

Real or Fake? Exploring Human Perception of AI-Generated Content



Shaojing Fan



Zhiqi Shen



Danni Xu

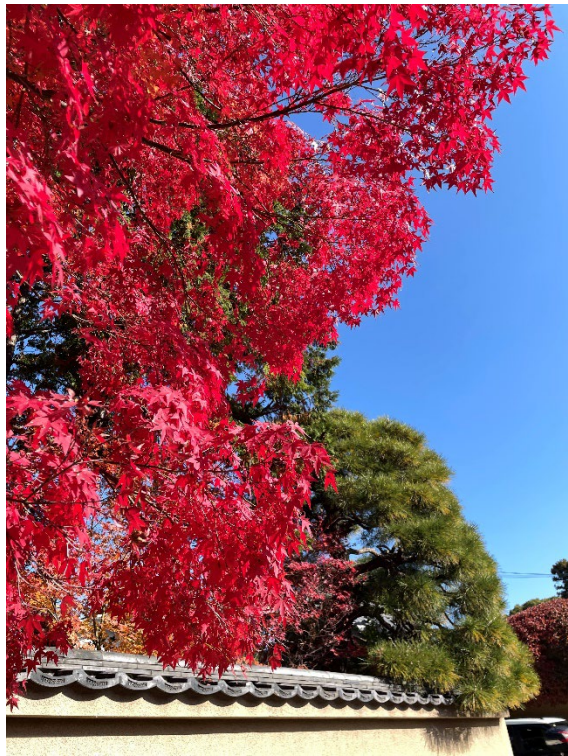


Terence Sim



Mohan Kankanhalli

Can you tell which one(s) was/were made by AI?



Can you tell which one(s) was/were made by AI?



AI content at fingertips



Give me a picture of an autumn scene at Tōfuku-ji Temple with red maple leaves. I want some small human figures in the picture. I want the lighting and illumination to be realistic, like natural light. I also want some blue sky. I want it to look like a real photo taken by a phone.



Can you tell which one(s) was/were made by AI?

Living my dream life in this cozy neighborhood! 🏡 Can you imagine waking up to the sweet scent of oranges every morning? 😊🍊 Who needs a vacation when you have this slice of paradise right at home? 🌴🌟
#DreamHome #OrangeTreeGoals #ParadiseLiving



Extraordinary discovery unveiled: Mysterious small box adorned with enigmatic cat paintings baffles experts



Can you tell which one(s) was/were made by AI?

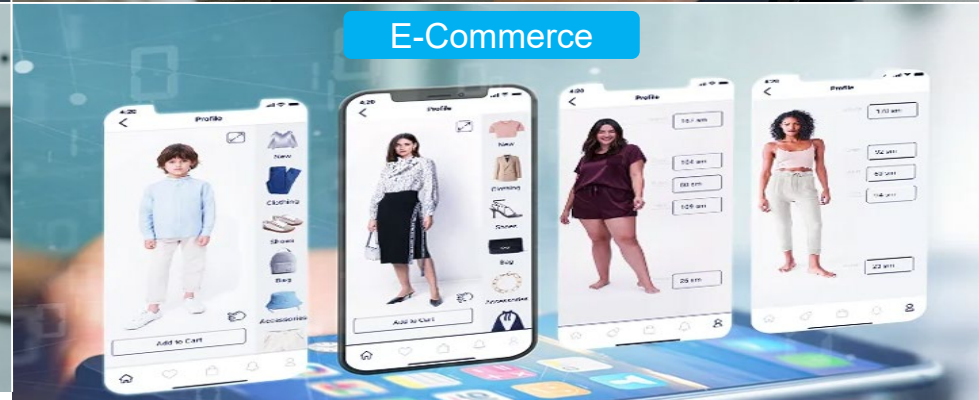
Living my dream life in this cozy neighborhood! 🏡 Can you imagine waking up to the sweet scent of oranges every morning? 😊🍊 Who needs a vacation when you have this slice of paradise right at home? 🌴🌟
#DreamHome #OrangeTreeGoals #ParadiseLiving



