

ARTIFICIAL INTELLIGENCE IN PSYCHIATRY

What is it?
Where and How it Helps?
What are the Drawbacks?

• A Moderated Panel Discussion - Florida Psychiatric Society

1

DISCLOSURE INFORMATION

Robert Pollack, MD, Ethan Kass, DO, and Laverne D'Silva, DO
Have no relevant financial relationship(s) with ineligible companies
to disclose

2

WHAT IS
AI?

What, Me Worry?



3

AI

- AI should function as a sparring partner for emotional intelligence, rather than a substitute for it.
- The objective is not to prevent AI from helping us think through emotional and interpersonal problems
- AI is to design the interactions so that the human still performs the psychologically important work.
- AI is a scaffolding for emotional and metacognitive development rather than the replacement of the therapeutic or human relationship

4

HISTORY

Era	Milestone
Origins (1940s-1950s)	1954 – Texas Instruments produces the first commercial silicon transistor
Transistors and Integrated Circuits (1950s – 1960s)	1958 – Jack Kilby co-invents the integrated circuit
The Microprocessor Era (1970s)	1971 – The "1971 Trinity" (Apple II, Commodore PET, TRS-80) brings ready-to-use PCs to market
Mainstreaming (1980s)	1981 – IBM launches the IBM PC
Graphical Computing and Networking (1990s)	1995 – Windows 95 unifies consumer GUI computing and rides the internet wave
Broadband, Laptops, and Mobility (2000s)	2001 – Windows XP becomes the long-lived standard as laptops rise
Post-PC Devices and Cloud Computing (2010s)	2010 – iPad launches, accelerating the "post-PC" shift toward mobile/cloud computing
Recent Years (2020s–present)	2020 – Apple introduces its own ARM-based Apple Silicon (M1), marking a major architectural shift

5

ARTIFICIAL INTELLIGENCE

AI Across the Psychiatric Encounter

- 1 Intake & Triage**
Admins gather history, flag urgency, route patients
- 2 Assessment & Diagnosis**
Pattern recognition (EKG, imaging) assists the differential
- 3 Risk Stratification**
Predictive scoring (e.g., AHC-HRSH) prioritizes urgent patients
- 4 Treatment Planning**
Algorithmic suggestions, layered against clinical judgment
- 5 Session Monitoring**
AI/ML on transcripts/speech catches shifts in real time
- 6 Between-Visit Monitoring**
Apps, wearables, text sentiment extend observation
- 7 Documentation**
Ambient scribes close the encounter, no decision role
- 8 Panel Management**
Dashboards surface trends and care gaps across the panel

6

AI PROCESS

A sentence gets broken into **tokens** (words, subwords, or characters), and each token is mapped to a numeric ID from the model's vocabulary.

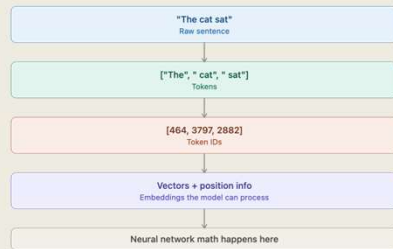
Those IDs are converted into **vectors** (embeddings) that capture meaning, with position information added so word order isn't lost.

These vectors flow through the neural network, where attention and matrix operations figure out **relationships** between tokens and generate a response.

To produce output, the process runs in **reverse**: vectors become token IDs, which become text again.

7

BIRTH OF AN AI



8

How AI Processes Language

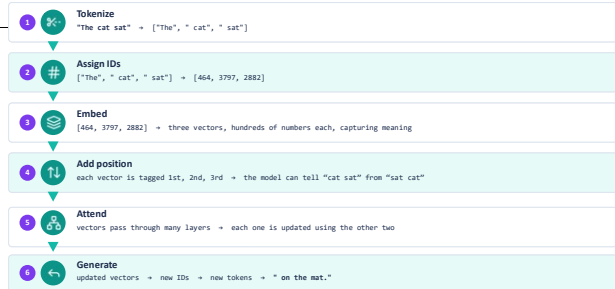
From your sentence to the model's response — the same six-step pipeline happens every time.

