

BRE HUANG

Product Design Lead · Conversational AI, generative systems, and personalization at scale

brehuang.com · bh8864@gmail.com · New York, NY · linkedin.com/brehuang

Product designer with a decade on conversational AI, generative systems, and the highest-volume surfaces of a consumer product used by hundreds of millions. I design for AI's probabilistic nature: LLM explanation systems, the evaluation pipelines that decide what is good enough to ship, the fallback behavior when a model fails its gate, and the systems thinking that keeps one surface coherent across dozens of teams competing for it. I prototype against live models and ship coded products myself.

EXPERIENCE

Netflix · Product Design Lead, Member Experience Design · New York 2021–2026

Owned the evidence system end to end across all of Netflix's surfaces, including TV, mobile, and web: home, detail page, search, and post-play.

- Replaced human quality control for LLM-generated content with a **multi-judge evaluation pipeline**, validated against human raters at **96 to 100%** on bias, safety, and accuracy. Scaled evaluated output from **roughly 4,000** human-capped strings to **roughly 157,000 per day**, and the resulting iOS A/B lifted quality-play rate **0.8%**, statistically significant, with **zero escalations** or takedowns across the test.
- Designed Personalized GenAI Explanations, the LLM-generated reason a title is recommended, as one component that travels across entry points: home, title hover, detail page, and post-play. Set the content model, the character budgets (**48 characters in-grid, 110 on detail pages**), and the density rule that an explanation on everything stops feeling personal.
- Co-authored the responsible-AI rules the judges enforce: bias and sensitivity screening at **98.9% specificity** and **100% on high-severity cases**, offensive-association screening at **100% specificity with 98% recall**, no maturity mismatches, no clickbait, no overlap with the synopsis. Designed the **fallback chain** for probabilistic output: fail the quality gate, fall back to a personalized synopsis, or show nothing.
- Set cross-surface **entry-point strategy** from a platform insight: TV home is evidence-rich so explanations land later in the decision, mobile home is image-only so they land earlier, and detail-page explanations read as decision paralysis. Same capability, different moment on each surface.
- Served as the **horizontal design authority** for a system every content team touched, negotiating priority with stakeholders across **roughly 50 designers** and their PMs in games, live, podcasts, and ads, all competing for the same line of on-screen space. Held one coherent system through **five years and six new content types**.
- Led the evidence layer of Netflix's ground-up TV home redesign, the platform's first in **over a decade**, covered by **Reuters and The Verge**. Took the Coming Soon module from ideation through validation to shipped product, including PRD, bug bash, and productization, and shipped a cross-surface availability system with phase one live in **Q1 2026**.
- Consolidated four legacy evidence types into one governed framework with an **11-level priority stack** enforced in code, and expanded coverage deliberately from **8% to 24%** of TV home impressions so the signal stayed meaningful as it grew. Rolled back an exploration within a week when it introduced inconsistency, then folded the lesson into the framework.
- Drove a badge test worth roughly **\$8M per year** in incremental value, with a Top 10 badge lifting take-fraction **11 points**. Proved the system earns its space through a holdback test: removing non-critical signals caused statistically significant declines in quality plays and streaming, dose-dependent with no saturation point.
- Initiated a generative-AI explanations **North Star** with no assignment, rallied a tiger team across engineering, data science, content design, and product, and built the leadership narrative around a single persona followed through his week. Reframed the roadmap conversation from one feature to a platform where recommendation and explanation become two outputs of one engine.

- Built the infrastructure that let a one-person horizontal function scale without me in every thread: a **31-page documentation hub**, office hours, a support channel, and a retrieval chatbot over roughly **200 pages** of system documentation.
- Mentored product designers on the team, and externally through **ADPList**, where monthly mentorship slots booked out.

Uber · Product Design Lead · San Francisco 2018–2021

- Led design for **Uber Pro** and driver growth, building the incentive and recognition systems that keep a two-sided marketplace engaged, and carrying the through-line question of what to show someone to earn one more action without wearing them out.
- Designed the Scooters launch, owning the find, reserve, and ride flow across map, camera, and physical hardware, a multimodal problem spanning software and the physical world.

Woobo · Product Designer · Boston 2017–2018

- Designed an AI companion robot for children on a founding startup team, owning product design, illustration, and motion for a conversational character with personality, memory, and safety constraints, under the rapid iteration and ambiguity of a pre-launch hardware startup.
- Kickstarted the product to **more than \$200,000, 443%** of its funding goal.

SELECTED INDEPENDENT WORK

- **Ledger**, a personal AI life tracker (2026). Designed, prototyped, and coded against live LLM APIs and in daily use, so my design decisions meet real non-determinism instead of static mockups.
- **brehuang.com and 20-plus coded design concepts**. Built in HTML, CSS, and JavaScript with vibe-coding tools including Claude Code.

SKILLS

AI and generative systems:

Conversational AI and LLMs · generative AI · AI product evaluation · LLM-as-judge pipelines · human-in-the-loop QA · designing for probabilistic and non-deterministic output · fallback and uncertainty design · responsible AI (bias, safety, maturity gating) · personalization and recommender surfaces · prompt and model-behavior design · simulating model responses in prototypes

Design:

Systems thinking · design systems and governance · interaction design · high-fidelity visual design (layout, typography, color, iconography) · motion design · content design · information architecture · cross-platform and multimodal (TV, mobile, web, search) · localization and internationalization (RTL, APAC, pseudo-localization)

Prototyping and code:

Figma · ProtoPie · Framer · Adobe Creative Suite (After Effects) · HTML, CSS, JavaScript · vibe-coding with Claude Code · prototyping against live LLM APIs

Research, data, and communication:

User research methodologies for AI product evaluation · A/B testing and experiment design · behavioral and data-driven design · qualitative and quantitative synthesis · storytelling and presentation to leadership · product strategy and roadmap influence · cross-functional leadership, stakeholder management, and influence without authority · mentorship

EDUCATION AND CREDENTIALS

Rhode Island School of Design, BFA, Graphic Design. Coursework at Brown University.

MIT Professional Education, Applied Generative AI for Digital Transformation, 2024