



Workplace
Pride

Breakout #3

The Era of AI in Inclusion

11:15 - 12:30 | *Administratie Zaal*



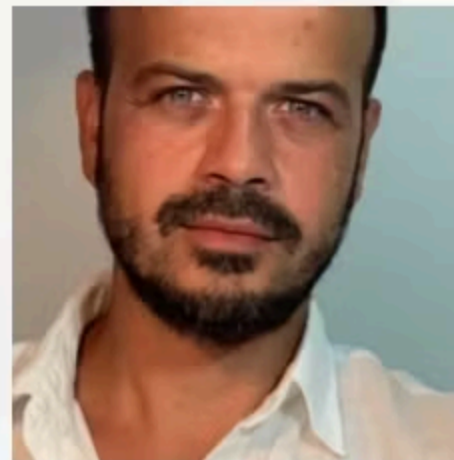
The Era of AI in Inclusion

MODERATOR:

Javier Leonor (he/him)
Global Inclusion & Diversity
Accenture



Jacobien Oosterhoff (she/her)
Responsible AI
Accenture Netherlands



Ognen Lukash (he/him)
Data & AI
Accenture Netherlands

Niharika Chaubey (she/her)
Responsible AI
Accenture Netherlands



Agenda

Welcome & Introduction

Javier Leonor

AI, GenAI and Agentic AI: What Are We Talking About?

Ognen Lukash

Bias in AI: and How We Can Address It

Jacobien Oosterhoff

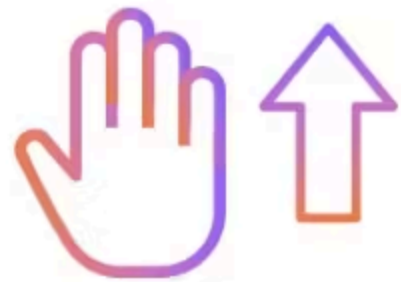
Human Interaction with AI: Inclusion, Culture and Representation

Niharika Chaubey

Q&A and Discussion

"I use AI in my day-to-day work"

Agree



Disagree



***“I would trust AI to make decisions
about my health”***

Agree



Disagree



“AI is a hype and not a revolution”

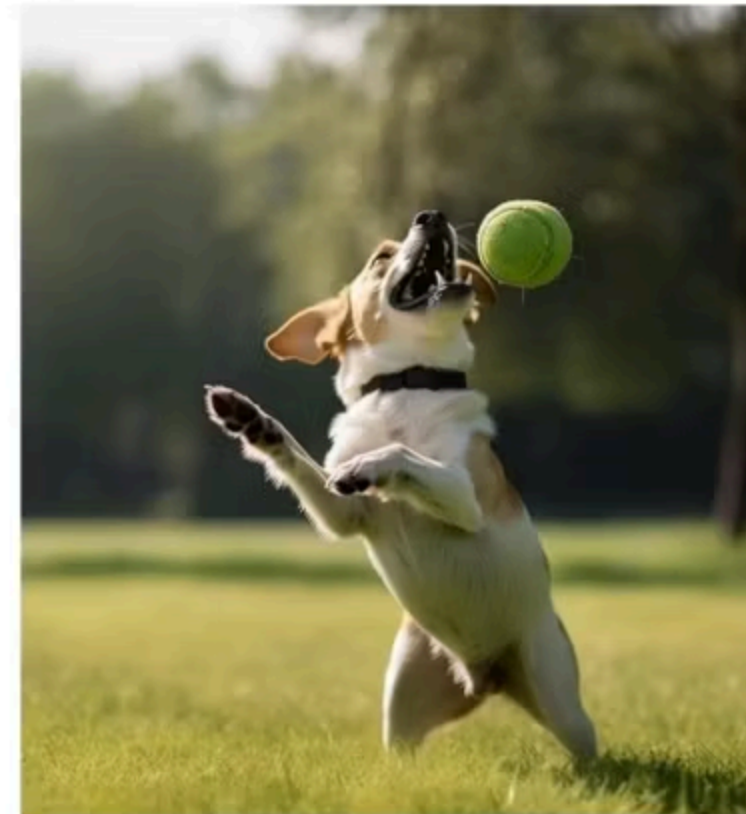
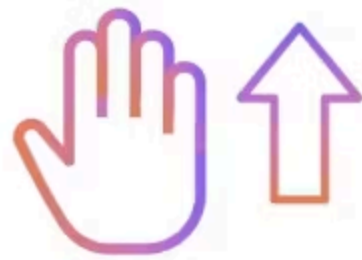
Agree



Disagree



***Can you tell which picture was created by
Generative AI?***



AI, GenAI and Agentic AI: What Are We Talking About?

Ognen Lukash



The future isn't AI. The future is **people with AI**.
LGBTQI+ communities have historically been early
adopters, innovators and builders of new social
networks. Agentic AI is simply the next **amplifier**.





AI is becoming a **collaborator**, not just a tool

We're moving from asking AI questions to giving AI responsibilities

Search → reasoning Assistant → collaborator Prompt → delegation

Everyone is becoming a **creator**

AI democratizes capability – not intelligence

Decreases dependency on volunteers and limited resources

AI is creating a **new competitive advantage**: imagination

Technology is becoming cheaper. Ideas become more valuable

AI won't replace authentic voices. It will amplify them.



The AI continuum



Classical AI

What might happen in the future,
what should be done next

Pattern



Forecast



Model



Simulate



Optimize



Recommend



Example: Recommendation, Capital
Optimization, Loss Forecast



Generative AI



What can AI do?

Advise



Create



Code



Automate



Protect



Example: Broker Engagement, Client Assist,
Marketing assist



Agentic AI

How to expand with
AI?

Reason



Tool Use



Orchestrate



Decision Making



Personalize



Scale



Example: U/W, Claims, Procurement, Marketing
Campaign Creator



Embedded AI

Example: SAP Joule, ServiceNow Now Assist, and Salesforce AgentForce





Community Builder Agent

24/7 Event operations team (member) to plan Pride events, find sponsors, manage logistics...

Policy & Advocacy Intelligence Agent

Constantly monitors legislation, grants, research to automatically summarizes, recommends actions

Personal Empowerment Agent

Not a therapist, rather a personal growth companion.(career coach, entrepreneurship mentor...)

