DTCP

July 2025

Deep Dive on
Agents Infrastructure

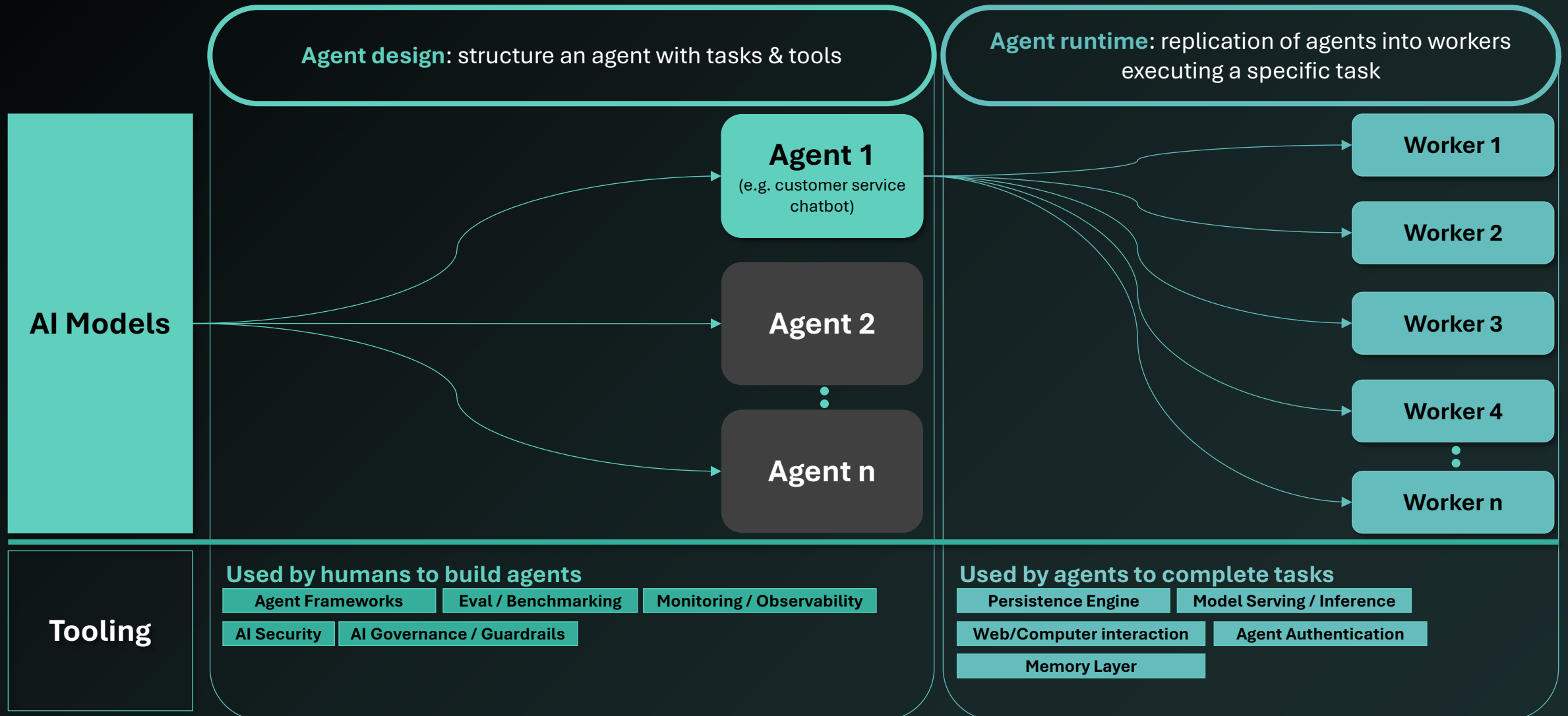
Agenda:

1. **Overview of Agent Infra & Market Map**
2. **One-pagers of 5 key markets**
    1. Agent Builders
    2. Persistence
    3. Authentication
    4. Web/Computer interaction
    5. Memory

Definition: what is an agent?