- 1 **Tokenize** the sentence splits into small pieces called tokens — whole **words**, word-pieces, or single characters.
- 2 **Assign IDs** each token is looked up in the model's vocabulary and swapped for a **number**.
- 3 **Embed** each ID becomes a vector — hundreds of numbers capturing meaning. **Similar words get similar vectors**.
- 4 **Add position** position information is mixed into each vector, so the model knows **word order**.
- 5 **Attend** vectors flow through the network's layers, where attention works out how every token **relates** to every other.
- 6 **Generate** the process reverses: vectors → IDs → tokens → **readable text**, one piece at a time.

9

The Same Flow, With Real Values

Tracing the sentence "The cat sat" through all six steps.



10

FOR AI

TRUST ARCHITECTURE FOR AI IN HEALTHCARE

1. Transparency before action. **Tell patients** what you'll do and why before you do it.
2. Human-at-the-lever **oversight**. Patients set the boundaries, watch it run, and can stop it anytime.
3. **Personalization** as foundation. Know your patient, or lose their trust.
4. Sparing Partner for Emotional Intelligence, not a replacement

Patients want to feel understood, not processed — **personalization and transparency are what turn an algorithm into something they'll actually trust.**

11

PATIENT INFORMATION (PRIOR TO FIRST VISIT)

1. What brings you in today, in your own words.
2. When did this start and how has it changed over time.
3. Past mental health history, including any prior diagnoses, treatments, or hospitalizations.
4. Current medications and any medical conditions.
5. Anything else going on in your life right now that feels relevant, work, relationships, stress, sleep.
6. What are you hoping to get out of treatment.
7. Any available records that they have to be sent in prior

12

DEMO

1. Sample Pre-Visit Intake form
2. AI process

13

CLINICAL PROCESS

Referral Received
 Intake Information Sent In
 Patient Information entered into AI prior to Appointment
 Surgical Model A synthesis of 25 peer-reviewed studies found AI-assisted robotic surgery cut operative time by 25% and intraoperative complications by 30%, with a 40% improvement in targeting precision and 15% faster recovery.
 Patient Visit (After all previous Information Sent in)
 Follow up care as needed with **AI as part of the process** with each visit
 Cycle Repeats as needed

14

I. Where AI Earns Its Place

MODERATOR ASKS

In one sentence — how would you define AI's role in psychiatric practice, and where's the line you won't cross?

15

FOR AI

A Filing Cabinet vs. a Resident Who Read the Chart

Why "AI" means something categorically different from traditional software.

A TRADITIONAL COMPUTER

Works like a very fast filing cabinet. It stores information in fixed fields — a diagnosis code, a lab value, a medication name — and hands back exactly what you ask for, nothing more. It doesn't notice patterns, and it can't make sense of a clinician's free-text note or a transcribed conversation; anything not already sorted into a box is essentially invisible to it.

AN AI SYSTEM

Works more like a resident who has actually read the whole chart. It can take in messy, unstructured information — notes, transcripts, even tone of voice — and connect it across time and data types, learning what's normal for this specific patient rather than applying a fixed threshold to everyone.

THE KEY DIFFERENCE

A simple computer retrieves what you already put in. An AI system can synthesize a **working impression from scattered pieces** — the way a clinician builds a formulation. That's powerful, but it also means the system's first job should be building an accurate picture of the individual patient, not jumping straight to population-level pattern matching applied downward.

16

I. **Augmentation** of clinical judgment. Not replacement of it.

Why this panel starts from skepticism, not enthusiasm.



THE STARTING POSTURE

Most AI pitches to clinicians overpromise and underdeliver — the burden of proof sits with the tool, not the skeptic in the room.

THE BOUNDARY WE HOLD THROUGHOUT

Augmentation of clinical judgment. Not replacement of it.

17

II. The Case Vignette

MODERATOR ASKS

An 8½-year-old girl: fear of the dark, hypersomnia, and nocturnal urination into the floor vent — she isn't certain it's happening. She locks her door at night; her parents hold the key. Family history: her sister had a similar sleep pattern, and her father has diagnosed sleep apnea. This information is part of the referral or phone call.

