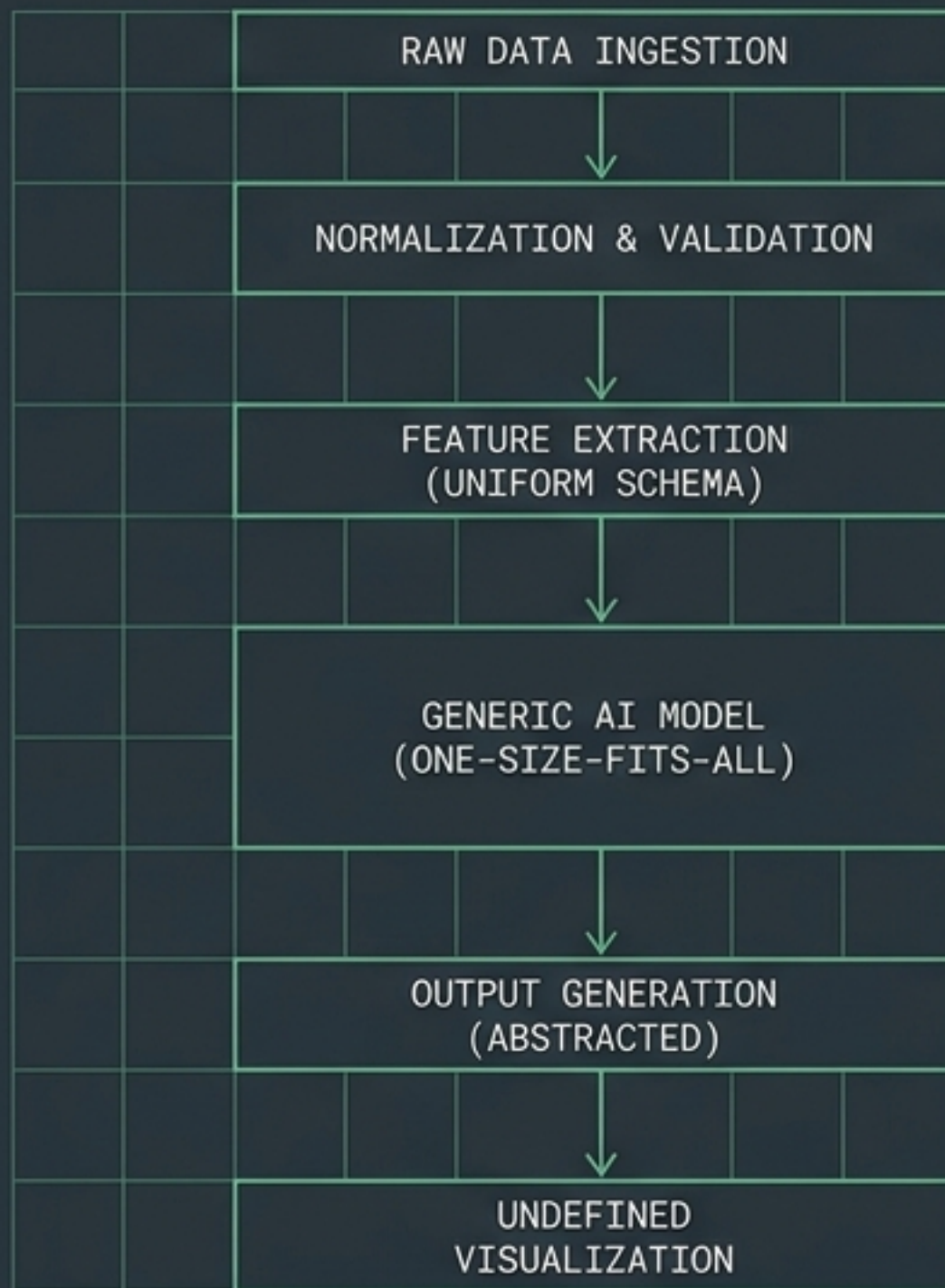