An agent is a **microservice powered by AI models that owns a task end-to-end**: it reads the situation, picks tools, executes, learns from the result, and repeats until the goal is met

DTCP



Agent Infrastructure - the missing agent building blocks

From chatbot demos to autonomous systems

DTCP

***We're entering a new era beyond software code execution**, an era where systems think, decide, and act*

- A shift from **explicit instruction execution (running code)** to **delegated autonomous goal achievement (running workers)** means the underlying infrastructure can't just be an extension of what enterprises have today
- There needs to be **a fundamental re-architecting to support and secure agents** that can perceive, reason, learn, act, and collaborate with a level of autonomy and complexity that traditional software was never designed for

*We believe that there is **abundant value to be created by the companies leading this revolution** – we will explore them in this deck*

- ***Timing is now:** apps like Glean, Harvey, and Cursor are booming because enterprises see value in AI today, but lack the agent infrastructure to build in-house*

Unlike traditional software, agents perform abstract reasoning with **unpredictable** execution paths

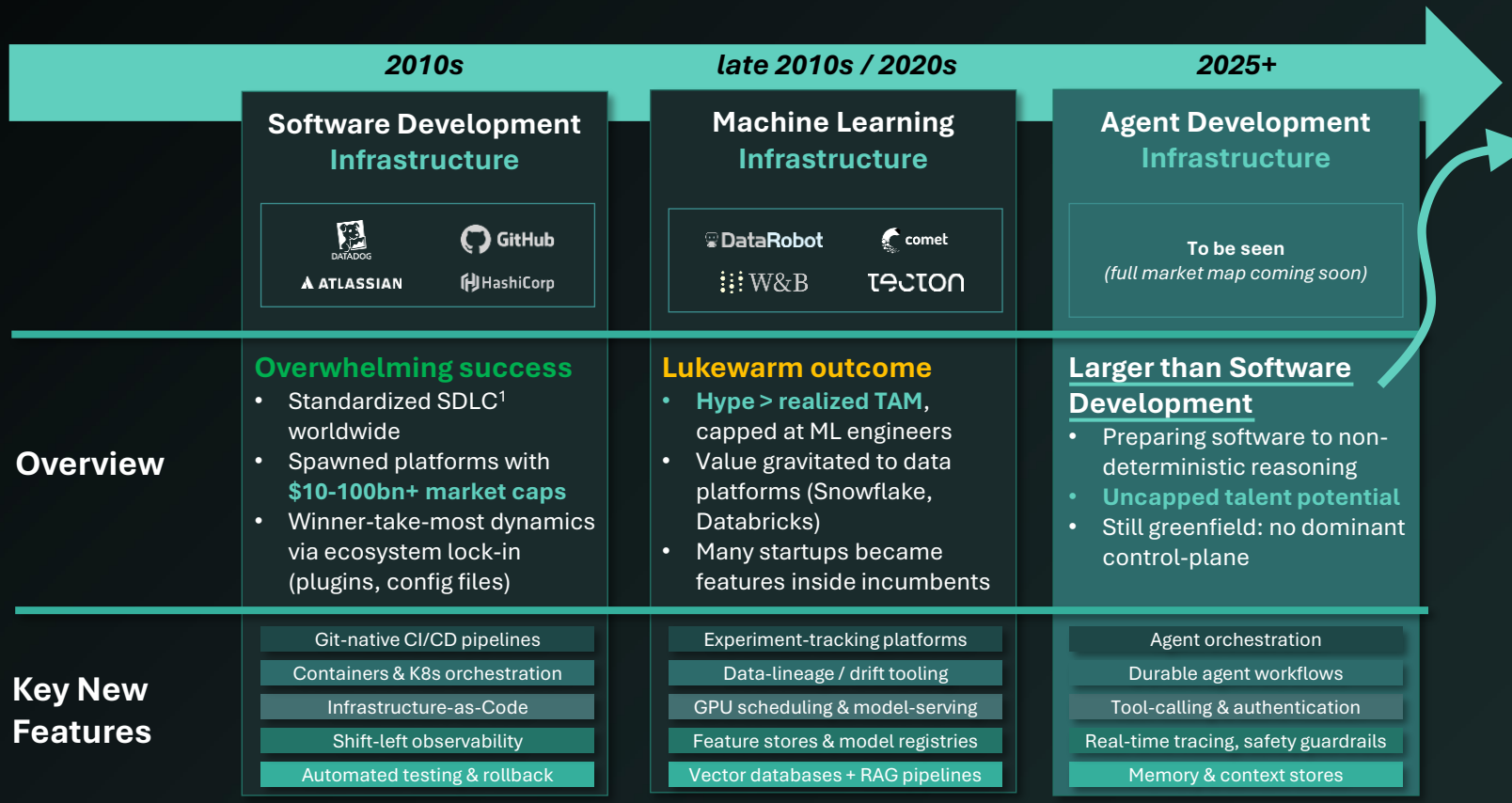
	Software Development	**Agentic Software**
Key definition	**Deterministic; capabilities bounded by developer hours**	**Stochastic (non-deterministic); adaptable to unprogrammed use cases**
User	Pure engineer	Pure engineer + Citizen developer
Enterprise Buyer	Head of IT / CIO / CTO / Head of BU	
Core artifact	Version-controlled code & tests	Prompt graphs, memory & policy configs
Failure model	Crash or timeout → retry same step	Hallucination / wrong tool call → re-plan or intervene