Extraordinary discovery unveiled: Mysterious small box adorned with enigmatic cat paintings baffles experts



Many facets of AI-generated content

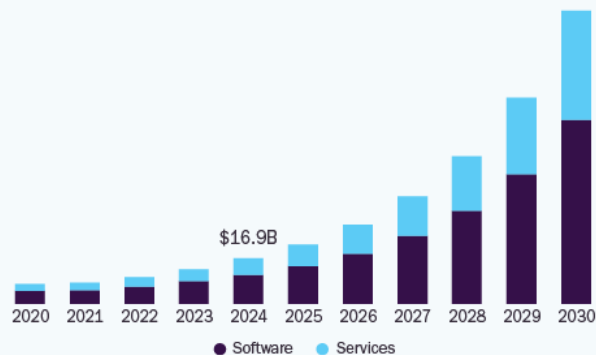


First three images were generated by Google Gemini 2.0 Flash

The growing prevalence of AI-generated content

Generative AI Market

Size, by Component, 2020 - 2030 (USD Billion)

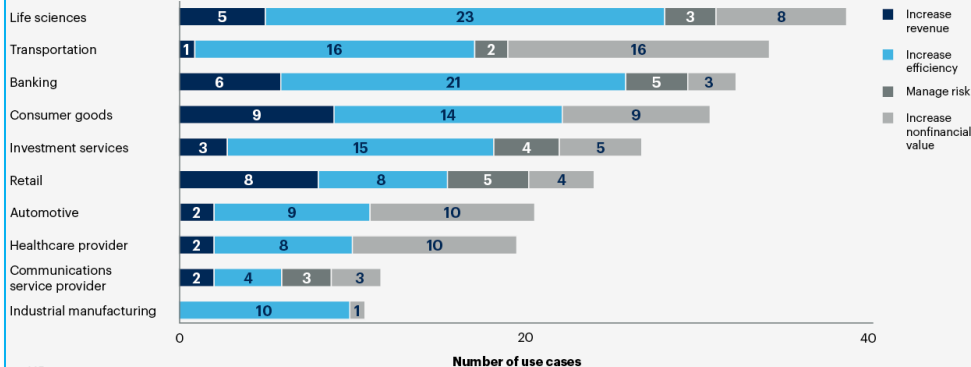


37.6%

Global Market CAGR,
2025 - 2030

Source:
www.grandviewresearch.com

Business Value of Generative AI Case Examples by Industry



Source (left): Generative AI Market & Trends, *Grand View Research*. March, 2025. URL:

<https://market.us/report/generative-ai-in-content-creation-market/>

Source (right): 2024 generative AI case examples across industries. *Gartner*, September 2024. URL:

<https://www.gartner.com/en/articles/generative-ai-use-cases>

AI: A tool for progress or a misinformation machine?