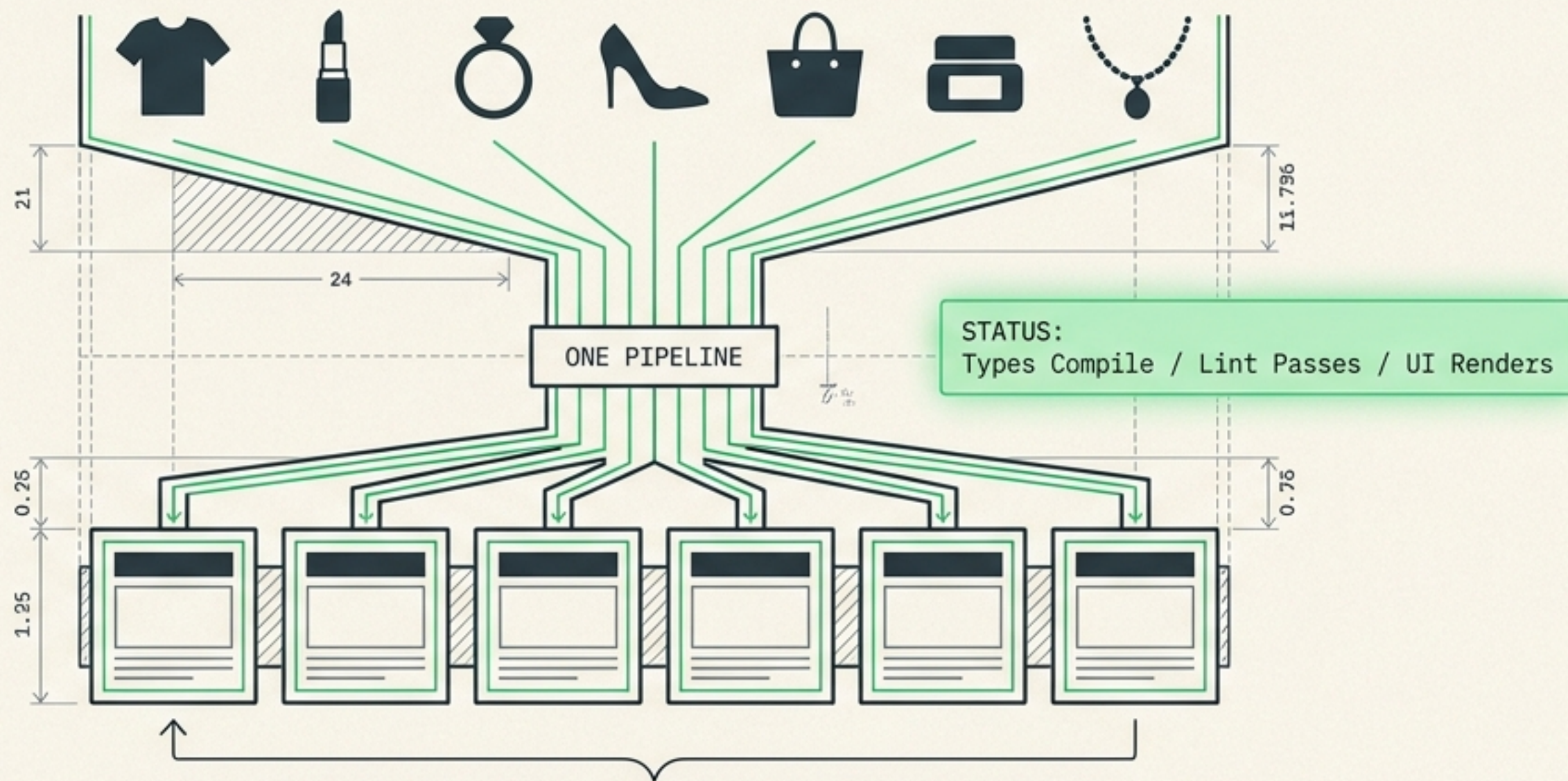