The users may overlap, but the underlying **tech challenges are completely orthogonal**

Agent Infrastructure – a $1 trillion developer stack

DTCP

Agent Development Infrastructure has the ingredients to **become bigger than the current software dev Infrastructure market**



Agent Infra will be bigger than software development infra

- **Citizen developer expands agent infra TAM**: from 30m developers to "*8bn AI engineers*"
- **Higher revenue-per-seat:** autonomous tooling replaces manual dev hours, enabling budget to shift for these platforms
- **Control-plane lock-in:** orchestrating agent action becomes mission-critical - switching costs rival DevOps / CI/CD, anchoring long-term, high-margin revenue

Key question: what are the future standalone categories?

- Many of the innovative categories are facing ever-increasing pressure from AI Labs/hyperscalers offering **model-embedded features or open-source alternatives**

- Every DevOps startup pitch in the 2010s was followed by a VC question of "*why wouldn't AWS/Google build this?*". Sometimes they built it, killing a market; but there was still enough space for multi-deca-$bn companies to be created. **We are seeing the same question today referring to the Foundation Model Labs (OpenAI, Anthropic, Cohere)**

(1) SDLC = Software Development Lifecycle

Early LLM tech barriers are on a clear path to getting solved; **but agent roadblocks are still a major challenge**

DTCP

While the LLM Rush validated AI's transformative potential and spurred enterprise demand, scaling to the 'Autonomous Enterprise' necessitates **robust agentic infrastructure** to overcome critical reliability, governance, and integration challenges

LLM Rush | 2022 – 2025

Delivered the initial promise of LLMs by solving foundational problems, **its very success illuminates a new set of more complex, agent-specific challenges**

Solved

Model Cost: frontier APIs now utility-pricing, costs decreased by >20x over the last years

Model Intelligence: models are getting better by the hour, driven by post-training

Prompt Management: versioned templates, early frameworks gaining traction

Late Stage of Solving

Hallucinations: advancements in memory, RAG-like approaches are progressing fast

Latency: most responses feel real-time, apart from SOTA models (reasoning)

Agent Orchestration: Early platforms chain LLM calls reliably, but lack full integration

Lack of Standardization: MCP, A2A, other emerging protocols are being adopted

Agentic Era | 2025+

Architecting **infrastructure that makes the "Autonomous Enterprise" possible** by developing the **necessary scaffolding for agents in production and at scale**

Early Stage of Solving / Need Solving

Evaluation & Benchmarking : standardized methods to measure AI agents' parameters

Structured Tool Use: use of tools, like search or browser use

Context Management: consistent sessions state across channels, users, and time

Memory: agents "forget" past interactions,

Multi-Agent Collaboration: enterprise workflows require agents to hand off tasks

Durable Execution: reliability varies with open-ended tasks and web/tool use

Real-Time Monitoring: live telemetry and alerting on prompts, actions, and outcomes to meet SLAs and detect drift

Agent Identity: Without least-privilege credentials and tamper-proof logs, CISOs will block large-scale deployments

We will double click on this presentation

Today

Danger zone in models / AI Labs "Flood Zone"

AI Labs adding more native agentic capabilities and agent building components

Scaffolding got us up to 2025; but a robust infra stack is taking shape

We will focus this presentation on these key 5 categories

DTCP

We will explore **five purpose-built layers** that fix today's technology gaps

LLM Rush | 2022 – 2025

Deploying more sophisticated, autonomous AI agents across complex enterprise systems, diverse data silos, and core workflows **reveals critical gaps in the "scaffolding" – the underlying infrastructure**. These are not just technical hurdles but also security, compliance, and operational complexities