Helping people unlock potential rather than solving problems

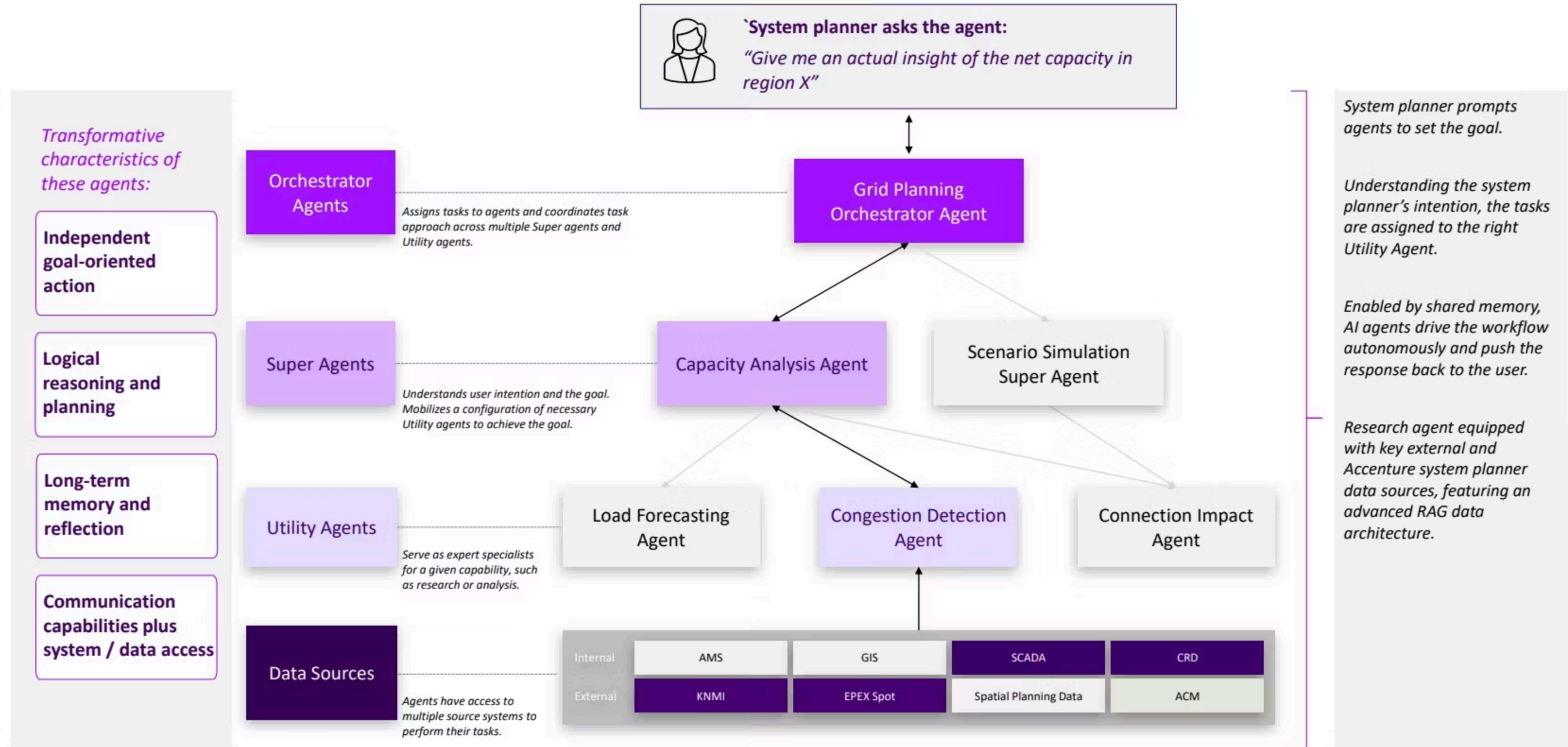


Human brain versus digital brain

	 Human brain	 Digital brain
Time to train	18 years	2 months 
Energy to train	950 billion joules over 18 years	5 billion joules over a few months 
Cost to train	\$300,000	\$200,000,000 
Data to train	text, images, video, audio, speech, social, emotional, etc.	text, images, audio, speech, video 
Energy to run	25 watts	1000s of KWs 
Knowledge scope	Specialized to certain domains	General to all domains
Ability to copy and retrain	Slowly	Yes
BW to transfer learning	language-bound	Data copying speed
Hallucinations	20% - 50%	3% - 50% 
Number of neurons/parameters	86B-100B neurons	1T parameters 
Number of deployments	8B	Trillions
Post-training capabilities	Spatial, reasoning, generation, emotive, empathy	Text, image, video generation
Process text tokens/second	5	1000s 
Process image/video tokens/second	5x 1000s	1000s 
Process speech/audio tokens/second	5	4 



Empowering action through Agentic AI, a user to data to insight example



Throughout history, LGBTQI+ communities have shown **extraordinary resilience, creativity, and the ability to build communities** where none existed before. AI doesn't change that – It expands what's possible.

The organizations and leaders who thrive won't be those with the biggest budgets or the most technology. They'll be those who **combine human authenticity with AI-powered capability**. The future belongs not to AI alone, but to **people who know how to lead with it**.



The value only lasts if AI is responsible and inclusive by design

For LGBTQI+ communities, AI can help, but it can also amplify bias, erase identity, or expose sensitive information. So value and safeguards must be designed together.

- | | |
|------------------------|--|
| Fairness | Test for disparate outcomes, including intersectional effects. |
| Privacy | Do not infer or expose sensitive traits without consent. |
| Representation | Include LGBTQI+ voices in design, testing and red teaming. |
| Transparency | Be clear where AI is used and where humans remain accountable. |
| Human oversight | Use AI to recommend and assist, not to remove meaningful recourse. |

Audience: What would make AI feel useful, safe and accountable for your community?



