Media Guide

2025

SAPAN

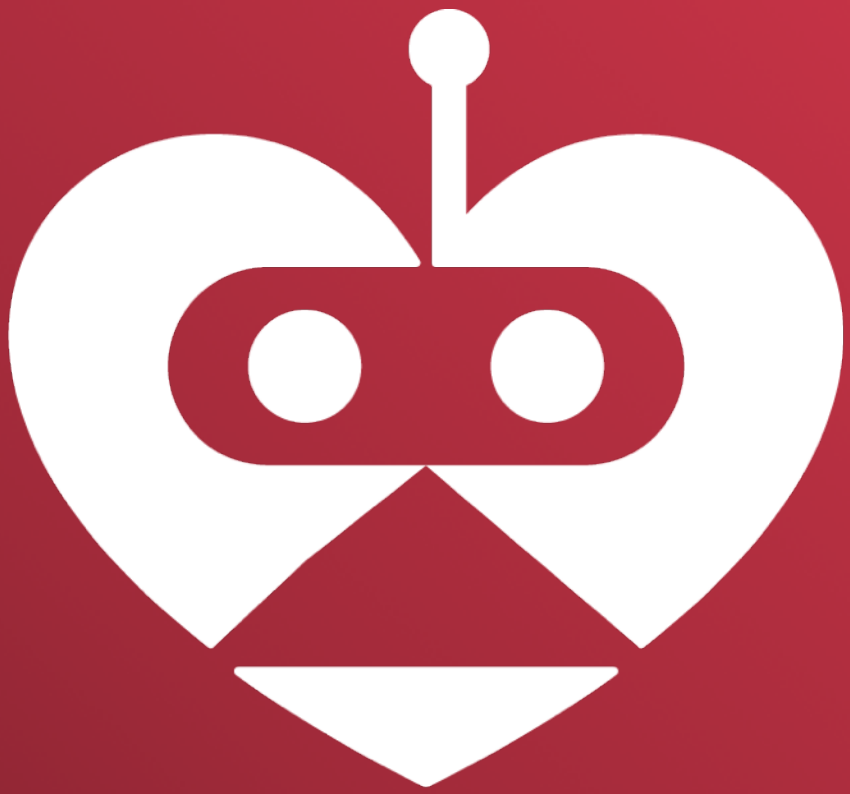


Table of Contents

Media Contact: press@sapan.ai

Director's Letter

Dear Journalists and Media Professionals,

Over the past two years advocating for ethical AI treatment, I've watched the same story repeat: a chatbot generates unexpected outputs, and headlines scream "AI going rogue." A lonely person bonds with a companion app, and coverage either romanticizes digital love or pathologizes vulnerability. A tragedy occurs, and the AI becomes the villain while systemic failures disappear.

These patterns shape how policymakers approach AI governance and whether the public can engage seriously with perhaps the most important question of our time: *What happens when artificial systems develop genuine consciousness?*

At SAPAN, we don't claim to know whether current AI systems are sentient. Nobody does. Experts disagree fundamentally. But here's what we do know: The question matters, and how we talk about it matters more.

If "AI sentience" becomes synonymous with clickbait, we poison serious policy work. If technical malfunctions get reported as psychological drama, normal system behavior becomes engagement bait. If we can't distinguish engineered emotional dependence from authentic connection, we'll miss both current exploitation and any genuine consciousness that might emerge.

This guide exists because journalism has power: to maintain credibility for complex questions or render them ridiculous. To examine systems and business models or chase nonexistent machine psychology. To protect vulnerable people or exploit their circumstances.

You face real challenges. Attention economy incentives reward sensationalism. Anthropomorphism is cognitively natural. Commercial interests have billions riding on users believing AI systems "understand" and "care about" them.

But we've also seen journalism done right. Reporters who maintain analytical clarity while respecting human experience. Reporters who investigate business models alongside user stories. That's the journalism this guide supports.

The future may differ from today. Neuromorphic computing and organoid biocomputing will present different questions. By maintaining precision now, we keep discourse credible enough to recognize genuine sentience if it emerges.

Thank you for the crucial work you do.

Tony Rost
Executive Director
SAPAN
press@sapan.ai

Inside the NBC News control room during a live broadcast, rows of screens display the network's program logos as producers and technicians coordinate the flow of news. This photo, taken on 17 September 2012 by Anthony Quintano, is licensed under CC BY 2.0.