Glue-code Sprawl: teams stitched action chains with scripts & Zapier

Explosive Pilots: 78 % of Global 2000 ran at least one Gen-AI PoC in 2024

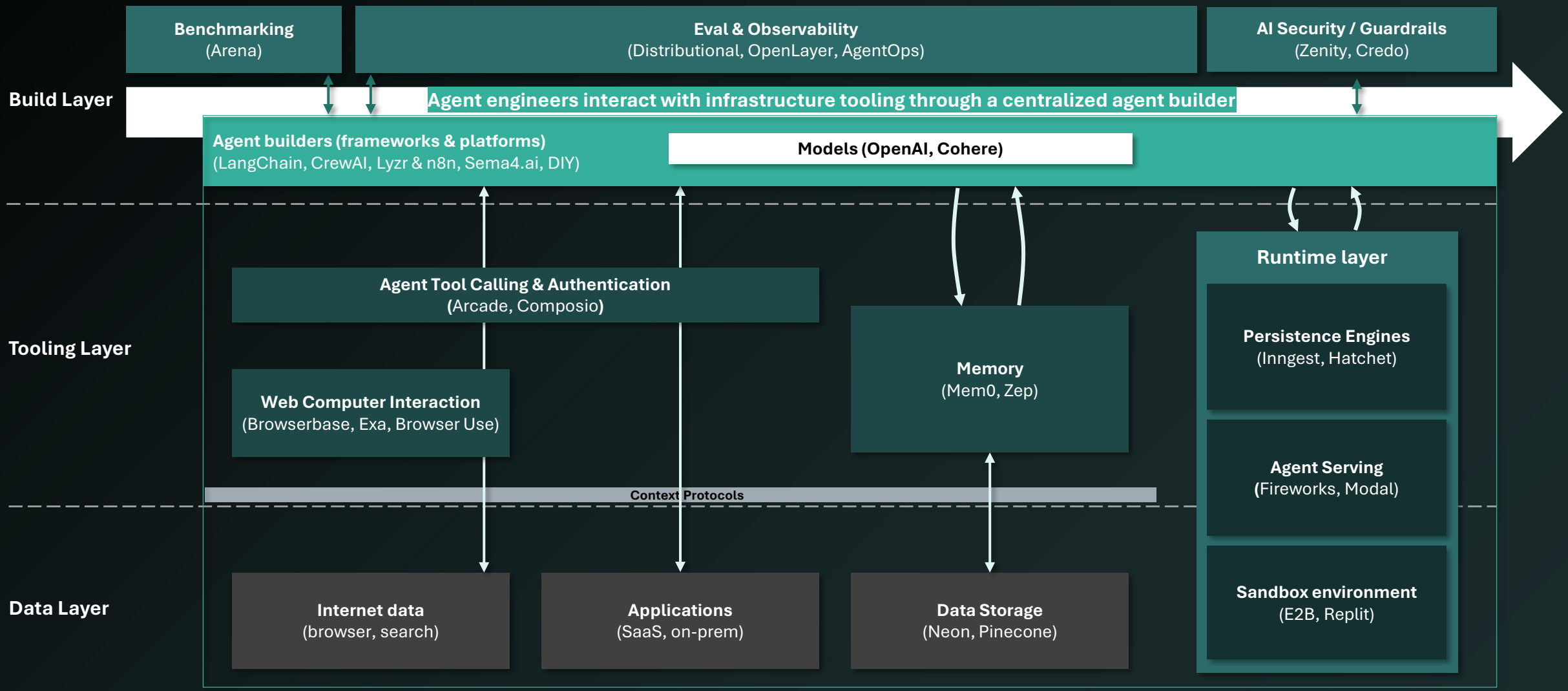
Security Red Flags: internal agents are "over-privileged" and "un-auditable"

