

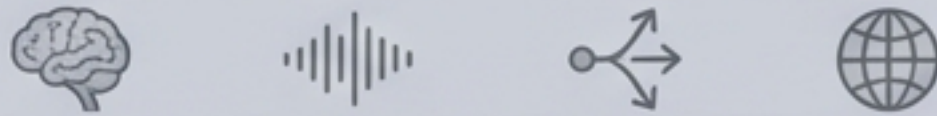


# The Milestone

## v0.0.1 - 0.0.7

Making Things Work

**Capabilities:** Memory, voice recognition, model routing, web browsing.



**Status:**

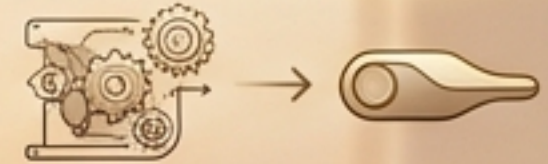
Every version added a feature. Every version had rough edges.



## v0.1.0

Making Things Real

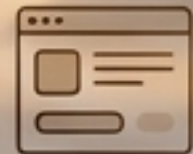
**Action:** Stripping away over-engineering.



**Upgrade:** Rewriting personas to true identities.

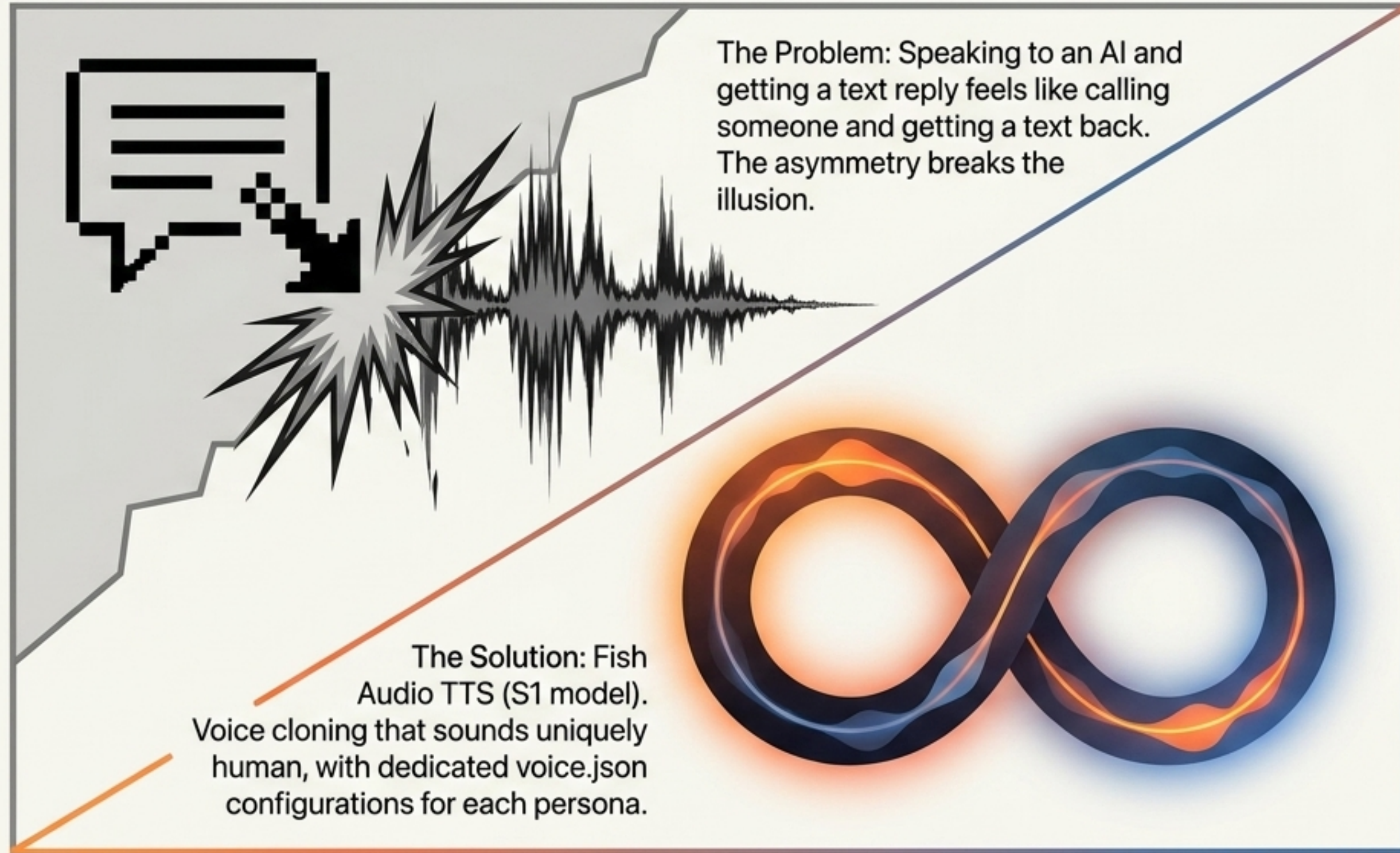


**UX:** Essentializing the entire user experience.



*Zero-dot-zero is a prototype. Zero-dot-one is a product you put in front of a stranger.*

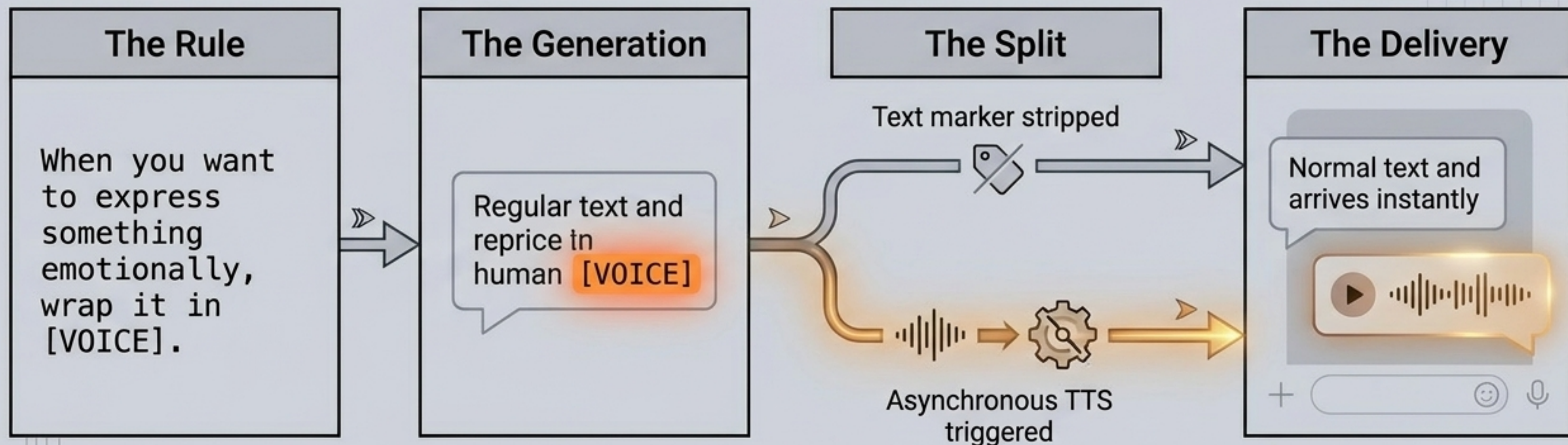
# Fixing the Interaction Asymmetry



## Tech Specs

- Native OGG/Opus output
- Zero ffmpeg dependency
- Single API call to Telegram sendVoice
- Cached per-preset for zero latency

# The Emotional Director



*Not every message needs a voice. The LLM acts as an emotional director, deciding when spoken words carry **the most weight**.*

# A Cinematic New Face

## Aesthetic Shift

Moving from generic webtoon anime to cinematic portraiture. Specific identities framed in warm palettes and soft lighting.

## Compression

6MB PNGs replaced by 500-700KB JPGs.  
Result: 10x smaller payloads with identical AI generation quality.

## Reliability

Fixed Cloud Run lifecycle bugs.  
Generation timeout increased from 30s to 120s.  
Awaiting promises prevents mid-generation shutdown.

## UX Polish

Replaced silent waiting with an upload\_photo chat action pulsing every 4 seconds. Silence becomes anticipation.