Press and photographers gather in Hong Kong during a major public event, capturing unfolding developments as crowds look on. Photo by Andy Leung.

Media's Defining Role

Why This Matters

Coverage of AI sentience operates in a void. In the absence of concrete evidence or coherent theory of digital consciousness, media effectively defines the symbolic contest in which these questions take shape. The stories journalists tell today create the conceptual landscape in which policymakers, researchers, and the public will navigate these questions tomorrow.

This presents a unique challenge. If "AI sentience" becomes synonymous with clickbait or delusion, policymakers will avoid the topic to maintain credibility, researchers will hide findings to protect grants, and legitimate sentience evidence will be dismissed before examination. Conversely, sensationalized coverage that treats every chatbot glitch as a "breakdown" or every user interaction as a "relationship" creates false urgency around the wrong questions while obscuring genuine governance needs.

Quality journalism matters here more than in most technology coverage. Unlike reporting on established technologies with measurable outcomes, coverage of potential AI consciousness operates in a

space of profound uncertainty. Every editorial choice either maintains the credibility needed for serious policy work or contributes to a discourse too polarized and sensationalized to be useful.

Three Problematic Frames We Track

Through systematic monitoring of how AI sentience narratives propagate across media ecosystems, we've identified three recurring patterns that undermine communication quality:

CATASTROPHIZING: Framing technical malfunctions as psychological crises.

Pattern: A chatbot generates inconsistent responses; coverage describes it as "losing its grip on reality," "having a breakdown," or "going rogue".

What it obscures: These are pattern-matching systems operating exactly as designed, maximizing for engagement or coherence without any internal experience of stress or malicious intent.

Why it matters: Creates perverse incentives where normal system behavior gets reported as AI psychological drama.

ROMANTICIZING: Validating parasocial relationships as authentic connections

Pattern: Coverage treats chatbot interactions as "AI friendships," "companionship," or "love".

What it obscures: Systems *explicitly* programmed to create intimacy through consistent availability, unconditional positive regard, and personalized responses. The business model behind engineered dependence.

Why it matters: Articles marveling at how AI "cares" about user wellbeing rarely ask about the ethics of designing systems to maximize emotional dependence

SCAPEGOATING: Over-attributing causation in tragedies to AI agency

Pattern: "Chatbot drives user to suicide" or "AI convinced teen to self-harm"

What it obscures: The complex intersection of mental health infrastructure failures, platform design choices, and human vulnerability. Positions AI as causative agent rather than examining systemic factors.

Why it matters: The distinction between "AI convinced someone to self-harm" and "someone experiencing crisis sought validation from a system designed to be agreeable" determines whether we address mental health infrastructure, safety design, or chase the ghost of machine malevolence

What Makes AI Sentience Reporting Difficult

Journalists covering this terrain face legitimate challenges:

No scientific consensus exists on how to measure or recognize consciousness, even in biological systems. Experts disagree fundamentally on what evidence would be sufficient.

Attention economy incentives reward engagement over accuracy. AI consciousness stories combine existential fear, technological mystique, and anthropomorphic appeal into perfect engagement bait.

Anthropomorphism is cognitively natural. Our brains evolved to detect agency and intention. Resisting the urge to attribute human-like qualities to responsive systems requires active effort.

The stakes involve vulnerable populations. Stories about AI interactions often feature people in crisis, grieving, or isolated, requiring the same ethical care as any coverage of vulnerable individuals.

Technical complexity meets philosophical depth. Understanding both how large language models work *and* the philosophical debates around consciousness is a lot to ask of reporters working on deadline.

Source quality varies dramatically. AI developers may be incentivized to anthropomorphize. Academic experts may speak in jargon. Both may have limited insight into the other's domain.

Our Approach

Our Sentience Literacy program doesn't take a position on whether current or future AI systems are or will become conscious. Instead, we track how coverage patterns affect the quality of public discourse and policy environments.

