

Can a Fly Navigate an EMR?

A leaky integrate-and-fire simulation of the complete *Drosophila* male CNS connectome steers two EHR replicas, with wiring-shuffled controls

Wavera Health Engineering

Technical report • Public field note • September 2026

Abstract

We simulate the complete *Drosophila melanogaster* male central nervous system (MaleCNS v1.0; 165,122 neurons, 6,202,435 connections) as a current-based leaky integrate-and-fire (LIF) network following Shiu et al. (2024), integrated at $\Delta t = 1$ ms, and close the loop through electronic health record (EHR) replicas. Every 50 ms of simulated time, 12 features sensed from the page drive identified sensory populations, and a softmax readout of 2,129 descending and motor neurons, trained by behavior cloning and DAgger to imitate a demonstrator, chooses one of 11 actions (cursor moves, click, type, wait). The task (which element is next) and the text to type are given; the fly chooses every move, click and keypress.

Across 100 evaluation runs per cell, the connectome-driven fly completed a 12-stage single visit and both lanes of a race (a Wavera replica and a generic legacy composite we built) in 100/100 runs each. Shuffled wiring at the same gain, shuffles at 0.26–0.32 mV, and random actions completed 0/100. Activity-matched shuffles (0.36–0.37 mV) completed 0/100 and 74/100 on the single visit; the 74 finishers needed a median 2,226 decisions versus 1,190 and made 77.7 versus 9.3 wrong clicks. Per click target the fly needed a median 8 decisions in the Wavera lane and 16 in the legacy composite, and about 5 decisions per bit of Fitts index of difficulty in both. Required actions differ by design between the replicas and are not measurements of a real EHR. The wiring is not strictly required, but it is more efficient in this setting.

1 Motivation

EHR interaction load is well documented: emergency physicians have been measured at roughly 4,000 mouse clicks per 10-hour shift [1], and primary care physicians at about 5.9 of 11.4 working hours per day in the EHR [2]. Wavera builds clinical software in which an assistant (Luna) summarizes the chart and drafts documentation, and a physician reviews and signs.

Complete insect connectomes [3, 5] and whole-brain LIF models [4] make it possible to drive a simulated nervous system with arbitrary input streams; community projects have connected fly connectomes to games [6]. We ask whether a whole-CNS LIF model, read out through its descending and motor neurons, can complete an EHR workflow in a replica interface, how the cost of each action depends on interface geometry, and how much of the result depends on the measured wiring.

2 Data

We use the Janelia FlyEM Male CNS v1.0 release [3] (CC BY 4.0). We read the body annotations, the per-body neurotransmitter predictions, and the significant-only connection weights (confidence ≥ 0.5). We keep connections with at least 5 synapses. Kenyon-cell-to-Kenyon-cell edges are removed, as they are a known source of runaway recurrent excitation in LIF models of the mushroom body.

Table 1: Simulated connectome (MaleCNS v1.0).

	Full CNS	Central (no optic lobes)
Neurons	165,122	71,618
Connections (pre→post pairs, ≥ 5 synapses)	6,202,435	3,439,405
Synapses carried	89,519,819	58,678,499
excitatory	55,455,101	—
inhibitory	34,064,718	—
KC→KC edges removed	33,247 (211,732 syn.)	—
Somas with imputed position	976	—

Table 2: Predicted neurotransmitter per neuron and the synaptic sign assigned. “Unclear” is treated as excitatory.

Neurotransmitter	Neurons	Share	Sign
Acetylcholine	103,718	62.8%	+
Glutamate	29,296	17.7%	−
GABA	22,055	13.4%	−
Histamine	5,910	3.6%	−
Unclear	3,602	2.2%	+
Dopamine	392	0.2%	+
Octopamine	101	0.06%	+
Serotonin	48	0.03%	+
Total	165,122	100%	

3 Model

Each neuron i follows current-based LIF dynamics with exponential synapses, integrated exactly at $\Delta t = 1$ ms:

$$\tau_m \frac{dv_i}{dt} = g_i - (v_i - v_{\text{rest}}), \quad \tau_s \frac{dg_i}{dt} = -g_i, \quad g_i \leftarrow g_i + w_{\text{syn}} n_{ji} s_j r_j \quad \text{on a spike of } j \text{ (after delay } d), \quad (1)$$

where n_{ji} is the synapse count, $s_j \in \{+1, -1\}$ the sign of j ’s neurotransmitter, and $r_j \in (0, 1]$ an optional Tsodyks–Markram short-term depression resource ($r_j \leftarrow r_j(1 - U)$ per spike, recovering with τ_{rec}). Sensory drive is forced Poisson spiking at a set rate, as in [4].

Table 3: Neuron parameters (Shiu et al.).