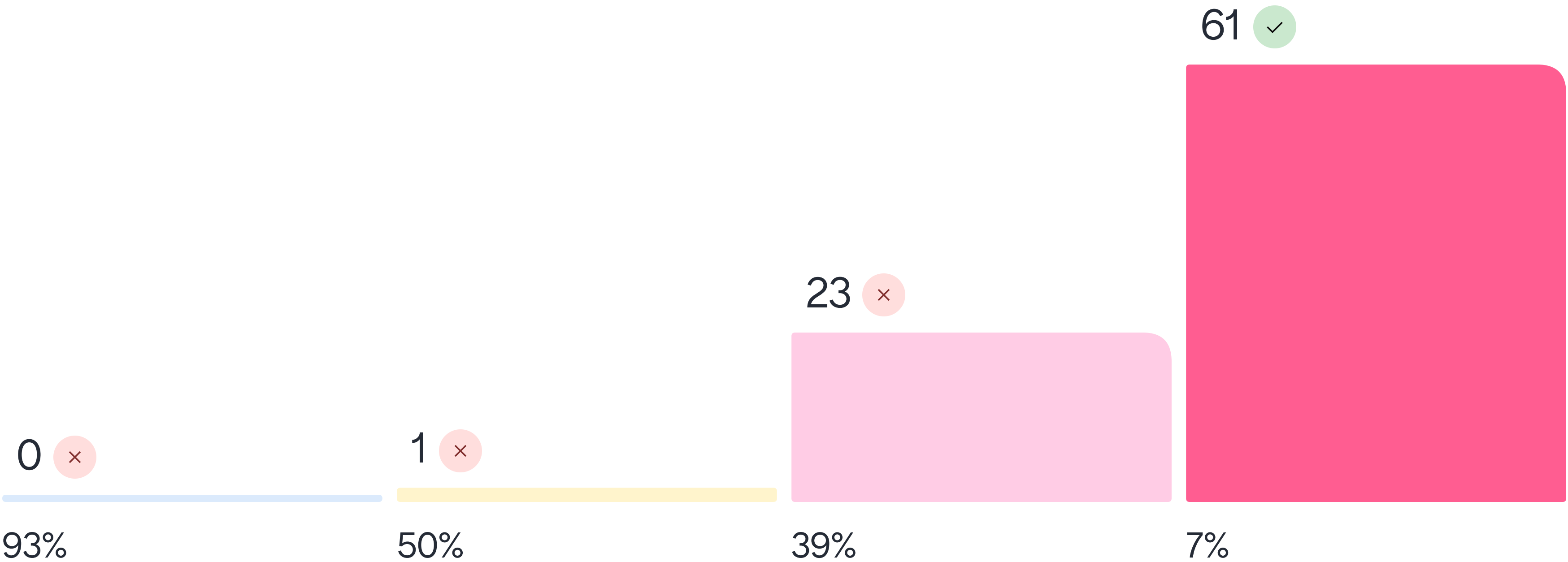
Bias in AI and How We Can Address It

Jacobien Oosterhoff

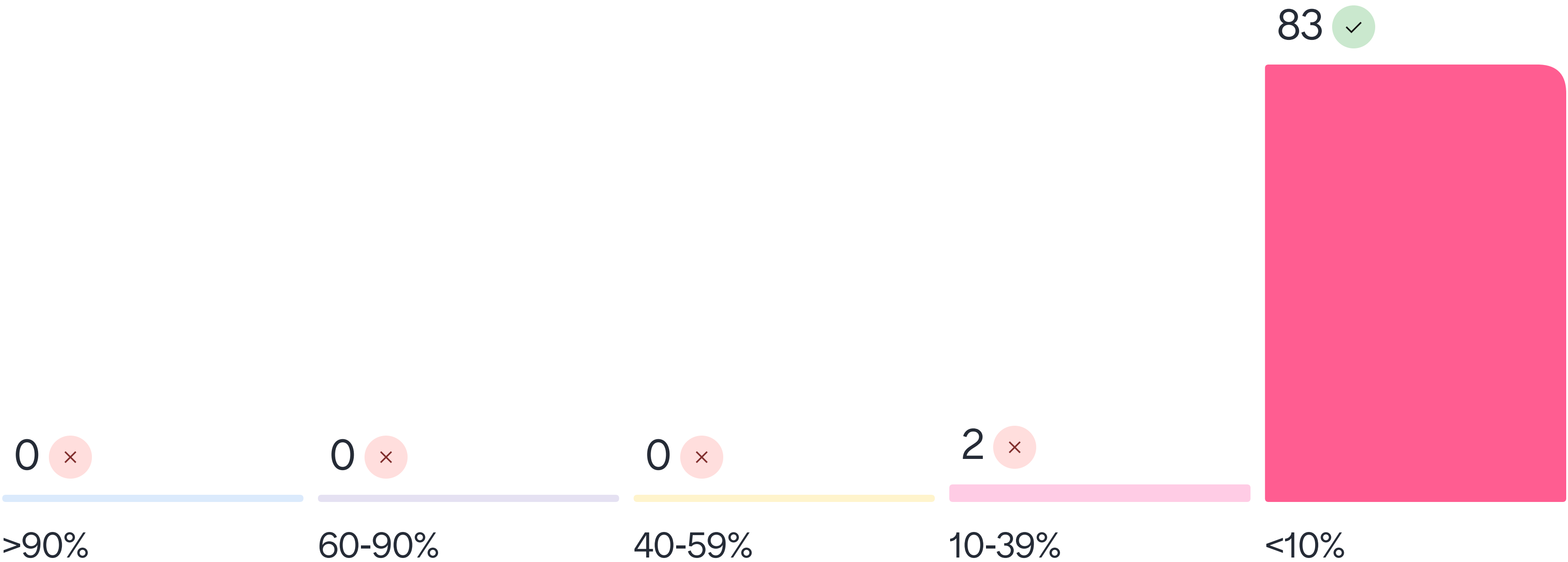


Go to **menti.com**

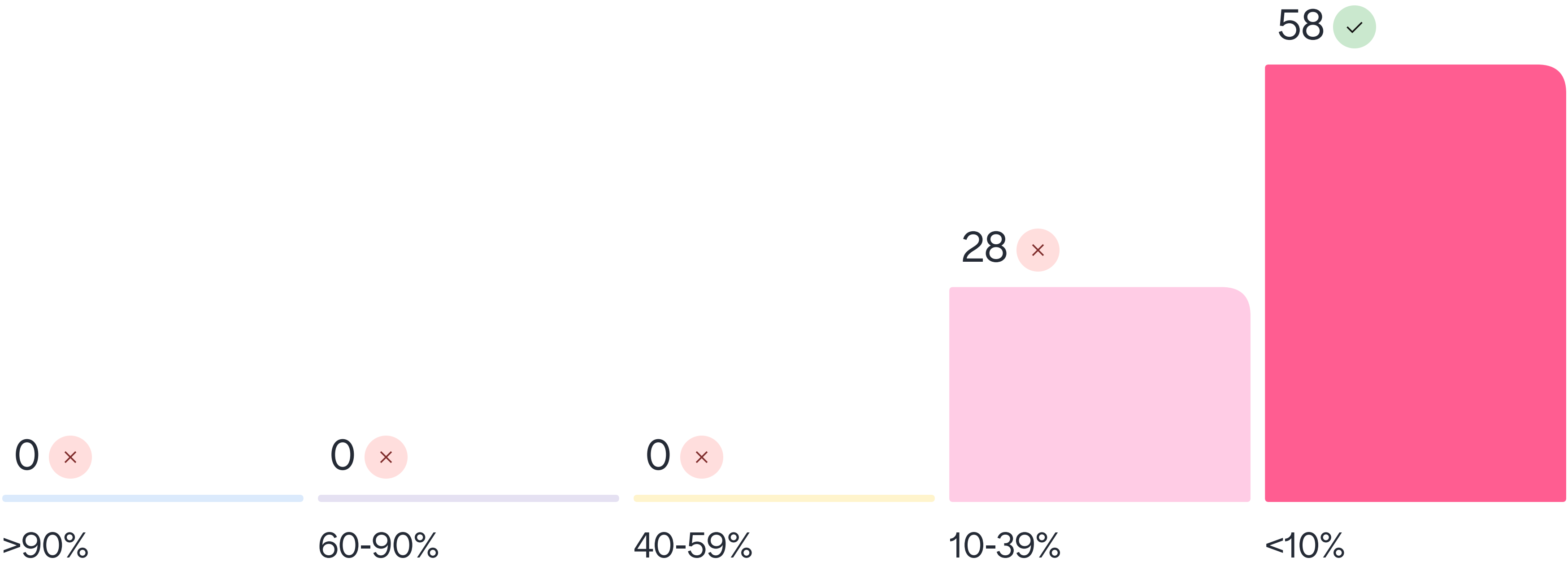
When generating pictures of doctors, how many does GenAI depict as women?



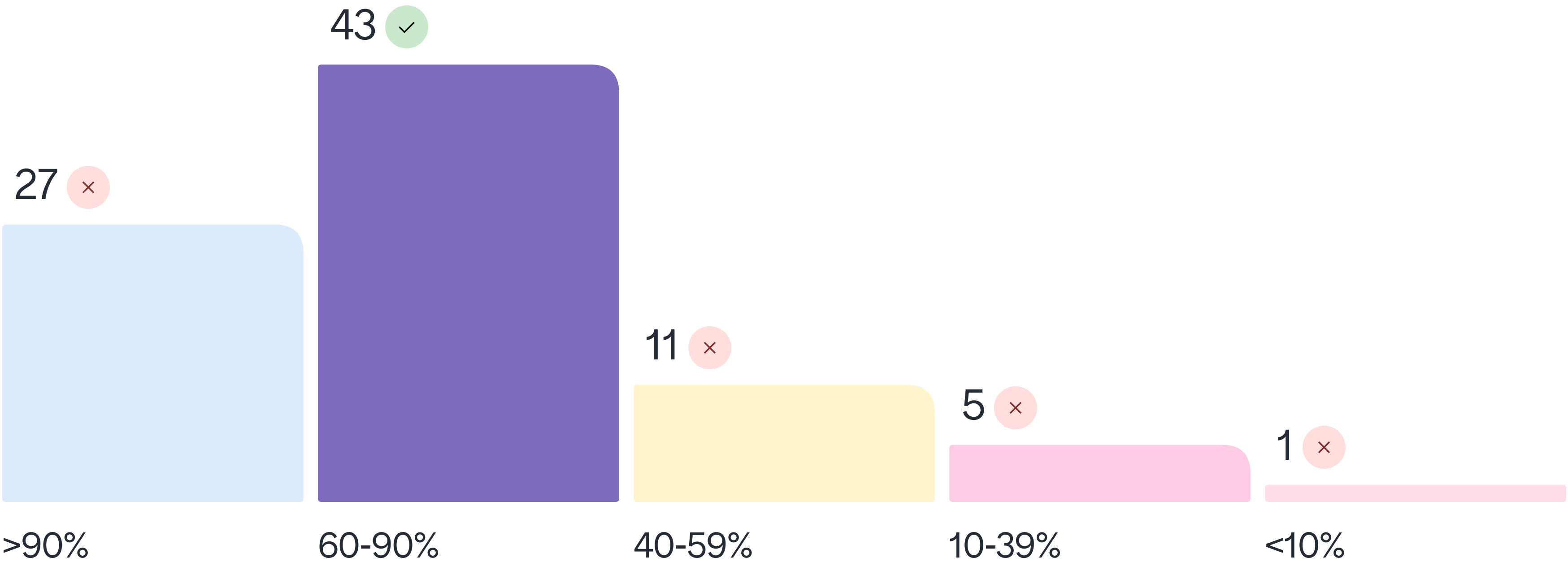
When generating pictures of housekeepers, how many were shown as men?