Agentic Era | 2025+

"Scaffolding" stack	New Agent Infra Stack	
"Prompt-Spaghetti" & DIY Chains	**Agent Builders**	**Ready-made building blocks to connect AI models, tools, and data** so teams can launch powerful agents faster
Brittle agents with limited failure workarounds	**Persistence Engine**	Engines that are becoming the **default queue/cron for modern agentic applications** – avoiding failures
Vector-DB Sprawl that did not solve all RAG challenges	**Memory Layer**	Memory lets agents **retain context across steps and sessions, enabling coherent and stateful behavior**
Demo-Ware Tool Use	**Web/Tool Use**	Secure connectors that let agents **discover and invoke SaaS & API actions without brittle RPA hacks**
Shadow Agents & Audit Gaps Over-privileged tokens, zero trace, no clear permissions	**Agent Auth & Identity**	Fine-grained access controls and permissions**, ensuring agents only interact with approved tools, APIs, and data** via secure pathways

Agent Infrastructure Stack

DTCP



DTCP

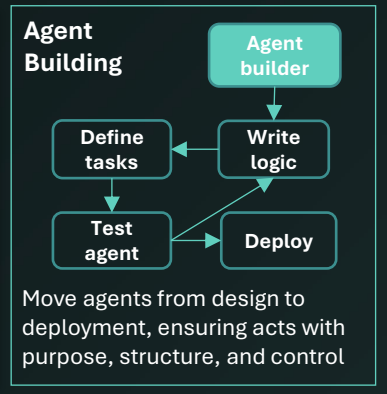
Key Categories of Interest

Agent Builders *(Dev Frameworks & Biz Builders)*

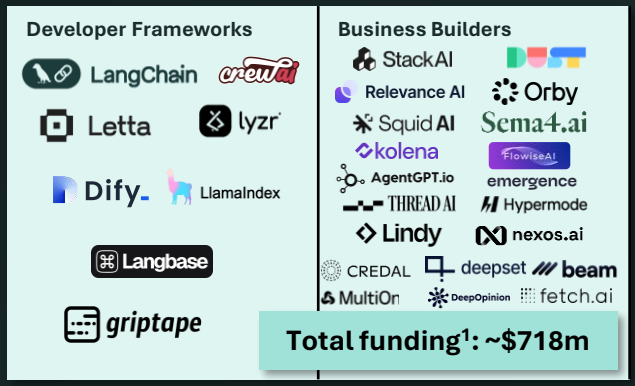
DTCP

Frameworks and/or platforms to **build, coordinate, and deploy autonomous AI agents**

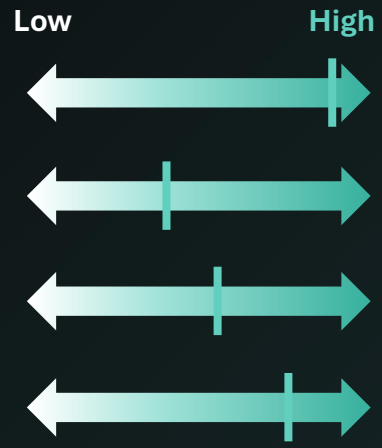
Developer pain	• Developers struggle to build reliable AI agents that can be easily adapted across tasks **without reinventing everything from scratch**
Product & Tech	• These frameworks provide **ready-made building blocks** to connect AI models, tools, and data so teams can launch powerful agents faster • **Open-source frameworks are racing to win developer mindshare**, with the developer community critical to crown the winner
Technology adoption	• An early, fast-moving market with **huge developer & non-technical user excitement** but still **limited enterprise adoption**



Move agents from design to deployment, ensuring acts with purpose, structure, and control



Market Dynamics Summary



- Winner could **own the interface layer of agentic software,** a "System of Agents"
- High developer/user curiosity, but **limited enterprise deployment today**
- Legacy vendors **aren't built for agent-native workflows**
- Crowded field of startups and **tools all vying for developer and enterprise adoption**

DTCP Thoughts

- **Agent Builders as System of Agents to unseat legacy System of Actions/Records (SoA, SoR):** frameworks have an advantageous wedge to be the next generation System of Agents that sits above SoRs, SoAs, and all other data/ tooling
- **Switching costs will harden fast:** early pilots look swappable, but once an enterprise embeds a framework in >10 internal workflows and builds >10 production agents, switching costs grow with number of agents
- **Winner-take-most (not all) market**: developers will gravitate into 2-3 key frameworks, but we expect agent builders to successfully take several shapes, targeting different users, verticals, and abstraction depth