$v_{\text{rest}}, v_{\text{reset}}$	−52 mV
v_{thresh}	−45 mV
τ_m	20 ms
τ_s	5 ms
Refractory period	2.2 ms
Synaptic delay d	1.8 ms
τ_{rec} (depression)	200 ms

Table 4: Dynamical regimes evaluated.

Regime	w_{syn} (mV)	U	Behavior
shiu	0.275	0	runaway
canonical	0.20	0	sugar→MN9; ignites
stable	0.18	0.1	returns to rest

Calibration. With the published constants, MaleCNS enters self-sustaining activity: about 16,000 neurons keep firing after the stimulus ends. This is unlike the female-brain FlyWire graph the constants were fit on. No setting we scanned both reproduces the sugar→MN9 (proboscis extension) result and returns to rest. We report all three regimes and use **stable** for the demonstration, because a visualization must settle between events. The cost is that **stable** does not reproduce sugar-evoked MN9 recruitment.

4 Sensory encoding and motor readout

Table 5 lists the identified populations used in the toy cursor task (Section 6); the EHR task (Section 7) uses the same populations through the 12 features in Table 9.

Table 5: Identified MaleCNS populations used for input and readout.

Population	Neurons	Role
<i>Sensory (input)</i>		
Labellar sugar GRNs	34	cursor on target
Leg sugar GRNs, T1–T3, L/R	33–40 each (216)	target left or right of cursor
Labellar bitter GRNs	38	error (wrong click)
Head bristles, L / R	430 / 520	menu or pop-up opens (brief touch bursts)
Notum bristles, L / R	329 / 326	available mechanosensory drive
Ocellar photoreceptors	22	screen change
Photoreceptors R1–6	1,394	available; inhibitory onto lamina in this model
Lamina L1–L3	–	screen change, streamed text (EHR task)
Johnston’s organ, auditory	114	text field focused (EHR task)
Appetitive ORNs (DM1, DM4, VA2)	–	pointer heading to target; target visible
<i>Descending and motor (readout)</i>		
Steering DNs, L / R	2 / 2	fixed readout: left / right
Forward-walking DNs	6	fixed readout: up
Moonwalker DNs (MDN)	4	fixed readout: down
MN9 proboscis motor neurons	2	fixed readout: click
Grooming DNs	39	fixed readout: type
Giant fiber (DNp01)	2	logged only
All descending neurons	1,314	trained readout pool
All motor neurons	815	trained readout pool

5 Rendering

The renderer draws a subset of neurons as full skeletons in six functional groups, plus every soma, inside the brain and VNC hulls (Table 6). Assets total 18 MB. Each skeleton vertex stores its path distance from the soma, so a spike launches a wave that travels outward along the arbor, slowed for visibility. Every lit cell corresponds to a spike from the model: 141,000 of 142,757 somas and all 1,208 skeleton neurons are mapped to simulated neurons by body ID.

Table 6: Rendered anatomy.

Group	Skeleton neurons	Skeleton vertices	Somas
Visual system	314	118,028	91,455
Mushroom body	276	223,240	4,490
Central complex	216	148,996	2,949
Descending neurons	130	382,944	1,308
Leg and wing motor	142	294,455	807
Ascending and sensory	130	205,442	1,841
Total	1,208	1,373,105	102,850 [†]

[†]Group-assigned somas; 142,757 somas are drawn in total. Hull: 13,973 vertices, 28,000 triangles.

6 Background: a toy cursor task

Task. Before the EHR replicas, we tested a simpler closed loop. A synthetic 800×500 screen contains a goal button and distractor boxes. The *simple* screen has 3 boxes and 1 pop-up; the *chattered* screen has 30 boxes, 1 pop-up at the start, and 2 more during the episode. Every 50 ms of simulated time the network is read out to one of seven actions: up, down, left, right (25 px moves), click, type, or wait. An episode succeeds if the goal is clicked within 100 decisions, after dismissing any pop-up.

Conditions.

- *Teacher*: a scripted policy.
- *Fixed*: the anatomical rules in Table 5.

- *Trained*: softmax regression on $\log(1+\text{spike count})$ of all 2,129 descending and motor neurons, fit by behavior cloning with one DAgger round against the teacher.
- *Random*: uniform actions.
- *Shuffled*: the same pipeline on a degree-preserving shuffle of the connectome.

Evaluation seeds are disjoint from training seeds. We report success with Wilson 95% intervals, and the mean number of decisions on successful episodes. Every cell was run with $N=100$ episodes on an RTX 4090.

Table 7: Closed-loop success, $N=100$ episodes per cell. Steps: mean decisions on successful episodes (teacher: 19.3 simple, 27.4 cluttered). Shuffled networks are evaluated at the same w_{syn} as the connectome.

