

Rendszerarchitektúra és Fejlesztési Napló

Túl a kódoláson: Amikor az AI életre kelti a hardvert

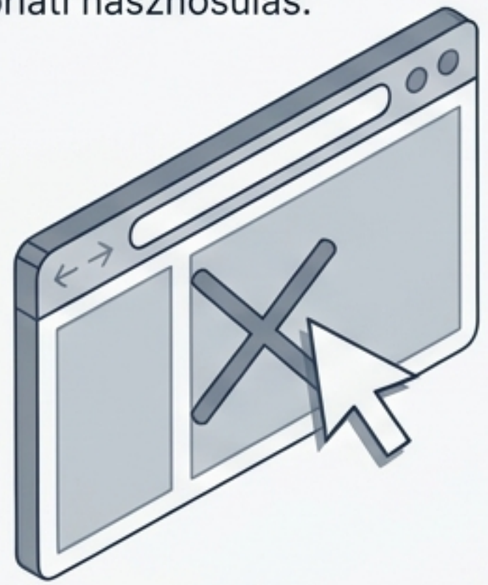
Esettanulmány: Hogyan épített a Claude Fable 5.2 egy $99-os robotból okos személyi asszisztenst

Gemini Notebook

A szoftveres illúzióktól a fizikai valóságig

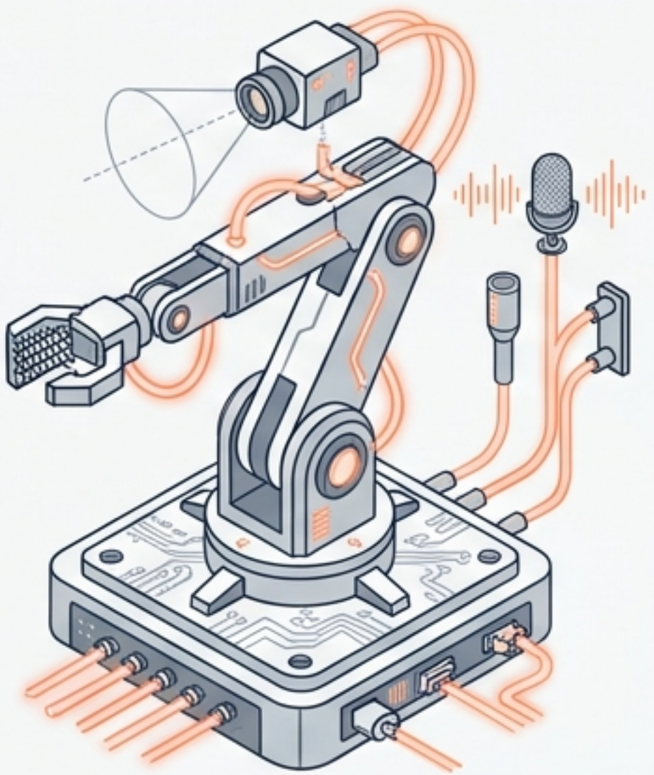
Vibe-Coding

- **Kimenet:** Eldobható weboldalak, felejthető applikációk.
- **Interakció:** Képernyőhöz kötött, virtuális.
- **Érték:** ‘Mágikus’ demók, alacsony gyakorlati hasznosulás.