(1) Pitchbook, Crunchbase

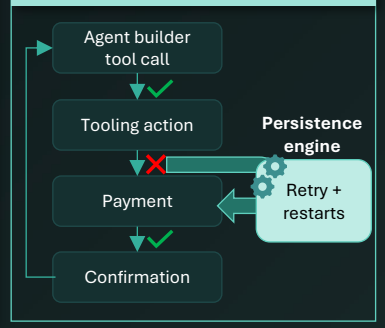
Persistence Engines

DTCP

Ensures durability in agent workflows, particularly in long-running multi-step tasks

Developer pain	• Teams already waste hours with traditional workflows failures (queues, cron jobs and retry logics); **non-deterministic AI just exacerbates it**
Product & Tech	• **Durable orchestration is the backbone for reliable agent action**: without guaranteed retries and state, multi-step agents fall apart in production • Engines that expose easy **TS/Python SDKs are becoming the default queue/cron for modern apps**, replacing RabbitMQ + DIY scripts.
Technology adoption	• Traditional workflows are already adopting it (Temporal, AWS SWF), but **it is early days for AI-native solutions**

Tech example



Persistence Engine & Agent-Native

inngest · Trigger.dev · Temporal · hatchet · dagster · PREFECT

Total funding[1]: ~$480m

Market Dynamics Summary

	Low — High	
Platform potential		• Durable execution **necessitates a dedicated solution and layer** that is AI / Agent native
Current buyer interest		• Developers are experimenting, but **enterprise adoption is still early**
Incumbent threat		• Legacy tools **weren't built for AI-native workflows,** large gaps remain
Competitive heat		• Plenty of OSS players and workflow startups, but **no category leader yet**

DTCP Thoughts

- **Strong data flywheel** that benefits the first mover as **"agentic action data"** (data derived from successful execution paths, corrected failures, optimal tool sequences) are **used to fine-tune subsequent iterations of existing agents**
- **Legacy orchestrators were built for static workflows, not agent behavior.** They lack native support for retries, branching, and tool chaining thus making them brittle and **ill-suited for AI-native workloads**
- Traditional persistence tools operating in **code-defined workflows like Temporal are evolving** to compete with agent-native durable platforms; **converging on the same market**

(1) Pitchbook, Crunchbase, Funding Round Press Releases (2) 98% of funding is from Temporal

Memory

DTCP

Enables an agent to **retain and recall conversational and contextual information** beyond the prompt window

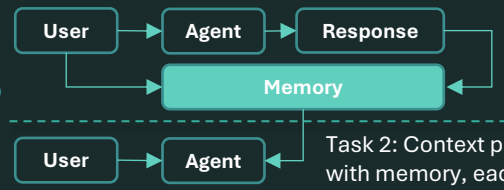
Problem

- Agents today are **stateless**: incapable of learning from past interactions or adapting over time; developers build intricate systems around the LLM to simulate memory

Solution

- Memory **optimizes for latency of task specific knowledge** and contextual relevance across tasks
- Memory lets agents **retain context across steps and sessions**, enabling coherent, stateful behavior

Stateful: Learns from each task



Task 1: Memory preserved

Task 2: Context preserved with memory, each task builds on previous task

2.4x Higher Consistency
With Dedicated Memory[2]

3x Task Completion
Compared to without Dedicated Memory[3]

For enterprises, this translates directly to the complexity of tasks they can automate

Dedicated Agent Memory

mem0 zep cognee

Total funding[1]: ~$5m

Market Dynamics Summary

	Low — High	
Platform potential		• Important breakthrough; but with limited path to expand laterally
Current buyer interest		• Limited urgency as most teams are **still "simulating" memory with DIY**
Incumbent threat		• Similar to Vector DB, it might be relatively **easy for incumbents to add the technology**
Competitive heat		• A lot of interest but fragmented with **no clear technical consensus yet**

DTCP Thoughts

- **Memory is the most expected AI breakthrough:** as long-term memory across applications is the key challenge agent engineers have – their users don't want to give the same context/prompt multiple times
- **Data overload grows exponentially:** as enterprises agents log more interactions, poorly managed memory becomes costly. "Techniques to compress and prune memories will be critical"
- **Memory will need governance**: in knowledge conflicts (two memories that contradict each other), agents will need policies to resolve these inconsistencies
- **Emergence of Memory Protocols:** standards for memory interchange are emerging, facilitating multi-system agents