Walking Back the Abstraction



[Post-Mortem: Shichuan Part 3 / AI Generation Templates for E-Commerce]

A clean abstraction built a false sense of safety.



The premise: Six slots acting as persuasion angles (First impression, professional breakdown, visual identity, quality proof, lifestyle context, decision info).

The infrastructure was mathematically elegant. It ran flawlessly for two weeks. Then, a merchant uploaded a denim jacket.

The Denim Gemstone Bug.



[Upload: Denim Jacket]

+



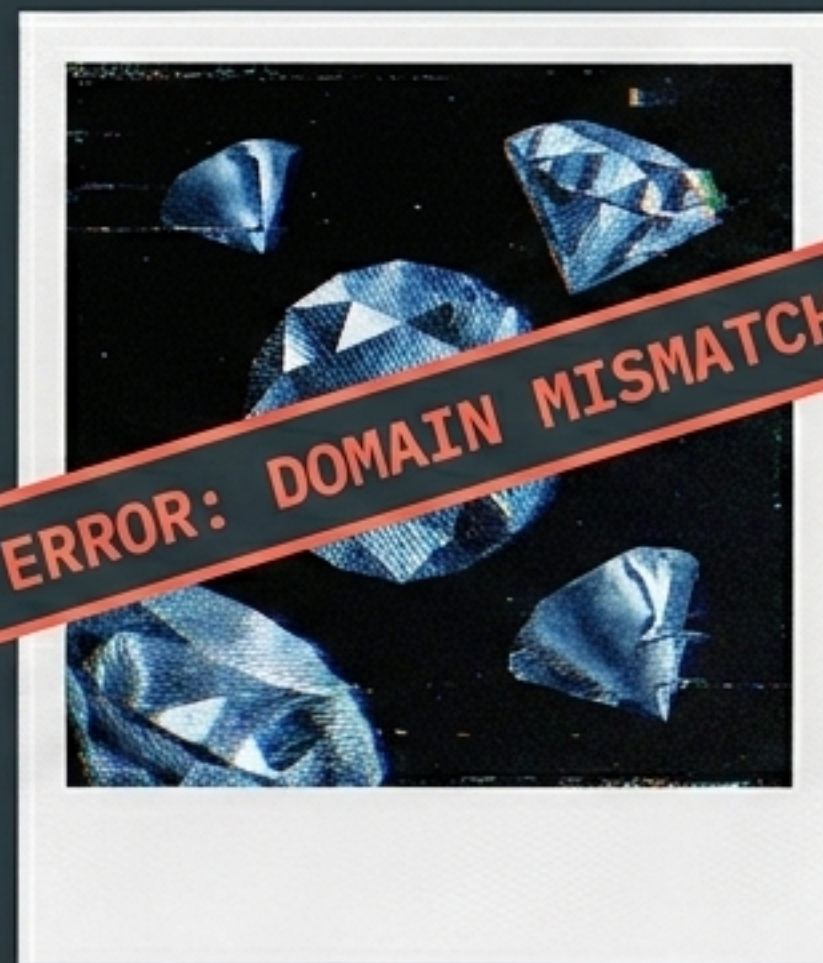
[UI Label]

×

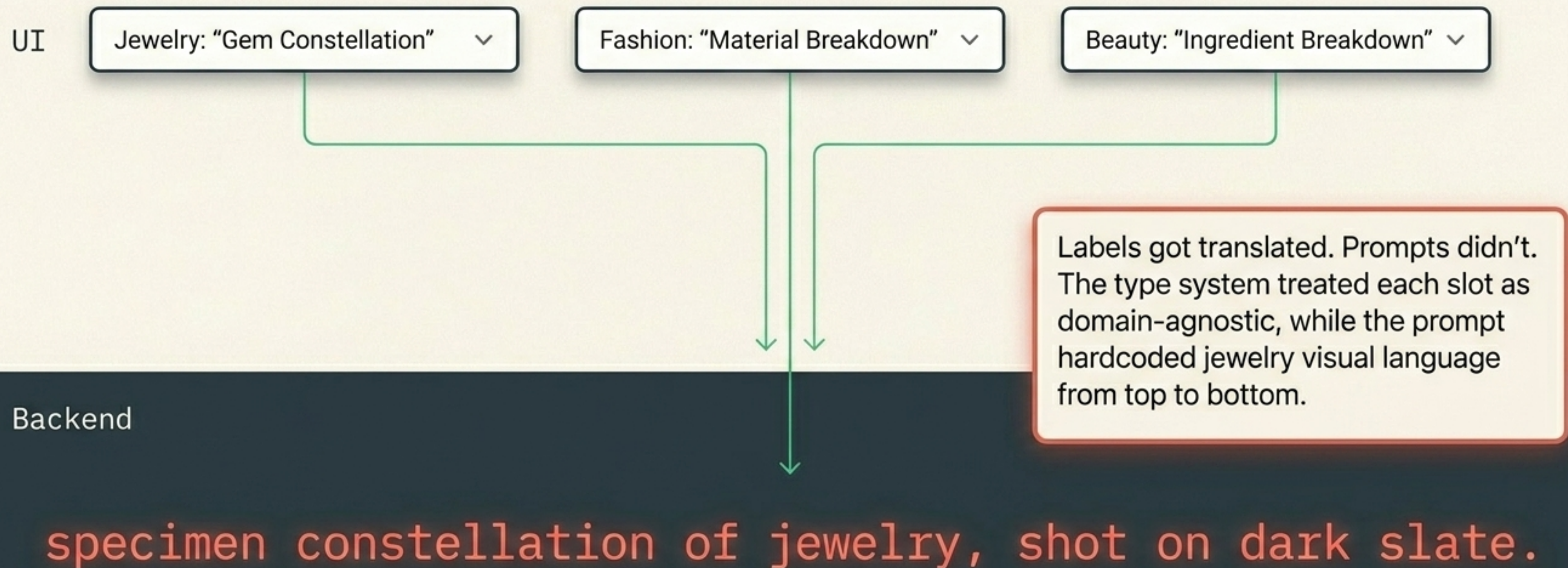


[Hardcoded Prompt]