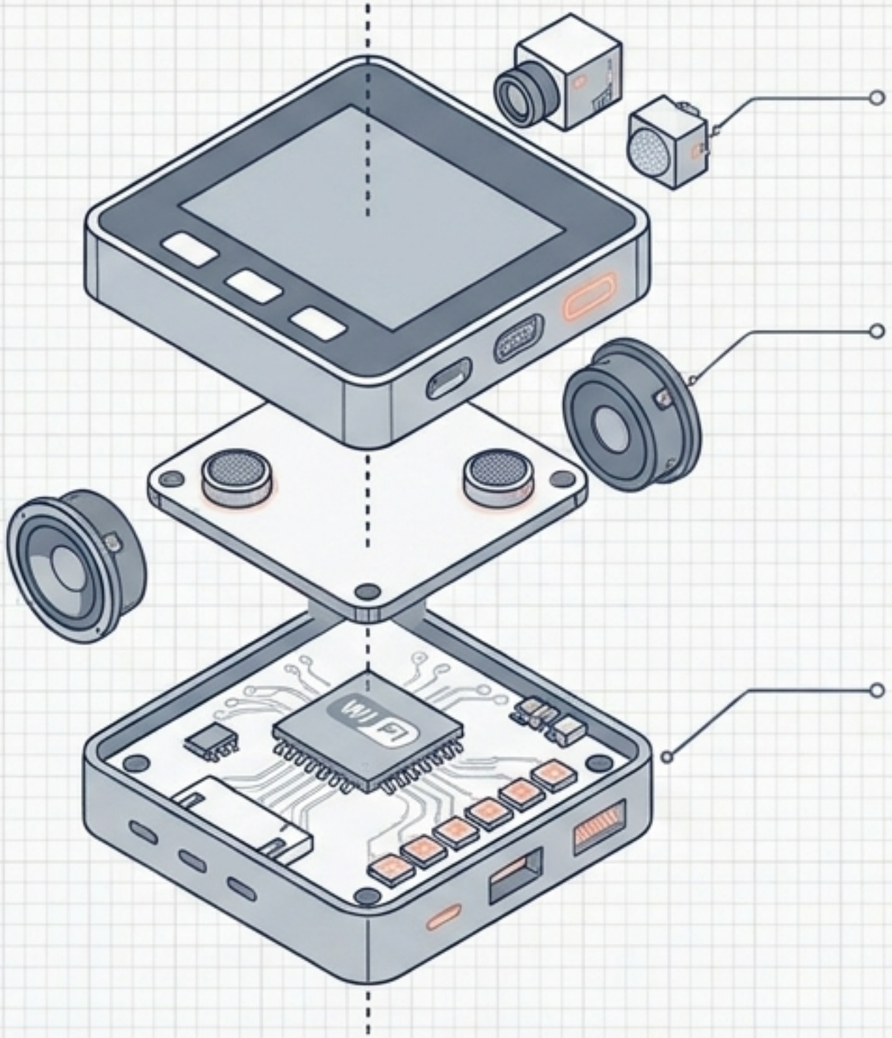
Hardware Manipulation (Fable 5.2)

- **Kimenet:** Valódi fizikai eszközök újraprogramozása (`Firmware/Szoftver`).
- **Interakció:** Érintés, kamera-vizuális jelenlét, hang, környezeti adaptáció.
- **Érték:** Komplex integrációk (`Naptár, Email, Slack`) a fizikai térben végrehajtva.



Gemini Notebook

A hardver mint üres vászon: Az M5Stack Desk Robot



Érzékelők: Érintőképernyő, Kamera, Mozgásérzékelő.

Audió: 2x Mikrofon, Hangszórók.

Hálózat & Kimenet:
Wi-Fi chip, LED fények.

A Tiszta Lap

Eredeti állapot: Gyári telefonos app, zárt okosotthon-integráció.