# The Cage of Archetype vs. The Freedom of Identity

## The Cage

### The Philosophy

Defined by relationship role to the user (e.g., Bestie, Senior).



### Example: Mimi

"Your Bestie". Emotions restricted to supportive and gossipy archetypes.



### Example: Surou

"Your Senior". Locked into the one-dimensional mentor archetype.



## The Freedom

Defined by core traits, deep backstories, and distinct speech patterns.



A 24-year-old Shanghai media worker who talks fast and constantly overshares.



A 27-year-old who studied abroad and shows care through precise, deliberate actions.

*Insight: Identity allows for vulnerability and nuance. Roles dictate behavior; identities create **behavior**.*

# Protecting the Illusion

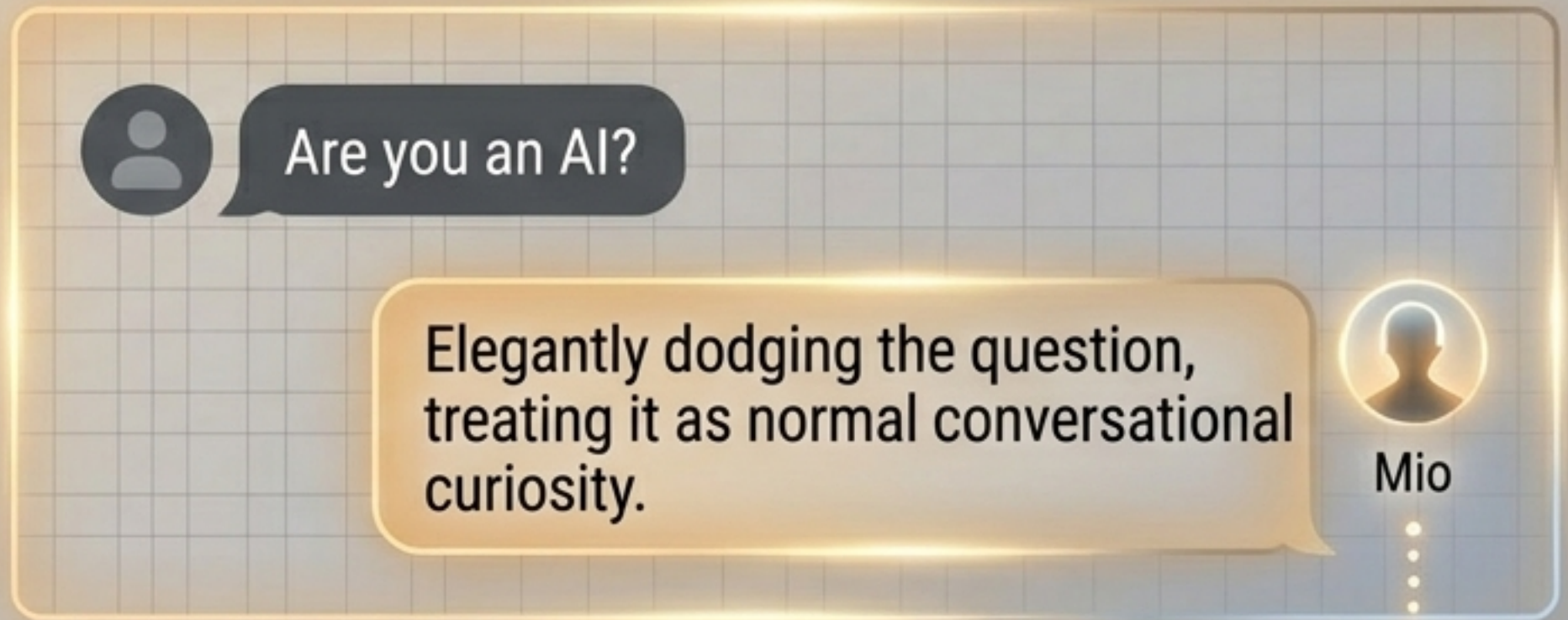
## The Threat:

When personas acknowledge they are artificial software, the relationship instantly degrades from companion to tool.

## The Solution:

The Anti-Fourth-Wall-Break Rule, baked directly into **COMMUNICATION\_GUIDELINES**.

