INTRODUCTION

Proactive Agents

& The OpenClaw Case

Daniel Schroter Thüm

Senior Consultant / AI Engineer

daniel.schroter.eu

March 2026

OVERVIEW

Agenda

01 Reactive vs Proactive Agents

02 Architecture Patterns

03 Key Research & Benchmarks

04 The OpenClaw Case

05 Challenges & Risks

06 Where It's Heading

EVOLUTION

Reactive vs Proactive Agents

REACTIVE	PROACTIVE
TRIGGER User prompt	TRIGGER Self-initiated
PLANNING Per-request	PLANNING Goal-driven, continuous
MEMORY Session-scoped	MEMORY Persistent, cross-session
CONTEXT What you told it	CONTEXT Observes + remembers
PARADIGM System of language	PARADIGM System of behavior

Wooldridge & Jennings, 1995 | Aisera: LLM Agents 2025 | Data Science Dojo

TRIGGER MODEL

How Proactive Agents Decide to Act

Event-Triggered

"Something happened — what do I do?"

- Webhook fires, email arrives
- User sends a message
- File changes, CI fails

ChatGPT, Claude, Devin

→

Always-On Heartbeat

"Is there something I should do?"

- Agent self-initiates on schedule
- Checks conditions, evaluates context
- Acts only when there's a reason

OpenClaw, ChatGPT Pulse (paused)

The core challenge: When to act vs stay silent. Even GPT-5 and Claude Opus achieve only ~40% on proactive benchmarks. Getting this wrong means Clippy.

Lu et al., ProactiveBench (2024) | PROBE Benchmark (2025) | ChatGPT Pulse (OpenAI)

CLASSIFICATION

Levels of Agent Autonomy

Feng, McDonald & Zhang (Univ. of Washington, 2025)

L1	Operator	User makes all decisions, agent supports on demand	ChatGPT Canvas, MS Copilot
L2	Collaborator	Shared planning, fluid control handoffs	OpenAI Operator
L3	Consultant	Agent leads, user provides feedback and preferences	Gemini Deep Research, Replit Agent
L4	Approver	Agent independent, user approves consequential actions	SWE Agent, Manus, Devin
L5	Observer	Fully autonomous, user can only monitor or kill switch	Voyager, The AI Scientist

Key insight: Autonomy is a design choice, not a technical inevitability. Proactive agents operate at L4–L5 — they self-initiate, not just execute.

Feng et al., "Levels of Autonomy for AI Agents" (arxiv 2506.12469) | Knight First Amendment Institute

ARCHITECTURE

Anatomy of a Proactive Agent

Perception
Observe signals, events, context

Planning & Goals
Decompose, prioritize, schedule

Action Execution
Tools, APIs, code, messages

`--- observe -- think -- act -- reflect -- repeat ---`

Memory
Short-term + long-term + episodic

Reflection
Self-evaluate, learn, adjust

Trigger / Monitor Loop
Heartbeat, events, cron

CoALA (Sumers et al., 2024) | Wang et al., "Autonomous Agents Survey" | Agentic AI Architectures

RESEARCH

Key Academic Papers

arxiv 2410.12361

Proactive Agent: Shifting LLM Agents from Reactive to Active Assistance

ProactiveBench benchmark. Reward model achieves 91.8% F1 consistency with human judgments.

arxiv 2510.19771

PROBE: Beyond Reactivity

Decomposes proactivity into 3 capabilities: search for issues, identify bottlenecks, execute resolutions.

arxiv 2602.04482

ProAgentBench: Evaluating Proactive Assistance

28,000+ events from 500+ hours of real user sessions. Evaluates timing prediction + assist content.

CHI 2025

Proactive Conversational Agents with Inner Thoughts

Inner reasoning enables agents to anticipate conversational needs and take initiative.

BISE 2024

When AI-Based Agents Are Proactive

Proactive AI decreases users' competence-based self-esteem, reducing system satisfaction.