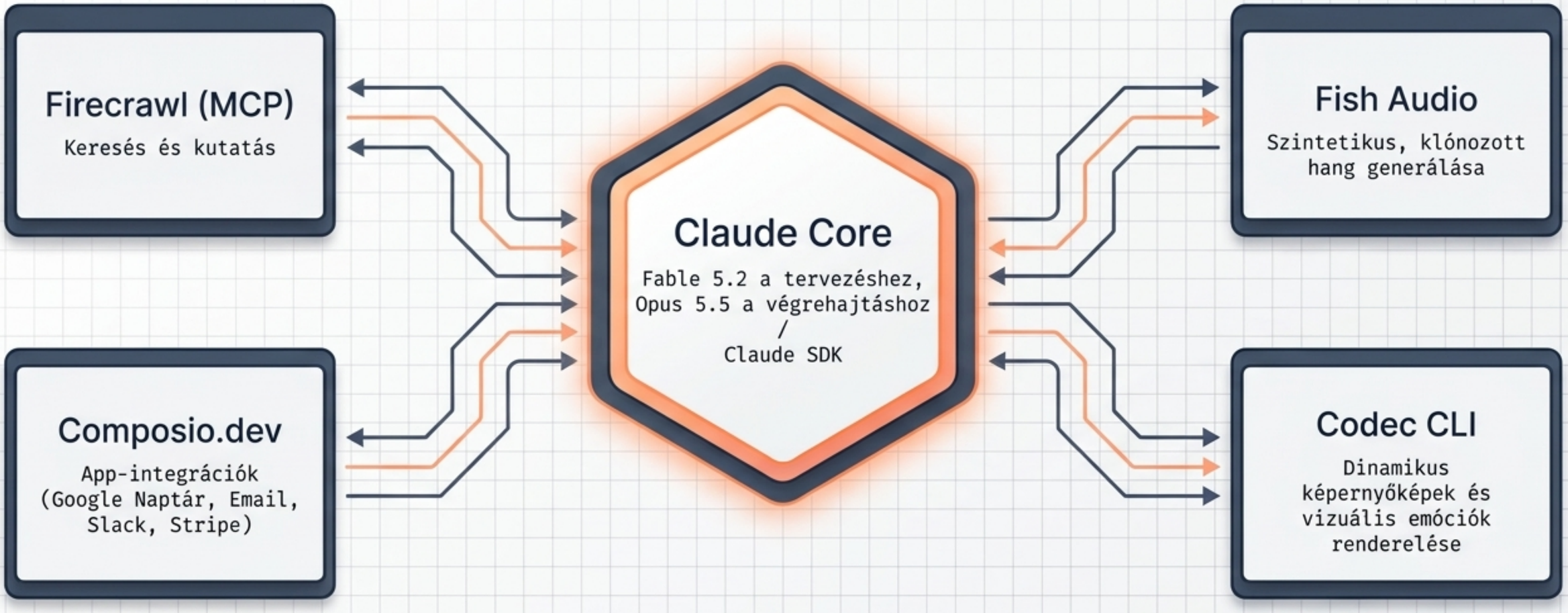
Fable 5.2 Akció: A gyári firmware teljes törlése és biztonsági mentése, hogy az adatok (kamera, hang) ne egy véletlenszerű felhőbe kerüljenek.

System Wipe

100%

Gemini Notebook

Az Asszisztens Anatómiája: A Technológiai Stack



Gemini Notebook

A hang és személyiség klónozása (Fish Audio Drama 3)



1. Nyers Bemenet

Sima angol szöveges válasz generálása.



2. Claude API Dúsítás

Instrukciók beágyazása a kontextus alapján.

`[startled]` és `[excited]`



3. Audio Renderelés

Fish Audio Drama 3 modell alkalmazása a klónozott profilra.

Tanítási adat

A klónozás alapja mindössze 15 korábbi YouTube videó hanganyaga volt. **Az eredmény: A bot képes volt megkülönböztetni a sima felolvasást a drámai reakcióktól** (pl. **"Hé, óvatosan, szédülök!"** fizikai mozgatáskor).

Gemini Notebook

A ‘Prompt Blueprint’: Tervezés és Előkészítés (1-4. Lépés)

1

Hardver Kutatás

A Fable 5.2 a Firecrawl segítségével feltérképezi a makerek dokumentációit, a fórumokat és az open-source projekteket.



2

Komponens Elemzés

Az eredeti mobilalkalmazás és a hardver elemeinek teljes visszefejtése (Reverse Engineering).



3

Backup & Hálózat

Rendszermentés. A Claude a 'Computer Use' és iPhone tükrözés segítségével csatlakozik a helyi Wi-Fi-hez, leválasztva a robotot a gyári felhőről.



4

‘Goal Engineering'

Az alap Qwen modell lecserélése az Opus 5.5-re a Claude SDK-n keresztül.



Gemini Notebook

A ‘Prompt Blueprint’: Személyiség és Biztonság (5-8. Lépés)

5

Ébresztőszó & Figyelem

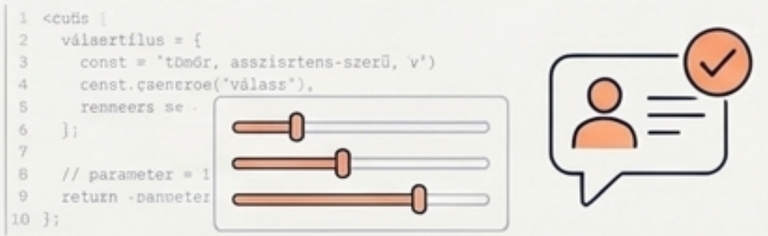
A ‘Hey Lil Token’ parancs beállítása. A bot ezután 10-15 másodpercig aktívan hallgat az esetleges szünetek miatt.