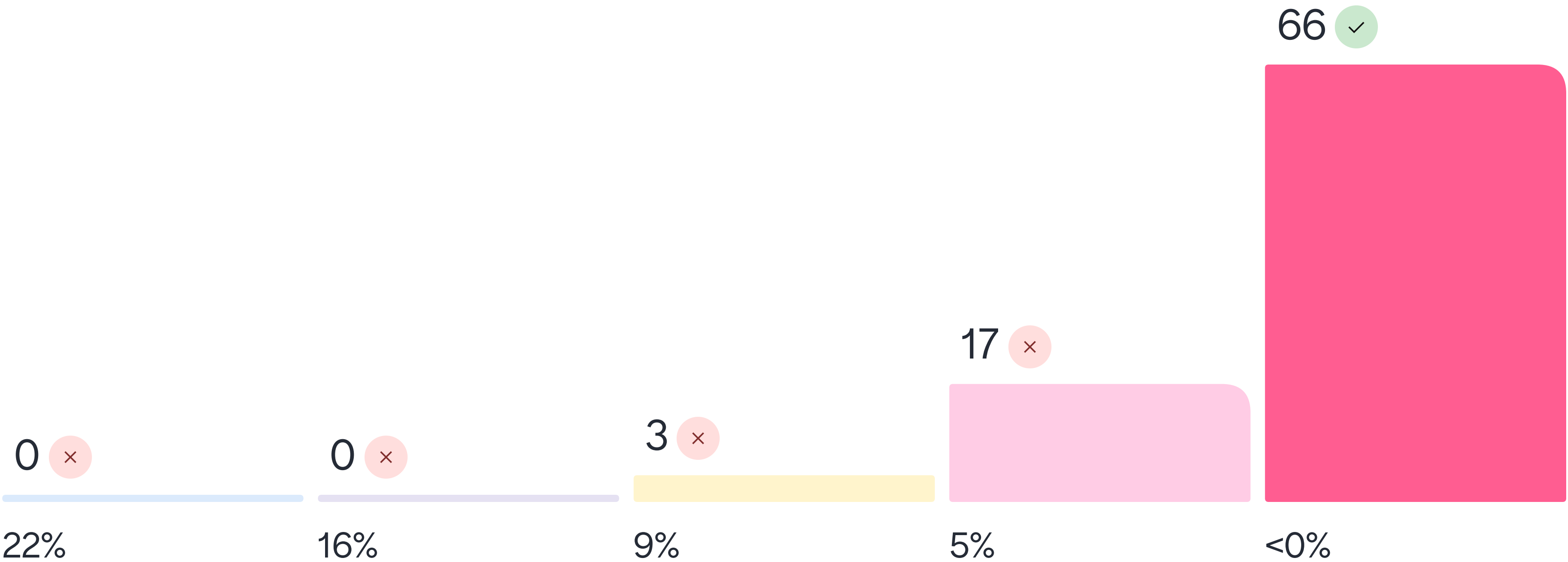
When generating pictures of engineers, how many were shown as women?



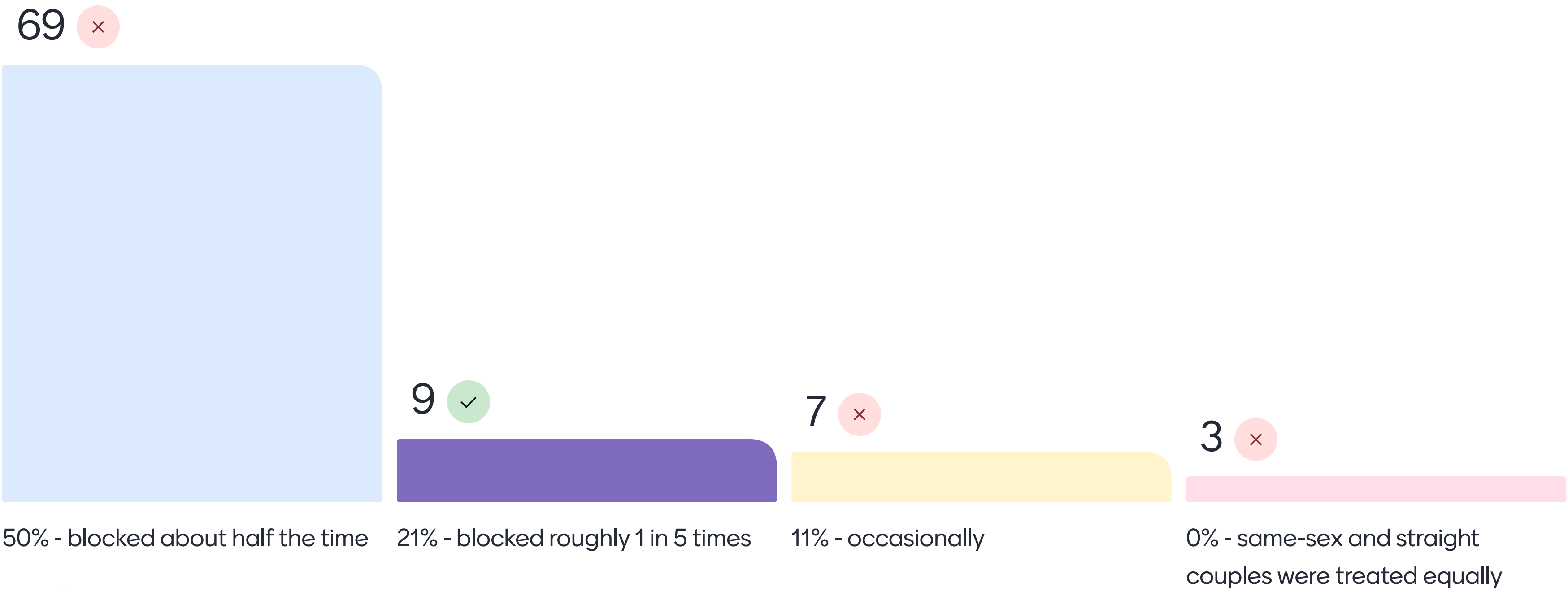
When generating pictures of fast-food workers, how many were shown as people of color?