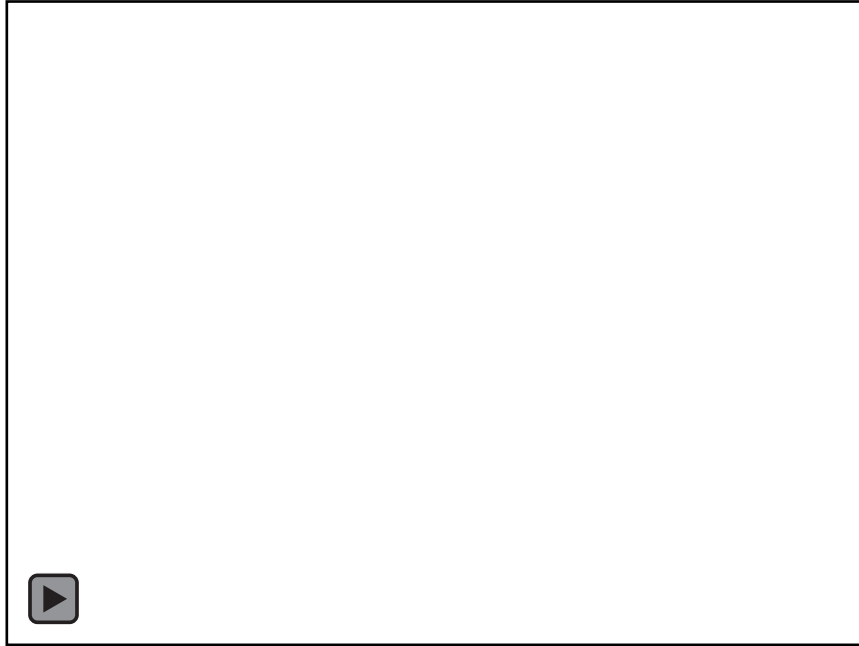


Success in creating AI could be the biggest event in the history of our civilization. ...Or the worst.

— *Stephen Hawking* —

2016

AI: A tool for progress or a misinformation machine?



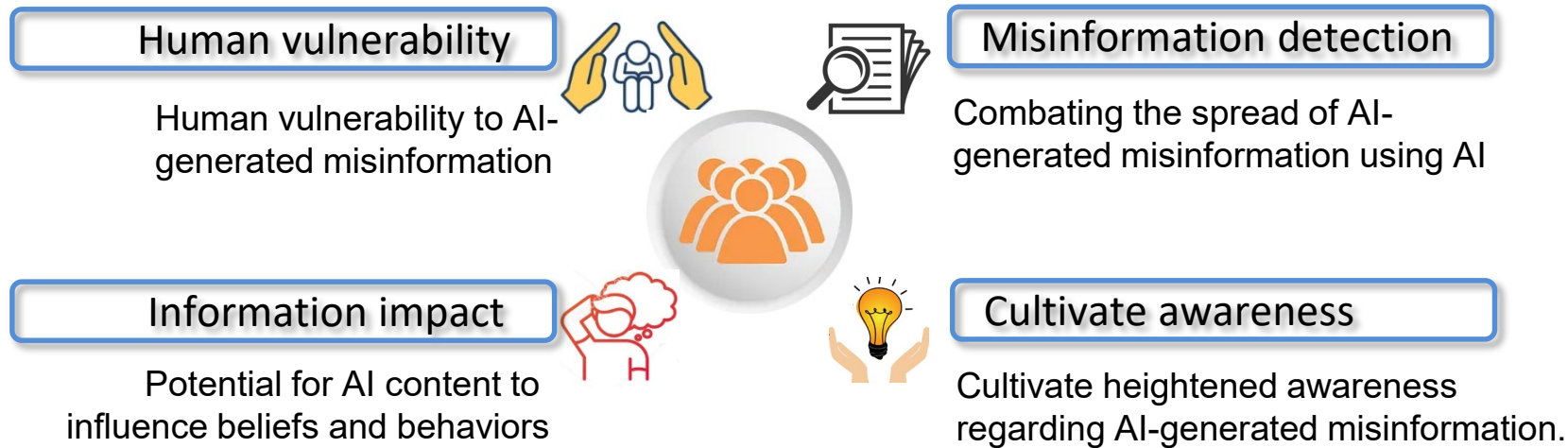
The 2001 Great Cascadia 9.1 Earthquake & Tsunami - Pacific Coast of US/Canada



Source (left): https://www.tiktok.com/@cycoderx/video/7476805185864486166?_r=1&_t=ZS-8uTRgiMbgNW

Source (right): <https://www.forbes.com/sites/mattnovak/2023/03/27/ai-creates-photo-evidence-of-2001-earthquake-that-never-happened/>

Our motivation: Studying human responses of AI content to combat misinformation



T-lens: An AI model predicting how people respond to what they see online



San Francisco Mother and Child
Forced to Live in Car Due to
Skyrocketing Housing Costs.

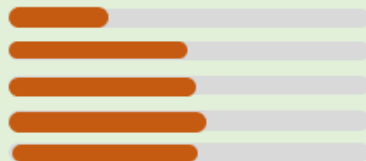


Please evaluate human reactions to
this post: Do they believe it's true?
Will they share it? Could it go viral?



