RAG9 Research Desk

Evidence Brief — Wearables as the Substrate of World Models

The Next Common Crawl? Wearables as the Substrate of World Models

Shift markets: from focused devices (wearables, ML, HCI) into a plausible, economical pathway to force economic discovery. Historical data was indexed through crawl. This brief explores if sensor-rich, high-fidelity inputs, fields, and a wristwatch of this triggers a second index.

Component	Weight	Score	Weighted
Data availability & quality	29%	72	18.4
AI & robotics maturity	20%	73	14.6
Business incentives & monetization	20%	86	17.2
Device proliferation & pricing	17%	72	12.7
Regulatory & privacy friction	10%	42	4.2
Public acceptance & ethics	10%	60	6.0
Total percent score			**74.3**

Preliminary Likelihood: 74%

Confidence: Medium — composite rubric. This score measures economic plausibility, weighted line.

Recommended immediate posture:

- Track key embodied datasets
- Trial new product avenues as test harnesses
- Explore privacy-preserving contextualization
- Seed builder frameworks and guidance
- Gradual but forward posture to reduce monopoly moat
- Prepare critical readout for investors/CxO's scroll.

High-signal evidence (selected):

- Egen (2025): academic video dataset showing usefulness for temporal/action labeling.
- Luma (2025): consumer synthetic data stacks lower marginal cost of training embodied models.
- Device Roundups / Meta Quest (2023–25).
- Multiple firms pushing consumer AR glasses and wearables with cameras/ML.
- Academic RLHF (preference, actual JGM-style representations emphasize human inputs).
- Acceleration of lifespan/gut/audio teams suggest platform interest in capture signals.

Top risks & blindspots:

- Privacy & bystander risks alongside install systems are re-identifiable.
- Misaligned signals: limited datasets, fuzzy/incomplete real-life captures.
- Bias: data tilting toward dense demographics who own hardware devices.
- Concentration: dominant vendors can gate access to embodied AI capabilities.

Watchlist (lead triggers):

- SDC confidential cloud upload via device processing.
- Meta acquisitions of lifelogging/capture-feature startups.
- Device pricing & self-disruption: consumer-friendly volume signals scale.
- Rule in: reposition vendor claims (consumer embedding as training source).
- Regulatory moves: GDPR rulings, FCC or national privacy legislation.