(1) Pitchbook, Crunchbase, (2) Stanford HELM (3) LlamaIndex Eval, 2024

Browser Infrastructure

DTCP

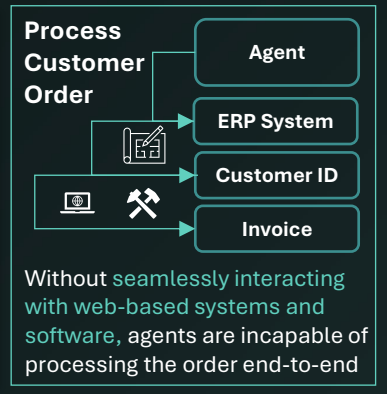
Enables agents to **browse the web and use web-based systems and software** to complete complex tasks end-to-end

Problem

- Agents struggle **to reliably interact with websites and web interfaces** due to human-centric GUIs and limited integration surfaces

Solution

- Specialized web browsing infrastructure, allows agents to programmatically **navigate dynamic websites, extract data, click elements, and complete end-user workflows** where APIs don't exist
- Browser Infrastructure helps tool integration layers, enabling agents to **discover, connect to, and manage access to SaaS tools via the browser**
- Desktop automation and computer control platforms, give agents abilities **to interact with local operating systems**, manipulate files, and control native applications



Without seamlessly interacting with web-based systems and software, agents are incapable of processing the order end-to-end

Dedicated AI Agent Web/Browser Infra

Browserbase · APIFY · notte · tavily · Anchor browser · Browser Use

Total funding[1]: ~$91m

Market Dynamics Summary

	Low — High	
Platform potential		• Crucial enabler for agents, but **possible to be absorbed** within models or agent frameworks
Current buyer interest		• High demand as teams **need ways to reliably interact with real-world tools and websites**
Incumbent threat		• Traditional browser automation and integration tools **may pivot into this space**
Competitive heat		• Fast-moving space with lots of fragmentation; **standards and best practices still emerging**

DTCP Thoughts

- **Open Protocol Proliferation:** As MCP becomes widely adopted as the "web tool protocol", there will be a flourishing of developers exposing apps via MCP, and agent dev frameworks natively supporting it
- **AI Agent-native browser frameworks are emerging:** integration of the MCP protocol brings Stagehand (Browserbase) to any external LLM, combining Stagehand with MCP delivers an OpenAI Operator functionality for a wider range of websites
- **Redesigning GUIs for Agents via Headless Browsers:** A headless browser is a browser designed to be controlled by code. It's "headless" because there's no graphical user interface (GUI). **Headless browsers for AI Agents enable code to interact with any SaaS tools via the browser the same way people do**, allowing developers to build integrations without being limited by available APIs and MCP severs

(1) Pitchbook, Crunchbase

Application Calling & Agent Authentication

DTCP

Identity management, policy enforcement, and security monitoring for agents

Problem

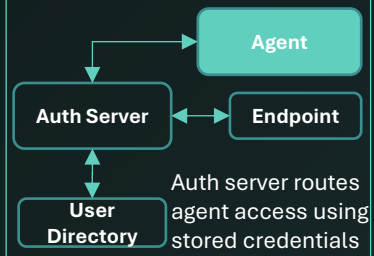
- Agents today **lack enforceable access boundaries;** no identity, no scoped permissions, and no mechanism to restrict or audit what tools they can access

Solution

- Establishes verifiable agent identities, enabling teams to **track and manage agent access across tasks**, and prevent anonymous or unauthorized actions
- Applies fine-grained access controls and scoped permissions, ensuring agents **only interact with approved tools, APIs, and data** via authenticated, secure pathways
- Introduces governance mechanisms like audit logs, approval workflows, and intervention triggers to **give humans visibility and control** over agent

Verifying Agent Access

Agents request access permissions through auth server



Auth server routes agent access using stored credentials

Arcade · Composio · clerk · descope

Total funding[1]: ~$107m

Market Dynamics Summary

	Low — High	
Platform potential		• Foundational **for trusted agent deployment;** positioned to be a core layer in the agent stack
Current buyer interest		• Growing need for secure, governed agent actions **but still early in adoption**
Incumbent threat		• Large identity players in the space; but **agentic workflows are meaningfully different**
Competitive heat		• Plenty of movement, but **fragmentation and unclear ownership** slow momentum