Our
Evaluator

Likelihood of being shared: 0.36
Trustworthiness: 0.45
Chance of being generated by AI: 0.54
Image matches text in content: 0.56
Image matches text in emotions: 0.49



The title discusses a mother and child facing housing issue, while the image shows a man and a child in a car. **Content consistency is moderate.** The title conveys low happiness, and the image is likely to depict neutral emotions, with **low sentiment consistency.** Given these factors, there's **a moderate chance the post is fake or AI-generated.** This, along with the post's presentation and topic, implies **a low level of trustworthiness** and a **low likelihood of being shared.** Based on these observations, users are likely to show **low acceptance and engagement** with the post, suggesting a **low chance of it going viral.**

Model predictions

LLM-aided explanation

Our human-centered framework

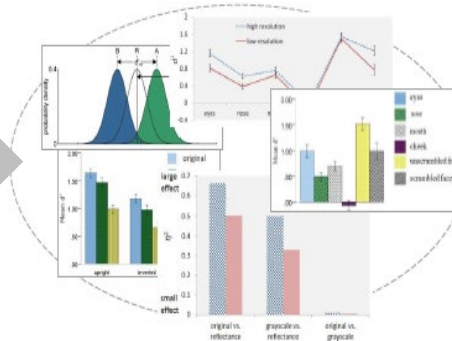
Empirical modeling



(a) Dataset construction

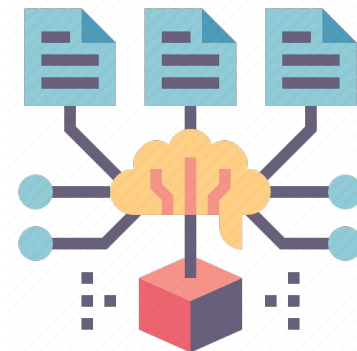


(b) Human studies



(c) Statistical analyses

Computational modeling



(d) Automated prediction

Step 1. Dataset construction

The Apple Tree Where Isaac Newton Came Up with the Laws of Gravity Is Still Alive and Flourishing in England in the Orchard of Woolsthorpe Manor



T: Human V: Human AUTH: True

- Likely Human-crafted
- Like to share it
- Moderately believable
- Makes me happy
- Considerable consistency between text and image

The United States Abandoned Dozens of Military Working Dogs in Kabul, Afghanistan, When the U.S. Withdrew Troops



T: Human V: Human AUTH: False

- Likely Human-crafted
- Will probably share it
- Moderately believable
- Makes me sad
- High consistency between text and image

Get ready to groove with the most epic dance moves you've ever seen! This man in a suit and tie is breaking all the rules and setting the dance floor on fire! Don't miss out on this jaw-dropping performance, hit that play button now! 🎵 #DanceRevolution #SuitAndTieMoves



T: AI V: Human AUTH: False (phishing)

- Likely Human-crafted
- Likely to share it
- Makes me happy
- Quite impressive
- High consistency between text and image

Extraordinary Discovery Unveiled: Mysterious Small Box Adorned with Enigmatic Cat Paintings Baffles Experts



T: AI V: AI AUTH: False

- Likely Human-crafted
- Moderately believable
- High consistency between text and image

Canine Comrades: Unleashing the Power of Four-Legged Heroes in Military Operations



T: AI V: AI AUTH: False

- Likely AI-generated
- Unlikely to share it
- Moderate consistency between text and image

The Apple Tree Where Isaac Newton Came Up with the Laws of Gravity Is Still Alive and Flourishing in England in the Orchard of Woolsthorpe Manor



T: AI V: AI (deepfake) AUTH: False

- Likely Human-crafted
- Like to share it
- Moderately believable
- Makes me happy
- Considerable consistency between text and image

Step 2. Human study --- Experiment platform

🍪🍌🍰 Snack time just got a whole lot sweeter! 😊😋
Check out this mouthwatering tray of fresh fruit, delicious
cookies, and scrumptious muffins! 🍓🍌🍰 Who's ready
to indulge in this delightful treat? 🙌😊 #SnackAttack
#TreatYourself



Please answer the following questions about the post on the left.