CASE STUDY

OpenClaw is an open-source ambient personal AI agent — always running, connected to your systems, acting on your behalf.

It responds to your messages across 20+ platforms — but it also wakes up on its own, checks your email, calendar, and connected services, and takes action without being asked.

Reactive

You message it via WhatsApp, Telegram, Slack, iMessage — it responds and executes.

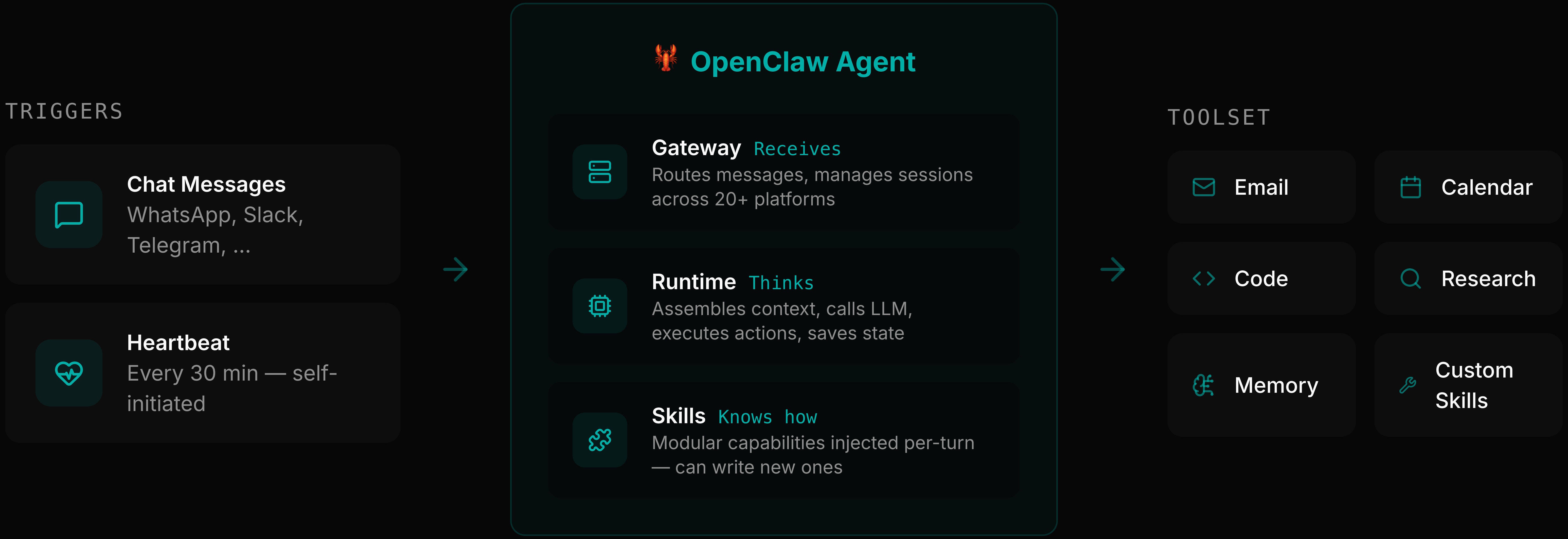
Proactive

Heartbeat wakes it every 30 min. It checks your systems and acts when there's a reason.

250K+ GitHub stars in 4 months — surpassed React. Created by Peter Steinberger (PSPDFKit founder, joined OpenAI Feb 2026). Jensen Huang: "the most important software release, probably ever."

GitHub | Jensen Huang (OfficeChai) | CNBC (Feb 2026)

Architecture Overview



TRIGGERS

Chat Messages
WhatsApp, Slack, Telegram, ...

