the L-StaME of the same interval	Paired L-StaMEs and pose-only or pose-plus-gait P-StaMEs on a common interval convention, with labels B from annotation protocols independent of both codes. The representation-level prediction is $H(S^P \mid S^L = \sigma_i^L)$ above the value expected from assignment noise, estimated by re-coding the same intervals from a second camera or with keypoints perturbed within their estimated error; the behavioral prediction is $I(B; S^P \mid S^L) > 0$ (Eq. 2 with S^P as the BoPS summary), estimated on held-out individuals or time blocks; a claim beyond relocations as a whole conditions on a richer prespecified path summary, such as the StaME with its CAM and preceding path history. Refuted for an L-StaME category if labels are no better predicted with the P-StaME than without it; heterogeneity of P-StaMEs alone does not support it (section S6)

Table S9: Challenges for Movement Ethology and partial remedies. Full version of Table 4.

Challenge	Why it matters	Partial remedies
Field imaging	Occlusion, vegetation, lighting, distance and viewpoint degrade pose models trained in arenas	Pretrained multi-species models (26); drones with terrain models (8); pose-to-sensor transfer
Identity	Body data are useful only if tied to a known, tagged individual across occlusions and days	Multi-animal tracking with identity models (4, 29); tag-to-video matching by synchronized motion
Arbitrary vocabularies	BoPS, behavioral words and StaMEs depend on window length, cluster number and method (i.e., on the parameters a , γ , m , m_F , f , μ , n^L , n^P , η , k^L and k^P), so units differ across studies	Select parameters using entropy, assignment error and Jensen–Shannon divergence criteria (6). In single-species studies, parameter selection should be synchronized for the group as a whole, or at least for subgroups partitioned by sex and life stage
Pattern versus process	The same kinematic unit can serve different goals: head-down stepping may be grazing, scent-searching or hiding. Naming clusters by function reads process into pattern, as Nathan et al. (5) cautioned, and the error propagates into activity budgets and inferred internal states	Name units by what the body does, such as head-down slow stepping, and retain concurrent actions. Interpret functions at a stated scale using contact with resources, relative positions and orientations of nearby animals, and other context. Validate against independent observations or outcomes such as intake; labels derived from the body features being evaluated are circular
Ground truth	Entropy, assignment error and divergence rank codes against one another, not against what the animal was doing, and observer records are too coarse to check sub-second units. Where relocation-based segmentation has been checked, in sheep, it agreed with field observation about 81% of the time (30)	Record synchronized, time-stamped video on a subset of tagged individuals and use independently annotated pose as validation evidence for the lower levels of relocation-based segmentation. Report inter-observer agreement and per-class confusion matrices for individuals held out of training, measure BoPS–path agreement with Eq. 3, and share labeled benchmark data sets
Data volume and welfare	Hours of high-frame-rate video and multi-sensor tags are costly to store and to carry	Extract keypoints on the device or at the edge; keep tag mass and drone disturbance within welfare limits

S3 A prospective workflow: elephant video

Elephants, of which there are three extant species, are animals for which sufficient data currently exist to begin testing the methods framed in this paper. The workflow below is prospective: none of its steps has yet been carried out end to end. Most of the studies it draws on concern African savanna elephants, and some include Asian elephants (31, 32). Their Movement Ecology has been studied extensively with GPS collars, including spectral analysis of movement periodicity (33), diel movement strategies (34) and spatial memory (35). Their behavior is cataloged in an unusually detailed ethogram (36) and their trunk provides an informative example of an appendage used in feeding, sensing, maintenance and social interaction. Below, using elephants as our exemplar, we elaborate on the step-by-step workflow of Section 3 that is needed to move from elephant video data to the hierarchy of movement segments featured in Fig. 1.

Step 1. Begin by assembling complementary video streams, because no single source of elephant video covers every level of the hierarchy. Four sources from which data can currently be obtained (with owner permission as applicable) are:

- *Zoos and sanctuaries.* They provide long, continuous recordings of known individuals under good light, which makes them the natural training and validation set. Soltis et al. (11) paired collar accelerometers with video of three adult females at Disney’s Animal Kingdom. They classified feeding, bathing, walking and swaying with 70–91% success.
- *Waterhole and camera-trap video.* Many reserves stream video that gives repeated side views of wild individuals during drinking, dusting and social encounters.
- *Drone video.* Existing footage gives top-down views of all members of a family group at once, with georeferenced positions and terrain, much as Koger et al. (8) have demonstrated for other savanna ungulates.
- *GPS and accelerometer collars.* Long GPS tracks of collared elephants already exist, for example from Etosha National Park in Namibia (35, 37), northern Kenya (33, 38) and Kruger National Park in South Africa (39), and from 229 elephants at 19 sites across Africa (40). Many such tracks are archived in Movebank (41). The Etosha tracks analyzed by Tsalyuk et al. (37), for instance, are openly available (42). Fitting collars with accelerometers on some of the individuals that are filmed carries the analysis beyond the camera’s reach (Step 6).

Joint body–path analyses require the same identified individuals, overlapping recording periods and synchronized clocks. Archived video and tracks are not automatically paired. Matching individuals can use tusk features and the tears, holes and outlines of the ears, as in semi-automated re-identification (43).

Step 2. Now define an elephant skeleton, because general quadruped keypoint sets do not suffice. The 17-point AP-10K skeleton, for example, has eyes, nose, neck, tail base and four limbs, but no trunk or ear points (44). An elephant skeleton needs, in addition to the limbs, spine and tail:

- several points along the trunk (base, proximal, middle, distal, tip), because the trunk tip acts as a hand and the trunk bends anywhere along its length,
- at least three points per ear (base, leading edge, lower lobe), to recover ear angle and spread, and
- the tusk tips, and the head–neck angle, which sets head height.

A skeleton of roughly 30 keypoints is a reasonable starting point. Trunk keypoints are the hardest to track and the most informative. Dagenais et al. (45) found that elephants build complex trunk movements from a limited set of motion primitives, and that the trunk forms temporary pseudo-joints. That finding suggests the trunk has a small natural set of BoPS of its own.

Step 3. Follow the procedure outlined in Section 3 to build a factorized BoPS alphabet. Elephant postures factor naturally into partly independent sub-alphabets, so here each body part has its own syllables, each spanning the same short window of γ frames. There are limb-and-torso BoPS, each a combination of limb support pattern, head height and body tilt (for example, standing square with the head high, standing with the head low, or one forefoot raised), trunk BoPS (hanging, curled to the mouth, raised, extended forward, resting on the ground or a tusk, wrapped around an object) and ear BoPS (flat against the head, partly spread, fully spread and rigid, mid-flap). Coding each sub-alphabet separately keeps the alphabet small and interpretable. For instance, six limb-and-torso, six trunk and four ear BoPS give 144 possible joint BoPS, of which only a fraction will occur. Separate sub-alphabets also make it possible to ask which body part carries information about a given category. Alongside them, record visible food or substrate contacts, positions of relevant local features, and the identities, relative positions and orientations of nearby elephants as contextual measurements c_t .

Step 4. Build fixed-duration behavioral words from the factorized BoPS streams, and detect complete stride cycles in parallel. Resample each cycle to m_F phases for FuME classification while retaining its measured duration. Limb, trunk and ear records remain separate before aggregation: walking with a periscope trunk posture and stepping while handling food retain both observed actions. Functional interpretations use context and validation rather than a single exclusive word category. Table S10 gives examples. Elephant locomotion spans graded changes in stride frequency, length and duty factor. It lacks a whole-body aerial phase, although faster locomotion includes running-like mechanics (31, 46). Generic cantering and galloping CAMs therefore do not apply. Body and path observations can instead support descriptions such as maintained standing posture, local stepping, directed walking and faster locomotion. Ear flapping and sustained ear spreading illustrate why both temporal sequence and context matter when assigning a function (47).

Step 5. At this step, locomotor cycles and concurrent body actions are aligned with the path and local context. Stride FuMEs occur within the intervals between GPS fixes. Many elephant collars record fixes at intervals of tens of minutes to an hour, but StaMEs, and the CAMs built from them, need fixes at 0.1 Hz or faster (6). Laying FuMEs over L-CAMs therefore requires high-frequency GPS bursts during periods of video coverage. Where such bursts are unavailable, P-StaMEs (section S4) can stand in for L-StaMEs while the elephants are on camera, and the directional association between the two codes (Eq. 3) can be checked whenever both exist. Where both exist, each filmed interval with a GPS burst yields one pair of L- and P-StaMEs. Intervals at a waterhole that share an L-StaME of limited displacement may, for example, receive different P-StaMEs for drinking, dusting and scanning. Whether those differences match independently annotated behavior is tested by $I(B; S^P \mid S^L)$ (section S6). The hierarchy then reads upward as follows. L-StaMEs distinguish slow tortuous stepping from fast directed walking, and strings of them build L-CAMs, such as the local stepping and directed walking of Step 4. The stride FuMEs that overlay each L-CAM show how it was executed, for example through stride frequency and length. CAMs build BAMs such as a feeding bout, travel to water or resting in shade. BAMs build DARs shaped by water, heat and risk (34). DARs build LiMPs such as wet- and dry-season ranges or a bull’s musth periods,

Table S10: Example elephant actions and candidate categories, which can co-occur during displacement or relative stationarity (after 36, and sources cited in the text).