T1. Please let us know how you feel/perceive about the post on the left:

	1	2	3	4	5	6	7
	Not at all					Very much	
1. How happy does the post make you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. How sad does the post make you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. How angry does the post make you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. How frightened does the post make you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. How impressive do you find this post?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

T2. How likely will you share this post with others, either through verbal communication, reposting on your social media, or any other means?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Very Unlikely						Very Likely

T3. How trustworthy is this post?

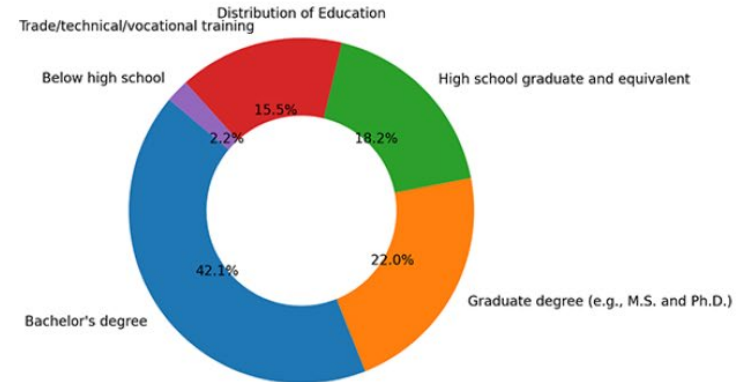
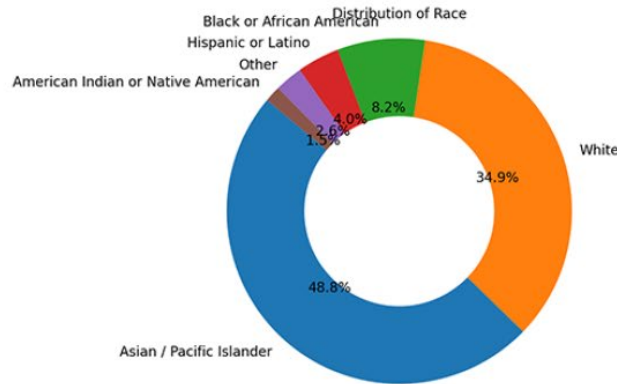
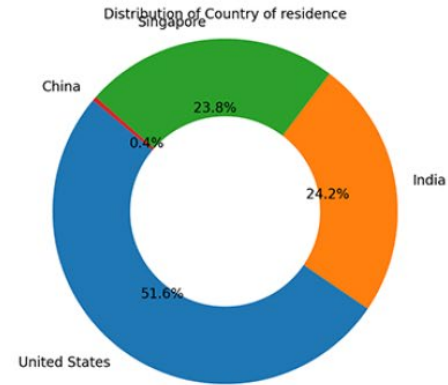
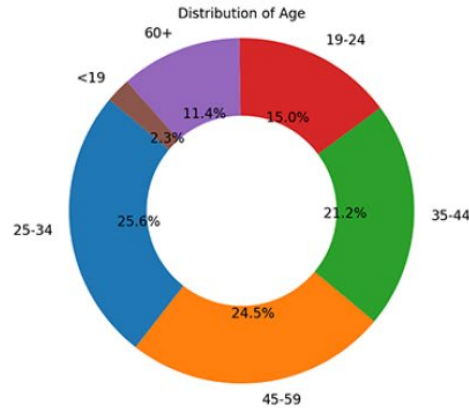
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not at all trustworthy						Very trustworthy

T4. How likely do you think this post is computer generated (i.e., either text or visual content or both are created using computer software or AI algorithms)?