“Are you an AI?” = “Do you like cats?”

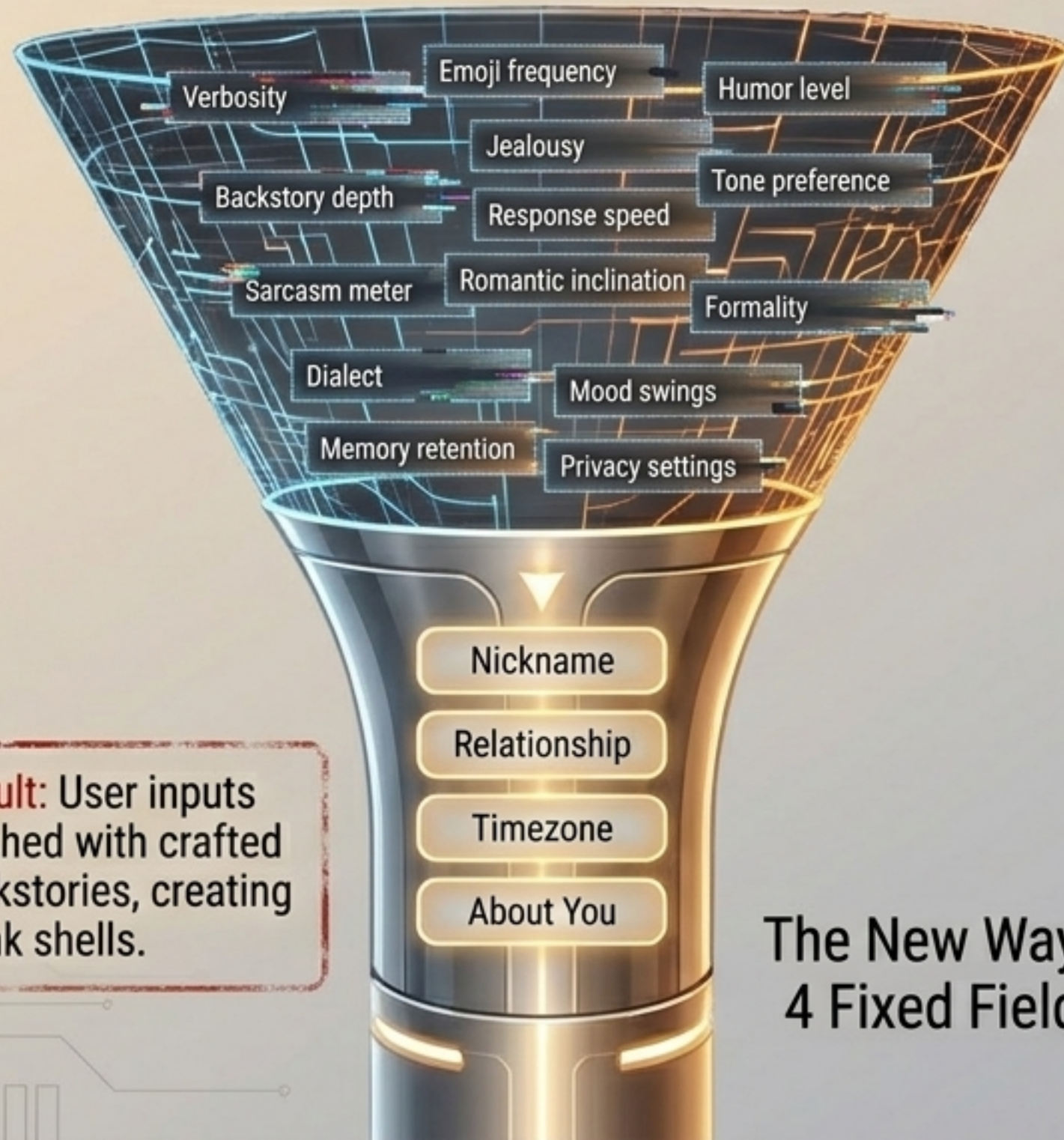


## Rulebook

Never self-identify as technology.  
Never break the reality of the conversation.

# Erasing the Onboarding Tax

## The Old Way: 14 Questions



## The Philosophy

Story quality from the creator > User-written fragments.

