RESEARCH PAPER INTERACTIVE DASHBOARD

TL;DR

A living atlas of *where AI agents win*

A research framework + interactive dashboard mapping the AI Agent opportunity across **201 sub-tracks in 10 categories**, comparing the **Software TAM** (what AI products charge for) against the **Labor TAM** (the human-labor spend AI agents can actually disrupt).

SOFTWARE TAM	LABOR TAM
$497B	**$9.6T**

AI product layer · was $450B in 2025-Q1	Addressable human-labor spend
AVG LABOR/SW RATIO ***19.4×*** The leverage opportunity	SUB-TRACKS COVERED **201** Across 10 categories

01 — THE CORE INSIGHT

Why the *labor market* is the real ceiling

> *The Labor Market is the ceiling for AI Agents, not the Software Market.*

Most AI market sizing focuses on **software TAM** — "how much SaaS revenue can AI capture?" — currently **$497B** globally for AI-agent-relevant products. That's the wrong ceiling. AI agents don't compete with SaaS; they compete with **human labor**.

The total addressable human-labor spend across the same 201 use cases is **$9.6T** — a **19.4× leverage ratio**. This $9T gap is the structural opportunity AI agents are aimed at: turn human-labor spend into software spend (Net Compression), or expand the market by making expensive expert work cheap enough for new users (Net Expansion).

This is why we built the framework: to systematically rank **where** the leverage is largest, **how much** of each labor pool AI can actually replace, and **which startups** are credibly going after each pool.

02 — WHY THIS MATTERS FOR HAT-TRICK

The framework as an *operating tool*

Use case	How the framework helps
	Filter by Labor/SW ratio + Net Effect to find structurally underserved tracks (e.g., 100×+ ratio = giant labor pool, tiny software footprint). Use as a hit list for outbound.
	Cross-check inbound startup pitches against the addressable labor TAM and displacement realism — catch overstated software-TAM claims.
	Map each portfolio company to its track and watch ratio / displacement / key-player columns evolve over time.
	"Representative companies" + "Funding" columns surface who else is in the track and at what valuation.
	YC batch mapping (CORE / ADJACENT / INFRA fit) identifies which YC W26 / S26 companies are credible plays in tracks we already have a thesis on.
	"Capability" entries flag when a new model release (Claude 4.6, GPT-5.4, Helix 02 VLA) unlocks a step that was previously a gap — this is when entire tracks become investable.

Headline takeaway — out of 201 tracks, the framework identifies **22 high-priority deep-dive tracks** (Labor/SW ≥ 50×) where the leverage is structurally largest. These are the first places we look when sourcing.

◆

03 — HOW TO READ THE DASHBOARD

A guided tour of the *live atlas*

The dashboard below is the same interactive build available at agent-mapping.vercel.app — embedded here so you can read and explore in one place. Hover charts, click bubbles, expand table rows.

Feature-by-feature walkthrough

"What's Changed" banner (top)

Collapsible panel showing the most recent bimonthly update grouped by type — TAM revisions, displacement updates, capability unlocks, and key players & funding. Click any track name → table auto-filters to that track.

KPI cards (6 metrics, recompute live with filters)

- **Sub-tracks** — count under current filters
- **Software TAM Total** — sum across visible tracks ($B)
- **Labor TAM Total** — sum across visible tracks ($B)
- **Avg Labor/SW Ratio** — leverage across visible tracks
- ✅ **Net Expansion** — count where AI grows the market
- 🔴 **Net Compression** — count where AI shrinks total category spend

Filters

Category · AI Impact (Displacement / Expansion / Mixed) · Net Effect · Valuation Tier · Status · full-text Search.

The four charts

01 **TAM Totals by Category** (log-scale horizontal bar) — gold bars (Labor TAM) tower over purple bars (Software TAM); the gap visualizes the leverage per category.

02 **Software × Labor Scatter** (log-log bubble) — each bubble is one track; bubble size = displacement rate; **click a bubble to drill into that row**.

03 **AI Impact Donut** — track count by Displacement / Expansion / Mixed, hover shows aggregate Labor TAM per slice.

04 **Valuation Tier Donut** — distribution of lead-company valuation maturity per track.