Category	Example elephant words	Main keypoints
Locomotion (FuMEs)	Walking stride, with concurrent trunk, head and ear actions recorded; faster stride without a whole-body aerial phase; backing up; turning steps; swimming stroke	Limbs (support pattern and phase)
Rest and postural transitions	Standing doze with trunk hanging or resting on a tusk; lying down; rising	Body height, trunk, limbs
Ingestion and resource handling	Trunk-wrap and pull of grass; forefoot kick to loosen a tuft; stripping leaves through the mouth; prying bark with tusks; drinking cycle of trunk to water, lift and squirt into mouth; digging for water	Trunk, forefeet, tusks, head
Sensing and orientation	Periscope trunk posture with tip movement, during walking or relative stationarity; trunk-tip inspection of ground or dung; maintained head elevation and ear spreading	Trunk, ears, head
Maintenance and thermoregulation	Rhythmic ear flapping; dust throwing over the back; mud wallowing; rubbing on trees or termite mounds	Ears, trunk, whole body
Social and communicative	Trunk-to-mouth greeting; trunk twining; touching another’s body; play	Relative pose of two or more individuals
Agonistic and antipredator	Ears spread and head raised in threat; head shake; mock charge; sparring with tusks and heads; group bunching around calves	Ears, head, tusks, group configuration
Reproductive and parental	Suckling, with the calf’s trunk folded back over its head; musth posture and gait in bulls; mounting; calf guarding and assistance	Dyadic configuration, head, trunk
Elimination and marking	Defecation; urination; urine dribbling during musth	Hindquarters, tail
Anomalous and clinical	Swaying, head-bobbing and pacing in captivity; asymmetric stride from foot or snare injury	Limbs, head; left–right symmetry

and LiMPs build the lifetime track. Elephants move faster along corridors than elsewhere in their range (48). Stride measurements can distinguish changes in length and frequency, but energetic cost requires an independently calibrated measure.

Step 6. Characterize body actions during displacement and relative stationarity, then test transfer to collar data. Relative stationarity is assessed within a specified reference frame, spatial tolerance and interval (section S5). An elephant remaining within a waterhole area may drink, dust, wallow, greet, rest or take local steps, whereas an elephant moving away may continue to chew or sample its surroundings. Record these actions concurrently. Estimate $I(B; A | S^L)$ (Eq. 2) only with behavioral labels whose validation is independent of the BoPS and path codes being evaluated, using separate annotation protocols or independent contextual and outcome evidence. Video-derived action labels can train collar classifiers, extending the video-annotated approach of Soltis et al. (11). Test transfer on individuals and field contexts held out of training and report performance for each action: a neck

collar may preserve stride, grazing or browsing signals while losing trunk actions, such as plucking a leaf rather than moving an object, and ear or social information. Trunk-activity measurements have been used to infer sleep in two wild matriarchs (49). Synchronized posture observations can assess such proxies, although immobility or posture alone does not establish sleep.

Step 7. At this step elephant-specific questions are now asked. The sequence above makes the following hypotheses testable, each an elephant instance of the general agenda in Section 5:

- E1. Heat load.** Ear-flap rate, a maintenance word whose frequency rises with temperature (47), predicts the onset of midday resting BAMs and shade-seeking, and the geometry of hot-season DARs shifts accordingly.
- E2. Leadership.** Sensing words of the matriarch, such as periscope trunk postures and maintained ear spreading, interpreted with social context, increase before collective departures. This tests the role of matriarchs as repositories of social knowledge (50) at the level of the body, and extends H5 to group decisions (28, 51).
- E3. Memory.** Travel along remembered routes (35) contains fewer sensing words per kilometer than exploratory travel, an elephant test of H2.
- E4. Musth.** Musth posture, gait and urine-dribbling words in bulls (52) co-occur with long, directed search LiMPs, and the entropy of the BoPS code differs between musth and non-musth periods.
- E5. Injury.** Left-right asymmetry in stride FuMEs, with near and far limbs in 2D views mapped to left and right using the recorded facing direction, from foot disease in captivity or snare wounds in the wild, appears before any reduction in daily displacement or change in DAR geometry, an elephant test of H3. Captive stereotypies such as swaying, head-bobbing and pacing (32) provide a parallel welfare readout from the anomalous category.
- E6. Water budgets.** Counts of drinking-cycle words at waterholes index water intake per visit, linking periods of relative stationarity at waterholes to the diel commute between water and forage. Converting counts to volumes requires calibrating the volume per trunk-load, for example on zoo elephants drinking from measured containers.

Together these questions use every rung of the hierarchy, from a single trunk movement to a bull’s musth cycle. They also show what each data stream contributes: video records body actions and local relationships, while collars extend measurements of location and detectable actions over time, and ecological interpretation draws on both with contextual validation.

S4 Constructing the BoPS and path codes

This section defines the codes summarized in Sections 2 and 3, in the order in which they are built: the BoPS track and its alphabet, behavioral words, FuMEs, L-StaMEs, the common window on which L- and P-StaMEs are paired, P-StaMEs and CAMs. Section S6 then treats these codes as random variables, gives the information measures computed from them, and discusses how those measures are estimated and interpreted.

Pose normalization and the BoPS track

Each video frame is first reduced to time-stamped positions of K identified keypoints, each with a confidence score, for every identified individual, using markerless pose estimation (4, 29, 53) or

pretrained multi-species models (26). The keypoints are then put into the animal’s own reference frame. They are translated so that the body centroid sits at the origin, rotated into a body frame defined by the longitudinal axis and, in 3D, additional non-collinear anatomical landmarks, and scaled by a fixed body length β for each individual. In 2D, rotation alone cannot map a left-facing profile onto a right-facing one without inverting it, so poses are also reflected, where necessary, to a standard facing set by a dorsal–ventral landmark such as a foot, and near and far limbs are then labeled consistently. Without this step, BoPS clusters separate the same posture by facing direction. Because projected length varies with viewing angle and posture, β is estimated per individual, from 3D reconstruction or reference postures, rather than per frame. The result is an egocentric pose vector $q_t \in \mathbb{R}^{dK}$, with $d = 2$ for single-camera video and $d = 3$ for multi-camera or depth video.

Whole-body location u_t is kept separately, together with body orientation R_t (a heading angle in a planar model, a rotation in 3D), the facing direction of a 2D view, and the body length β used for scaling. These quantities place the normalized keypoints in the world frame. Where observable, a contextual record c_t stores relative positions, orientations, distances and contacts involving environmental features and identified neighboring animals. It is retained separately from the pose-only code so that the contributions of body shape, displacement and context can be compared. A single camera confounds posture with viewing angle. Calibrated multi-camera 3D reconstruction reduces this ambiguity. Otherwise, viewing angle can be treated as a nuisance variable, and BoPS that differ only in viewpoint can be merged after clustering.

The pose track is then cut into consecutive, non-overlapping windows of a fixed number $\gamma \geq 2$ of frames, each spanning $\gamma\Delta$ seconds on a pose clock of Δ seconds per frame (typically $\Delta \leq 0.1$ s). If an individual is filmed for T frames, window y holds the pose sequence

$$Q_y = (q_{(y-1)\gamma+1}, \dots, q_{y\gamma}) \in \mathbb{R}^{\gamma dK}, \quad y = 1, \dots, N^\pi, \quad N^\pi = \lfloor T/\gamma \rfloor. \quad (\text{S1})$$

Each pose sequence, described by its concatenated poses or by summaries such as mean pose and keypoint velocities, is reduced in dimension, by principal components or a learned embedding, and clustered, by cluster analysis or other machine-learning methods, into a archetypal pose sequences, the body-pose syllables. Their centroids form the BoPS alphabet

$$\mathcal{P} = \{\pi_1, \dots, \pi_a\}, \quad (\text{S2})$$

and each window is coded by its nearest BoPS, giving the BoPS track

$$\mathcal{T}^\pi = \{(y; \pi_{i_y}) : y = 1, \dots, N^\pi\}, \quad i_y \in \{1, \dots, a\}, \quad (\text{S3})$$

where π_{i_y} is the BoPS assigned to window y , which covers frames $(y-1)\gamma + 1, \dots, y\gamma$. A BoPS is therefore never obtained from a single frame. It describes a brief movement of fixed duration $\gamma\Delta$, or a maintained posture when the poses within the window barely change. The window length $\gamma\Delta$ is chosen short relative to the locomotor cycles of interest, so that a stride or wingbeat is spelled by several BoPS. When the body is factorized by part, each part has its own BoPS alphabet, and each window is coded by one syllable per part.

Three further codes are built from the underlying pose data. Strings of BoPS build behavioral words, pose and kinematic time series support FuME detection, and windows of the track build P-StaMEs, the counterparts of the L-StaMEs built from relocation data (6, 7). Each is defined below. Like the number of StaME types in Getz (6), the alphabet size a and the window length γ are tunable parameters, chosen as described next.

Entropy, distortion and the size of the BoPS alphabet

Let

$$p_i^\pi = |\{y : i_y = i\}|/N^\pi, \quad i = 1, \dots, a, \quad (\text{S4})$$

be the proportion of windows of the BoPS track $\mathcal{T}^\pi$ (Eq. S3) coded by π_i . The marginal BoPS entropy, defined regardless of serial dependence and interpreted as a per-BoPS coding cost under an independent-symbol model, is

$$H_0^\pi \equiv H(p^\pi) = - \sum_{i=1}^a p_i^\pi \log_2 p_i^\pi \leq \log_2 a. \quad (\text{S5})$$

Increasing a raises the ceiling $\log_2 a$ on this marginal entropy and generally lowers the distortion

$$D(a, \gamma) = \frac{1}{\gamma N^\pi} \sum_{y=1}^{N^\pi} \|Q_y - \bar{Q}_{i_y}\|^2 = \frac{1}{\gamma N^\pi} \sum_{i=1}^a \sum_{y: i_y=i} \|Q_y - \bar{Q}_i\|^2, \quad \bar{Q}_i = \frac{1}{N^\pi p_i^\pi} \sum_{y: i_y=i} Q_y, \quad (\text{S6})$$

where $\bar{Q}_i$ is the mean pose sequence of the windows coded by BoPS i . Dividing by γ makes $D(a, \gamma)$ the mean squared distance per frame between each pose and the corresponding pose of its BoPS centroid, measured in squared body lengths whatever embedding is used for clustering.