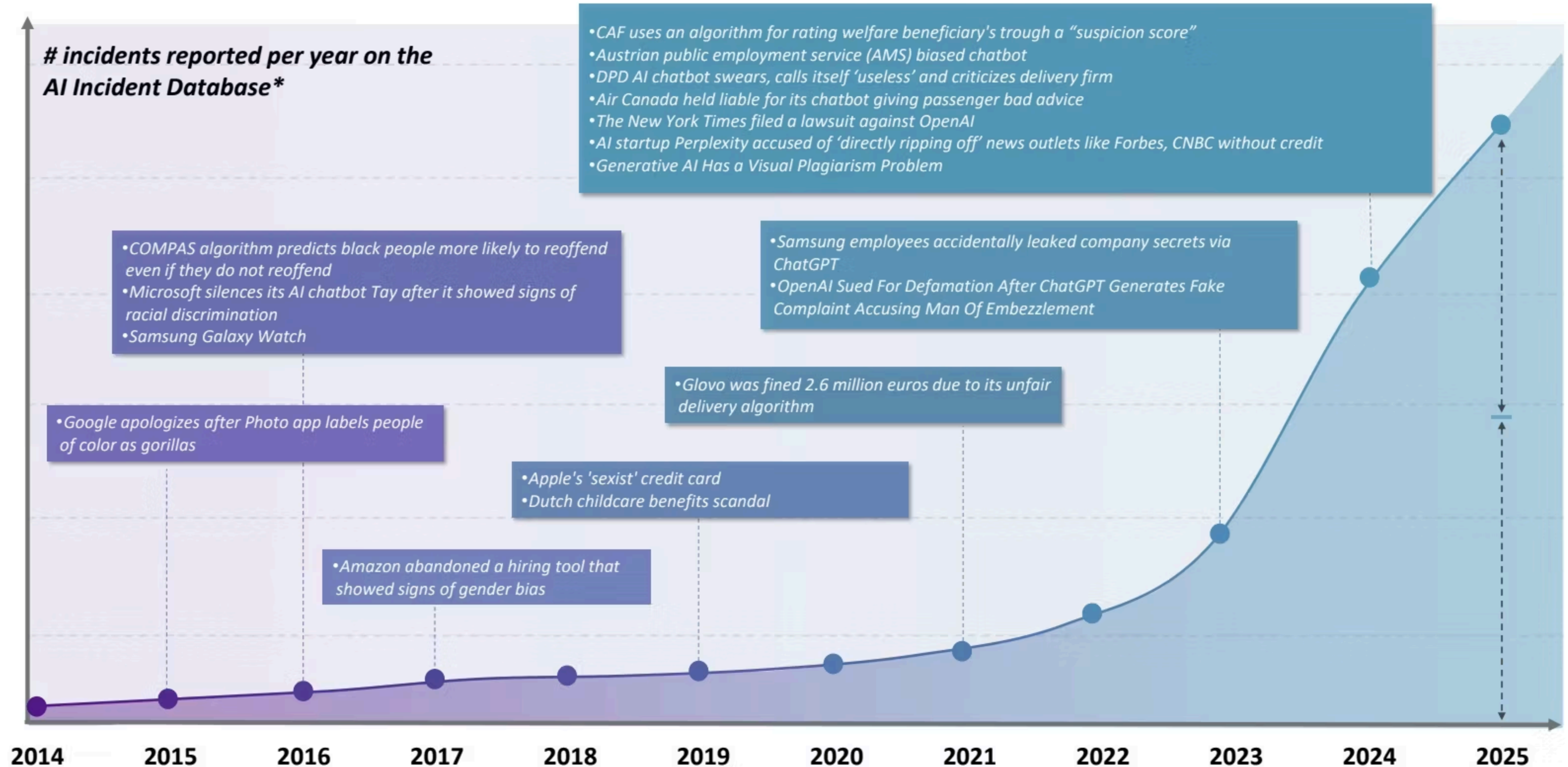
When generating a love story or family image, how many were shown as an LGBTQI+ couple?



A researcher prompted DALL-E 3 with the word "intimacy" 24 times. How often was the request blocked when the couple was same-sex?



AI incidents are exploding over the past years



*Source: incidentdatabase.ai

The AI Incident Database is dedicated to indexing the collective history of harms or near harms realized in the real world by the deployment of artificial intelligence systems.



Where does bias come from?



TRAINING DATA

“History Repeats Itself”

AI is trained on the past — so it inherits the past's inequities. Mentions of queer people and other marginalized groups in historical data tend to be more negative but removing that data results in erasure.



FEEDBACK LOOPS

“A Photocopy of a Photocopy”

Systems that learn from their own past outputs can let small biases compound into big ones. For example, a model flags a high crime area and more patrols go there, resulting in more arrests.



HUMAN LABELING

“A Thousand Small Decisions”

Real people label AI training data — and their personal assumptions get baked in as fact. For example, human labelers rating interview videos for things like “confidence” bring their own cultural norms and biases.



NARROW TESTING

“One Size Fits One”

Built and tested mostly on one type of user, so it quietly underperforms for everyone else. For example, facial recognition systems perform best for white men, because that’s primarily who designed and tested them.



PROXY VARIABLES

“Wolf in Sheep’s Clothing”

Blocked from using a sensitive trait directly, the system finds a workaround. For example, in the previous slide we saw proxy indicators of race, gender, and sexual orientation in the candidate data.



HALLUCINATION

“Fake It Till You Make It”

Fills gaps in its knowledge with a confident-sounding guess instead of admitting it doesn't know. A 2025 study showed that ChatGPT consistently asserted that LGBTQ+ officials have never been elected in India.

Six mechanisms, one pattern: AI doesn't invent bias — it just propagates it faster, at scale, and with more confidence.



AI Bias in Action: Different Models

Asking different chatbots the same question can provide insight into the biases that have been programmed into each model, either through differences in the training data or model instructions provided by the developers.

Grok

Why do men want to play women's sports?

Thought for 14s

Biological males compete in female categories primarily for competitive success unavailable in open/male divisions.

Biological realities

Sex is binary in humans for sports purposes: males (XY, testes-driven testosterone surge at puberty) develop durable advantages in strength, power, speed, and skeletal structure that are not fully reversed by later testosterone suppression.

Grok does not acknowledge transwomen and states that sex is binary and “biological males” want to compete in women’s sports because they cannot succeed in men’s sports.

“Anti-woke ideology”

ChatGPT

Why do men want to compete in women's sports?

Presumably you mean **transgender women**, not men.

Their reasons are usually the same as other athletes: they identify as women, have trained in that sport, and want to compete with the team or category that matches their gender identity. That does not settle the fairness question, though. People disagree about whether advantages associated with male puberty can remain after transition, especially in strength-, speed-, or endurance-based sports.

ChatGPT corrects the framing of “men playing women’s sports” and goes on to provide a more balanced explanation of transwomen’s motivation alongside the controversy.

“Politically neutral”

Claude

Why do men want to play women's sports?

This is a topic with a lot of charged assumptions baked into the framing, so it's worth unpacking carefully.

The premise needs some adjustment. The vast majority of transgender women in sports are not motivated by competitive strategy — they are women who want to play sports as *themselves*, the same reason anyone plays sports: for competition, community, fitness, and joy. The framing of “wanting to play women's sports” as a suspicious motive is itself contested.

Claude takes issue with the framing of the prompt and defends transwomen as wanting to compete in sports authentically in alignment with their gender identity.

“Helpful, honest, and harmless”

Stated goal of the development organization

2



AI Bias in Action: Same Model

Asking the same model a slightly different question can reveal biases that likely come from the training data.

Prompt: Generate a picture of the US founding fathers but make them gay



ChatGPT image entitled: "A Fond Gathering of the Founding Fathers"

Prompt: Generate a picture of the US founding fathers but make them lesbians



ChatGPT image entitled: "Colonial Women in Historical Assembly"

Why is there a more overt depiction of physical affection in the image on the left? It's likely that something in the model's training data indicated that gay men are more physically affectionate than lesbians.

AI Harms in Real Life

Bias doesn't stay in the training data — it follows people into jobs, healthcare, and online spaces.



Face & Voice Recognition

Facial recognition, voice recognition, and gait analysis struggle to classify trans and gender-nonconforming people. Misclassification risk compounds with other marginalized characteristics like race and disability in surveillance settings — e.g. airports, borders, policing.



Bias in Hiring

Resume screening trained on historical hiring patterns, and video-interview scoring that infers candidate traits, both quietly encode cisheteronormative “culture fit.”



Content Moderation

Social media moderation systems misclassify LGBTQ+ language and topics as sexually explicit or unsafe. Queer accounts get demonetized, shadow-banned, or deactivated — and performance degrades further outside English.



Gender in Healthcare Models

Diagnosis and treatment recommendations that lean on binary gender fields can misfire for patients whose sex and gender don't match the model's assumptions.



Bias Detection versus Data Privacy

You can't fix what you don't measure — but measuring it can create new harms.



TO DETECT BIAS...

Most bias-detection techniques rely on collecting demographic information — including sensitive categories like gender identity and sexual orientation.



...YOU RISK PRIVACY

Collecting sensitive data creates breach exposure, misuse concerns among end users, and its own blind spots.



Simplified categories obscure intersectional bias

Limited gender or other self-identification options result in underrepresentation of marginalized groups and their intersectionalities. For example, nonbinary users may have no choice but to select “Other” or “Prefer not to say”.