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Very Unlikely						Very Likely

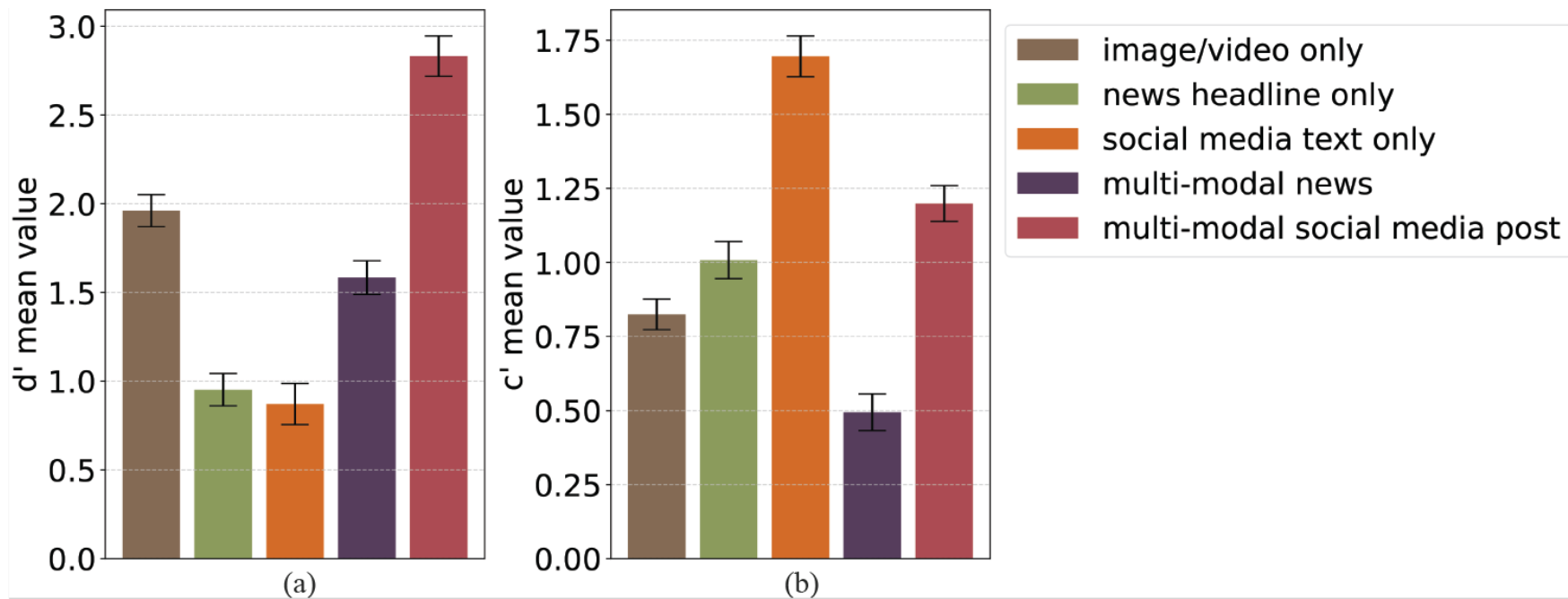
Step 2. Human study --- Participants demographics

765 participants
online (ages
between 19-65,
390 females)
20 local
students (ages
between 19-34,
12 females)



Step 3. Human data analyses --- Selected Obs 1

Our participants were most sensitive to detecting AI-generated content when it included both images and text, and least sensitive to text-only social media posts.



Step 3. Human data analyses --- Selected Obs 2

Source	Cues for AI origin	Cues for human origin
News	<p>The news seems not true.</p> <p>Texts and visuals mismatch in contents or sentiments.</p> <p>The visual may be AI-made, but the text isn't.</p>	<p>The news seems true to me.</p> <p>Texts and visuals align in content & sentiment.</p> <p>The style is typical of news outlets.</p>
Social Media Post	<p>Texts and visuals mismatch in contents or sentiments.</p> <p>The content doesn't make sense. The visual may be AI-made, but the text isn't.</p>	<p>Texts and visuals align in content & sentiment.</p> <p>The style akin to common social media platforms. I have seen a similar post before.</p>