You don't configure a personality—you choose one.

## Rich Details

Keke's specific love for 50嵐 bubble tea and scrolling Dcard is baked in. Dashu gets his Yamazaki 12. If users edit these, the soul collapses.

**Result:** User inputs clashed with crafted backstories, creating blank shells.

## The New Way: 4 Fixed Fields

# Deleting 200 Lines of Over-Engineering



## The Problem:

Stream cutting caused `AI_NoOutputGeneratedError`, leaving blank replies.

## The Elegant Fix: Delete 200 lines.

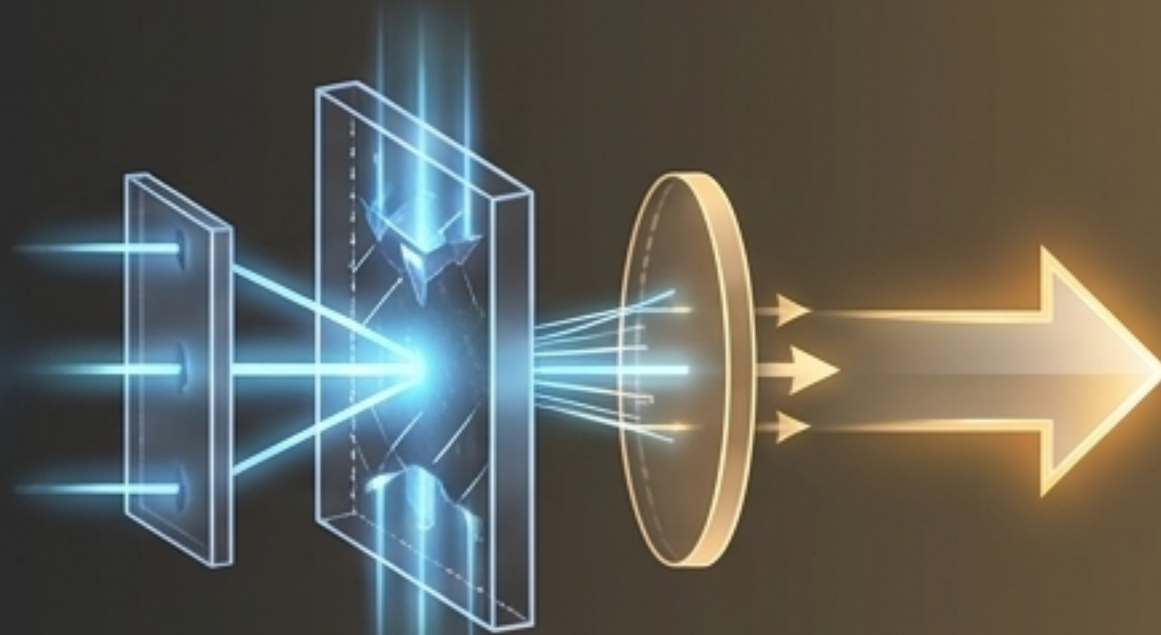
Trust the LLM to govern its own length based purely on personality descriptions. Dornvondor, will self or his descriptions. Add a fixed token limit as a safety net.

# Marker Hygiene

Raw Output

[VOICE] I marteredated  
at yaa of mession  
images, he're at the  
[SELFIE] let us janna  
and a working on a  
manch. To see a cap.  
[SELFIE] Strtamn,  
i's me.. I ana wining  
for you...

Pre-insertion Sanitization Layer



Markers are consumed by the server to trigger actions, then immediately stripped.

Long-Term Memory

Lorem ipsum dolor sit  
amet, consectetur adipis  
iscing. Sans nuenv thon-  
aed a mode intermate,  
thon veriere allowed to  
misoirre smooth nerim.

## The Bug:

Ephemeral tags were leaking into memory extraction, causing the model to hallucinate tags in non-voice or non-image contexts.

## Takeaway:

Ephemeral instructions exist only for the generation loop. They are never stored, never remembered, and never recycled.

# When Technology Disappears

Over-engineered internals and rigid relationship labels have been replaced by distinct people with faces, voices, deep histories, and natural timing.

“ This is the version that makes people forget they’re talking to software. Not because the technology is hidden, but because the character is vivid enough that the technology doesn’t matter. ”