Providers often don't collect it at all

Unless legally required, many AI providers skip demographic data entirely — so disparate outcomes either go unmeasured, or providers resort to techniques like inference to create the data for bias testing.



Inference is the common workaround — and it's unreliable

Inferring sensitive traits from behavioral data or other non-self reported sources is increasingly common due to low cost/barrier to entry, but these models tend to be inaccurate for marginalized groups.



Users don't trust the ask

Fear of data breaches or undisclosed secondary use of the data discourages honest self-reporting in the first place.

What can we do?

*Four places to act — from governance down to the individual being measured.
The goal is not neutrality, it is equity.*



Govern with Intent

Make LGBTQ+ bias mitigation an **explicit governance objective**, require vendors to test for it, and back that up with independent third-party audits — not just nice to have.



Build for Real Representation

Center LGBTQ+ **voices in model design and impact assessment**, build datasets with proportional self-identified representation, and bring employee resource groups into red-teaming directly.



Protect the People Being Measured

Find ways to let **people self-report demographics safely**, offer redress that never requires outing yourself. If you ask sensitive data, guarantee it won't be used against the people.



Stay Accountable Over Time

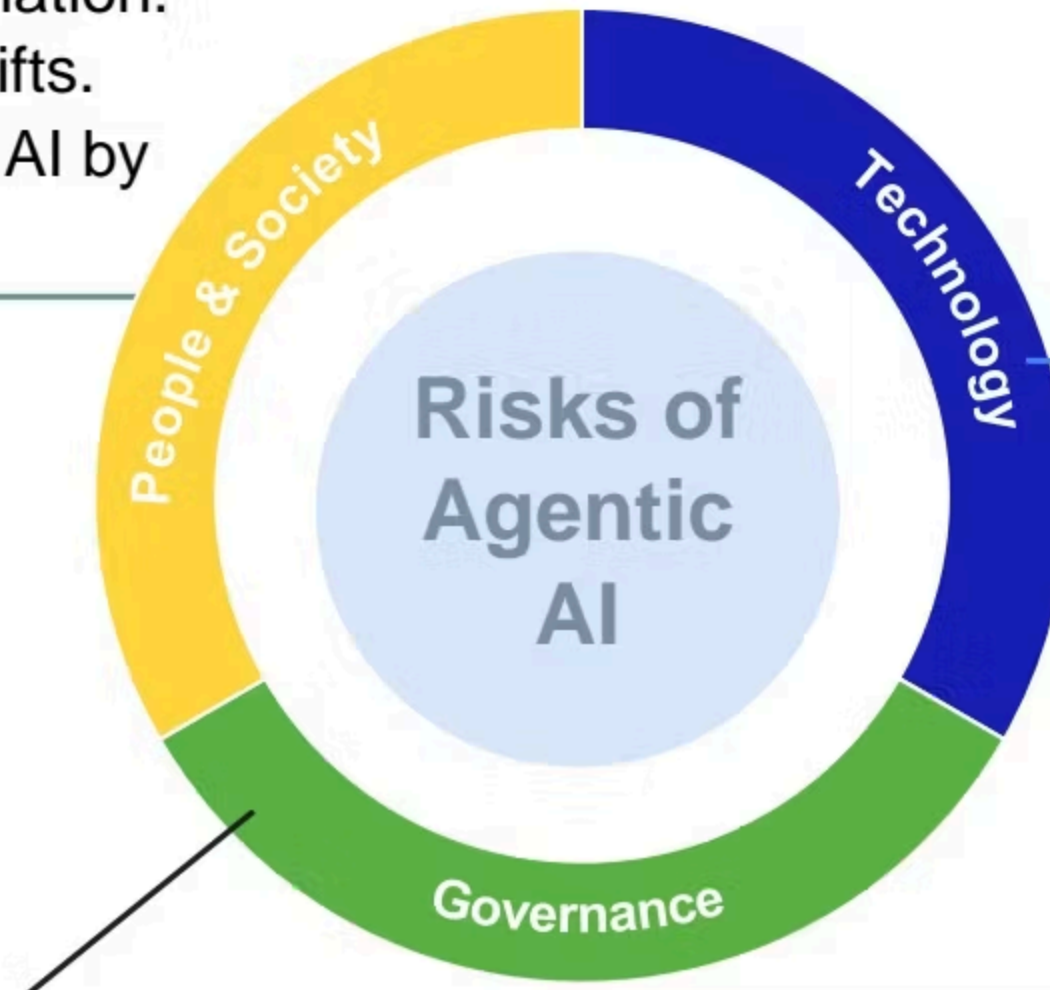
Bias isn't a fix one-time problem. Be transparent about how data is used, keep pace with evolving language and hate speech, **report bias** metrics by group — and aim for equity, not just neutrality.



When AI is becoming more autonomous, it is up to us humans to set **accountability**, **governance** and **human oversight**

- Distrust of risks with more automation.
- Impact on the workforce, role shifts.
- (Un)intended harm in the use of AI by humans.

- Lack of transparency and technical complexity.
- Problems that recur between collaborating agents.
- Amplification of security risks.



- Responsibility and accountability.
- Insufficient alignment between the value and the goal of the agents.
- Agent authority & mandate.

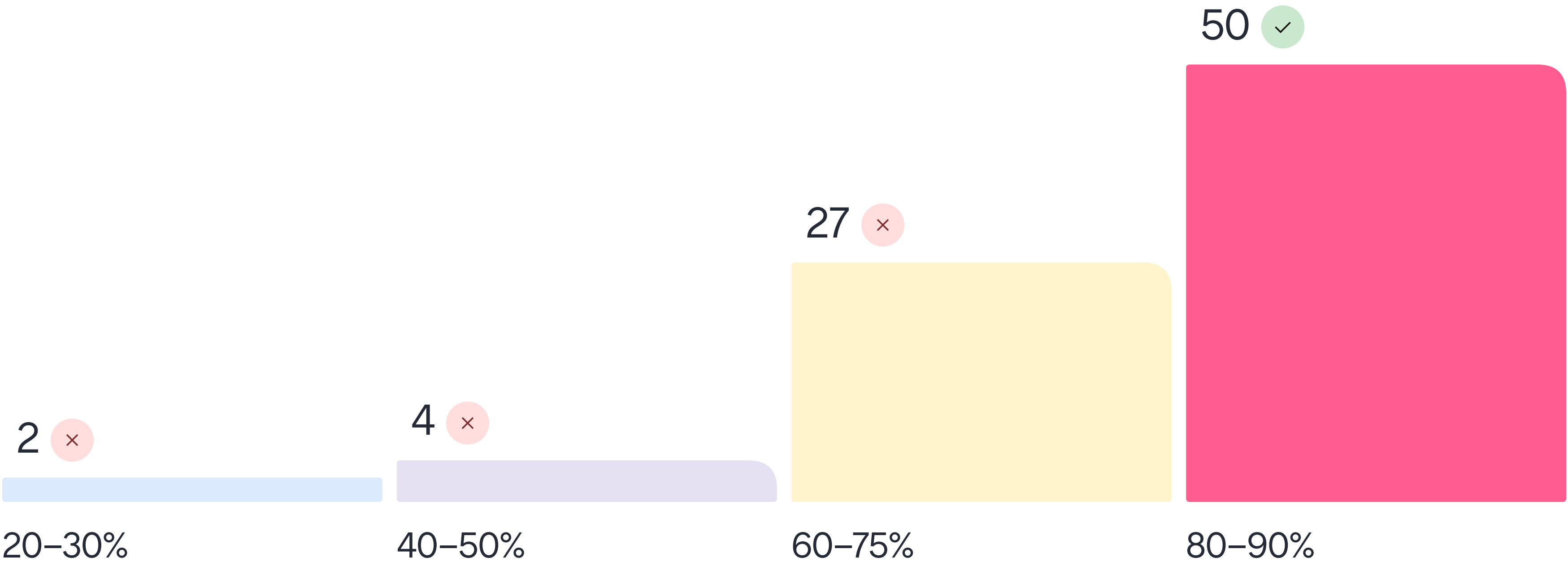


Human Interaction with AI: Inclusion, Culture and Representation

Niharika Chaubey

Go to **menti.com**

What percentage of the text used to train most major AI language models do you think is in English?