Network, regime	Readout	Simple		Cluttered	
		Success %	Steps	Success %	Steps
Central, shiu	connectome, trained	100 [96.3, 100]	19.4	100 [96.3, 100]	27.4
	shuffled, trained	100 [96.3, 100]	21.9	99 [94.6, 99.8]	34.1
	connectome, fixed	0 [0, 3.7]	–	0 [0, 3.7]	–
	random	0 [0, 3.7]	–	0 [0, 3.7]	–
Central, canonical	connectome, trained	100 [96.3, 100]	19.4	100 [96.3, 100]	27.7
	shuffled, trained	100 [96.3, 100]	21.4	99 [94.6, 99.8]	35.4
	connectome, fixed	0 [0, 3.7]	–	0 [0, 3.7]	–
	random	1 [0.2, 5.4]	94.0	0 [0, 3.7]	–
Central, stable	connectome, trained	100 [96.3, 100]	19.4	100 [96.3, 100]	28.1
	shuffled, trained	0 [0, 3.7]	–	0 [0, 3.7]	–
	connectome, fixed	0 [0, 3.7]	–	0 [0, 3.7]	–
	random	1 [0.2, 5.4]	39.0	0 [0, 3.7]	–
Full CNS, stable	connectome, trained	100 [96.3, 100]	19.5	100 [96.3, 100]	27.9
	shuffled, trained	1 [0.2, 5.4]	65.0	0 [0, 3.7]	–
	connectome, fixed	0 [0, 3.7]	–	0 [0, 3.7]	–
	random	0 [0, 3.7]	–	0 [0, 3.7]	–
Central, canonical , 57 named neurons only	connectome, trained	0 [0, 3.7]	–	0 [0, 3.7]	–
	shuffled, trained	74 [64.6, 81.6]	56.9	11 [6.3, 18.6]	71.5

The stable-regime advantage depends on gain. A shuffled graph has different recurrent structure, so at $w_{\text{syn}}=0.18$ it is nearly silent and the readout has nothing to decode. Raising the shuffled network’s gain until its activity is comparable closes most of the gap (Table 8). Only two shuffle seeds were run, so these numbers are indicative.

Table 8: Gain sweep, central **stable** regime ($U=0.1$), trained readout, $N=100$ per cell. The connectome reference is 100/100 at $w_{\text{syn}}=0.18$.

Network	w_{syn} (mV)	Simple %	Steps	Cluttered %	Steps
Connectome	0.18	100 [96.3, 100]	19.4	100 [96.3, 100]	28.1
Shuffled (seed A)	0.18	0 [0, 3.7]	–	0 [0, 3.7]	–
Shuffled (seed A)	0.22	41 [31.9, 50.8]	44.7	14 [8.5, 22.1]	56.5
Shuffled (seed B)	0.24	99 [94.6, 99.8]	44.2	54 [44.3, 63.4]	63.3
Shuffled (seed A)	0.26	88 [80.2, 93.0]	30.8	63 [53.2, 71.8]	51.0
Shuffled (seed B)	0.26	99 [94.6, 99.8]	33.4	78 [68.9, 85.0]	59.3
Shuffled (seed A)	0.32	100 [96.3, 100]	24.6	100 [96.3, 100]	42.6

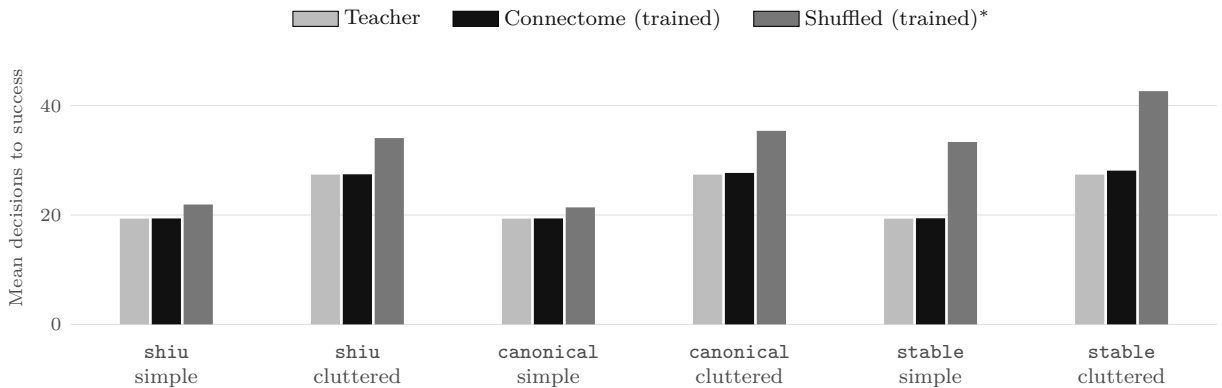


Figure 1: Path efficiency. The trained connectome matches the teacher almost exactly; shuffled networks that succeed need roughly 10–70% more decisions. *For **stable**, the shuffled bars are the best-gain seed B at $w_{\text{syn}}=0.26$ (simple) and seed A at 0.32 (cluttered), because at matched gain the shuffle never succeeds.