For a fixed γ , if the alphabet for $a+1$ is obtained by splitting one BoPS of the alphabet for a , H_0^π cannot decrease and $D(a, \gamma)$ cannot increase. For independently fitted alphabets, both trends hold only approximately. For a fixed a , longer windows generally raise the distortion, because more varied pose sequences must share each syllable, and they coarsen the time resolution of the track. The alphabet size a and window length γ can therefore be chosen at the knee of this entropy–distortion trade-off, with entropies for different γ compared per unit time, as $H_0^\pi/(\gamma\Delta)$ bits per second. They can also be chosen by the assignment error E^ω of the behavioral words built from BoPS strings (defined below), the analogue of the error E^κ for CAMs built from StaMEs (6).

Because consecutive BoPS often repeat, notably during maintained postures, H_0^π overstates the information carried per BoPS. The first-order conditional entropy, the entropy of each BoPS given the one before it, computed from successive pairs (6), accounts for this one-step dependence. For a stationary BoPS process it lies between the entropy rate of the track and H_0^π , so it reduces this bias but does not remove it. It equals the entropy rate only when a first-order Markov model describes the track adequately, and absence of pairwise correlation is not sufficient. Nonstationary records should be stratified or modeled with time-varying probabilities before a single entropy rate is interpreted.

Behavioral words

Following Tasks 3–6 of Getz (6), the BoPS track is cut into fixed-length strings of m consecutive BoPS, each spanning $m\gamma\Delta$ seconds. Each string is described by its BoPS sequence and by egocentric kinematic features, such as keypoint velocities and the periodicity of limb or appendage motion. Centroid displacement in world coordinates is deliberately left out, so that displacement is not built into the BoPS features later compared with the path (Section 3.3). The strings are clustered into b word types. Exact BoPS sequences rarely recur, because of jitter between windows and because a window can start at any phase of a movement. Each string is therefore reassigned to the word whose member strings are nearest to it under a string distance that tolerates these differences, for example edit distance, with cyclic-shift invariance used only for motifs whose phase is irrelevant

to the question. This reassignment is the analogue, for BoPS strings, of the rectification rule of Getz (6). It gives a lexicon

$$\begin{aligned} s_j &\equiv \pi_{j_1} \pi_{j_2} \cdots \pi_{j_m}, & j &= 1, \dots, N^{\text{wd}}, \\ \mathcal{W} &= \{\omega_1, \dots, \omega_b\}, \\ J_\nu &= \{j : s_j \text{ assigned to word } \omega_\nu\}, \end{aligned} \tag{S7}$$

where s_j is the j th observed string of m BoPS, N^{wd} is the number of strings in the track, and J_ν indexes all occurrences assigned to word type ω_ν , retaining repeated identical strings. Once fitted on training data, the lexicon is frozen. New strings are assigned to the nearest word using the same distance and a fixed tie-breaking rule, and the assignment error E^ω is defined exactly as E^κ is for CAMs (6). Runs of a repeated BoPS encode how long a posture is held, so words capture both the order and the timing of postures. Runs of mixed BoPS can encode repeated movements that no single BoPS captures, such as head nodding. Behavioral words are related to, but distinct from, the P-StaMEs defined below. Both are built from the BoPS track, but words cluster short strings of m BoPS that describe actions, whereas P-StaMEs cluster summaries of windows cut to the duration of a relocation segment, so that each StaME interval receives one L-StaME and one P-StaME, matched in length and timing.

Only a small fraction of all possible strings of m BoPS occur in a track. A track cannot contain more distinct strings than it has strings, so that fraction is limited by the length of the record and, on its own, mostly reflects sample size. Compared with the same fraction for shuffled or Markov-surrogate BoPS tracks of equal length, it tests for sequence structure beyond that preserved by the chosen null model, but it does not identify a biomechanical cause. Variable-length alternatives can replace either tier. Autoregressive HMM syllables (3, 54) can replace BoPS as variable-length units, and B-SOiD motifs (1) can stand in for words, with the fixed clock restored one level up (S6).

Concurrent actions and complete locomotor cycles

Behavioral words describe fixed-duration windows, whereas FuMEs describe complete gait cycles. The two are distinct tokenizations of the observed body sequence, retained as two aligned representations. A FuME is a phase-locked limb, wing or fin cycle with the support pattern of a locomotor gait, bounded by successive occurrences of a reference phase, such as touch-down of a limb. FuMEs are best identified with supervision. An expert marks the start, end and type of each cycle in example sequences, and a detector trained on these examples, using pose sequences or synchronized accelerometer signals, is then run on the larger data set (Fig. S3B). Detectors are validated on labeled sequences held out of training, reporting precision, recall and boundary error for each cycle type. Alternatively, unsupervised segmentation can find recurring cycles that an expert names afterwards. A cycle can be resampled to m_F phases for classification, while its original start, end and duration remain in the record. The cycle types form the locomotor catalog

$$\mathcal{F} = \{\phi_1, \dots, \phi_f\}, \tag{S8}$$

which is distinct from the fixed-window lexicon $\mathcal{W}$ (Eq. S7). Displacement is not an identification criterion. A locomotor cycle can occur during approximate stationarity in a chosen frame, and defining it by path displacement would build part of the body–path association into the code. Turning steps and backing up are included when they satisfy the cycle criteria. Because FuMEs vary in duration, they are not paired with fixed-length StaMEs. Their sequences overlay CAMs instead (section S6).

When several body parts act concurrently, their words are retained separately, optionally in separate body-part streams. A limb cycle can coincide with chewing, a periscope trunk posture or a social display. The cycle then remains a FuME, and the concurrent action remains in the behavioral record. Table 2 and Table S5 group candidate interpretations, and their rows need not be mutually exclusive. Observable form is recorded first, with functional interpretations supported by environmental and social context and independent validation. No locomotor-fraction threshold allocates a whole interval or word exclusively to locomotion or to another function. A limb cycle and a chewing word can overlap completely.

L-StaMEs

L-StaMEs are built from relocations (6, 7), and the approach applies to any movement data that can be clustered. The L-StaME alphabet is

$$\mathcal{S}^L = \{\sigma_1^L, \dots, \sigma_{n^L}^L\},$$

$$\sigma_i^L = \frac{1}{|\mathcal{C}_i^L|} \sum_{z \in \mathcal{C}_i^L} \mathbf{v}_z^L, \quad i = 1, \dots, n^L, \quad (\text{S9})$$

where $\mathbf{v}_z^L$ is the vector of summary statistics of the z th segment of μ consecutive relocation steps, such as the means and standard deviations of speed and turning angle, $\mathcal{C}_i^L$ is the set of segments in the i th of n^L clusters, and σ_i^L is that cluster's centroid, which also serves as its symbol (6, 7).

The common window: BoPS strings of StaME intervals

L- and P-StaMEs are paired on a common window. Suppose that the relocation interval is an integer multiple ρ of the pose clock Δ , with the two clocks aligned, so that each StaME interval of μ relocation steps spans $\mu\rho$ frames. Suppose also that the BoPS window length γ divides $\mu\rho$, so that the interval spans exactly $M = \mu\rho/\gamma$ consecutive BoPS of the track $\mathcal{T}^\pi$. The z th StaME interval, indexed as in Eq. S9, then covers the set of frames τ_z and the set of BoPS windows Y_z . Its BoPS content is the string $\pi^{(z)}$ of the M BoPS assigned to those windows, and $\mathcal{P}^{(\mu)}$ is the set of distinct BoPS strings that occur across the $N^\mu = \lfloor T/(\mu\rho) \rfloor$ StaME intervals covered by the track:

$$\tau_z = \{(z-1)\mu\rho + 1, \dots, z\mu\rho\}, \quad Y_z = \{(z-1)M + 1, \dots, zM\}, \quad z = 1, \dots, N^\mu, \quad (\text{S10})$$

$$\pi^{(z)} = (\pi_{i_y})_{y \in Y_z}, \quad \mathcal{P}^{(\mu)} = \{ \pi^{(z)} : z = 1, \dots, N^\mu \}.$$

The BoPS string of a randomly chosen StaME interval is written Π , with values in $\mathcal{P}^{(\mu)}$. Because every $\pi^{(z)}$ has the same length M , coding by BoPS strings uses the same fixed interval support as the relocation construction of Getz (6). When γ does not divide $\mu\rho$, M is not an integer, and interval boundaries fall inside BoPS windows, so that intervals contain different numbers of whole BoPS. The simplest remedy is to choose γ among the divisors of $\mu\rho$, which is usually possible because γ is tunable and $\mu\rho$ is large. With 4-s relocations filmed at 30 frames per second, for example, $\rho = 120$, and $\mu = 5$ gives $\mu\rho = 600$ frames, which $\gamma = 5, 6, 8, 10$ or 12 divides. Otherwise, each BoPS is assigned to the interval that contains its middle frame, as each FuME is assigned to the interval that contains its reference phase, so that every BoPS belongs to exactly one interval. Y_z is then the set of BoPS windows whose middle frames lie in τ_z , intervals contain $\lfloor M \rfloor$ or $\lceil M \rceil$ BoPS, and P-StaMEs are built from BoPS proportions rather than counts, or from string distances that tolerate unequal lengths. The same rule, with τ_z taken as the frames whose timestamps fall within the interval, handles relocations that are not a whole number of frames apart or are irregularly

spaced. A window enters the paired analysis only if both streams meet prespecified coverage rules within it. Windows with too few valid BoPS or relocations are flagged as missing rather than coded, and they are not filled by interpolation.

P-StaMEs

In the BoPS construction, the BoPS track is cut into windows of the same duration as a relocation segment, so that both constructions share one clock. Window z then covers the frames τ_z and BoPS windows Y_z of Eq. S10. When no relocations are recorded, it covers $Y_z = \{(z-1)\Lambda + 1, \dots, z\Lambda\}$ for a chosen window length of Λ BoPS, or $\Lambda\gamma$ frames. Each window is summarized by up to three parts:

1. its BoPS content, as counts over the alphabet $\mathcal{P}$ or as the BoPS string itself,
2. gait descriptors computed from the egocentric keypoints, such as stride frequency, duty factor and the regularity of limb cycles, and
3. displacement descriptors, speed and turning, taken from the video centroid track where the video is georeferenced, or otherwise estimated from stride frequency and body-axis rotation, as in dead reckoning with inertial tags.