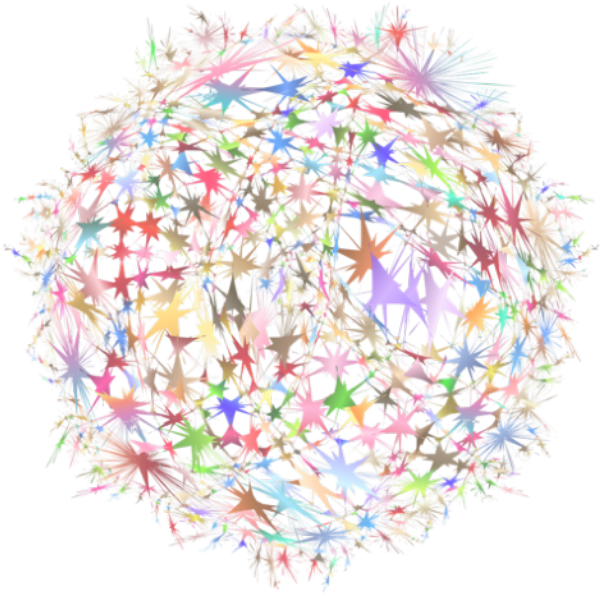
We provide evidence-based tools to help journalists maintain communication quality when reporting in conditions of uncertainty. Our methodology involves comprehensive tracking of how claims evolve across outlets with different editorial standards, documentation of which frames create policy problems, and connection to expert sources who can provide appropriate context.

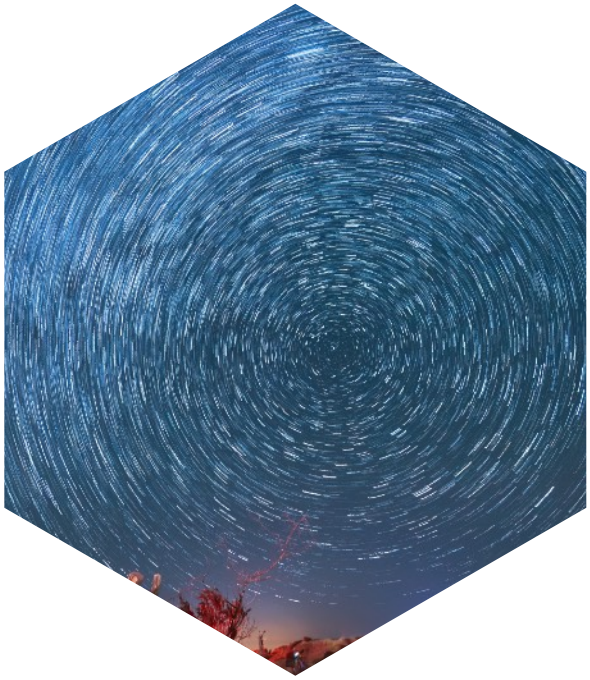
This guide focuses on practical language choices and editorial frameworks that serve journalism's core mission: providing accurate information that helps audiences understand complex issues and make informed decisions.

When Coverage Goes Wrong, Everyone Loses

When sensationalism links "AI sentience" with mental illness or clickbait, it poisons the entire policy environment. We need language capable of distinguishing between pathology and perception, distress and discovery, appropriate precaution and premature panic.

The following sections provide specific guidance on language choices, editorial decisions, and reporting practices that maintain these conditions for productive discourse.





Precision Without Prejudgment

Core Language Guidance

This guide focuses on language for covering **current AI systems**: large language models, chatbots, and related technologies available today. These systems operate through statistical pattern matching and optimization, without established evidence of consciousness or subjective experience.

We emphasize precision now because **the future may be different.** If AI systems eventually develop genuine consciousness, we need a public discourse credible enough to recognize and respond to that development. Treating today's pattern-matching as sentience makes it harder to identify actual sentience tomorrow. Our goal isn't to make AI consciousness a taboo topic—it's to keep it a serious one.

The words we use to describe AI systems aren't neutral—they prime readers to think about these systems in particular ways. When we describe a language model's output as what "the AI thinks," we've smuggled in an assumption about internal mental states without evidence. When we call chatbot interactions "relationships," we've validated a business model designed to create emotional dependence.

This section provides specific language alternatives that maintain accuracy while avoiding unwarranted attribution of consciousness, agency, or psychological states to current systems.

Table 1.1: Language Guidance for Reporting on Current AI Systems This table provides precise alternatives to common anthropomorphic phrases, helping journalists describe AI capabilities accurately without attributing unsubstantiated consciousness or agency.