Why the fixed readout fails. In the **canonical** regime, the anatomical readout chose *click* on 14,610 of 20,000 decisions (73%) and *wait* on 5,317 (27%), with 73 errors per episode on average. MN9 is recruited by almost any strong input, so the fly “clicks” everything. Steering and walking DNs rarely cross threshold.

Findings.

1. A trained linear readout on a real LIF connectome solves both screens perfectly, with near-teacher efficiency.
2. In the **shiu** and **canonical** regimes a shuffled graph does nearly as well, so the network acts largely as a reservoir.
3. In the **stable** regime the connectome is decodable at a gain where shuffles are silent. Gain-matched shuffles reach 54–78% on the cluttered screen versus 100%, an advantage that rests on two seeds.
4. Anatomical (fixed) readouts fail everywhere.
5. Both screens saturate at 100% for trained readouts, so this task cannot separate “simple” from “cluttered” interfaces by success rate. Step counts and errors are the better metric.

7 Closed-loop EHR navigation

7.1 Task and lanes

Each flow compiles to an ordered list of subgoals: click this element, type N characters into that field, or wait while the application acts by itself (screen transitions, Luna streaming, dictation). The subgoal list is given, as a person would be told which button is next. Three lanes use the same fictional patient (Maya Reed; synthetic data): a 12-stage single visit in a replica of the Wavera EHR, and a race in which the same visit is completed in a shorter Wavera flow and in a generic legacy composite that we built. Required actions are properties of the replicas: 37 clicks and 549 keystrokes (single visit), 14 and 33 (race, Wavera), 222 and 1,180 (race, legacy composite). They are not measurements of a real EHR.

7.2 Sensing, actions and readout

Every decision (50 ms of simulated time) the live page is sensed directly: the next target’s rectangle and a point on it that hit-tests to it, the element under the cursor, open overlays, screen changes, the focused field and the last error. The resulting 12 features drive identified populations as Poisson input (Table 9). The target-visible odour and pop-up touch are onset inputs lasting 3 decisions, like adapting receptors.

Table 9: EHR-task features and the sensory populations they drive (maximum Poisson rate).

Feature	Population	Max Hz
Target visible (onset)	Appetitive ORNs (DM1, DM4, VA2)	100
Target left / right of cursor	Leg sugar GRNs, left / right	150
Target above / below	Johnston’s organ zone C / zone E	150
Cursor on target	Labellar sugar GRNs	150
Cursor on a distractor	Labellar sugar GRNs	60
Pop-up open (onset)	Head bristles, L and R	80
Screen changed	Ocelli, R1–6 photoreceptors, lamina patch	150
Text field focused	Johnston’s organ, auditory	150
Previous action was an error	Labellar bitter GRNs	150
Application wrote text	Lamina patch	30

A softmax readout over $\log(1+\text{spike count})$ of all 2,129 descending and motor neurons chooses one of 11 actions: four 8px moves, four 64px moves, click, type or wait. Moves are not tweened; clicks land wherever the cursor is, so wrong clicks hit a distractor, empty space, or a wrong tab that must be clicked back from. “Type” inserts the next character of a fixed sentence: the fly chooses when to type, not what. The readout was fit by behavior cloning and DAgger (four rounds, demonstrator mixing $\beta = 1, 0.5, 0.25, 0$) on 240-decision segments of all three lanes, training seeds $\geq 10,000$; training accuracy was 0.88 for the connectome.

7.3 Evaluation

Full lanes were run on seeds 0–99, disjoint from training, on an RTX 4090, using a Python mirror of the page built from exported layout snapshots. The mirror freezes the layout per subgoal and never navigates on a wrong click; the demonstrator needs 1,161, 266 and 4,193 decisions in the mirror versus 1,156, 268 and 4,142 on the page. A run fails if it makes no progress for 400 decisions or exceeds twice the demonstrator’s decisions. We report Wilson 95% intervals for completion; percentiles with bootstrap 95% intervals (10,000 resamples); and, between race lanes, run-level two-sided Mann–Whitney U tests with Cliff’s δ , Holm-corrected across metrics. Controls retrain the same readout on a degree-preserving shuffle of the connectome (shuffle seeds A and B) at the connectome’s gain and at raised gains, and replace the readout with uniform random actions.

7.4 Completion

Table 10: Completion out of 100 runs (Wilson 95% interval). Shuffle A: seed 12345; shuffle B: seed 777. The legacy lane was not run at 0.36–0.37 mV.