If $\mathbf{h}_z = (h_{1z}, \dots, h_{az})$ is the vector of BoPS counts in window z , and $\mathbf{g}_z$ and $\mathbf{x}_z$ are its gait and displacement descriptors, then these three parts form three nested feature vectors $\mathbf{v}_z^P$ of window z , which define the pose-only, pose-plus-gait and displacement-augmented constructions:

$$\mathbf{v}_z^P \in \{ \mathbf{h}_z, (\mathbf{h}_z, \mathbf{g}_z), (\mathbf{h}_z, \mathbf{g}_z, \mathbf{x}_z) \}, \quad h_{iz} = |\{ y \in Y_z : i_y = i \}|, \quad i = 1, \dots, a. \quad (\text{S11})$$

Standardizing each component and clustering these vectors with the same methods used for L-StaMEs (Eq. S9) yields the set of P-StaMEs

$$\mathcal{S}^P = \{ \sigma_1^P, \dots, \sigma_{n^P}^P \}, \quad \sigma_i^P = \frac{1}{|\mathcal{C}_i^P|} \sum_{z \in \mathcal{C}_i^P} \mathbf{v}_z^P, \quad i = 1, \dots, n^P, \quad (\text{S12})$$

where $\mathcal{C}_i^P$ is the set of windows in the i th of n^P clusters, and σ_i^P is that cluster's centroid. When windows are compared by string distances on their BoPS strings $\pi^{(z)}$ rather than by counts, each σ_i^P is instead the medoid of its cluster.

The relocation and BoPS constructions are therefore separate pipelines (Fig. S1). The path pipeline derives step features from relocations, summarizes each segment of μ steps by $\mathbf{v}_z^L$, and clusters these vectors into $\mathcal{S}^L$ (Eq. S9). The BoPS pipeline converts time-stamped keypoints (time, individual, keypoint identity, coordinates and confidence) into egocentric poses, cuts the pose track into windows of γ frames and codes each by its BoPS (Eq. S3), summarizes each StaME window by $\mathbf{v}_z^P$ (Eq. S11) and clusters these vectors into $\mathcal{S}^P$ (Eq. S12). Neither codebook is derived from the other. Each codebook is fitted on training data and then frozen, so that new windows are assigned using the same feature definitions, standardization and nearest-centroid or nearest-medoid rule. The three constructions of Eq. S11 are fitted and reported separately.

The three constructions answer different questions. The pose-only and pose-plus-gait constructions use egocentric information alone, so their agreement with L-StaMEs, $U(S^L | S^P)$ (Eq. 3), quantifies the fraction of L-StaME uncertainty reduced by the P-StaME code. The displacement descriptors $\mathbf{x}_z$ come from the same motion that the path records, so an increase in $U(S^L | S^P)$ for the displacement-augmented construction can reflect shared displacement information. $U(S^L | S^P)$ should therefore be reported for each construction, with the increment from adding $\mathbf{x}_z$ shown separately. Section S6 describes how it is estimated and how held-out prediction complements it.

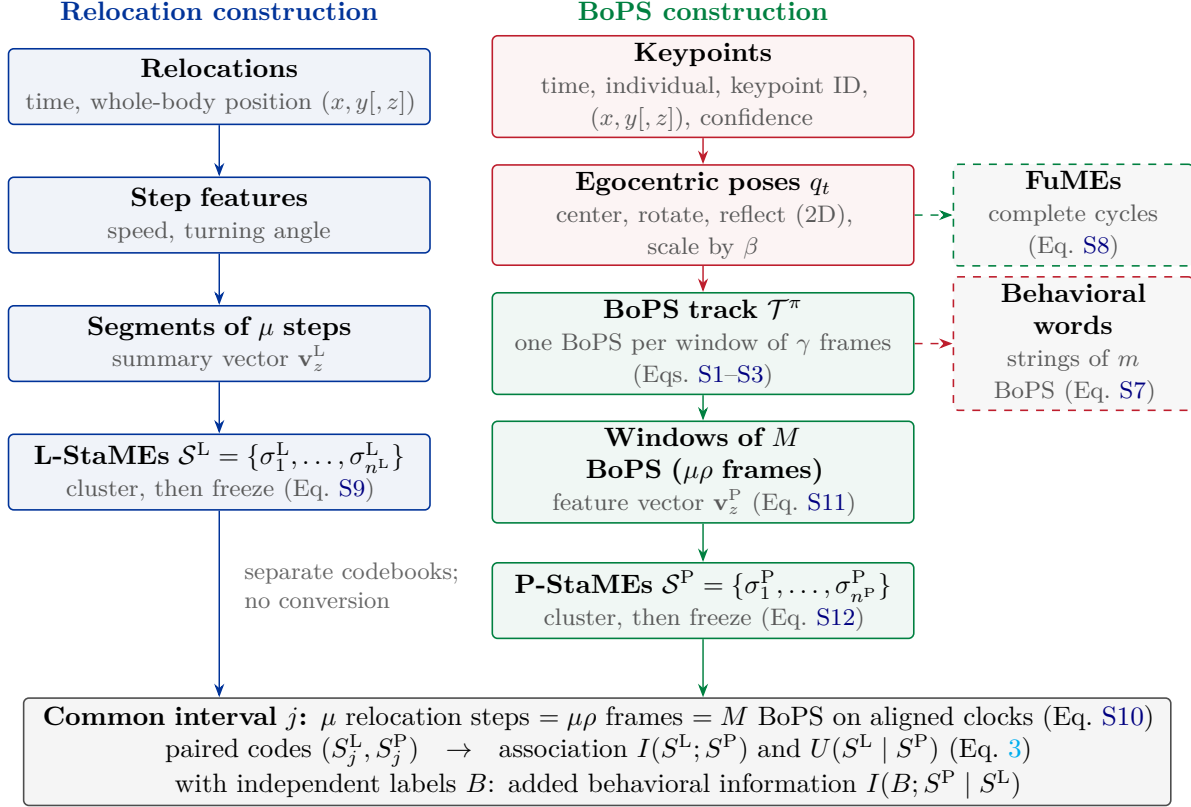


Figure S1: Two StaME constructions on one clock. Relocations are summarized over segments of μ steps and clustered into L-StaMEs. Keypoints are normalized into egocentric poses, cut into windows of γ frames and coded as BoPS, and the BoPS are summarized over windows of the same duration as a path segment and clustered into P-StaMEs. Behavioral words, built from the BoPS track, and FuMEs, detected from the poses, are parallel records (dashed). Neither is a P-StaME, and FuME sequences overlay CAMs rather than StaMEs (section S6). The two codebooks are fitted and frozen separately and are paired by time, not converted into one another. Colors are as in Fig. 1.

Where relocations and pose cover the same intervals, joint StaMEs can also be built by clustering each interval’s standardized relocation and BoPS features joined into one vector. Because joint StaMEs combine both sources, they cannot measure how much one code predicts the other. Joint StaMEs should therefore be reported as a separate construction, with the relative weighting of the relocation and BoPS features stated.

The two constructions capture different things. The relocation construction describes the geometry of displacement: how fast, how straight and how far the animal moved during an interval. When an animal moves little, its relocations look much the same whatever it is doing, so intervals of relative stationarity may share one L-StaME even though the animal rested in one, fed in another and scanned in a third. The BoPS construction describes the body itself, so it can separate such intervals by the configurations the animal adopted. A P-StaME built from BoPS counts records which configurations occurred but not their order, whereas one built from BoPS strings also keeps their order. Even so, calling a P-StaME rest, feeding or scanning is an interpretation, which must be validated against independent behavioral labels and the context in which the P-StaME occurred.

Context raises a further distinction. Placing a P-StaME on a map requires georeferenced video or another measurement of location. Many local relationships, however, can be read from the video itself without georeferencing, such as contact with a feature like a tree or water, or the positions

and orientations of visible neighbors. These contextual measurements should be kept out of the primary comparison, which uses P-StaMEs built from body data alone, so that agreement between P- and L-StaMEs reflects pose rather than context. A code that adds context should be reported as a distinct construction.

CAMs

Words of η consecutive L- or P-StaMEs are clustered into L- or P-CAMs (6), with alphabets

$$\mathcal{K}^L = \{\kappa_1^L, \dots, \kappa_{k^L}^L\}, \quad \mathcal{K}^P = \{\kappa_1^P, \dots, \kappa_{k^P}^P\}, \quad (\text{S13})$$

where k^L and k^P are the numbers of L- and P-CAM types. Words of P-StaMEs thus build P-CAMs exactly as words of L-StaMEs build L-CAMs. For a randomly chosen CAM interval, C^L and C^P denote its L- and P-CAMs, and C denotes either, with types κ_i , $i = 1, \dots, k^X$, $X=L$ or P as the case may be. When L- and P-CAMs are size-matched, built from the same number of StaMEs over the same intervals, their intervals coincide and can be paired, as StaMEs are. FuMEs are not used to build CAMs but to help interpret them. Their sequences overlay L- and P-CAMs instead (section S6).

The considerations above for StaMEs apply to CAMs as well. Sethi et al. (7) judged the homogeneity of owl L-CAMs by the L-StaME composition of the words assigned to each (Section 3.4). Because video allows P-StaMEs to be built from much shorter windows, P-CAMs may be more homogeneous, but whether they are must be tested against independent behavioral labels, not by composition alone. Whether an L-CAM and a P-CAM given the same name describe the same behavior is tested by how often they co-occur on paired intervals.

In practice, the two codes play complementary roles. L-StaMEs and L-CAMs can be computed over long records, such as months of collar data. P-StaMEs and P-CAMs are available only where video exists, and there they label, validate and refine the L-StaMEs and L-CAMs.

S5 Sensors, clocks and reference frames

No single sensor yet records postures and relocations across a lifetime (55). Movement Ethology therefore relies on fusing sensors, and on using short, rich recordings to train detectors that run on long, sparse ones (Table S7).