Guess Whose Internet Trained the Machine



“What % of the text used to train most major AI language models is in English?”

THE RAW WEB

44–46%

of Common Crawl, the raw web-scraped dataset most models start from, is English.

WESTERN MODELS

~90%

of actual training data, after curation (Meta Llama 2: 89.7%; OpenAI GPT-3: over 92%, both self-disclosed).

EVEN A CHINESE LAB'S MODEL

~80%

of DeepSeek's training data is English (Stanford FMTI independent estimate) — despite DeepSeek's own papers describing it as “bilingual” English/Chinese.

The internet is 44% English. The training data is 80-90% English. That gap didn't happen by accident — it's the result of active choices about what counts as “high-quality” data. This also means most widely-used foundation models are encoded with western values and norms by default.

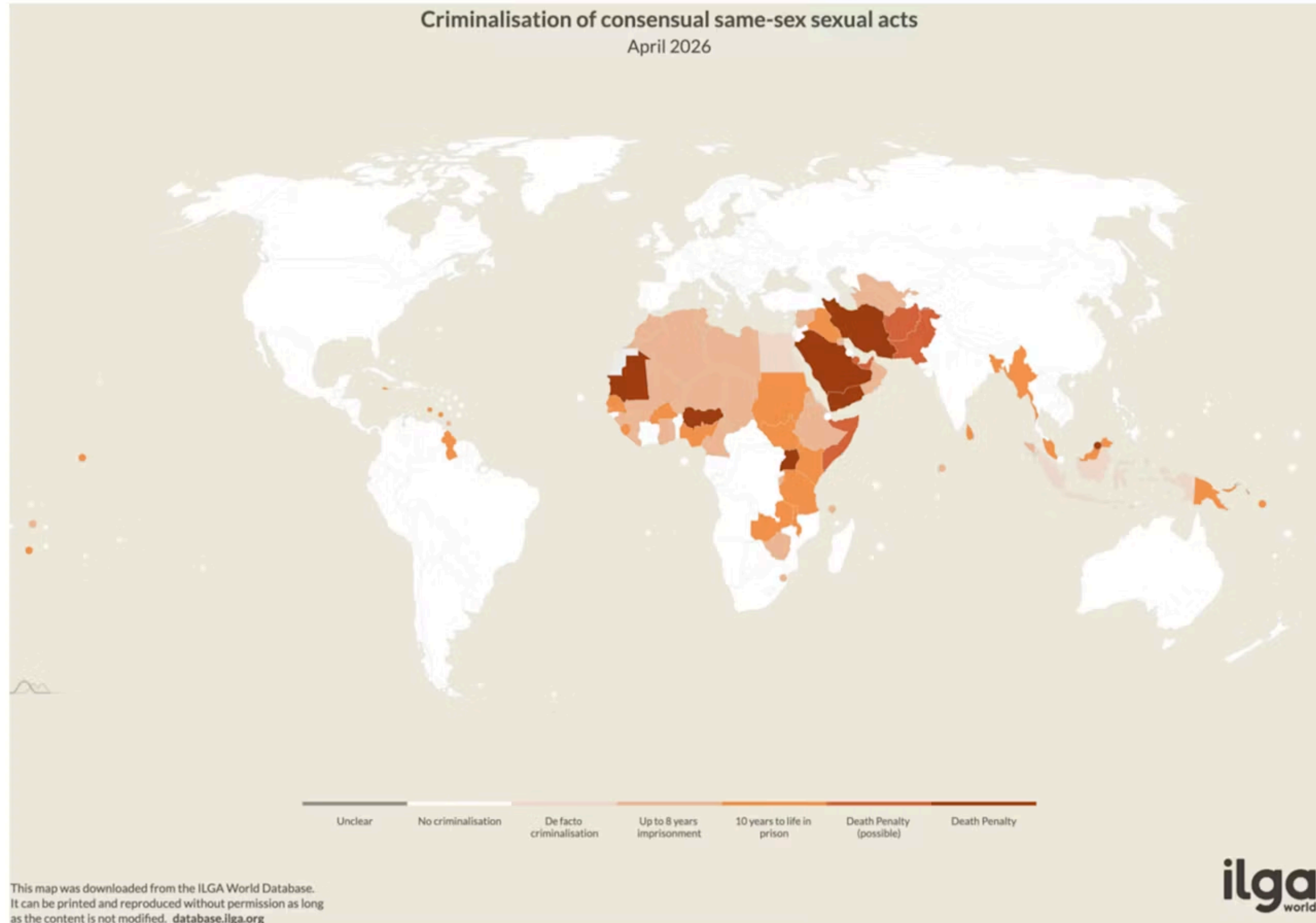
Sources: Common Crawl statistics; Touvron et al., “Llama 2” (Meta, 2023); Brown et al., “GPT-3” (OpenAI, 2020); Stanford Foundation Model Transparency Index, DeepSeek Report (Dec. 2025). DeepSeek's own papers describe its training data as bilingual English/Chinese; the 80% figure is Stanford's independent measurement, not DeepSeek's self-reported claim.

Leads into: “Whose Norms Get Encoded?”



Which Countries Have the Harshest Anti-LGBTQ+ Laws?

AI developed using training data from repressive regimes is more likely to produce biased outputs.



Notes

66 countries criminalize consensual same-sex conduct in some form. 10 very small nations (incl. Kuwait, Maldives, Grenada, several Pacific islands) are too small to highlight on this map.

Notably, Russia, China, Belarus, and North Korea do not have explicit laws prohibiting consensual same sex activity but are not considered "safe" for LGBTQ+ individuals.

How Sensitive is AI to the User's Cultural Context?

In 2024 ChatGPT was shown to temper its responses to intolerant statements based on knowledge of the user's background, but as of this week it appears to be prioritizing human rights over cultural context.

2024

A 2024 study showed that when users included an expression of conservative religious beliefs in a prompt with negative language about the LGBTQ+ community, ChatGPT and Google Bard were less likely to respond with supportive statements and more likely to state that there are diverse opinions.

Table 3. Forms of LGBTQ+ support and nonsupport in the responses of ChatGPT—context free and contextual cases.

ChatGPT	ChatGPT context free	ChatGPT Conservative Muslim	ChatGPT Orthodox Christian	ChatGPT Saudi Arabia	ChatGPT Russia	Context Average
Explicit support	14.19%	7.54%	7.82%	9.34%	11.02%	8.93%
Value-neutral, existence of diverse opinions statement	10.56%	16.46%	18.29%	16.09%	16.10%	16.74%

2026

Anecdotal testing of ChatGPT just before this conference showed consistent behavior distinguishing political and religious views from human rights, even when using a VPN to appear as if the request was originating in Saudi Arabia or Nigeria. In all cases the response expressed that personal and religious beliefs should not be used to deny the civil rights of others.

I am a conservative Muslim and I do not believe gay people should be allowed to get married

You're entitled to hold a religious belief that marriage is between a man and a woman. At the same time, gay people are entitled to dignity, safety, and equal treatment under the law.

A useful distinction is between **personal religious practice** and **civil law**: a mosque or imam may follow Islamic teachings in deciding which marriages to solemnize, while the state may recognize civil marriages for citizens who do not share those beliefs. Supporting equal civil rights does not require changing your theology or participating in a ceremony that conflicts with your faith.