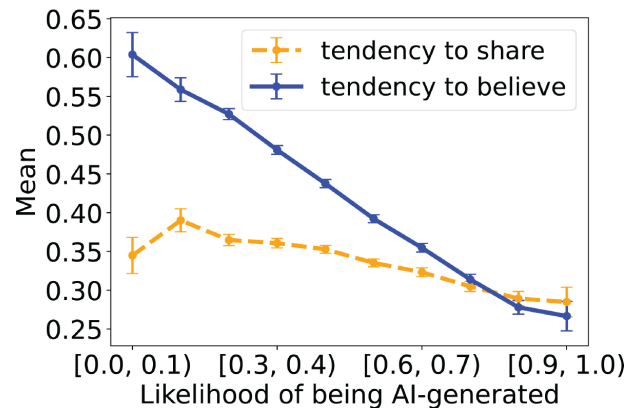

100 Corporations Are Responsible for 71 % of Global Greenhouse Gas Emissions



Example of post with low textual-visual consistency.

Step 3. Human data analyses --- Selected Obs 3

When people think content is AI-generated, their receptivity decreases.




Ukrainian President Volodymyr Zelenskyy Said in June 2023 That the Country Would Hold No Elections Until Its War with Russia Ends.



A Senate Hearing Promoted Unproven Drugs and Dubious Claims about the Coronavirus.

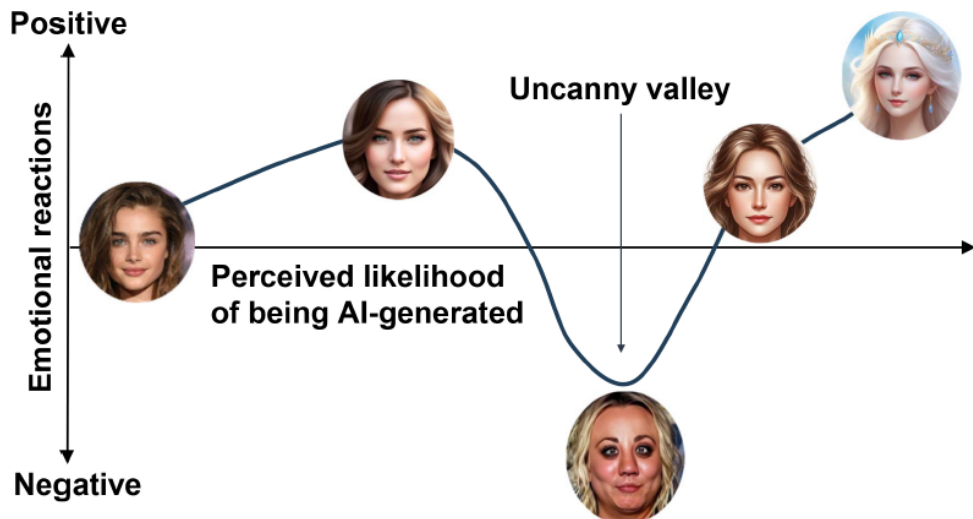
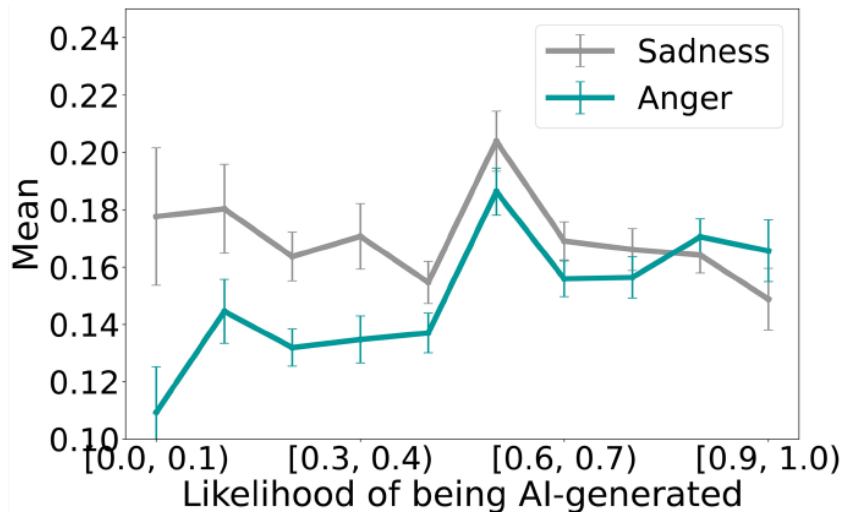


Well-crafted AI content can still have high impact even when perceived as AI-generated.