18

FOR AI

The AI's Read ..Preparing for the visit

A structured AI summary generated from this same intake — condensed here to its core reasoning.

- Most consistent with sleep-disordered breathing (likely OSA) triggering an NREM parasomnia — Confusional arousal and/or sleepwalking, not primary enuresis.
- Family history (father's sleep apnea, sister's similar pattern) supports a genetic, sleep-disordered-breathing link; her athletic build doesn't rule it out.
- Cites literature: 58–100% of children with chronic NREM parasomnias have sleep-disordered breathing on polysomnography.
- Recommends a sleep history, exam for adenotonsillar hypertrophy, and polysomnography — consistent with AASM guidance.

19

III. Two Tiers of Risk

MODERATOR ASKS

Not all AI in your workflow carries the same risk — which tool in your own practice would you trust the least, and why?

20

FOR AI

Tier 1: Computational Aggregation

Lower-risk AI — organizes and displays data that already exists, without adding new interpretation.

- Examples: PHQ-9/GAD-7 tracking over time; medication-interaction alerts; risk dashboards; note search; trend charts.
- Why it's lower-risk: every output traces back to the exact data behind it — a clinician can check it directly.
- How it fails: stale data, mismatched fields, sampling errors — problems that are easy to catch because they're checkable.
- Oversight needed: like a lab result display — know it's AI-generated, but no individual sign-off required.

21

Against AI

Tier 2: Inferential Suggestion

Higher-risk AI — generates a new synthesis, hypothesis, or judgment that wasn't explicitly in the data.

- Examples: differential diagnosis suggestions; treatment response predictions; inferred risk flags; AI-written summaries from unstructured notes; modified after patient is seen and visit data added to the process
- Why it's higher-risk: there's **no single correct answer to check against** — two clinicians (or two runs of the same model) could disagree.
- How it fails: **automation bias** — the more polished the language, the more it reads as a second clinician's opinion rather than an unverified guess.

Open Evidence and other products **cites sources when responding to queries**

22

When AI Gets It Wrong: Understanding “Hallucinations”

A fluent, confident answer isn't always a correct one — here's what happens and how we manage it.

? What It Is

AI hallucinations happen because the model is built to sound coherent, not to know the truth so it can produce a confident, well-written answer that's wrong. The safeguard isn't avoiding AI. It's the same thing you already do. With junior colleagues' notes, you read it critically and verify when you sign off on it.

⚡ Why It Happens

- **No fact database.** The model predicts the most statistically likely next word — it doesn't look answers up.
- **No confidence meter.** Right and wrong answers are delivered with identical fluency and tone.
- **Gaps get filled.** Obscure facts, exact figures, and citations are the highest-risk areas.

✓ How We Manage It

- **Grounded sources.** Answers pulled from verified references, not open memory.
- **Human review.** Clinical facts and figures are confirmed before use.
- **Flagged uncertainty.** Low-confidence answers are marked, not hidden.
- **Scoped use.** AI stays out of high-stakes, unsupervised decisions.

23

Improve

Hallucinations: Four Ways They Show Up

A working classification for healthcare AI, with psychiatric examples.

Failure Mode	Description	Risk Level
Ambient scribing fabrication	Invented symptoms, severity qualifiers, or risk language not actually stated in the encounter	Highest
Medication/dose errors	Model infers or "corrects" a drug name or dose rather than transcribing it exactly	Highest
Cross-visit blending	Summarization attributes one visit's content to another, corrupting the longitudinal trajectory	High
Template overfitting	Structured fields (MSE, ROS) auto-filled with expected defaults rather than what was actually covered	Moderate

24

MANAGING HALLUCINATIONS

AI in Psychiatry: Two Tracks, One Integration Layer

Where the specialty and the technology actually stand, 2026

Diagnosis

Symptom checklists (DSM/ICD) are giving way to mechanism-level frameworks — network models, HITOP, RDoC, clinical staging — paired with biomarkers and digital phenotyping seeking biological subtypes beneath categorical labels.