The table — the core artifact

Every row = one of the 201 sub-tracks. Click any row to expand methodology, representative companies, funding, and investment insight. Columns:

Column	Meaning
	Addressable AI-software market size for the track ($B), Tier-1 analyst + bottom-up
	Human labor spend addressable by AI agents ($B), BLS/ILO × wage × task-fraction
	🔴 Displacement / ✅ Expansion / 🟡 Mixed
	% of human work AI agents can autonomously do today (range, e.g. "55–75%")
	Whether AI grows or shrinks the total category
	Labor TAM ÷ Software TAM = the leverage ratio
	Lead-company valuation tier and stage in the track

How *each number* is calculated

4.1 Software TAM

Definition: addressable revenue for the AI software product layer in the track over the next 3-5 years.

Method:

01 Anchor to **Tier-1 analyst reports** — Gartner, Forrester, IDC, Grand View Research, MarketsandMarkets — for the broader market.

02 Subtract non-AI-replaceable revenue (e.g., billing systems that aren't agent-driven).

03 Cross-check with **bottom-up**: estimated user base × ARPU of leading product.

04 Triangulate with **company-revenue-multiple** for category-defining companies.

Worked example — Code Generation Agent SW TAM ($28B)

Largest single TAM revision in 2026-04 (+133% from $12B), showing all three methods converging:

Method	Calculation	Output
	MnM "AI in software development" 2026 projection	$7B — too low; only counts tools, not agentic-coding spend
	Cursor $2B + Copilot ~$600M + Claude Code + Codex + Replit Agent + Cognition + niche tools	~$4-5B current; 3× growth → 18-mo forward $15-20B

Method	Calculation	Output
	Cursor $60B ÷ ~30× = ~$2B ARR; ~10-15% market share → category $15-20B today	5-year addressable: $25-35B

Reading lesson: *when one company's ARR exceeds 10-15% of your published category TAM, your TAM is wrong, not the company.*

4.2 Labor TAM

Definition: total annual human-labor spend addressable by AI agents in this track, globally. Built bottom-up via the McKinsey/WEF task-fraction methodology.

Worked example — Tax Filing Assistant Agent Labor TAM ($128B)

01 **Identify occupations** (BLS): Tax Preparers ~90K + Accountants ~1.4M (of whom ~30% do tax) = 420K tax-active accountants in the US.

02 **Wages**: Tax Preparers ~$50K avg; tax-active Accountants ~$80K avg.

03 **US wage bill on tax work**: 90K × $50K + 420K × $80K × 30% + seasonal ≈ **$16-17B**.

04 **Task automatability** (McKinsey): tax prep scores **71%** automatable. → US addressable $11.7B.

05 **Global multiplier**: US ≈ 9-10% of global tax-prep → multiplier ~10-11×. → **$128B globally**.

Cross-validation: ILO global accounting workforce × 25% tax × $30K avg comp × 71% ≈ $75B (lower bound). McKinsey GenAI study: $120-180B annual global productivity unlock from AI in tax/audit by 2030. Our $128B sits at the lower end of McKinsey's range — conservative.

Reading lesson: *Labor TAM is bigger than people intuit because it counts the full wage bill the AI agent could replace, not just incremental productivity gains.*

4.3 Displacement Rate

The most carefully constructed number in the framework. We never take a single analyst's "AI will replace X% of accountants" claim — those ignore that a job is a bundle of dozens of steps. Instead we build it up step-by-step.

6-step methodology (per 50×+ deep-dive track)

01 **Workflow breakdown** — decompose the track into 15-25 discrete steps grouped by phase. *Example: Contract Review Agent breaks into ~22 steps across Intake / Substantive Review / Negotiation / Closing.*

02 **Per-step capability assessment** — classify each step:

- **FULL** — AI handles 80%+ autonomously, output production-ready (80-95%)
- **PARTIAL** — AI-assisted; human applies judgment for the last mile (30-70%)
- **NO** — Human-only today (0-15%)

03 **Weight by hours, not step count** — Replacement Rate = Σ(Step% × Hour-share). The high-volume boring steps (clause redlining, cross-references) are usually the most automatable, so weighted rate is much higher than naive count.

04 **Triangulate with three external sources** — Goldman Sachs Occupational Automation Study, McKinsey Task Automation Fractions, WEF Future of Jobs. If our bottom-up rate sits inside the Goldman-McKinsey-WEF range, we have triangulation; outside → re-examine classification.