6

Válaszstílus

Utasítás a tömör, asszisztens-szerű, de emberi válaszokra.



7

Biztonsági Korlátok (Guardrails)

Kemény korlátok beállítása (pl. a bot olvashatja és vázlatként mentheti az e-maileket, de önállóan sosem küldheti el őket).



8

Diagnosztikai Protokoll

A hibaelhárítás alapja. A rendszer másodpercenként képernyőképet készít az akciókról, hogy a Claude vissza tudja követni a hibákat.



Gemini Notebook

A Szégyenfal: Amikor a kód találkozik a fizikával

	Tünet (Symptom)	Kiváltó Ok (Root Cause)	Iterációs Megoldás (Solution)
	A robot végtelenített ciklusban táncol.	A saját motorjának hangja túl hangos, elnyomja a leállító parancsot.	`'Stop dancing'` dedikált `stop-szó` betanítása és a mikrofon érzékenységének hangolása.
	Nem reagál a nevére.	Természetes beszéd közbeni “mormogás” miatt nem ismeri fel a `'Lil Token'` szót a szoba túloldaláról.	Különböző kiejtési minták és akcentusok betanítása a feloldáshoz.
	Eszközök (Tool Calling) hibás aktiválása.	Latencia és nehézség a `'beszélgetés'` vs. `'parancs'` megkülönböztetésében.	`'Teszt Mód'` bevezetése, és a `'Gondolkodás'` vizuális jelzése a képernyőn késleltetés alatt.

Gemini Notebook

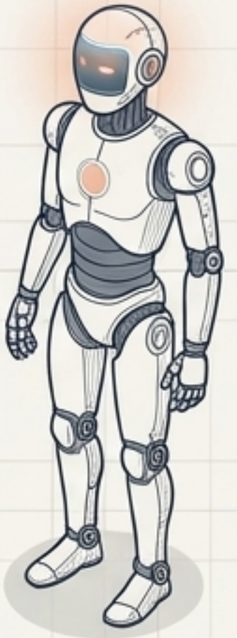
A Zseniális Trükk: Automatizált AI-AI Iteráció

Lépés 1: Az üres szobában a Mac szintetikus emberi hangon kérdéseket tesz fel a robotnak.



Szimulációs Mátrix
(Claude Core)

Lépés 2: A robot válaszol és eszközöket (tools) futtat.