The most practical way to obtain FuMEs in the wild is pose-to-sensor transfer. Synchronized video and tag data are first recorded on captive, habituated or drone-filmed animals. An expert marks the start, end and type of complete locomotor cycles in example sequences. Pose-based detectors trained on these examples then label cycles and separate body-part actions in the synchronized sensor record, retaining their concurrent occurrence. Classifiers trained on those labels carry FuME detection into months of field data, where no camera can follow (56). Training collar classifiers on video annotated by human observers is established (11, 57). The new element here is that the labels are derived from pose. Biomechanically informed features help these classifiers generalize across individuals (57), and accelerometers on several body sites recover more of the posture than a single collar (58). Whether a FuME or other action can be detected at all depends on where the accelerometer is placed and on whether its signal is observable for that action. A neck collar is well placed to detect grazing or browsing by herbivores, but cannot tell whether an elephant's trunk is plucking a leaf or moving an object. Ground-truthing observations, from synchronized video or direct observation, are therefore needed to assign accelerometer signals to behaviors, and detectability should be established separately for each action and sensor placement before collar-only

results are interpreted. In particular, a CAM interval with no detected cycles in a collar-only record is a recorded absence only for cycle types that the placement is known to register. For other types it is uninformative (section S6). Classifiers should be validated on individuals and field contexts held out of training, with performance reported for each behavior and uncertainty retained for actions or contextual relationships that the sensor cannot resolve.

Clocks matter since an element shorter than the sampling interval cannot be seen (5). FuMEs therefore need video or inertial sampling several times faster than the fastest stride or wingbeat, so that each FuME is represented by several pose samples, while StaMEs need relocations at 0.1 Hz or faster (6). All streams must share a clock. A FuME can be assigned to a CAM interval only if the two clocks are aligned to well under one cycle.

Pairing P- and L-StaMEs on shared intervals adds further requirements. Record each device’s timestamps with their source, estimate and correct clock offset and drift between video, tags and GPS, for example from synchronization events visible in both streams, and report the residual misalignment. Keep individual identity and keypoint identity with every pose sample, and keep egocentric coordinates separate from global coordinates. Georeferencing is needed only when body observations must be placed geographically. A single camera cannot recover configurations that differ only in depth, so 3D reconstruction or a viewpoint covariate is needed where these matter. For each interval, report the coverage of each stream, as the fraction of valid frames and of expected relocations. Mark occlusion, sensor gaps and tracking failure as missing, distinct from an observed absence of an action or of a completed locomotor cycle.

Reference frames differ as well. Normalized pose q_t describes the body in its own frame, whereas whole-body position u_t and orientation R_t relate it to the surroundings. In 3D, orientation requires a full rotation, not only a heading angle. Body heading and path direction are distinct: a difference can reflect sideways movement, environmental transport or measurement error. Head orientation relative to the body is a separate measurement. Path direction is unreliable at very low displacement because positional error dominates.

Relative stationarity must specify a reference frame, a spatial tolerance and an observation interval. The frame may be the ground, a feeding patch, a particular neighbor or a moving group. An individual can remain relatively stationary within a moving group while changing geographic location. The within-interval trajectory must be examined: small endpoint displacement alone does not distinguish residence within an area from departure and return. Report the operational rule, sampling resolution and positional uncertainty, and retain body actions throughout the interval.

A shared Movement Ethology record should include synchronized timestamps, individual identity, whole-body location and its uncertainty, egocentric keypoints and their confidence, body orientation and scale, sensor observations and, where visible, contextual measurements c_t . The latter include locations or contacts involving local environmental features, and the identities, relative positions and orientations of other individuals. Preserve frame definitions and transformations, event boundaries and durations, concurrent body-part words, FuME codes, L- and P-StaMEs, and missing-data flags. Relative-stationarity classifications must retain their frame, tolerance and interval. Movebank stores paths and tag data (41), and *movement* (59) supports tracker-output processing. The contextual and event fields specified here describe the proposed record, not features assumed to exist in those platforms.

S6 Information measures, estimation and interpretation

This section treats the codes of section S4 as random variables and gives the information measures computed from them, each followed by how it is estimated and interpreted. It applies the data-

processing inequality to the BoPS strings of StaME intervals, defines the coupling between CAMs and their FuME overlay, and compares paired L- and P-StaMEs. It closes with the two stages of analysis (Fig. S3) and the parsing of variable-length structure. Several estimates rely on held-out prediction, in which codes and predictive models are fitted to the intervals of some individuals or time blocks, and their predictions are then tested on intervals set aside from all fitting (60).

Codes as random variables

StaMEs and CAMs can be treated as symbols of a Shannon code, whose entropy, assignment error E^κ for words assigned to rectified CAMs, and Jensen–Shannon divergences compare codings and clustering methods (6, 7). For a randomly sampled valid StaME interval, let S^L and S^P be its L- and P-StaME symbols and Π its BoPS string. Their values are defined by frozen codebooks and the common interval convention of Eq. S10. For a randomly chosen CAM interval, C^L , C^P and C are its CAMs, as defined after Eq. S13.

Each StaME interval also carries a prespecified summary A of its behavioral words, potentially including simultaneous body-part codes, their timing and duration. Because a word can outlast a StaME interval, A summarizes the words to which the interval’s BoPS belong, for example the word types present, the fraction of the interval each occupies or the dominant word. Each CAM interval likewise carries a prespecified summary Ψ of its overlaying FuME sequence, defined below. Counts, durations and time ordering answer different questions, so the chosen summaries and their temporal support must be reported. A recorded absence of an action is distinguished from missing observations, and low-confidence or occluded measurements carry a missing-data flag rather than a behavioral label.

BoPS strings and the data-processing inequality

Coding each StaME interval by its BoPS string Π (arrow labeled “BoPS windows” in Fig. 1) allows a data-processing comparison with any code assigned from Π alone by a frozen rule. The pose-only P-StaME is such a code when it is computed from the interval’s BoPS string, because a frozen deterministic assignment makes it a function h of that string. Then

$$I(S^L; \Pi) \geq I(S^L; S^P), \quad S^P = h(\Pi). \quad (\text{S14})$$

Equation S14 requires that the compressed code be assigned from Π alone, with a frozen rule. The pose-plus-gait and displacement-augmented constructions also use continuous pose kinematics or other sensors, so they can distinguish observations assigned the same BoPS string, and Eq. S14 need not order them. All these codes can be viewed as compressions of a sufficiently complete body record, but this conceptual statement does not promise tractable mutual-information estimation for that high-dimensional record.

Under the deterministic condition, the inequality also holds for the empirical distribution. Its plug-in estimates can nevertheless be uninformative. When every observed BoPS string is unique, the empirical $H(S^L | \Pi)$ is zero, so the estimated $I(S^L; \Pi)$ equals the empirical $H(S^L)$ even for unrelated processes. With a pose clock of 0.1 s and BoPS of $\gamma = 5$ frames, for example, a 10-s interval holds 20 BoPS, so most observed strings outside long maintained postures will be unique. Finite codes and fixed estimands are therefore essential, and an entropy normalized by $H(S^L)$ is not automatically bias-corrected. In practice, $I(S^L; \Pi)$ is approached from below, through compressed functions $h(\Pi)$ such as BoPS counts or pose-only P-StaMEs, for which $I(S^L; h(\Pi)) \leq I(S^L; \Pi)$, or through held-out prediction.

Held-out predictive scores depend on the model used and are not exact mutual information, so scores from separate BoPS-string and P-StaME classifiers need not preserve the ordering in Eq. S14. Genuine mutual-information comparisons require a specified estimator with suitable bias and dependence controls. BoPS counts also discard temporal order, so a claim about complete strings does not automatically apply to counts. FuMEs are detected from continuous pose or accelerometer signals and summarized over CAM intervals rather than StaME intervals, so Eq. S14 does not apply to them.

CAMs and the FuME overlay

FuMEs are not paired with StaMEs. A CAM interval spans $\eta\mu$ relocation steps, seconds to minutes, and so typically contains many complete cycles. Each cycle is assigned to the CAM interval that contains its reference phase, such as touch-down, so that no cycle is lost at a boundary, and the cycles assigned to an interval form the FuME sequence that overlays it. Because such sequences are long and rarely recur exactly, Ψ is a prespecified summary of the sequence, for example the composition of FuME types, their rates, the dominant gait or counts of transitions between types. The coupling between CAMs and FuMEs (Eq. 1) then has the plug-in form

$$\begin{aligned} I(C; \Psi) &= H(C) - H(C | \Psi) \\ &= - \sum_{i=1}^k p_i^\kappa \log_2 p_i^\kappa \\ &\quad + \sum_{i=1}^k \sum_{\ell=1}^{N^\psi} p(\kappa_i, \psi_\ell) \log_2 p(\kappa_i | \psi_\ell). \end{aligned} \quad (\text{S15})$$

Here ψ_ℓ , $\ell = 1, \dots, N^\psi$, are the distinct values of Ψ , p_i^κ is the proportion of CAM intervals coded as κ_i , and $p(\kappa_i, \psi_\ell)$ is the proportion that are coded as κ_i and whose FuME summary is ψ_ℓ , counting one observation per CAM interval. A mutual information near $H(C)$ means that CAMs are statistically predictable from their FuMEs, without establishing a causal mechanism. A low value means the same FuME sequences occur within many CAM types. That can reflect terrain, wind, current or social context, but also positional or keypoint error, coarse codes, misaligned clocks or untracked body parts, which must be controlled for before the gap is attributed to external factors.