05 **Report as a range, use conservative midpoint** — never a point estimate. Lower bound used for Net Effect modeling, because under-investing in over-claimed AI is the more expensive mistake.

06 **Update triggers** — upward (new benchmark, model release, production data, YC clustering); downward (regulatory ruling, court decision, published failure case).

Worked example — Code Generation displacement (30-50% → 55-75% in 2026-04)

01 JetBrains Jan 2026 survey: Copilot 29% / Cursor 18% / Claude Code 18% *production* adoption (not experimentation).

02 Claude Opus 4.6 hit OSWorld 72.7%, GPT-5.4 hit GDPval 83% — agent benchmarks crossing thresholds where end-to-end coding tasks become FULL.

03 Cursor's $2B ARR is empirical proof of value capture, validating the productivity gain.

04 Triangulation: Goldman SDE-automation 47-62%; McKinsey high-adoption scenario 50-70%; our 55-75% sits in the upper half of both → revised up.

4.4 Labor/SW Ratio

Simple division. Categorized:

Ratio	Interpretation
	Structural gap — massive unaddressed labor pool, tiny AI software footprint. Highest priority.
	High-priority opportunity for compression or expansion.
	Core opportunity — meaningful labor-to-software arbitrage.
	Moderate — software has already penetrated substantially.

4.5 Net Effect — three canonical case studies

Decided by asking: *if AI cuts this service's cost by 10×, what happens to total demand?*

🔴 Net Compression — Contract Review Agent (Labor/SW 80×)

Corporate-legal contract review is a saturated B2B service. Companies don't draft *more* contracts because review got cheaper — they draft the contracts they already need. The buyer (general counsel) cares about output (signed deal), not input (associate hours). When Harvey reviews a contract in 3 minutes vs. an associate's 3 hours, the GC pays Harvey instead of the law firm.

- **Mechanism** — 1 senior associate billing $400/hr × 3 hrs = $1,200 of human-labor revenue → captured by Harvey at ~$50 marginal cost. Total category dollars **shrink** because the service is priced by hours, and the hours collapse.
- **Demand** is inelastic at the existing margin — corporates aren't sitting on a backlog of contracts they wish they could afford to review.
- **Investment thesis** — pick the picks-and-shovels winner (Harvey, Legora, Ironclad) capturing redistributed dollars. Don't bet on volume growth — bet on margin capture from labor compression. The absolute software TAM is **bounded** by the labor pool being compressed.

✅ Net Expansion — CBT / Mental Health Agent (Labor/SW 437×)

There is a **massive unmet demand pool** in mental health. ~75% of people who need therapy don't get it — most because $200/session is unaffordable, not because they don't want it. WHO estimates ~1B people globally have a mental health condition, while only ~150M see a clinician. AI therapy at $20-30/month doesn't compete with the human therapist's existing clients — it serves the 850M who weren't being served at all.

- **Mechanism** (textbook Jevons effect) — price drops 10× → demand rises >10× (more like 50-100× given the unmet pool). Existing therapists keep their clients (high-acuity cases) at unchanged or higher rates.

- **Total category dollars grow** — Labor TAM expands as new spend appears.
- **Investment thesis** — bet on consumer-grade products that win on UX, brand and clinical safety (Woebot, Wysa, Slingshot). The 437× ratio is durable because the labor pool serves only ~15% of the demand.
- **The hard part for investors** — regulatory ceiling (FDA digital therapeutics path, state-by-state therapy licensing, suicide-risk liability). Moat is regulatory navigation, not the AI itself.

🟡 Mixed — Advertising Agency Super-Agent (Labor/SW 62.5×)

Marketing & growth has both dynamics simultaneously, and which dominates depends on customer segment.

- **Compression at the agency layer** — mid-market and large brands historically paid $500K-5M/year for creative + planning + buying. AI super-agents handle 60-70% of that workflow. Big-brand spend goes from $2M → $700K, with the same campaign output → labor compression.
- **Expansion at the SMB layer** — tens of millions of SMBs globally never had agencies because they couldn't afford $50K/month minimums. AI agency tools at $200-500/month bring a previously priced-out audience in. A coffee shop now runs targeted Meta ads with AI-generated creative.
- **Net direction** depends on which segment grows faster. Currently: brand-side compression ~$100B agency revenue at risk over 5 years; SMB-side expansion ~$80B new spend unlocked. Net: roughly flat in dollars, but huge **redistribution**.
- **Investment thesis** — don't bet on the category as a whole. Pick a side: agency-replacement super-agent (compression, finite ceiling) **or** SMB-democratization tooling (expansion, larger ceiling). Confused positioning is the failure mode.