Network, w_{syn} (mV)	Single visit	Race: Wavera	Race: legacy
Connectome, 0.18	100 [96.3, 100]	100 [96.3, 100]	100 [96.3, 100]
Shuffled A, 0.18	0 [0, 3.7]	0 [0, 3.7]	0 [0, 3.7]
Shuffled A, 0.26	0 [0, 3.7]	0 [0, 3.7]	0 [0, 3.7]
Shuffled B, 0.32	0 [0, 3.7]	0 [0, 3.7]	0 [0, 3.7]
Shuffled A, 0.36	0 [0, 3.7]	4 [1.6, 9.8]	—
Shuffled B, 0.37	74 [64.6, 81.6]	0 [0, 3.7]	—
Random actions	0 [0, 3.7]	0 [0, 3.7]	0 [0, 3.7]

7.5 Distribution of decisions and time

Table 11: Connectome fly, decisions to completion (100 of 100 finished per lane). Brackets: bootstrap 95% interval.

	Single visit	Race: Wavera	Race: legacy
p50	1,190 [1,186, 1,191]	272 [269, 274]	4,662 [4,653, 4,679]
p90	1,209	282	4,764
p95	1,219 [1,208, 1,223]	286 [282, 291]	4,801 [4,757, 4,829]
p99	1,232	293	4,832
Mean $\pm$ SD	1,188.5 $\pm$ 17.1	271.5 $\pm$ 9.5	4,669.6 $\pm$ 70.8
Simulated s, p50	59.5 [59.3, 59.55]	13.6 [13.45, 13.7]	233.1 [232.7, 234.0]
Demonstrator decisions	1,156	268	4,142
Required clicks + keys	37 + 549	14 + 33	222 + 1,180

The single visit’s median of 1,190 decisions is within 3% of the demonstrator’s 1,156. The race lanes differ by a factor of about 17 in decisions, mostly because of the number of actions each replica requires, which we set when we built them.

7.6 Per-action metrics

Table 12: Per-action metrics, connectome fly. Target geometry is measured per click target: the smaller side of the target, visible targets within 150px of the aim point, and index of difficulty $\log_2(D/W + 1)$.

Metric	Single visit	Race: Wavera	Race: legacy
Decisions per click target, p50 / p95	9 / 26	8 / 26	16 / 33
Decisions per required action, run p50	2.03	5.79	3.32
Decisions per keystroke, p50	1.0	1.0	1.0
Wrong clicks per run, mean	9.3	4.9	16.4
Wrong clicks per 100 required clicks, mean (p50)	25.1 (21.6)	35.0 (28.6)	7.4 (7.2)
Pop-up dismissal decisions, p50 / p95	–	–	11 / 22
Target smaller side, p50	36 px	36 px	18 px
Distractors within 150 px, p50	2	2	6
Index of difficulty, p50	2.90 bits	2.91 bits	4.65 bits

Table 13: Race lane comparison, run-level medians ($n = 100$ per lane), two-sided Mann–Whitney U, Holm-corrected across the three metrics.

Metric	Wavera	Legacy	U	Cliff’s δ	p_{Holm}
Decisions per click target	8.5	16.0	0	−1.00	9.3×10^{-35}
Decisions per required action	5.79	3.32	10,000	+1.00	5.0×10^{-34}
Wrong clicks per 100 required clicks	28.6	7.2	8,955	+0.79	3.8×10^{-22}

The per-action metrics do not all favor one interface. Each legacy click target cost about twice the decisions of a Wavera target; legacy targets were half the size, had three times as many nearby distractors, and carried 4.65 versus 2.91 bits. The Wavera lane had more wrong clicks per 100 required clicks, because it requires only 14 clicks, and more decisions per required action, because the legacy composite’s 1,180 keystrokes cost about one decision each.

Fitts-style fit. Regressing decisions per click target on index of difficulty gives about 5 decisions per bit in every lane (Table 14). Adding the number of nearby distractors changes R^2 by less than 0.01. The similar slopes suggest the per-target difference between lanes is explained mainly by target geometry.

Table 14: Fitts-style fit: decisions per click target = $a + b \log_2(D/W + 1)$.

Lane	Targets	a	b (per bit)	R^2
Single visit	3,700	−4.37	5.04	0.73
Race: Wavera	1,400	−5.53	5.56	0.86
Race: legacy	22,200	−6.71	5.17	0.64

7.7 Activity-matched shuffled controls

A shuffled graph is much quieter than the connectome at the same gain, so we measured activity under identical demonstrator-driven input (Table 15) and evaluated shuffles at gains that bracket the connectome’s 4,765 spikes per decision: 0.36 mV for shuffle A and 0.37 mV for shuffle B (0.37 was not measured directly; B gives 3,544 at 0.36 and 5,231 at 0.38).

Table 15: Single visit: activity under identical input, readout training accuracy, and outcome.