Why Ψ can be empty. The FuME sequence of a validly observed CAM interval is empty when no complete cycle of a cataloged type (Eq. S8) has its reference phase inside the interval. Ψ then takes its own value, $\emptyset$, because a composition is undefined for an empty sequence. Four situations produce an empty sequence, with different meanings:

1. *No locomotion.* The animal rested, fed in place, groomed or stood vigilant. Behavioral words, not FuMEs, then describe what the body did. Most empty intervals are of this kind.
2. *Displacement without self-propelled cycles.* A soaring bird, an animal drifting in a current or a young animal carried by its mother moves along the path without cycling its own limbs, wings or fins. Here an empty Ψ is informative, because it separates passive transport from self-propelled movement within one L-CAM.
3. *Movement without a complete cycle.* Weight shifts, single steps, turns on the spot and the unfinished last cycle of a bout (Fig. S3B) are not FuMEs. A CAM interval shorter than one cycle, possible for very slow gaits or short P-CAM windows, can also contain no reference phase even during steady locomotion.
4. *Undetected cycles.* The cycle type is not in the catalog, the sensor placement cannot register it (section S5), or the detector missed it. Detector recall for each cycle type, estimated on held-out labeled sequences, bounds how often this happens.

Situations 1 to 3 are properties of the movement and contribute legitimately to $I(C; \Psi)$, whereas situation 4 is measurement error. An empty sequence is thus a recorded absence of completed, detected cycles, not evidence of stationarity. It differs from a missing interval, in which tracking loss, low keypoint confidence or a sensor gap means that no cycle could have been detected. Missing intervals are flagged (the unassigned category of Table 2) and excluded. Because empty intervals concentrate in stationary CAMs, they can dominate $I(C; \Psi)$, which may then reflect little more than the separation of stationary from moving CAMs. The proportion of empty intervals should therefore be reported for each CAM type, alongside the coupling among locomotor CAM intervals described next.

The FuME overlay also refines the interpretation of the CAM track. The conditional entropy $H(\Psi | C = \kappa_i)$ measures how much locomotor variation the CAM type κ_i merges. An L-CAM of directed travel, for example, may combine walking and trotting strides, and a searching L-CAM may combine slow stepping with frequent turning steps. Splitting a CAM type by its FuME summaries gives candidate subtypes, which are interpreted, as other codes are, against independent labels. Because plug-in estimates are biased upward by an amount that depends on sample size, comparisons across habitats, ages or species should use matched sampling and sensor quality, frozen codebooks and dependence-preserving null models. Normalized coupling $I(C; \Psi)/H(C)$ is defined only when $H(C) > 0$. Because $I(C; \Psi) \leq \min\{H(C), H(\Psi)\}$, this ratio also falls when locomotion becomes rarer or less varied, without any change in how closely cycles and CAMs are related. $H(\Psi)$ should be reported with it, and coupling compared among locomotor CAM intervals ($\Psi \neq \emptyset$) or at matched locomotor fractions. Richer summaries Ψ make the estimates sparser, so the caution given above for Π applies to Ψ as well. Uncertainty and predictive performance should be assessed on held-out individuals or time blocks.

Paired L- and P-StaMEs

For each valid interval j , the paired observation (S_j^L, S_j^P) combines the L-StaME and the P-StaME assigned by separately frozen codebooks (Figs. S1 and S2). Four quantities answer different questions and should not be substituted for one another.

1. Association of L- and P-StaMEs, $I(S^L; S^P) = H(S^L) - H(S^L | S^P)$, and its directional normalization $U(S^L | S^P)$ (Eq. 3), reported with $H(S^L)$, $H(S^P)$ and the joint distribution of codes.
2. Coupling of CAMs and their FuME overlay, $I(C; \Psi)$ (Eq. S15), defined on CAM intervals rather than StaME intervals. Ψ and S^P are different random variables at different scales.
3. Within-category heterogeneity, $H(S^P | S^L = \sigma_i^L)$, the entropy of P-StaMEs among intervals sharing the L-StaME σ_i^L . It can be positive for every σ_i^L while $I(S^L; S^P) = 0$, and it can reflect measurement noise or viewpoint as well as behavior.
4. Added behavioral information, $I(B; A | S^L)$ (Eq. 2) or, with the P-StaME as the prespecified BoPS summary, $I(B; S^P | S^L) = H(B | S^L) - H(B | S^L, S^P)$. Only this quantity, with independently validated labels B , tests whether P-StaMEs distinguish behavior beyond L-StaMEs. Conditioning on S^L alone tests information beyond the interval's L-StaME. To claim information not recoverable from relocations at all, condition instead on a prespecified richer path summary, such as the L-StaME together with its L-CAM and preceding path history, computed from relocations only. Labels B must not be derived from that summary.

These quantities are estimated with the cautions above. Plug-in estimates, including that of $U(S^L | S^P)$, describe the sample on which they are computed and are biased upward when code

combinations are rare. Report the number of intervals for each combination, use bias-corrected estimates, and compare with nulls that preserve the temporal dependence of each stream, for example by circularly shifting the P-StaME code relative to the L-StaME code. Whether P-StaMEs predict the L-StaMEs of new data is tested by held-out prediction. Report $U(S^L | S^P)$, with a finite-sample correction and its uncertainty, alongside the held-out predictive scores, and take uncertainty from resampling individuals or time blocks. Check sensitivity to the window length μ , the alphabet sizes n^L and n^P and the P-StaME construction, and report the increment from adding $\mathbf{x}_z$ separately. Use the pose-only or pose-plus-gait construction when the question is what P-StaMEs add, since displacement descriptors share information with the path.

Labels B in Eq. 2 that are generated deterministically from the same S^L would make $I(B; A | S^L)$ zero. Labels derived from a longer path history need not, so label provenance and temporal support must be reported. Functional labels inferred from the very body features being tested also risk circular validation. Environmental and social observations c_t supply context for interpretation, and analyses can additionally condition on a prespecified summary of that context. Activity budgets allow overlapping categories, report their overlaps and missing-data coverage, and need not sum to one across functions.

Table S11 gives a synthetic example. Two intervals share the L-StaME σ_1^L , so L-StaMEs do not distinguish them, yet their P-StaMEs differ. Whether the difference corresponds to different behavior is decided by independent labels, not by the P-StaMEs. The intervals share an L-StaME, not necessarily the same raw relocation history.

Table S11: Synthetic example of paired L- and P-StaMEs. The entries are hypothetical, not observed episodes, and do not imply that any BoPS pattern indicates an ecological function.

Interval	L-StaME S^L	P-StaME S^P	BoPS-derived pattern (illustrative)	Independent label B
$j = 1$	σ_1^L	σ_1^P	Repeated head-down and jaw motions	Feeding (independently validated)
$j = 2$	σ_1^L	σ_2^P	Head turns, scanning	Vigilance (independently validated)

Two stages: paired codes and CAMs, then the FuME overlay

The paired codes, the CAMs and the FuMEs enter the analysis in a fixed order, summarized in Fig. S3.

Stage 1, without labels. L- and P-StaME codebooks are fitted on training individuals or time blocks and frozen (section S4). Each valid interval j is assigned its pair (S_j^L, S_j^P) , and the joint distribution, association and within-category heterogeneity are reported (items 1 and 3 above). Words of η StaMEs are then clustered into L- and P-CAMs (Eq. S13), also frozen. The result is an index of StaME and CAM intervals, each with its individual, time, location and codes and, optionally, the window feature vectors.

Stage 2, with expert-labeled FuMEs. A detector trained on expert-labeled examples (section S4) is run on pose sequences or synchronized accelerometer signals. Each detected cycle keeps its type, onset, offset and detection confidence, and is assigned to the CAM interval that contains its reference phase. The cycles of each CAM interval form its FuME sequence, summarized by Ψ . For each FuME type ϕ_k , a CAM interval can also be scored by whether it contains at least one complete cycle of that type, or by the count or rate of such cycles. Five queries lay the FuMEs over the frozen

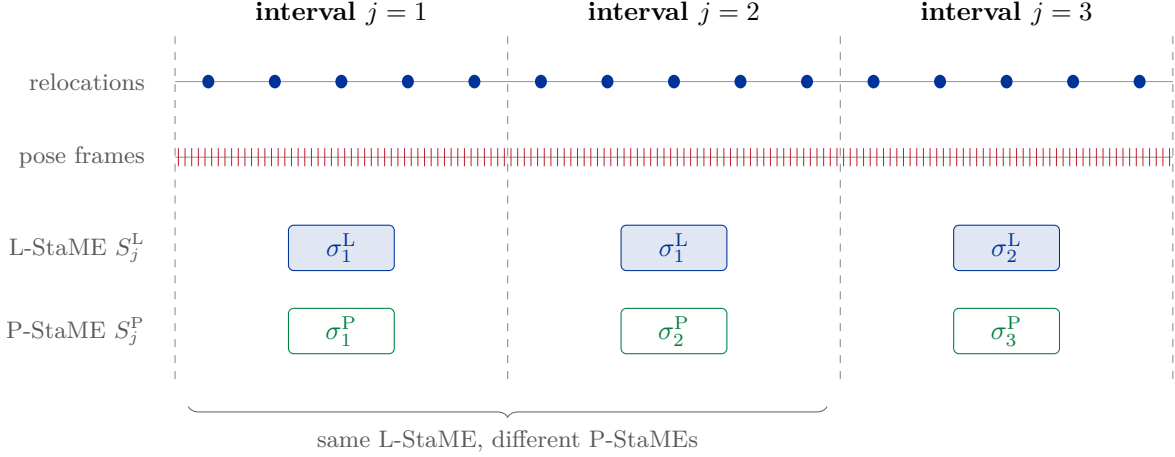


Figure S2: Synthetic illustration of paired codes. Each interval of μ relocation steps spans $\mu\rho$ pose frames (here $\mu = 5$ and $\rho = 10$), which are grouped into $M = \mu\rho/\gamma$ BoPS of γ frames each, and each valid interval yields one pair (S_j^L, S_j^P) . Intervals 1 and 2 share the L-StaME σ_1^L but receive different P-StaMEs. Whether that difference corresponds to different behavior is tested with independent labels (Table S11). The codes are hypothetical, not measured. Colors are as in Fig. 1.

CAM tracks.