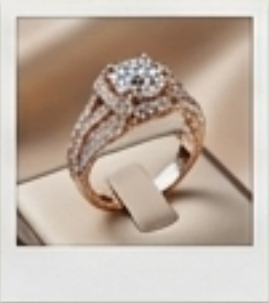
=



Abstraction stopped at the label layer.

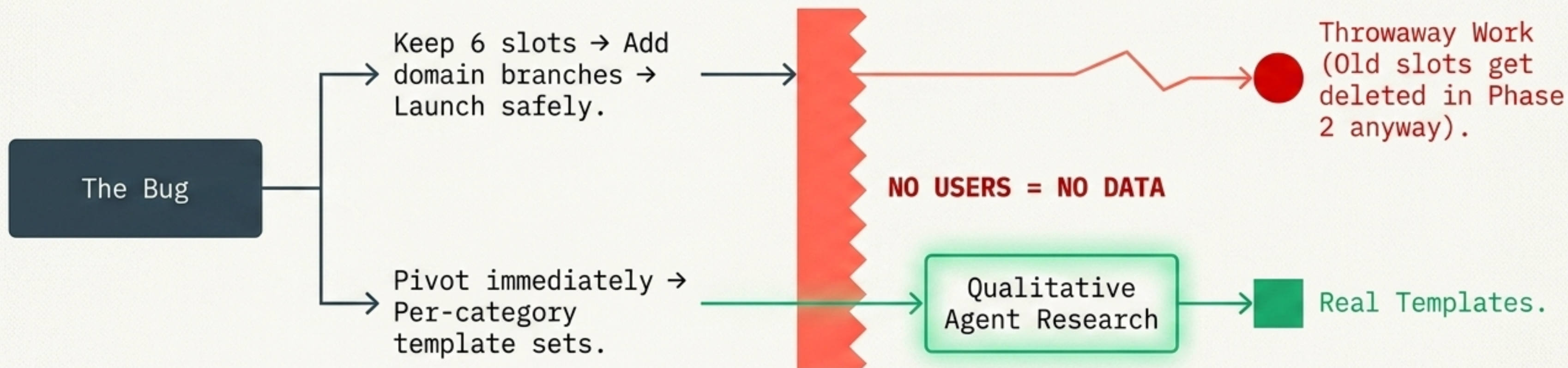


You cannot bridge visual grammar with a label swap.

Category	Top-Converting Format	Sizing Strategy
Fashion 	OOTD (Full-body model, 60%+ frame)	Text overlay on model ('163cm / 48kg wearing M')
Beauty 	Swatch Grid (5-in-1 composite)	Product next to applicator
Jewelry 	Studio Macro	Coin-scale comparison