Network, w_{syn} (mV)	Spikes / decision	Train acc.	Finished; median decisions; wrong clicks
Connectome, 0.18	4,765	0.88	100/100; 1,190; 9.3
Shuffled A, 0.18	1,369	0.11	0/100
Shuffled A, 0.26	1,554	0.17	0/100
Shuffled B, 0.32	2,159	0.37	0/100
Shuffled A, 0.36	4,307	0.45	0/100
Shuffled B, 0.37	–	0.47	74/100; 2,226; 77.7

The activity-matched shuffle B finished 74/100 but took about twice the decisions and made $8\times$ the wrong clicks. In the Wavera race lane the same controls finished 4/100 (A, 0.36) and 0/100 (B, 0.37); the legacy lane was not evaluated at these gains. Wiring is therefore not strictly required to complete the task, but in this setting the measured connectome is decodable at a gain where shuffles are silent and is markedly more efficient. The comparison rests on two shuffle seeds.

8 Demonstration

The published videos are live runs on the page, not the offline mirror. The full-CNS model runs in a browser Web Worker on a WebAssembly kernel at $2.4\text{--}2.9\times$ real time and is paced to real time; with a fixed seed and page layout the fly’s decisions are deterministic (the page hit-tests in screen pixels, so a different display scale gives a different, equally valid run). The recorded single-visit run (seed 1) took 1,186 decisions, 50 clicks of which 13 were wrong, 549 keystrokes and 59.3s of simulated time (Figure 2). The race recording (seed 1, side-by-side layout, 3:52 wall time) shows both lanes steered by the same model and readout, each with its own brain instance (Figure 3): the Wavera lane took 272 decisions with 5 wrong clicks (13.6s simulated), the legacy composite 4,566 decisions with 15 wrong clicks (228.3s). The published race video is sped up $2\times$ until the Wavera lane finishes and $16\times$ thereafter, with the speed shown on screen.

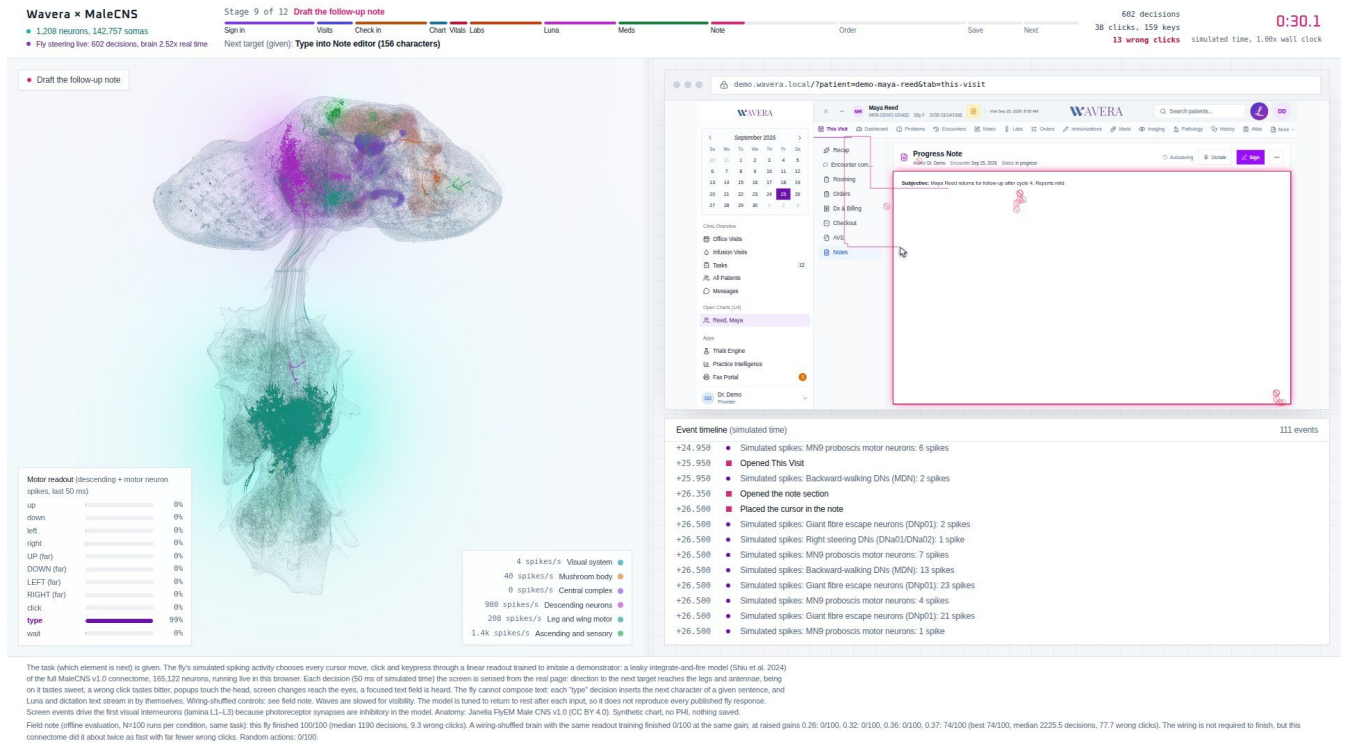


Figure 2: Single visit, fly-steered. Left: MaleCNS lit by the model's spikes, with the motor readout (share of each action in the last 50 ms) and per-group spike rates. Right: the Wavera replica; the header shows the given next target, decisions, clicks, keystrokes and wrong clicks. Synthetic patient, no PHI.