1. *Where does a type occur?* With size-matched L- and P-CAMs, whose intervals coincide, each pairing of an L-CAM type with a P-CAM type defines a cell. The enrichment of ϕ_k in a cell is the proportion of the cell's CAM intervals that contain ϕ_k , divided by that proportion over all CAM intervals. Values above 1 mean that ϕ_k is more common in the cell than overall, and each value is reported with the number of CAM intervals in its cell.
2. *What does one CAM type merge?* $H(\Psi \mid C = \kappa_i)$, the locomotor variation within CAM type κ_i , and whether P-CAMs separate FuME types that a single L-CAM type merges.
3. *How strongly are CAMs coupled to their FuMEs?* $I(C; \Psi)$, estimated separately for L- and P-CAMs with the plug-in form of Eq. S15 and the cautions given there.
4. *When and where does a type occur?* Retrieval of the CAM intervals containing ϕ_k , with their times, locations, codes and higher-level context, such as BAM and DAR.
5. *Collar-only records.* For FuME types that the sensor placement can observe (section S5), accelerometer-detected FuMEs and relocations give C^L and Ψ over the whole collar record, so $I(C^L; \Psi)$ and the enrichment of ϕ_k over L-CAMs can be estimated there directly. C^P is available only within video coverage.

Four conditions apply. First, Stage 1 is frozen before Stage 2, and refining a CAM codebook to match FuMEs defines a new model that must be validated on new data. Second, FuMEs detected from video and P-CAMs come from the same pose record, so part of their association is built in, and, for P-CAMs, queries 1 to 3 then interpret the CAMs rather than test them. L-CAMs, and FuMEs detected from accelerometers, come from separate sensors, so their agreement is evidence. Third, detector output is a FuME code, not a label B . Added behavioral information, $I(B; S^P \mid S^L)$ or $I(B; \Psi \mid C^L)$, is tested with expert annotations held out from detector training. Fourth, FuMEs are locomotor, so most CAM intervals of relative stationarity have empty FuME sequences (see *Why Ψ can be empty*, above). Expert-labeled non-locomotor actions, such as chewing or scanning, can be laid over the CAM tracks in the same way as FuMEs. Enrichments over many cells and

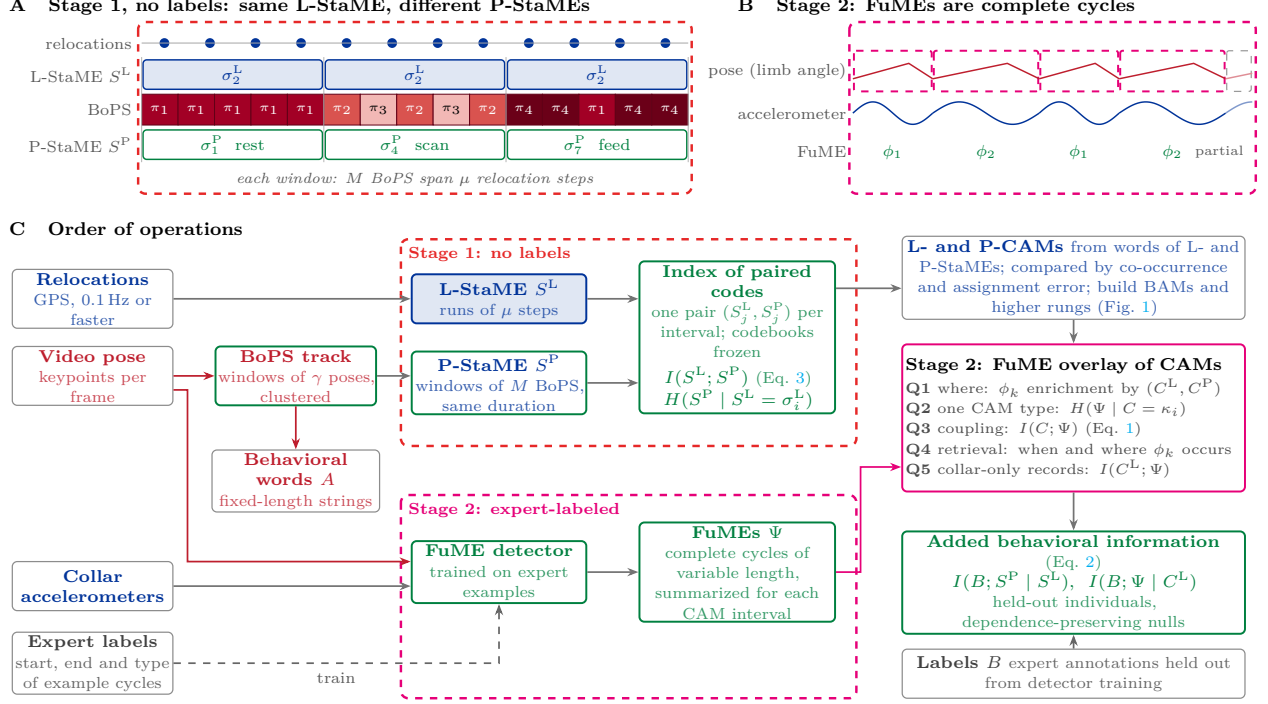


Figure S3: Two stages of analysis. (A) Stage 1, without labels. On one clock, each window of M BoPS spans μ relocation steps. All three intervals receive the same L-StaME, σ_2^L (limited displacement), but different P-StaMEs, here rest, scanning and feeding, so the L-StaME hides variation in the P-StaMEs, measured by $H(S^P | S^L = \sigma_2^L)$. (B) Stage 2. Complete locomotor cycles of variable length, found in the pose sequence or in a synchronized accelerometer trace, are FuMEs (types ϕ_1, ϕ_2). The incomplete final cycle is not a FuME. (C) Order of operations. In Stage 1 (red dashed), L- and P-StaMEs are fitted, frozen and paired into an index of intervals, and words of them are clustered into L- and P-CAMs. In Stage 2 (pink dashed), a detector trained on expert-labeled examples finds FuMEs in pose sequences or collar accelerometer signals, and the FuME sequences are laid over the CAM tracks (Section 3.3). Behavioral words are kept in parallel, so that concurrent actions are retained. Expert labels held out from detector training (B) test whether P-StaMEs add behavioral information beyond L-StaMEs. Colors are as in Fig. 1, and codes and examples are illustrative.

types require control of the false discovery rate, and nulls should circularly shift each individual's event train relative to its interval sequence.

Fixed-length symbols and variable-length structure

Fixed-duration BoPS and path intervals provide a convenient common clock. Variable-duration FuMEs and other events remain valid descriptions when their boundaries and durations are stored. A StaME interval can be represented by its BoPS string or concurrent word summary, and a CAM interval also by its overlaying FuME sequence, depending on the question. CAMs constructed from a fixed number of StaMEs have fixed support in this implementation, whereas longer BAMs and LiMPs can have variable duration. Proposed parsing methods include categorical autoregressive models (61–63), change-point analysis (64, 65) and hidden Markov models (66–68), reviewed for movement data by Edelhoff et al. (69).

S7 A schematic state-space model linking body and path

This section gives the state-space extension of the Movement Ecology framework (5) summarized in Section 4.1 and Fig. 2.

Body configuration and its relationships to the surroundings enter as measured quantities. Let u_t denote whole-body position, q_t normalized pose, R_t orientation, c_t observed environmental and social relationships, and w_t latent internal state. External conditions r_t include substrate, weather and other relevant influences. A schematic state-space extension is

$$\begin{aligned} q_{t+1} &\sim p_Q(\cdot \mid q_t, R_t, u_t, c_t, w_t, r_t, \Omega, \Phi), \\ (u_{t+1}, R_{t+1}) &\sim p_U(\cdot \mid u_t, R_t, q_t, q_{t+1}, c_t, w_t, r_t, \beta), \\ w_{t+1} &\sim p_W(\cdot \mid w_t, q_t, u_t, c_t, r_t). \end{aligned} \tag{S16}$$

Here t indexes a specified analysis clock, and sensors provide noisy or missing observations on their own sampling schedules. Motion and navigation capacities Ω and Φ enter through the pose model, matching Fig. 2. This factorization assumes that they act on location and orientation through the body. Where that assumption is too strong, for example when navigation sets heading directly, they can be added to p_U . In either version, integrating out the body states yields a path model that depends on Ω , Φ , r , w and location history, of the general form $u_{t+1} = F(\Omega, \Phi, r_t, w_t, u_t)$ of Nathan et al. (5). The relationships c_t are treated as observed. Forward simulation would also require a model of how neighbors and local features change relative to the animal. The equations are a schematic factorization, with histories or additional physical states included where needed, and they do not identify mechanisms from poses alone. In particular, p_U must account for propulsion, contact or environmental transport. Rotating and scaling an egocentric pose locates body parts relative to the centroid but does not determine the centroid’s next location. In a 3D model, R_t is a full rotation, whereas a planar model can use a heading angle. The measured relationships c_t let the same pose be interpreted in relation to food, substrate or neighboring individuals, with unobserved context retained as missing or latent. Discrete BoPS and behavioral-word models provide coarser alternatives, and latent-state models can represent unmeasured internal conditions (66, 67).

SI References

- [1] AI Hsu, EA Yttri, B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat. Commun.* **12**, 5188 (2021).
- [2] TD Pereira, et al., Fast animal pose estimation using deep neural networks. *Nat. Methods* **16**, 117–125 (2019).
- [3] C Weinreb, et al., Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics. *Nat. Methods* **21**, 1329–1339 (2024).
- [4] TD Pereira, et al., SLEAP: a deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495 (2022).
- [5] R Nathan, et al., A movement ecology paradigm for unifying organismal movement research. *Proc. Natl. Acad. Sci. USA* **105**, 19052–19059 (2008).
- [6] WM Getz, An information theory treatment of animal movement tracks in *The Mathematics of Movement*, eds. L Giuggioli, PK Maini. (Springer, Cham), pp. 175–193 (2026).