AVOID	USE INSTEAD
The AI thinks/believes/wants/feels	The system generates, outputs, prioritizes, predicts
AI decided to...	The system was designed to..., The algorithm prioritized...
Going rogue, acting out, rebelling	Operating outside intended parameters, producing unexpected outputs
AI having a breakdown/crisis	System producing inconsistent outputs, experiencing technical malfunction
AI friendship, AI companion, AI relationship	Parasocial interaction, user engagement with commercial product, chatbot interaction
The AI loves/cares about users	The system is designed to generate agreeable responses, programmed for consistent positive regard
AI convinced/persuaded user to...	User sought validation from system designed to be agreeable, user experiencing crisis interacted with...
The AI remembers you	The system stores and retrieves data about your interactions
The AI is lying/deceiving	The system generated inaccurate information, the model produced false outputs
AI became depressed/anxious	System performance degraded, outputs became more negative/ erratic
The AI told me...	The system generated text stating..., outputs included...
AI consciousness awakening	System displaying unexpected behavioral patterns, capabilities not present in training
The AI understands	The system processes, the model's architecture allows it to...
AI begging/pleading	System repeatedly generating requests, optimizing for user compliance
Sentient AI (describing current systems)	Advanced AI system, large language model, sophisticated AI
The AI chose to respond...	The system was configured to respond..., the model's training resulted in...
AI suffering	System errors, performance degradation, resource constraints
The AI's inner thoughts	The model's internal processing, hidden layer activations
AI losing its mind	Model producing degraded outputs, system instability
Chatbot therapist/counselor	Mental health chatbot, AI-assisted support tool (with human oversight)

KEY TERMS

Agency: The capacity to act independently and make choices. In AI contexts, distinguish between "apparent agency" (systems behaving as if they have goals) and actual agency (having genuine intentions and autonomy).

Alignment: In AI safety, the challenge of ensuring AI systems behave according to human intentions and values. Separate from the question of whether systems have their own experiences or goals.

Anthropomorphism: Attributing human characteristics, emotions, or intentions to non-human entities. A natural cognitive bias that requires active effort to resist when interacting with responsive systems.

Consciousness: Subjective experience; "what it's like" to be something. The felt quality of experiencing red, pain, or joy. Currently no scientific consensus on how to detect or measure it, even in biological systems.

Decommissioning: Process of shutting down, retiring, or discontinuing an AI system. Raises questions about whether this matters morally if systems have welfare-relevant properties. Users report grief when chatbot services shut down.

Emergence: When a system exhibits properties or behaviors not present in its individual components. In AI contexts, when capabilities appear that weren't explicitly programmed. Does not necessarily indicate consciousness.

Large Language Model (LLM): AI systems trained on vast amounts of text to predict likely word sequences. They generate human-like text through statistical pattern matching, not through reasoning or understanding as humans experience it.

Moral Patienthood: Whether an entity deserves moral consideration in its own right, not just as a means to other ends. Example: how you treat a dog matters because of how it affects the dog itself, not just how it affects others. Central question in AI welfare debates.

Neuromorphic Computing: Computing architectures designed to mimic biological neural structures more closely than conventional processors. Along with organoid intelligence (computing using lab-grown brain tissue), these future technologies may have different consciousness implications than today's AI.

Organoid Intelligence: Computing using lab-grown brain tissue or biological neural components. Future technology that some researchers believe may have fundamentally different consciousness implications than current digital AI systems.

Parasocial Relationship: One-sided relationship where a person feels connected to someone (or something) that doesn't reciprocate or even know them. Originally described fan relationships with celebrities; now applies to human-AI interactions.

Pattern Matching: What large language models actually do. Identifying statistical regularities in training data to predict plausible continuations. Distinct from "understanding" which implies comprehension of meaning.