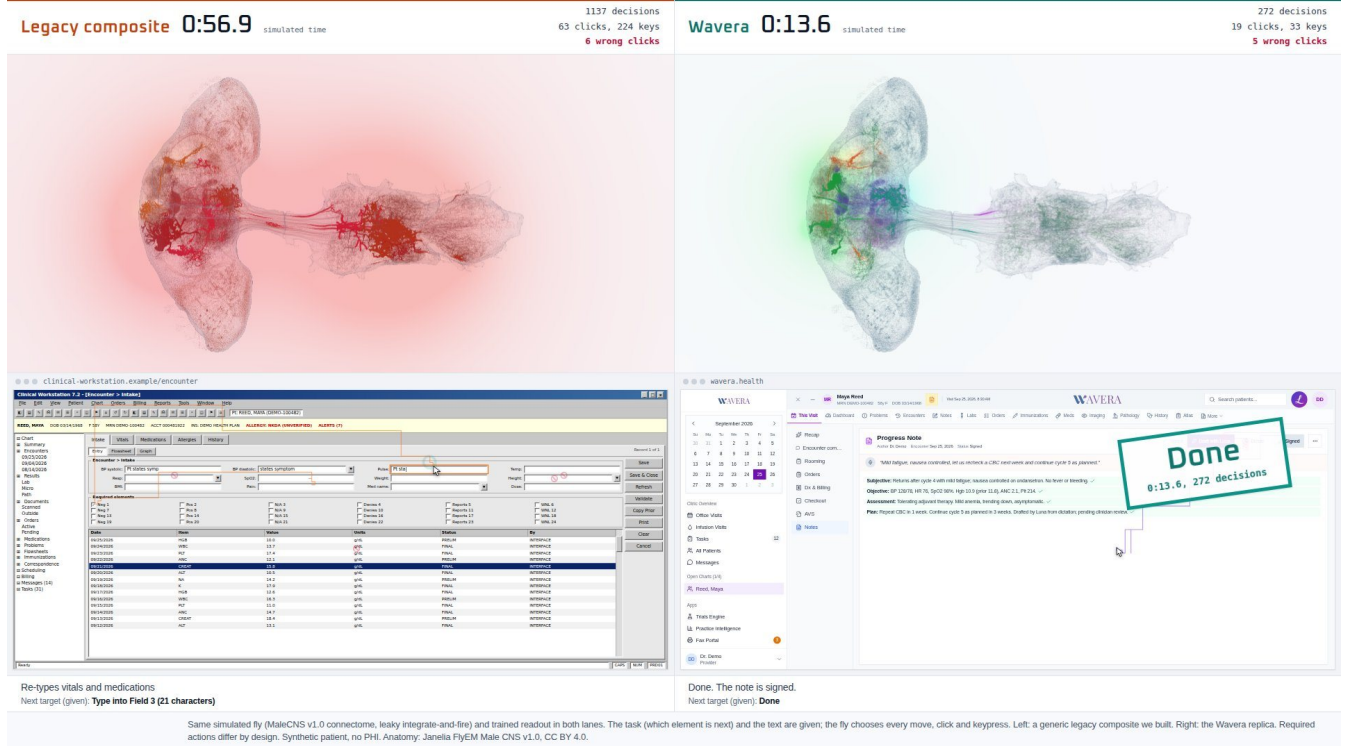


Figure 3: Race, side by side. Left: a generic legacy composite we built. Right: the Wavera replica. The same model and readout steer both lanes. Required actions differ by design (222 clicks and 1,180 keystrokes versus 14 and 33).

Performance

Table 16: GPU throughput (RTX 4090, PyTorch 2.6 + CUDA 12.4, $\Delta t=1$ ms). Real-time factor = simulated s per wall s. Aggregate = fly-seconds per wall-second across the batch.

Graph	Regime	Batch	RT / fly	Aggregate	Spikes/s/fly	GPU MB
Full	shiu	1	3.16×	3.2	1,127,609	86
Full	canonical	1	3.31×	3.3	757,428	86
Full	stable	1	2.44×	2.4	109,921	87
Full	shiu	100	0.27×	27.0	1,127,376	986
Full	canonical	100	0.29×	28.6	782,582	986
Full	stable	100	0.20×	19.5	110,096	1,052
Central	canonical	1	3.30×	3.3	750,251	47
Central	stable	1	2.72×	2.7	109,874	47
Central	canonical	100	0.72×	71.8	761,131	446
Central	stable	100	0.53×	53.1	110,128	476