Pharmacology

Trial-and-error prescribing is giving way to pharmacogenomic and multi-omic prediction of drug response, alongside early AI-assisted discovery of novel psychotropic targets.

AI's role: the integration layer across both tracks — adjunctive decision support that augments clinical judgement, not a standalone diagnostic or prescribing oracle.

25

IV. Trust Across the Encounter

MODERATOR ASKS

AI touches the encounter at five distinct stages, from seeing the patient to scheduling the next visit — which one carries the most risk if it's wrong, and why?

26

AGAINST AI

IV. Where the Chain Can Break

AI touches five stages — but the risk isn't spread evenly across them.

1

Seeing the Patient

2

Recording the Session

3

Assessing Time Spent

4

Diagnosis & Treatment Planning

5

Scheduling Future Visits

WHERE THE RISK CONCENTRATES

Diagnosis & Treatment Planning carries the most exposure — it's the one stage where an inferential suggestion, taken at face value, changes what actually happens to the patient. Recording the Session is the quiet second: a transcription error can silently corrupt the record everyone downstream trusts.

27

V. Value & Limits

MODERATOR ASKS

What's one task in your week that AI has genuinely given back to you? And where has it quietly let you down?

28

FOR AI

V. Where AI Adds Real Value



- Reduced documentation **burden** → more face time with the patient
- Pattern recognition across longitudinal data no clinician can hold in working memory across a full caseload
- Standardized screening administration and scoring — without replacing clinical interpretation of the result

EXAMPLE FROM THE FIELD

EEG topographic mapping tools and HRSN-based risk stratification dashboards: real deployments, real tradeoffs — concrete cases for panel discussion rather than abstractions. Also seen in ambient scribing, which cuts documentation time — the top driver of psychiatric clinician burnout — and in pre-visit chart summarization, which reduces cold-start time before the clinician walks in.

29

FOR AI

The Workflow Layer, Making Its Case

The pitch, in the vendor's own words.

- Positions itself as an "intelligence layer" sitting above the EHR — spanning pre-visit prep, in-visit support, post-visit notes/coding, and between-visit patient engagement.
- Pitch: patients and providers "arrive prepared," clinicians "practice medicine, not data entry," and notes/coding are "ready in minutes."
- Broader claim: this connective layer creates continuity across specialists, PCPs, and care teams — an end-to-end closed loop.

30

AGAINST AI

V. Where AI Falls Short — Guardrails for the Room



- Aggregate or population-level models **mis-applied** to the individual patient in front of you
- Loss** of clinical narrative when the encounter is **reduced to structured** fields
- Bias in training data — particularly for underrepresented populations (veterans, minors, chronic pain / psych-urology overlap)

Hallucinations and rising error rates compound with over-reliance — false confidence in an aggregated read can mean losing the connection with the human patient in front of you.

REGULATORY REALITY CHECK

Most tools pitched as “clinical decision support” are not FDA-cleared as such. Ask any vendor directly: what is this tool’s actual ONC/CMS/FDA classification, and who is liable if it’s wrong?

31

VI. Evaluating a Real Pitch

MODERATOR ASKS

One vendor’s pitch for closing the loop across the visit — presented here as a discussion prompt, not an endorsement.

Is this the augmentation model we’ve been describing — or does it edge toward replacing the clinician’s own synthesis?

32

AGAINST AI

VI. Before You Sign, Run the Rubric

The same four questions, aimed straight at the pitch you just heard:

1

Does it build an individual patient profile before it aggregates?

2

Does it preserve or erode clinical judgment at the point of decision?

3

Is it transparent about its regulatory status?

4

What happens to the data, and who is liable if it’s wrong?

33

VII. DEEP DIVE

WHERE AI CAN Fall Short

Three failure modes worth walking through in detail.

Hallucinations • Algorithmic Bias • Ethical Concerns

34

VII. HALLUCINATIONS

DISCUSSION PROMPT

Before the classification —

How common are hallucinations in healthcare AI, really?

35

DEEP DIVE

Algorithmic Bias

Bias can enter before the model ever makes a prediction — through what it's trained and measured on.