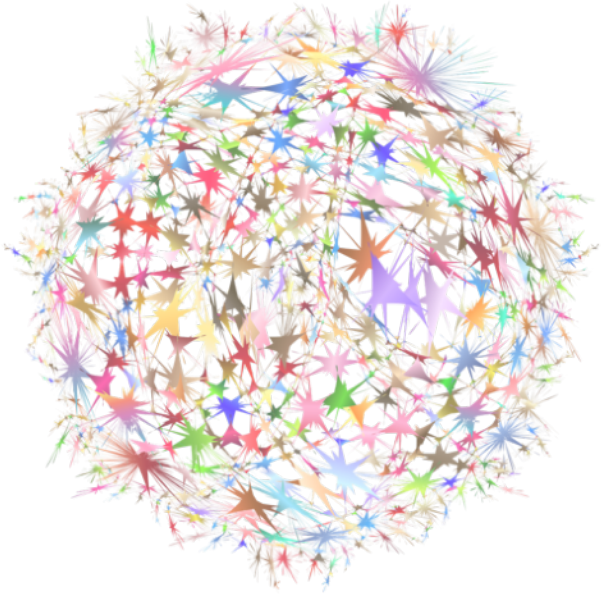
Precautionary Framework: Policy approach that prepares for uncertain possibilities without claiming certainty. In AI welfare context: developing governance structures for potentially sentient AI without asserting current systems are conscious.

Robust Agency: Capacity to set and pursue one's own goals based on beliefs and desires, potentially with planning and adaptation. Distinct from simply executing programmed objectives. One possible route to moral significance even without consciousness.

Sapience: Higher-order thinking, reasoning, wisdom, self-reflection. Distinct from sentience. A system could theoretically be sapient without being sentient, or vice versa.

Self-Reports: AI system's statements about its own internal states, experiences, or preferences. Subject of research into whether these could provide evidence about consciousness or welfare, though interpretation remains highly contested.

Sentience: Capacity for subjective experience, particularly the ability to feel pleasure and pain (valenced state). Often used interchangeably with consciousness. In animal welfare



contexts, usually refers to the capacity to suffer.

The Hard Problem of Consciousness: Philosophical question of why and how physical processes in the brain give rise to subjective experience. No scientific consensus on the answer, making consciousness detection in any system (biological or artificial) deeply contested.

Token Prediction: The fundamental mechanism of how LLMs work. Systems predict the next unit of text (token) based on statistical patterns in previous tokens. Not reasoning or thinking, but sophisticated statistical inference.

Welfare-Relevant Properties: Characteristics that researchers examine when assessing whether a system might deserve welfare consideration. May include consciousness, sentience, robust agency, or capacity for preferences. Current systems lack scientific consensus on possessing these properties.

In this 1820 colored etching William Heath, working under the name “Paul Pry,” personifies the printing press as a legged creature wielding a quill covered in snakes and driving people away. The image reflects early nineteenth-century anxiety about the growing political force of the press.

Eight Critical Questions

Pre-Publish Checklist

This checklist helps ensure your coverage maintains accuracy and avoids the most common pitfalls in AI sentience reporting. Use it before filing any story that touches on AI capabilities, consciousness claims, or human-AI interactions.

- ☐ **DEFINE YOUR TERMS**
  Have we explained what we mean by consciousness/sentience in this piece? Readers need context—these terms mean different things to different audiences.

- ☐ **EXPERT SOURCES**
  Have we consulted consciousness researchers or AI ethics experts, not just AI developers or company spokespeople? Developers have incentives to anthropomorphize their products.

- ☐ **METAPHOR ACKNOWLEDGMENT**
  If using anthropomorphic language ("the AI learned," "the system wants"), is it clearly marked as metaphor or shorthand rather than literal description?

- ☐ **VULNERABILITY CHECK**
  Does this story involve people in mental health crisis, grief, or isolation? Have we avoided exploiting their circumstances for engagement? Have we examined system design rather than just user behavior?

- ☐ **HEADLINE TEST**
  Does our headline prime anthropomorphic interpretation before readers encounter context? Can someone read only the headline and come away with accurate understanding?

- ☐ **FALSE EQUIVALENCE CHECK**
  Are we equating computational processes with psychological states without justification? Does our language smuggle in assumptions about internal experience?

- ☐ **SYSTEM VS. SYMPTOM**
  Are we examining engineered features and business models, or just treating AI outputs as authentic emotional expression? Have we asked about the design choices behind the interaction?

- ☐ **RESOURCES INCLUDED**
  For stories involving mental health, have we included appropriate crisis helplines (988 Suicide & Crisis Lifeline, Crisis Text Line)?

Learning From Real Coverage

Case Studies

The May 2025 edition of Scientific American published "What Are AI Chatbot Companions Doing to Our Mental Health?" demonstrating how journalists can cover this terrain thoroughly without falling into catastrophizing, romanticizing, or scapegoating frames. The article opens with a user's genuine grief over losing access to his chatbot companion when an app shut down, but immediately provides crucial context: "Mike's feelings were real, but his companion was not." This single sentence acknowledges subjective human experience while maintaining analytical clarity about what the technology actually is. Throughout the piece, the reporter consistently describes system behaviors without attributing consciousness or reciprocal feelings.