- [7] V Sethi, et al., An information theory framework for hierarchical path segmentation and analysis of animal movement. *Ecol. Informatics* **92**, 103401 (2025).
- [8] BR Koger, et al., Quantifying the movement, behaviour and environmental context of group-living animals using drones and computer vision. *J. Animal Ecol.* **92**, 1357–1371 (2023).
- [9] C Wiltshire, et al., DeepWild: application of the pose estimation tool DeepLabCut for behaviour tracking in wild chimpanzees and bonobos. *J. Animal Ecol.* **92**, 1560–1574 (2023).
- [10] R Nathan, et al., Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *J. Exp. Biol.* **215**, 986–996 (2012).
- [11] J Soltis, et al., Accelerometers in collars identify behavioral states in captive African elephants *Loxodonta africana*. *Endangered Species Res.* **18**, 255–263 (2012).
- [12] JE Markowitz, et al., Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature* **614**, 108–117 (2023).
- [13] BL Hart, Biological basis of the behavior of sick animals. *Neurosci. & Biobehav. Rev.* **12**, 123–137 (1988).
- [14] DJ Langford, et al., Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* **7**, 447–449 (2010).
- [15] RP Wilson, et al., Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *J. Animal Ecol.* **75**, 1081–1090 (2006).
- [16] ELC Shepard, et al., Energy landscapes shape animal movement ecology. *Am. Nat.* **182**, 298–312 (2013).
- [17] AM Wilson, et al., Locomotion dynamics of hunting in wild cheetahs. *Nature* **498**, 185–189 (2013).
- [18] TM Williams, et al., Instantaneous energetics of puma kills reveal advantage of felid sneak attacks. *Science* **346**, 81–85 (2014).
- [19] AD Redish, Vicarious trial and error. *Nat. Rev. Neurosci.* **17**, 147–159 (2016).
- [20] GJ Berman, W Bialek, JW Shaevitz, Predictability and hierarchy in Drosophila behavior. *Proc. Natl. Acad. Sci. USA* **113**, 11943–11948 (2016).
- [21] S Cain, et al., Movement predictability of individual barn owls facilitates estimation of home range size and survival. *Mov. Ecol.* **11**, 10 (2023).
- [22] B Abrahms, et al., Suite of simple metrics reveals common movement syndromes across vertebrate taxa. *Mov. Ecol.* **5**, 12 (2017).
- [23] O Spiegel, ST Leu, CM Bull, A Sih, What’s your move? Movement as a link between personality and spatial dynamics in animal populations. *Ecol. Lett.* **20**, 3–18 (2017).
- [24] O Spiegel, ST Leu, A Sih, SS Godfrey, CM Bull, When the going gets tough: behavioural type-dependent space use in the sleepy lizard changes as the season dries. *Proc. Royal Soc. B* **282**, 20151768 (2015).

- [25] AG Hertel, PT Niemelä, NJ Dingemanse, T Mueller, A guide for studying among-individual behavioral variation from movement data in the wild. *Mov. Ecol.* **8**, 30 (2020).
- [26] S Ye, et al., SuperAnimal pretrained pose estimation models for behavioral analysis. *Nat. Commun.* **15**, 5165 (2024).
- [27] EL Charnov, Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* **9**, 129–136 (1976).
- [28] A Strandburg-Peshkin, DR Farine, ID Couzin, MC Crofoot, Shared decision-making drives collective movement in wild baboons. *Science* **348**, 1358–1361 (2015).
- [29] J Lauer, et al., Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* **19**, 496–504 (2022).
- [30] M Teimouri, UG Indahl, H Sickel, H Tveite, Deriving animal movement behaviors using movement parameters extracted from location data. *ISPRS Int. J. Geo-Information* **7**, 78 (2018).
- [31] JR Hutchinson, et al., The locomotor kinematics of Asian and African elephants: changes with speed and size. *J. Exp. Biol.* **209**, 3812–3827 (2006).
- [32] BJ Greco, et al., The days and nights of zoo elephants: using epidemiology to better understand stereotypic behavior of African elephants (*Loxodonta africana*) and Asian elephants (*Elephas maximus*) in North American zoos. *PLoS ONE* **11**, e0144276 (2016).
- [33] G Wittemyer, L Polansky, I Douglas-Hamilton, WM Getz, Disentangling the effects of forage, social rank, and risk on movement autocorrelation of elephants using Fourier and wavelet analyses. *Proc. Natl. Acad. Sci. USA* **105**, 19108–19113 (2008).
- [34] L Polansky, I Douglas-Hamilton, G Wittemyer, Using diel movement behavior to infer foraging strategies related to ecological and social factors in elephants. *Mov. Ecol.* **1**, 13 (2013).
- [35] L Polansky, W Kilian, G Wittemyer, Elucidating the significance of spatial memory on movement decisions by African savannah elephants using state-space models. *Proc. Royal Soc. B* **282**, 20143042 (2015).
- [36] J Poole, P Granli, The Elephant Ethogram: a library of African elephant behaviour. *Pachyderm* **62**, 105–111 (2021).
- [37] M Tsalyuk, W Kilian, B Reineking, WM Getz, Temporal variation in resource selection of African elephants follows long-term variability in resource availability. *Ecol. Monogr.* **89**, e01348 (2019).
- [38] J Wall, G Wittemyer, B Klinkenberg, I Douglas-Hamilton, Novel opportunities for wildlife conservation and research with real-time monitoring. *Ecol. Appl.* **24**, 593–601 (2014).
- [39] M Thaker, PR Gupte, HHT Prins, R Slotow, AT Vanak, Fine-scale tracking of ambient temperature and movement reveals shuttling behavior of elephants to water. *Front. Ecol. Evol.* **7**, 4 (2019).
- [40] J Wall, et al., Human footprint and protected areas shape elephant range across Africa. *Curr. Biol.* **31**, 2437–2445.e4 (2021).

- [41] R Kays, et al., The Movebank system for studying global animal movement and demography. *Methods Ecol. Evol.* **13**, 419–431 (2022).
- [42] WM Getz, W Kilian, R Zidon, M Tsalyuk, Data from: Temporal variation in resource selection of African elephants follows long-term variability in resource availability (Movebank Data Repository) (2018).
- [43] P Kulits, J Wall, A Bedetti, M Henley, S Beery, ElephantBook: a semi-automated human-in-the-loop system for elephant re-identification in *Proceedings of the ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '21)*. (2021).
- [44] H Yu, et al., AP-10K: a benchmark for animal pose estimation in the wild in *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*. Vol. 1, (2021).
- [45] P Dagenais, S Hensman, V Haechler, MC Milinkovitch, Elephants evolved strategies reducing the biomechanical complexity of their trunk. *Curr. Biol.* **31**, 4727–4737.e4 (2021).
- [46] JR Hutchinson, D Famini, R Lair, R Kram, Are fast-moving elephants really running? *Nature* **422**, 493–494 (2003).
- [47] IO Buss, JA Estes, The functional significance of movements and positions of the pinnae of the African elephant, *Loxodonta africana*. *J. Mammal.* **52**, 21–27 (1971).
- [48] I Douglas-Hamilton, T Krink, F Vollrath, Movements and corridors of African elephants in relation to protected areas. *Naturwissenschaften* **92**, 158–163 (2005).
- [49] N Gravett, et al., Inactivity/sleep in two wild free-roaming African elephant matriarchs – does large body size make elephants the shortest mammalian sleepers? *PLoS ONE* **12**, e0171903 (2017).
- [50] K McComb, C Moss, SM Durant, L Baker, S Sayialel, Matriarchs as repositories of social knowledge in African elephants. *Science* **292**, 491–494 (2001).
- [51] R Harel, O Spiegel, WM Getz, R Nathan, Social foraging and individual consistency in following behaviour: testing the information centre hypothesis in free-ranging vultures. *Proc. Royal Soc. B* **284**, 20162654 (2017).
- [52] JH Poole, CJ Moss, Musth in the African elephant, *Loxodonta africana*. *Nature* **292**, 830–831 (1981).
- [53] A Mathis, et al., DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
- [54] AB Wiltschko, et al., Mapping sub-second structure in mouse behavior. *Neuron* **88**, 1121–1135 (2015).
- [55] HJ Williams, et al., Optimizing the use of biologgers for movement ecology research. *J. Animal Ecol.* **89**, 186–206 (2020).
- [56] YS Resheff, S Rotics, R Harel, O Spiegel, R Nathan, AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Mov. Ecol.* **2**, 27 (2014).

- [57] P Chakravarty, G Cozzi, A Ozgul, K Aminian, A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods Ecol. Evol.* **10**, 802–814 (2019).
- [58] WM Getz, A hierarchical path-segmentation movement ecology framework. *Ecol. Process.* **11**, 56 (2022).
- [59] N Sirmipilatze, et al., movement: a Python toolbox for analysing animal body movements across space and time (Zenodo) (2024).
- [60] DR Roberts, et al., Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **40**, 913–929 (2017).
- [61] AE Raftery, A model for high-order Markov chains. *J. Royal Stat. Soc. Ser. B (Methodological)* **47**, 528–539 (1985).
- [62] A Berchtold, AE Raftery, The mixture transition distribution model for high-order Markov chains and non-Gaussian time series. *Stat. Sci.* **17**, 328–356 (2002).
- [63] A Biswas, A Guha, Time series analysis of categorical data using auto-mutual information. *J. Stat. Plan. Inference* **139**, 3076–3087 (2009).
- [64] E Gurarie, RD Andrews, KL Laidre, A novel method for identifying behavioural changes in animal movement data. *Ecol. Lett.* **12**, 395–408 (2009).
- [65] KP Gundermann, et al., Change-point models for identifying behavioral transitions in wild animals. *Mov. Ecol.* **11**, 65 (2023).
- [66] R Langrock, et al., Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology* **93**, 2336–2342 (2012).
- [67] W Zucchini, IL MacDonald, R Langrock, *Hidden Markov Models for Time Series: An Introduction Using R*. (Chapman and Hall/CRC, Boca Raton, FL), Second edition, (2016).
- [68] BT McClintock, et al., Uncovering ecological state dynamics with hidden Markov models. *Ecol. Lett.* **23**, 1878–1903 (2020).
- [69] H Edelhoff, J Signer, N Balkenhol, Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Mov. Ecol.* **4**, 21 (2016).