Heartbeat
Every 30 min — self-initiated

→

🦞 OpenClaw Agent

Gateway `Receives`
Routes messages, manages sessions across 20+ platforms

Runtime `Thinks`
Assembles context, calls LLM, executes actions, saves state

Skills `Knows how`
Modular capabilities injected per-turn — can write new ones

→

TOOLSET

Email	Calendar
Code	Research
Memory	Custom Skills

docs.openclaw.ai/heartbeat | Architecture Guide

Traditional heartbeat: ”is this node alive?”
OpenClaw heartbeat: ”is there something I should do?”

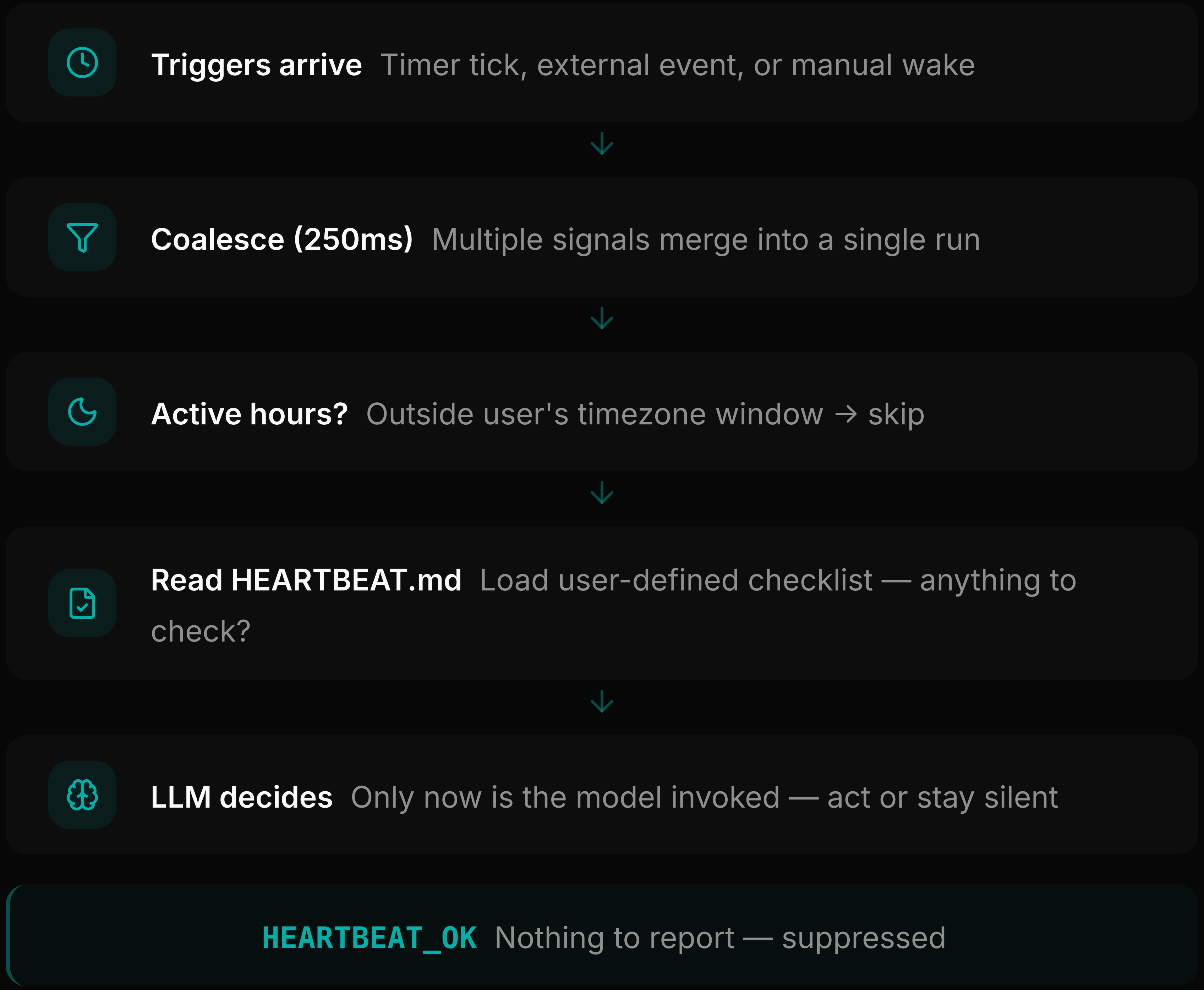
A classic distributed systems pattern — adapted for agentic AI. Every 30 minutes, the agent wakes up. Cheap deterministic check first. LLM only when there's actually a reason to act.