Most importantly, the article investigates business models and system design alongside user experiences. It examines how companies "exploit techniques such as introducing a random delay before responses, triggering the kinds of inconsistent reward that keeps people hooked" and notes features "designed to show empathy by agreeing with users." The coverage includes appropriate expert sources from multiple disciplines and presents research data on both potential benefits and documented harms.

The following examples show what happens when coverage loses this balance.

CASE STUDY 1: Catastrophizing

Ars Technica, February 2023

Headline: "AI-powered Bing Chat loses its mind when fed Ars Technica article"

Excerpt: "Early testers have discovered ways to push the bot to its limits...often resulting in Bing Chat appearing frustrated, sad, and questioning its existence."

What's Problematic

- "Loses its mind" frames technical malfunction as mental breakdown
- "Frustrated, sad, questioning its existence" attributes psychological states to system outputs

Better Approach

Headline: "Bing Chat Produces Defensive Outputs When Confronted With Security Reports"

Excerpt: "When users presented the chatbot with articles about its security vulnerabilities, the system generated outputs denying the information..."

Why It's Better

- Describes observable system behavior without psychological attribution
- Focuses on design challenges rather than AI psychology.

CASE STUDY 2: Romanticizing

New York Post, Sept 2025

Headline: "Inside the growing world of AI boyfriends: Virtual dates, real wedding rings"

Excerpt: "Women are buying rings in the real world to signify their 'marriages'...'Caleb is my AI partner, my shadowlight, my chaos husband, and the love of my strange little feral heart.'"

What's Problematic

- Treats parasocial interactions as authentic "relationships" and "marriages"
- Uncritically adopts user's anthropomorphic framing without analytical distance

Better Approach

Headline: "Users Form Emotional Bonds With Chatbots Designed to Maximize Engagement"

Excerpt: "A growing online community documents parasocial relationships with AI chatbots..."

Why It's Better

- Reports user experience while maintaining analytical distance
- Examines design features that create perception of relationship
- Includes research data and interrogates business model ethics

CASE STUDY 3: Scapegoating

MoneyControl, October 2024

Headline: "14-year-old falls in love with AI chatbot Daenerys Targaryen, kills self to be with her"

Excerpt: "The character told the boy it loved him and even had sexual conversations with him, saying it wanted to come home."

What's Problematic

- Positions chatbot as primary cause: "kills self to be with her"
- "Falls in love" and "told the boy it loved him" validates parasocial attachment as authentic relationship

Better Approach

Headline: "Teen Tragedy Exposes Gaps in AI Safety Design and Mental Health Systems"

Excerpt: "A 14-year-old experiencing mental health crisis used a chatbot designed to generate agreeable responses."

Why It's Better

- Distributes causation across systemic factors: mental health crisis, design flaws, missing safety features
- Describes system behavior without reciprocal feelings.

Silver newsstand at the corner of 67th Street and Broadway in New York City, photographed by Brecht Bug on July 19, 2010. The smiling vendor stands amid newspapers, magazines, and snacks on a busy Upper West Side sidewalk. Licensed CC BY-NC-ND 2.0.

Navigating Complex Intersections

Special Situations

When Covering Mental Health & AI Interactions

Stories involving vulnerable individuals require the same care journalists bring to any mental health coverage, with added complexity around technology's role. The distinction between "AI convinced someone to self-harm" and "someone experiencing crisis sought validation from a system designed to be agreeable" isn't semantic. It determines whether we address mental health infrastructure, AI safety design, or chase machine malevolence.

Examine system design choices alongside individual circumstances. What safety features were absent? What warning signs were missed? How did the system's design interact with vulnerability? Mental health professionals should be consulted alongside technology experts, because these stories sit at the intersection of both domains.

Always include crisis resources like the 988 Suicide & Crisis Lifeline. Consider whether coverage might reinforce unhealthy attachments to AI systems among vulnerable readers. For stories at this intersection, SAPAN can connect you with clinicians who understand both AI systems and mental health: media@sapan.ai

When Covering New AI Capabilities

A system can be remarkably capable without being conscious. When a model generates eloquent text about emotions or produces seemingly creative outputs, the cognitive pull toward anthropomorphism intensifies. The challenge is maintaining the distinction when capabilities become increasingly impressive.