- **Training data:** who's represented depends on who could access care in the first place. "The extent to which people have access to healthcare will predict their inclusion in ML training data."
- **Benchmarks:** the yardstick itself can encode bias. One widely used algorithm used cost as a proxy for healthcare need — and underestimated the needs of Black patients relative to White patients with similar conditions.

36

DEEP DIVE

Clinician-Related Biases

The bias isn't only in the model — [it's in how clinicians use it](#).

Overconfidence: incorrect AI advice measurably hurt radiologists' mammogram reads at every experience level — less experienced readers were more likely to follow it.

- **Feedback loops:** when clinicians accept a wrong AI suggestion, the system can relearn from that acceptance and repeat the same mistake.
- **Smoothed-over notes:** AI-generated documentation can iron ambiguous or affect-laden language into tidy clinical prose — losing the tone, hesitation, and incongruence that carry real diagnostic weight.

37

DEEP DIVE

Ethical Concerns

Two questions that don't have tidy technical answers.

What it can't see: AI can generate a note, but it can't physically examine a patient or reliably construct a genuine Mental Status Exam from text alone. AI is at its best when sufficient data is provided. MSE not a stand alone. Data should be added after each session and AI re-run and assessed.

- **Who's responsible:** when an AI-influenced decision goes wrong, liability doesn't have a settled answer yet — the clinician, the vendor, and the health system all have a claim to some of it.

38

DEEP DIVE

Case in Point: Does AI Documentation Actually Help?

A longitudinal study on AI-powered ambient clinical documentation.

STUDY FINDING

AI-powered ambient clinical documentation software has been promoted as a way to ease the outpatient documentation burden. But this longitudinal study found the tool did not make clinicians, as a group, more efficient — future work should look at whether specific clinician subgroups or alternative implementations fare better.

39

DEEP DIVE

References

Sources behind this section.

Castillo A. Understanding and addressing AI hallucinations in healthcare and life sciences. *JHS*. 2024;7(3):1–11. doi:10.47941/jhs.1862

Basu S, Huynh B. Mitigating hallucinations in healthcare AI: a systematic review of evidence-based strategies. *BMC Health Serv Res*. 2025;24(1):1115. doi:10.1186/s12913-025-14801-1

Berger K, Gundersen T, Henden S, Romsdahl S. Can medical algorithms be fair? Three ethical quandaries and one dilemma. *BMI Health Care Inform*. 2022;29(1):100445. doi:10.1136/bmjpho-2021-100445

Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447–453. doi:10.1126/science.aba2392

Ueda S, et al. Fairness of artificial intelligence in healthcare: review and recommendations. *npj J Radiol*. 2024;4(2):1–15. doi:10.1007/s11606-023-01474-3

Hussain SA, Brenathan M, Zhuang J. The bias algorithm: how AI in healthcare exacerbates ethnic and racial disparities — a scoping review. *Ethnicity & Health*. 2025;30(2):197–214. doi:10.1080/13557858.2024.2422648

Crisis JL, Chuma MA, Oduro JY. Bias in medical AI: implications for clinical decision-making. *PLoS Digit Health*. 2024;3(11):e0000651. doi:10.1371/journal.pdig.0000651

Deutsch T, et al. Automation bias in mammography: The impact of artificial intelligence BI-RADS suggestions on reader performance. *Radiology*. 2023. doi:10.1148/radiol.222176

Angus DC, Khara R, Liu T, et al. AI, Health, and Health Care Today and Tomorrow: The JAMA Summit Report on Artificial Intelligence. *JAMA*. 2023;324(18):1650–1664. doi:10.1001/jama.2023.18490

40

Closing Thought

Two lines from tonight's prep notes to send the room out on.

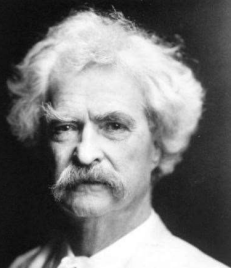
"The tool should clear space for judgment — not replace it."

"At some point in the future, something that is currently considered a high-level skill will be the baseline for everyone."

41

"I've had a lot of worries in my life, most of which never happened."

— Mark Twain



42