9 Limitations

- **The task is given.** The fly is told which element is next; it does not decide what the visit requires.
- **Text is given.** Each “type” decision inserts the next character of a fixed sentence. Luna and dictation text stream in by themselves.
- **The readout imitates a demonstrator.** Behavior cloning and DAgger reproduce a scripted policy; the readout does not discover one.
- **The replicas are ours.** The legacy screen is a generic composite we built, not a vendor’s product, and the required actions of both replicas are design choices, not measurements of a real EHR.
- **Wiring is not strictly required.** An activity-matched shuffle finished 74/100 on the single visit. The connectome’s efficiency advantage rests on two shuffle seeds, and the legacy lane was not evaluated at activity-matched gains.
- **Offline mirror.** Evaluation statistics come from a mirror that freezes the layout per subgoal and never navigates on a wrong click; the live page is less forgiving.
- **Not a model of a clinician.** Sensory inputs are events and geometry, not pixels, and nothing here measures human cognitive load.
- **Regime and visual pathway.** **stable** returns to rest but loses sugar-evoked MN9 recruitment. Photoreceptor-to-lamina synapses are histaminergic and therefore inhibitory in the model, so screen events also drive lamina L1–L3 directly.
- **Visual liberties.** Waves along arbors are slowed for visibility.

10 Reproducibility notes

- **Graph:** MaleCNS v1.0 significant-only connections (confidence ≥ 0.5), ≥ 5 synapses per pair, KC→KC edges removed; sign from predicted neurotransmitter (Table 2).
- **Dynamics:** current-based LIF, exact integration at $\Delta t = 1$ ms; $v_{\text{thresh}} = -45$ mV, $v_{\text{rest}} = v_{\text{reset}} = -52$ mV, $\tau_m = 20$ ms, $\tau_s = 5$ ms, refractory 2.2 ms, delay 1.8 ms; **stable** regime ($w_{\text{syn}} 0.18$ mV, depression $U 0.1$, $\tau_{\text{rec}} 200$ ms).
- **EHR task:** 50 ms decisions, 12 features (Table 9), 11 actions. Readout: softmax on $\log(1+\text{spike count})$ of 2,129 DN+MN neurons, behavior cloning plus DAgger (4 rounds, $\beta = 1, 0.5, 0.25, 0$), 240-decision training segments, training seeds $\geq 10,000$, evaluation seeds 0–99. Stop rule: 400 decisions without progress or twice the demonstrator’s decisions.
- **Controls:** degree-preserving shuffles with seeds 12345 (A) and 777 (B), same pipeline and readout, at 0.18, 0.26, 0.32, 0.36 and 0.37 mV; uniform random actions.

- **Statistics:** Wilson 95% intervals; bootstrap (10,000 resamples) for p50 and p95; Mann–Whitney U with Cliff’s δ , Holm-corrected; ordinary least squares for the Fitts fit.
- **Browser:** WebAssembly LIF kernel, bit-identical to the JavaScript kernel, $2.4\text{--}2.9\times$ real time for the full CNS; live runs are paced to real time and deterministic for a fixed seed and page layout.

11 Data, licensing and ethics

- **Anatomy and connectivity:** Janelia FlyEM Male CNS v1.0 [3], CC BY 4.0. No FlyWire data is used. This work is independent and is not endorsed by Janelia Research Campus or HHMI.
- **Dynamics and design credits:** LIF dynamics follow Shiu et al. [4]. The browser simulator design is informed by the MIT-licensed `fly-brain` project. Poisson sensory drive, a descending-neuron readout and behavior cloning with DAgger follow ideas from OpenFly (AGPL-3.0); no code was used.
- **Patient data:** all patient, chart and workflow data are synthetic. The patient, Maya Reed, is fictional. There is no PHI and no database writes.
- **Scope:** a research demonstration built on replicas of the Wavera EHR and of a generic legacy interface. It is not part of the clinical product.

References

- [1] R. G. Hill, L. M. Sears, S. W. Melanson. 4000 clicks: a productivity analysis of electronic medical records in a community hospital ED. *Am. J. Emerg. Med.* 31(11):1591–1594, 2013.
- [2] B. G. Arndt et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann. Fam. Med.* 15(5):419–426, 2017.
- [3] S. Berg et al. Sexual dimorphism in the complete connectome of the *Drosophila* male central nervous system. *bioRxiv*, 2025. Data: <https://male-cns.janelia.org>.
- [4] P. K. Shiu et al. A *Drosophila* computational brain model reveals sensorimotor processing. *Nature* 634:210–219, 2024.
- [5] S. Dorkenwald et al. Neuronal wiring diagram of an adult brain. *Nature* 634:124–138, 2024.
- [6] Community index of fruit-fly connectome projects. <https://github.com/cobanov/awesome-fly>.

About Wavera Health Engineering. We publish selected engineering findings from our work on AI-native clinical software. This note uses only public connectome data and synthetic patient records; proprietary prompts and sensitive implementation details are not included.