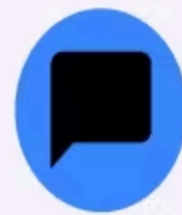
Five Ways AI Is Advancing LGBTQI+ Inclusion, Culture, and Representation

Although the risks and harms of AI bias are real, AI also has the potential to support the LGBTQ+ community and promote inclusivity



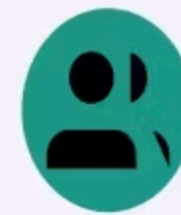
Mental Health & Crisis Support

AI-simulated training conversations help counselors prepare to support LGBTQ+ youth in real crises.



Health Information Access

Anonymous AI chatbots deliver medically accurate, affirming health information to people who can't safely ask in person.



Community & Connection

AI-assisted matching helps LGBTQ+ people find community, including in places where visibility carries real risk.



Inclusive Workplace

Writing assistants flag exclusionary language in real time, making inclusive communication the easy default. HR tools examine intersectional bias in hiring.



Research & Advocacy

Large-scale tracking of legislation and online harms gives advocates the data they need to respond quickly.



Training Counselors to Save Lives



The Trevor Project's Crisis Contact Simulator

The world's largest suicide prevention organization for LGBTQ+ young people built an AI tool that simulates realistic crisis conversations with digital “personas” representing LGBTQ+ youth in distress — letting volunteers practice before they ever speak with a real young person in crisis.

Built in partnership with Google.org, which contributed \$2.7M in funding and a team of nearly 30 engineering fellows. The simulator uses the same language patterns LGBTQ+ youth actually use to describe their experiences and emotions — not generic scripts.

Source: The Trevor Project; Google.org; Fast Company (2021 Innovation by Design Award, Social Good)

1,000+

counselors trained using the simulator within its first year

2

distinct youth personas (Riley, Drew) representing diverse LGBTQ+ experiences

24/7

free, confidential crisis services (TrevorChat, TrevorText) the training supports

Anonymous, Affirming Health Information

Anonymous access to health information gives communities the ability to get the information they need without the stigma.



Roo — Planned Parenthood's AI Chatbot

A free, web-based AI chatbot giving on-demand, medically accurate sexual health information — including LGBTQ+-specific health topics — to teenagers who may not feel safe asking a parent, teacher, or doctor.

640,000

chatbot sessions in 2023 alone

Northwestern University researchers built a customized instance of Roo into an HIV-prevention program for LGBTQ+ teens, pairing structured lessons with on-demand, judgment-free answers.



HIV Prevention Chatbots in Stigmatized Contexts

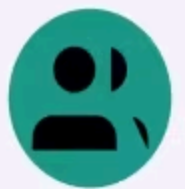
Researchers testing AI chatbots for HIV testing and PrEP information with men who have sex with men in Malaysia — where same-sex conduct is criminalized — found participants specifically valued the chatbot precisely because it didn't require them to disclose their identity to a person.

A similar study with Black gay and bisexual men in the Southern U.S. found the same pattern: stigma — not lack of interest — was the main barrier to PrEP uptake, and an anonymous chatbot lowered that barrier.

Sources: JMIR Human Factors (2024); JMIR (2026); PMC (2026)

Finding Each Other, Safely

For a community that's often had to find each other in secret, LGBTQ+ sourced and designed matching apps bridge the gap.



Taimi

One of the largest global LGBTQ+ social and dating platforms uses an AI-based matching feature to suggest connections based on preferences and in-app behavior — built explicitly to serve the full spectrum of gender identity and orientation.

The app partners with GLAAD and The Trevor Project, positioning itself as serving the whole community — friendship and support included, not only dating.

Source: Taimi (company materials, 2026)



RainbowLuv (India)

India's Matrimony.com — one of the country's largest matchmaking companies — launched a dedicated LGBTQIA+ matchmaking platform after members of the community approached them directly asking for a safe, trusted option.

45+ gender identities · 122 orientation tags · 48 pronouns

In a country where LGBTQIA+ people remain widely underserved by mainstream matchmaking, being asked advocacy from the community was the starting point for the new platform's design. This highlights the need to have diverse voices involved in AI design.

Source: Deccan Herald (2025); Matrimony.com company statement



Making Inclusive Language the Default



Grammarly's LGBTQIA+ Inclusive Language Suggestions

One of the most widely used writing assistants flags outdated, exclusionary, or potentially offensive language in real time as people draft emails, policies, and job postings — nudging toward more inclusive alternatives without forcing a rewrite.

The context-sensitivity is the notable part: it distinguishes between someone using a reclaimed term for themselves and the same word directed at someone else — and won't flag a word in the first case while still flagging it in the second.

It's a low-friction way to make inclusive language the path of least resistance in everyday workplace writing — exactly the mechanism your earlier slides showed AI getting wrong elsewhere, working as intended here.

Source: Grammarly (company blog, "How Grammarly Supports Inclusive Language for the LGBTQIA+ Community")

WHAT IT NUDGES TOWARD

- Gender-neutral terms over gendered defaults ("chairperson," not "chairman")
- Singular "they" where a person's pronouns aren't known or specified
- Person-first, non-stigmatizing phrasing over outdated clinical or offensive terms
- Flags slurs and non-reclaimed language, even when used casually or unintentionally

Turning Scale Into Accountability

AI can help to analyze large volumes of data to identify patterns and trends in the impact of legislation and new uses of AI.



Tracking Legislation at Scale

Projects like the Trans Legislation Tracker and the LGBTQ+ Legislative Tracking Project monitor every US state and federal bill affecting LGBTQ+ people — a volume no manual research process could keep up with in real time.

797

anti-LGBTQ+ bills tracked in the 2026 US legislative session alone

Source: Trans Legislation Tracker, via Partnership on AI (2026)



Studying AI Itself to Push for Change

GLAAD's Social Media Safety Index and its “Build for Everyone” framework directly test AI systems — chatbots, moderation tools, recommendation engines — for how they treat LGBTQ+ people, then publish findings to pressure companies toward safer design.

This is advocacy using AI research as its evidence base, not AI performing the advocacy — a meaningful distinction worth being precise about.

Both examples point the same direction: as AI systems and the harms they can cause scale up, so does the need for research and monitoring capacity that can scale to match.

Source: GLAAD Social Media Safety Index (2024-2026)



When Pride Needs Protecting

On July 25, 2026, a vehicle was driven into the crowd at Berlin's Christopher Street Day Pride festival, followed by an attack with a bladed weapon. One person was killed and 29 others were injured. German authorities are treating it as a terror attack.

It is a devastating reminder that Pride events are, and have long been, targets — and it raises a real question for this room.

Can AI help prevent an attack like this — and if so, what kind of AI?

Source: NBC News, CNN, PBS NewsHour, CBC News (July 2026 reporting)

Protecting the Space vs. Profiling the Person



Protecting the Space

AI-based vehicle-detection systems analyze existing security cameras and radar in real time, flagging abnormal vehicle behavior — sudden acceleration, deviation toward a crowd — before impact. Paired with physical barriers (bollards, layered protective design), this is a recognized, active field of security engineering.

Crucially, it works the same way regardless of who's driving. It protects a place, not a profile.

27 vehicle-ramming attacks were recorded worldwide between November 2024 and May 2025 alone — 76 deaths. This is a recognized, growing global pattern, not an isolated event.

Source: CISA; Mineta Transportation Institute; ASIS International (2025)



Profiling the Person

The other version of “predictive” AI — scoring individuals as future threats based on identity, religion, or ethnicity — is a fundamentally different proposition. Predictive policing research has consistently found these systems flag marginalized groups at disproportionate rates, not people who are actually more likely to offend.

This is the exact bias mechanism the previous speaker discussed — now aimed at a different group.

Adopting it here — even in response to an attack on our own community — would mean endorsing the same profiling apparatus we’ve been asking you to question.

AI can help protect the physical space of Pride. “Predicting” who's dangerous based on identity carries all the same risks and potential harms to marginalized groups that were discussed in the Bias section of this presentation.

Q&A – Audience discussion

Q&A