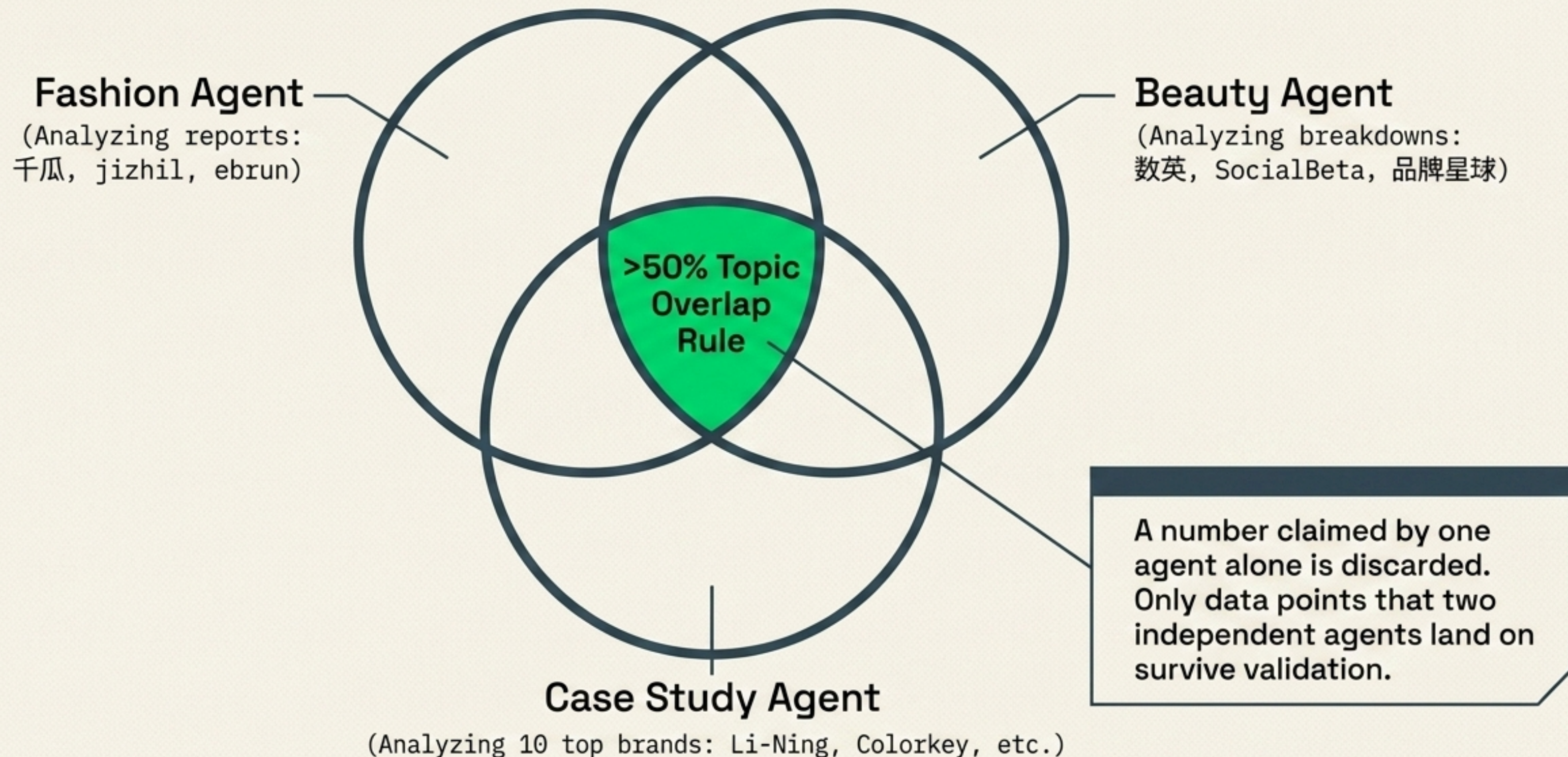
Patching a single prompt with domain branches fails. These are different images native to their specific Xiaohongshu vocabulary, not different labels on the same image.

Phased rollouts without data generate throwaway work.



Without data, Phase 1's guess quality equals Phase 2's guess quality.
Phased rollout without data isn't risk management.
It's a feeling of risk management.

Parallel agents required a strict cross-validation constraint.



The visual grammar proven by top-tier merchants



OOTD = 60-70% of fashion merchant content on Xiaohongshu



Li-Ning moved 300K units & ¥100M GMV using 'scene + product + model' composition



Outfit grids drive +50% repost rates (10 outfits x 1 denim jacket)



Swatch grids are the dominant format; Colorkey built a brand accounting for ~45% of output via this template



Skincare routine videos capture 51.32% of beauty category engagement

This killed the default assumption: ~~top-tier merchants rarely post standalone hero studio shots~~

Cutting templates requires harder evidence than picking them.

Rejected (Reasonable but Unproven)

~~Stitch Detail / Craft Macro~~

(Only high-end designers use it; underperforms for 90%).

~~Height-grid model shots~~

(Blogger niche, merchants lack multi-model budgets).

~~Outfit moodboard collage~~

(Blogger identity tool, not a merchant sales tool).

~~3-Column Ingredient Infographic~~

(AI garbles diagrams; merged into Benefits Card).

The Rule

The Hard Constraint:
Exactly six templates per category.

The Litmus Test: If research can't surface at least one top-tier merchant using the template in a proven viral post, it does not deserve a slot.