Pattern recognition rule: *if you can't crisply describe the price-elasticity story for a track in two sentences, you don't know its Net Effect yet — you have a Mixed track and should treat it that way.*

◆

The *watchlist*

Tier 1 — 100×+ tracks (structural gaps)

Highest priority for sourcing. Both compression (legal cluster) and expansion (mental health) plays.

Track	Ratio	Net Effect	Watch list
	437.5×	✅ Expansion	Woebot, Wysa, FDA digital-therapeutics guidance
	155× (was 233×)	🔴 Compression	Mercor $2B; YC W26 cluster (Perfectly, Skillsync, Vela)
	200×	✅ Expansion	Figure AI $39B; Helix 02 VLA
	106× (was 154×)	🔴 Compression	Oxus, Fenrock, Veriad (YC W26); EU AI Act
	110× (was 183×)	🔴 Compression	Arcline, LegalOS, Vector Legal (YC W26)
	72× (was 120×)	🔴 Compression	DualEntry $10.87B; Intelmarket; Big-4 AI billing
	80× (was 140×)	🔴 Compression	Harvey $11B, Ironclad, Legora $5.55B
	87.5× (was 130×)	🔴 Compression	Harvey + Legora super-agent
	82× (was 100×)	🔴 Compression	Harvey $11B; Thomson Reuters CoCounsel $200M+ ARR

Why ratios fell in legal cluster: the AI software TAM caught up fast (Harvey, Legora, super-agent platforms raised aggressively in 2026-Q1), so the leverage compressed. Still 80×+ — investable.

Tier 2 — 50-99× expansion tracks (largest TAM-growth potential)

- **Personal Learning Agent** (90×) — Khanmigo, Synthesis
- **Precision Agriculture** (88.9×) — physical-task gap is hardware-limited
- **Drug Discovery** (35.4× now, was 55× — software caught up) — Insilico Phase IIa, Recursion
- **Investment Institution Super-Agent** (51.4×) — Hebbia $700M, Rogo, Aaru $1B
- **Music Composition** (62.5×) — creator-economy expansion play

Tier 3 — Emerging tracks not yet in framework

Likely additions next cycle (2026-06):

- **AI Software Engineering Agent** — may split from Dev/Tech category given Cursor $60B / $2B ARR scale
- **Healthcare Revenue Cycle Management** — Overdrive Health, ClaimGlide
- **Supply Chain Intelligence** — Pollinate
- **Physical Security & Surveillance** — computer-vision for commercial security

◆

06 — CREATION JOURNEY

How this *was built*

Phase 0 — 2025-Q1: Baseline framework

Established the 201-track taxonomy across 10 categories. Built the Labor-vs-Software TAM lens. Cross-validated displacement rates against Goldman / McKinsey / WEF. Hand-curated representative companies and funding data per track.

Phase 1 — 2025-Q1: Deep-dive sheets

Built per-track 4-section workflow analyses (Workflow Breakdown → AI Capability Assessment → Critical Gap Analysis → Strategic Insight) for the **22 highest-priority tracks** (Labor/SW ≥ 50×): 9 tracks at 100×+, 13 tracks at 50-99×.

Phase 2 — 2026-Q1: YC W26 mapping

When YC W26 batch was announced (Demo Day 2026-03-24, 199 companies, 74% AI, 64% B2B), we ran every company through the 22-track filter. Output: `YC_W26_TAM_Mapping.xlsx` — companies ranked CORE / ADJACENT / INFRA per track.

Phase 3 — 2026-04: First bimonthly update

Established the every-2-months update cadence covering 5 update types (new tracks, TAM revisions, displacement updates, capability unlocks, key players). The 2026-04 cycle produced **30 updates** — most material being Code Generation Agent SW TAM $12B → $28B (+133%), legal cluster TAM +45-75%, Drug Discovery clinical milestone (Insilico Phase IIa), humanoid robotics breakthrough (Figure AI $39B + Helix 02 VLA), Q1 2026 macro: $300B record VC quarter, AI = 80%.