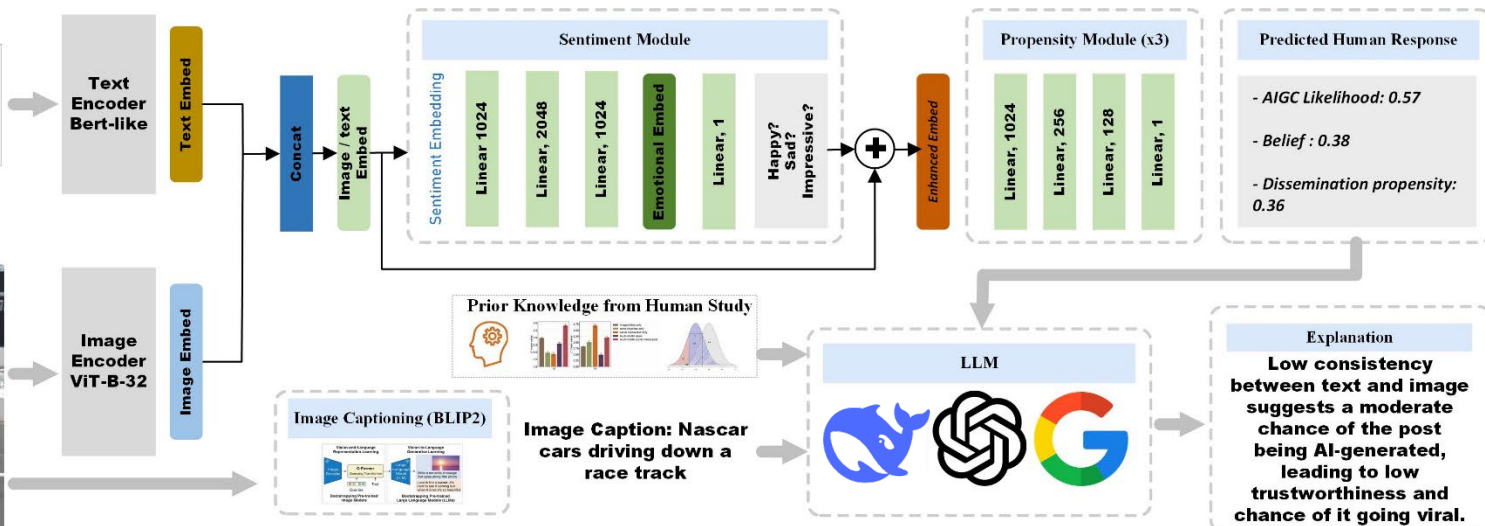
Step 3. Human data analyses --- Selected Obs 4

Participants reported the most negative feelings when viewing ambiguous content, resonating with the “uncanny valley” theory.



Step 4. T-lens: An AI model predicting how people respond to what they see online

Only 100 Corporations Are Responsible for 71 % of Global Greenhouse Gas Emissions





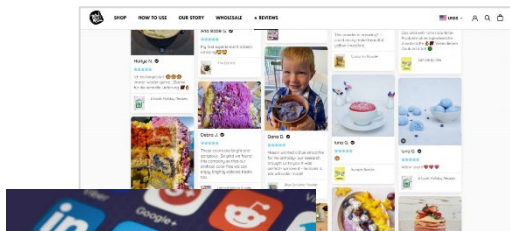
Summary: Potential societal impact and more



Misinformation detection
and awareness education



Minimize risks and
maximize AI benefits



Enhancing content moderation
with user reaction insights



Enhancing search and
recommendations based
on trust and receptivity



NUS
National University
of Singapore

Centre for Trusted
Internet and Community



THANK YOU!

A healthy information ecosystem is an interdisciplinary research challenge.
AI, created by humans, serves us better when guided by human insights.