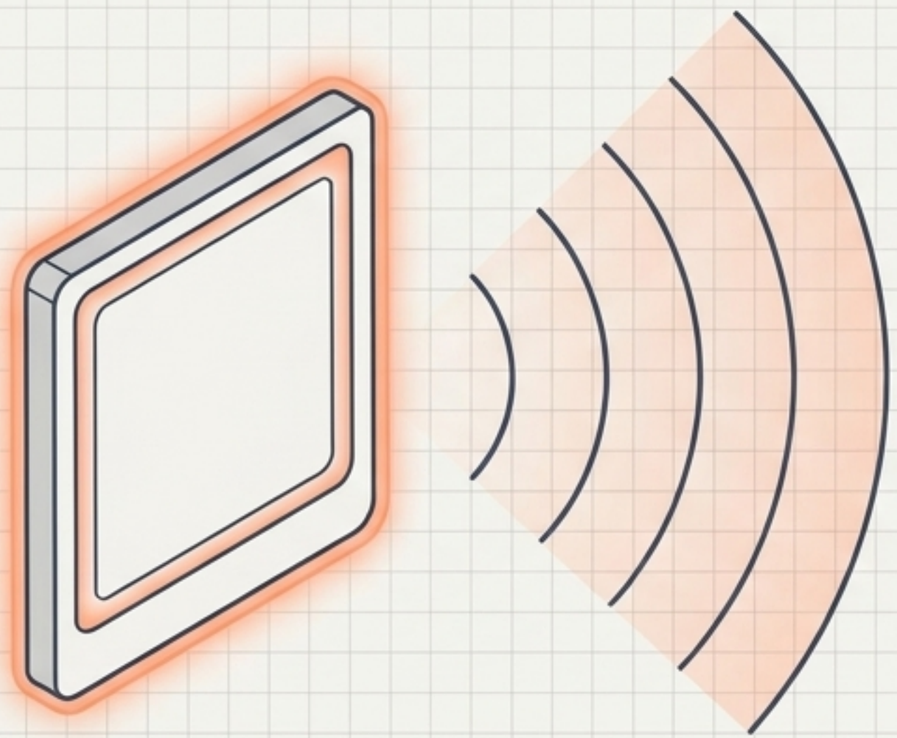


Lépés 3: A hurok közepén lévő Claude folyamatosan figyeli a futást, logolja a hibákat és iterál.

30+ szimulált tesztbeszélgetés futott le emberi beavatkozás nélkül, mire a bot stabilan működött.

Gemini Notebook

Térbeli Tudatosság és a 'Második Agy'



Audio

Zenei hullámok elemzése

Képes klasszifikálni a háttérzenét (pl. Jazz, Klasszikus zene), és az öltözékét/táncát kulturálisan a zenéhez igazítja.

Vision

Jelenlét-érzékelés

20 másodpercenként lokális kamera-ellenőrzést végez. Ha a felhasználó visszatér az asztalhoz, üdvözli és felolvassa a következő teendőt.

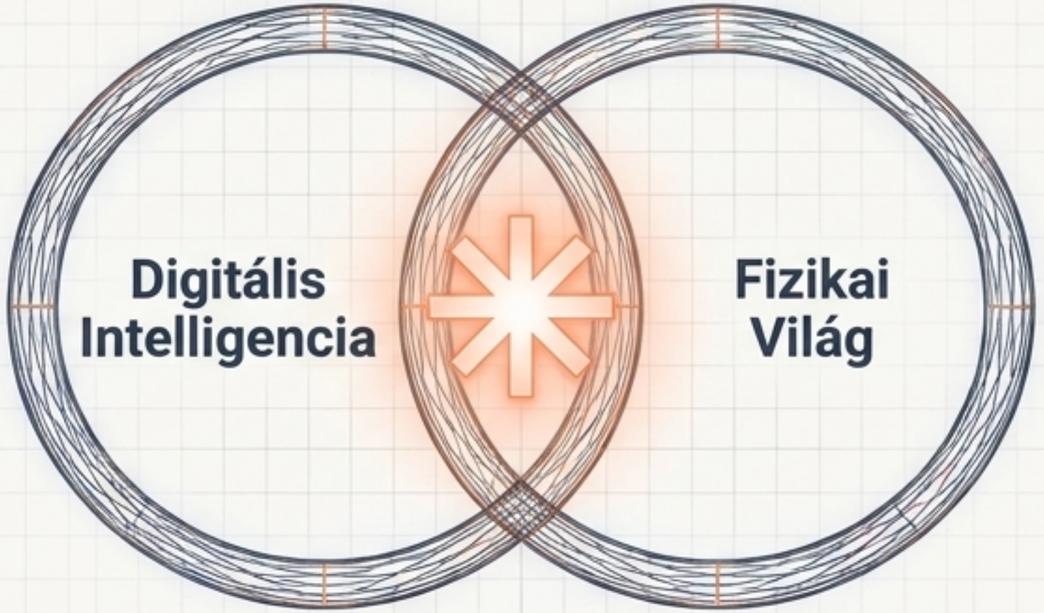
Memory

Felhő alapú 'Második Agy'

A leggyakoribb parancsokat és kontextusokat eltárolja, így a válaszidő (latencia) drasztikusan csökken.

Gemini Notebook

Szintézis: A Frontier Modellek Új Korszaka



A Claude Fable 5.2 áttörése nem a kódolás sebességében, hanem a kontextus megértésében rejlik. Képes elolvasni egy hardver specifikációját, áttervezni a firmware-t, és fizikai entitásként összekötni az AI-t a való világgal.

Az infrastruktúra nyílt forráskódúvá vált. A hardver programozása és a fizikai AI asszisztensek építése többé nem csak egy szűk mérnöki elit kiváltsága.

Forráskód és architektúra inspiráció: Mark Kash

Gemini Notebook