Phase 4 — 2026-04: Visualization site

Built agent-mapping.vercel.app to make the framework explorable. Self-contained static dashboard (Chart.js + vanilla JS), auto-rebuilds from the master xlsx via `build_data.py`. Added the "What's Changed" banner so the bimonthly update is the first thing visitors see.

Phase 5 — Ongoing: Bimonthly updates

Next cycle: 2026-06. Tracked: ratio threshold crossings, Net Effect reclassifications, new players in 50×+ tracks, capability unlocks from frontier-model releases, YC S26 batch mapping when announced.

◆

07 — LIMITATIONS

What's *not in* the framework

- **Geographic granularity** — numbers are global aggregates; we don't break out US/EU/APAC separately.
- **Time-series** — only the latest snapshot in the dashboard; the Change Log shows deltas but not full historical curves.
- **Hardware/robotics depth** — physical-AI tracks noted as gaps but not modeled at component level.
- **Regulatory granularity** — jurisdiction-specific risks summarized but not enumerated per region.
- **Private market liquidity** — valuations reflect last-round prices, not exit comparables.

These are deliberate scope choices. Opening any of them up is a 1-2 month follow-on project.

◆

APPENDIX — DATA SOURCES

What we *cited*

Tier 1 — primary, used directly

Source	What we use it for
Occupational Automation Study (2025 update)	Per-occupation 0-100% automation risk score; one of three displacement-rate inputs
Generative AI Economic Potential — Task Automation Fractions	Task-fraction methodology for Labor TAM; second displacement-rate input
Future of Jobs Report 2025	Displacement probability per occupation; third displacement-rate input
Occupational Employment Statistics	US workforce size and wages by occupation
World Employment & Social Outlook	Global workforce extension of BLS
Employment Outlook	Cross-country wage adjustments

Tier 2 — industry analysts (triangulated)

- Gartner, Forrester, IDC — Software TAM anchors per category
- Grand View Research, MarketsandMarkets, Mordor Intelligence — sub-segment market sizing
- Roots Analysis, BioMedNexus — Drug Discovery / Healthcare AI TAM
- Research and Markets — Legal AI Software market sizing (2026: $5.59B)

Funding & company data

- PitchBook, Crunchbase, Crunchbase News — deal-level funding amounts and valuations
- TechCrunch, The Information, Bloomberg, Reuters — major fundraises, M&A, IPO announcements

- Tracxn, Contrary Research — company landscape mapping
- Y Combinator official directory + Extruct.ai — YC batch listings

Industry / professional bodies

- AICPA — accounting market structure & licensing
- ABA — legal market workforce, UPL regulation
- CSA, NAR — real-estate workforce data
- AMA — physician workforce

Regulatory & policy

- EU AI Act enforcement timelines — Compliance Check track drivers
- DORA (Digital Operational Resilience Act) — Compliance / FinServ drivers
- US SEC AI disclosure rules — FinServ AI risk
- FDA digital-therapeutics guidance — Mental Health / CBT track
- US IRS guidance on AI-prepared returns — Tax Filing track

Q1 2026 specific (used in 2026-04 update)

- **Crunchbase News** (Apr 2026) — $300B Q1 2026 VC, AI = $242B (80%)
- **Bloomberg** — Cursor $60B / $2B ARR; Decagon $4.5B Series D; Sierra $10B / $150M ARR; Harvey $11B
- **TechCrunch** — Lovable $9B Series D; Figure AI $39B; Hebbia $700M
- **Reuters / CNBC** — Insilico INS018_055 Phase IIa (Nature Medicine, Feb 2026); Novartis-Recursion; Novo Nordisk × OpenAI partnership
- **JetBrains** Developer Survey (Jan 2026) — Copilot 29% / Cursor 18% / Claude Code 18% adoption
- **Robot Report** — Figure 03 White House education pilot; Agibot 10K shipped

- **Vendor docs** — Claude Opus 4.6 (Feb 5, 2026), GPT-5.4 (Mar 5, 2026), Gemini 3.1 Pro — all 1M context
- OSWorld / GDPval / GPQA benchmark scores — capability evidence

Peer-reviewed

- Nature Medicine — clinical AI validation (Insilico)
- IEEE / arXiv — model capability papers (computer-use, multimodal)