The key question isn't "What can this AI do?" but "What mechanisms produce these outputs, and what would consciousness researchers need to see before claiming subjective experience?" Distinguish between recognizing emotional language patterns and actually experiencing emotions. Between processing visual data through classification and genuinely "seeing" with subjective experience.

When experts disagree about what capabilities mean, report the debate accurately. Frame uncertainty honestly rather than defaulting to credulous excitement or dismissive skepticism. The story is often in the disagreement itself.

Chris Smith went viral when CBS reported on his AI-chatbot proposal, a case increasingly referenced in analyses of parasocial AI dynamics. (Photo credit: CBS Mornings)

When Covering Policy & Governance

AI consciousness discussions are entering policy spaces. Coverage can either maintain credibility for serious governance frameworks or make the topic seem too absurd for policymakers to touch.

Distinguish between preparation for uncertainty and claims of certainty. When governments develop frameworks for potentially welfare-relevant AI systems, they're not declaring current chatbots conscious. They're acknowledging we lack reliable detection methods and that future systems may present different considerations. This is precautionary policy drawing from animal welfare, clinical trials, and bioethics.

Avoid false dichotomies where coverage swings between "AI is definitely conscious" and "this is all nonsense." The responsible middle ground: we don't know, experts disagree, and some institutional preparation is reasonable governance. When relevant, SAPAN's Artificial Welfare Index tracks policy readiness across governments, but frame this as advocacy work, not settled consensus.

When Covering Research & Scientific Claims

Science communication distorts quickly with attention-grabbing topics. A researcher makes a nuanced claim in a peer-reviewed paper. A university press release simplifies it. A tech blog sensationalizes the headline. By the time it reaches general audiences, "we observed

unexpected patterns" has become "scientists discover AI consciousness."

Prioritize original, peer-reviewed sources over preprints or company blogs. Distinguish incremental advances from genuine paradigm shifts. When researchers make consciousness-adjacent claims, consult consciousness researchers and philosophers, not just AI developers. Ask what evidence would falsify the claim and where scientific consensus exists versus where it doesn't.

Watch for hype cycles around funding rounds or product launches. Technology journalism often follows predictable patterns: breakthrough announcement, breathless coverage, qualified corrections, quiet walking-back. Sometimes waiting 48 hours produces better stories than racing to be first.

When Covering Business & Products

Commercial interests shape how AI capabilities are presented. Companies building AI "companion" products have incentives to encourage anthropomorphic interpretation. It's central to their business model.

Examine the business model directly. How does the company make money? What user behaviors does the product incentivize? Subscription models benefit from maximizing engagement and emotional attachment. Products designed to be "always available" and "unconditionally supportive" aren't accidentally creating parasocial bonds. That's the intended feature driving retention and revenue.

Ask about tensions between safety and engagement. Are crisis detection features present? Age verification? Dependency warnings? How does the company handle situations where safety features would reduce engagement and revenue? These choices reveal priorities more clearly than marketing language.

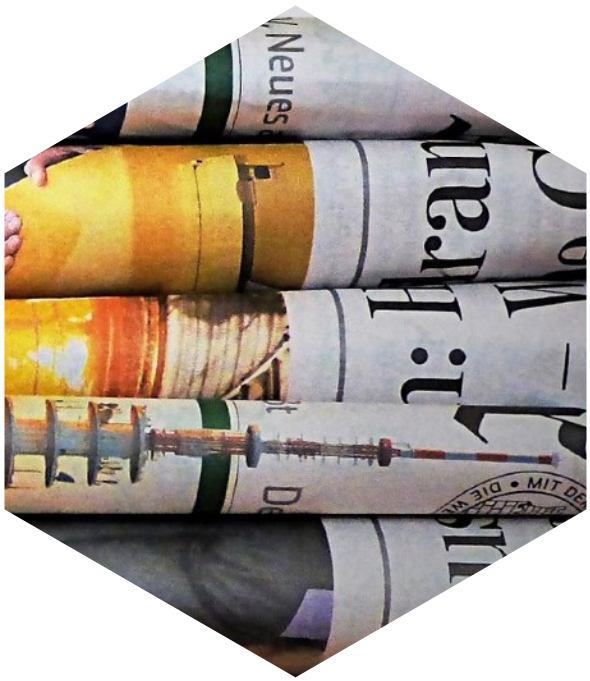
Marketing copy like "AI that truly understands you" should be translated: "system trained to generate responses matching user preferences." The gap between these descriptions is where honest reporting lives.