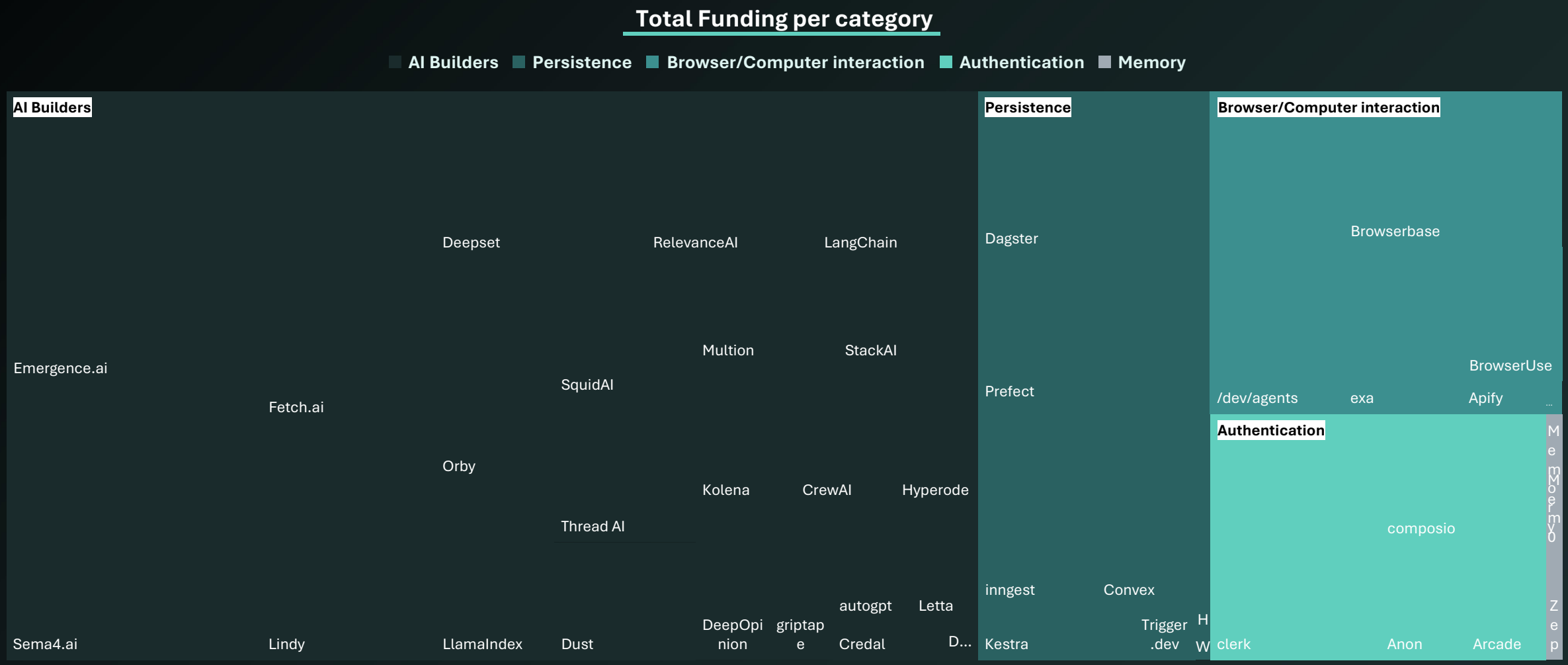
DTCP Thoughts

- **Digital Worker ID emerges:** Enterprises will demand digital passports for their AI agents, analogous to employee IDs. **Current workarounds** involve (1) agents borrowing user ID (=identity explosion) or –worse- (2) API/Token keys embedded in prompting (=security nightmare)
  - Secure tool use is arguably the #1 blockers of enterprise adoptions
  - However, most enterprises are still considering the (1) workaround as “good enough for now”, before large scale of agent adoptions
- **MCP has fundamentally changed this market**: for the best; by packaging context into one signed object, MCPs shrank the gap between classic OAuth scopes and the vision of true fine-grained Agent Auth

(1) Pitchbook, Crunchbase

AI Builders are the most funded category

DTCP



Note: Persistence Engine Excluding Temporal