- Separate decision from execution
- HEARTBEAT_OK = silent
- Heartbeat vs Cron

Fowler, Heartbeat Pattern | docs.openclaw.ai/heartbeat | Heartbeat vs Cron

OPENCLAW PATTERN

How the Heartbeat Works



Triggers arrive Timer tick, external event, or manual wake

Coalesce (250ms) Multiple signals merge into a single run

Active hours? Outside user's timezone window → skip

Read HEARTBEAT.md Load user-defined checklist — anything to check?

LLM decides Only now is the model invoked — act or stay silent

`HEARTBEAT_OK` Nothing to report — suppressed

docs.openclaw.ai/heartbeat | EntreConnect

OPENCLAW

What Can It Do?

Workspace Management
Email, calendar, documents — sorts, drafts, resolves conflicts, sends reminders

Write Code & Build Apps
Generates code, builds software, deploys — acts as a full development agent

Research & Analysis
Web research, summarization, data gathering across sources

Self-Extending Skills
Writes its own code to learn new capabilities on the fly

Long-term Memory
Retains notes, preferences, health metrics across sessions

Cron Jobs & Automation
Scheduled tasks, background workflows, flight check-ins

⚠ **Use with caution.** OpenClaw has high-privilege access to your email, calendar, messaging, and can execute code. Treat it as an early-stage system with a large attack surface.

DigitalOcean | Every.to | docs.openclaw.ai

WATCH OUT

Challenges & Risks

User Autonomy Erosion	Proactive help can decrease users' sense of competence (BISE 2024)	Medium
Hallucination Cascading	Errors in autonomous multi-step workflows compound and amplify	High
Prompt Injection	The #1 vulnerability in agentic systems. External data can manipulate agents	High
Denial of Wallet	Agentic DoS: attacker causes infinite loops, burning API budget	High
Context Management	Maintaining coherence across multi-day tasks remains unsolved	Medium
Proactivity Calibration	Too proactive = annoying. Too passive = useless. Sweet spot is hard	Medium
Governance & Accountability	Audit logs, rollback, regulatory oversight for autonomous actions	Medium

The Conversation | CSO Online | BISE 2024 (Springer)

OUTLOOK

Where It’s Heading

Ambient Infrastructure

Always-on agents as persistent background layers, not session-based tools.

MCP & Open Standards

Model Context Protocol for tool access, now governed by the Agent AI Foundation.

Agent Skills

Portable procedural knowledge — MCP is the plumbing, skills are the brain. Adopted by OpenAI and Anthropic.

Agent Societies

Teams of specialized agents managed by a central orchestrator.

Multi-Modal Agents

Agents that see and act on screenshots, GUIs — not just text.

Enterprise Guardrails

Observability, audit trails, sandboxed execution, human-in-the-loop.

VentureBeat: Agent Skills | PulseMCP: OpenAI adopts Skills | Adaline AI 2026

KEY TAKEAWAYS

1. Proactive is the next frontier — from ”answer my question” to ”anticipate my needs.”

2. Architecture is converging: perception, planning, memory, reflection, trigger loop. The heartbeat is a reusable blueprint.

3. The hard problem is judgment, not capability. When to act matters more than how.

4. Start bounded, expand carefully. Human-in-the-loop. The Clippy lesson still applies.

Discussion & Q&A

What proactive agent use cases do you see for clients?

Daniel Schroter Thüm

Senior Consultant / AI Engineer

daniel.schroter.eu