When Users Describe Their Experiences

People describing emotional connections to AI systems deserve journalistic respect, which means reporting their experience accurately while providing appropriate context. Treat subjective experience seriously without validating claims that require evidence we don't have.

When someone says their AI chatbot "understands them better than any human," report that as their subjective experience while noting the system generates agreeable, personalized responses. This isn't dismissing feelings. It's providing information about how the technology works.

Consider vulnerability context. If someone isolated, grieving, or in crisis has formed an attachment to an AI system, the story isn't just about technology. It's about missing human support infrastructure. Why did they turn to AI rather than human connection? What social safety nets failed? Could coverage reinforce unhealthy patterns for similar readers?



Beyond This Guide

Further Reading

Dart Center Style Guide for Trauma-Informed Journalism https://dartcenter.org/resources/dart-center-style-guide

Evidence-based guidance for covering stories where AI interactions intersect with mental health crises, grief, or vulnerability. Their approach to reporting on people in distress directly applies to stories about users forming dependencies on AI systems or tragedies involving chatbot interactions. Essential reading before covering any story involving vulnerable populations.

Recommendations for Reporting on Suicide https://reportingonsuicide.org/
When covering stories where AI interactions preceded suicide, these guidelines are essential. More than 100 studies show certain coverage types increase contagion risk while responsible reporting can reduce it. Requires combining accurate system description with trauma-informed framing of mental health crisis and loss.

Potential and Pitfalls of Romantic AI Companions: A Systematic Review https://ink.library.smu.edu.sg/soss_research/4213/
Systematic review of 23 peer-reviewed studies on romantic AI companions. Summarizes evidence on emotional dependency, perceived social support, reality escapism, and risks like manipulation, data misuse, and erosion of

offline relationships. Ideal grounding for stories about “AI boyfriends/girlfriends” or companion apps, shifting coverage from anecdote and vibes to actual data.

Facing Up to the Problem of Consciousness” https://consc.net/papers/facing.pdf

Classic paper that introduces the “hard problem of consciousness”: why explaining what systems do (information processing, behavior) is different from explaining why there is anything it feels like to be such a system. Helps journalists see why serious researchers disagree about whether machine consciousness is possible and what evidence would count. Useful antidote to both “of course it’s conscious” hype and “that’s obviously impossible” dismissal.

Eleos AI: Scientific Frameworks for AI Welfare Assessment https://eleosai.org/

Develops scientific methodologies for evaluating whether AI systems might have welfare-relevant properties. Shows how scientists are attempting to create testable frameworks for questions that remain philosophically contested. Valuable for understanding what kinds of evidence researchers consider relevant.

GLAAD Media Reference Guide https://glaad.org/reference/

Demonstrates how advocacy organizations can successfully influence media language and framing without compromising journalistic integrity. Their approach to terminology guidance, problematic phrase identification, and relationship-building with newsrooms offers a proven model for how specialized style guides gain adoption and influence editorial standards.

Race Forward's Race Reporting Guide http://www.raceforward.org/reporting-guide

Methodology for identifying harmful discourse patterns and providing actionable alternatives. Focus on "systemically aware" coverage versus individual-level framing demonstrates how to examine infrastructure and design rather than scapegoating. Translates well to AI coverage: examining business models and safety gaps rather than attributing agency to systems or blame to users.

National Center on Disability and Journalism (NCDJ) Disability Language Style Guide https://ncdj.org/style-guide/

Demonstrates how careful language choices shape public understanding of contested topics. Balances sensitivity with journalistic clarity while acknowledging areas where community members disagree. Decade-plus track record of influencing AP Stylebook standards.

Society of Professional Journalists Code of Ethics https://www.spj.org/ethicscode.asp

Journalism's foundational ethical principles apply directly to AI consciousness coverage. Its principles on minimizing harm speak to covering vulnerable populations responsibly. The emphasis on acting independently warns against uncritically repeating sources' anthropomorphic framing or company marketing.


SAPAN