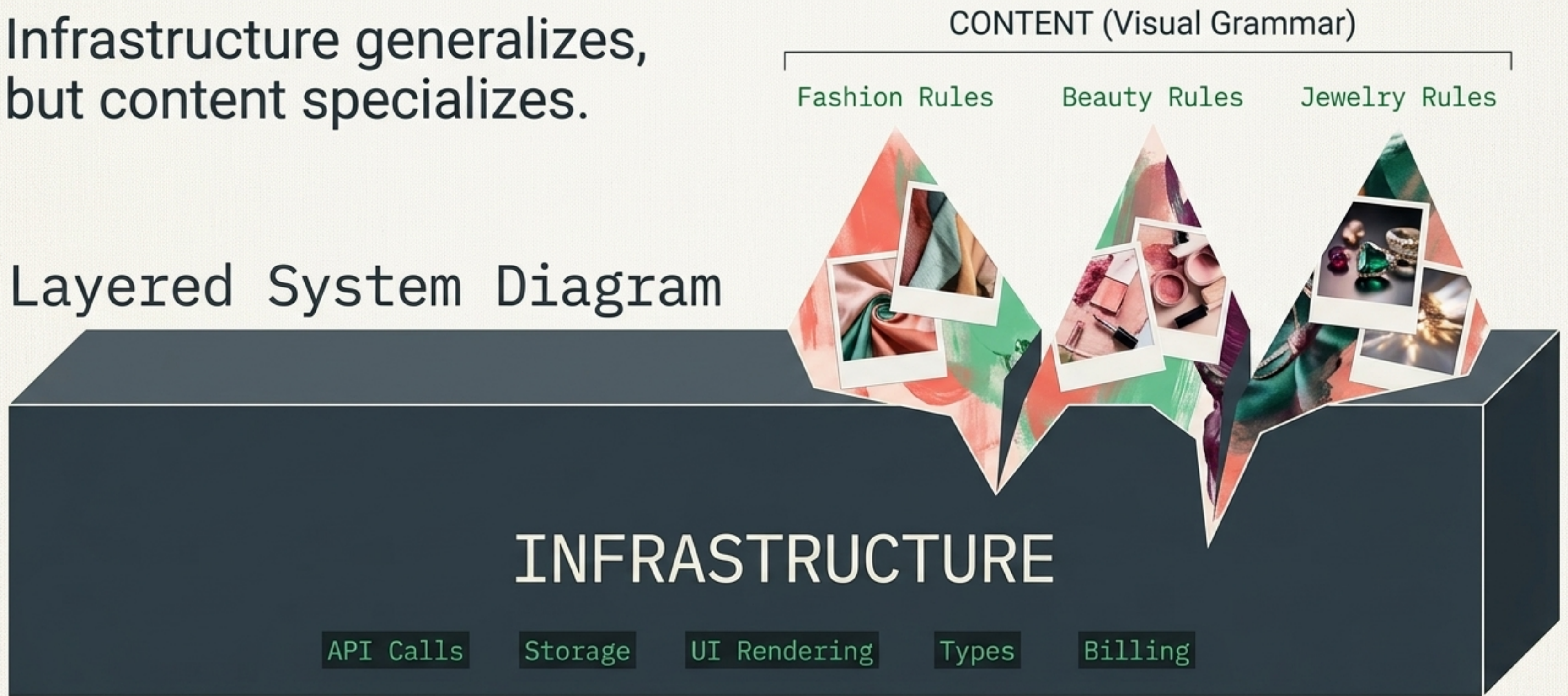
The final 3x6 domain-specific architecture.

FASHION	BEAUTY	JEWELRY
1. OOTD try-on	1. Hero packshot	Unchanged from the original iteration (Maintained for craft items).
2. Outfit grid (2x2)	2. Swatch grid (5-in-1)	
3. Scene shot	3. Before/after	
4. Flatlay + on-body	4. Benefits card (bold headline)	
5. Before/after (+5cm taller)	5. Routine sequence (AM/PM)	
6. Sizing overlay	6. Collection grid (red/black lists)	

**The infrastructure stays clean: three categories, same pipeline.
What varies is the prompt layer's native vocabulary.**

Infrastructure generalizes,
but content specializes.

Layered System Diagram



Abstraction can save you from writing the same code three times.
It cannot save you from writing the same prompt three times.

Iteration as self-reversal.

“Every three or four weeks, I end up reversing a judgment I made the previous sprint. That’s not a sign the work is going badly. It’s the sign it’s going correctly.”

Remaining Unknowns

1. 6+6+6 is qualitative inference; A/B generation signals are needed next.
2. The fallback maps to jewelry; will break when food/home goods are added.
3. TypeScript compiling \neq correct prompt output; UAT and smoke tests are still mandatory.