UNIVERSITY OF TWENTE.

AI IN MEDICINE AND PSYCHIATRY

Y. J. Erden, Philosophy

What do you already know about AI in relation to medicine / psychiatry?

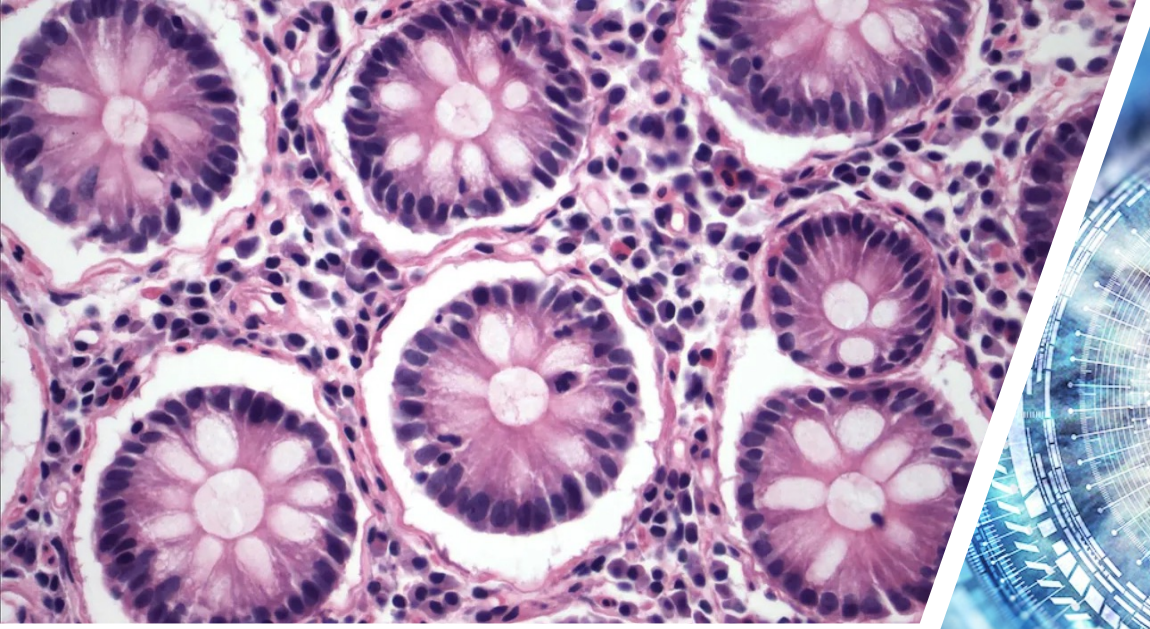
Did you finish the reading?

Thoughts / comments?

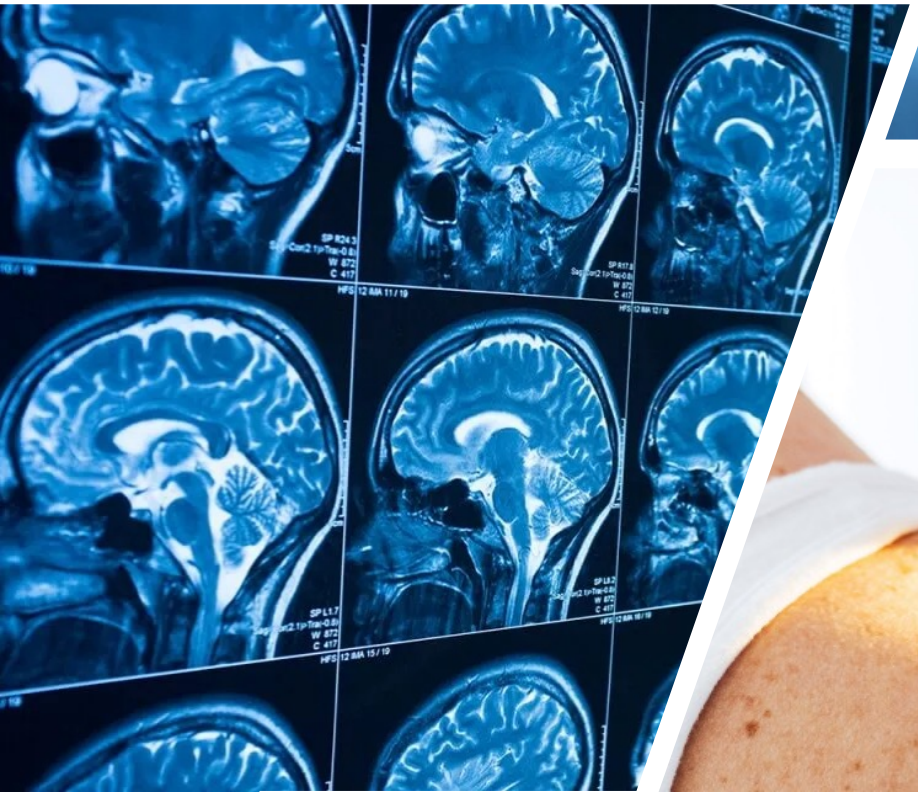

Why might
iCub
struggle
to learn
to share?

Outline

- AI for medicine: making 'good decisions'
- The language of psychiatry
- AI in and for psychiatry
- Some core principles



Imaging



Do you notice the image again?

AI for medicine

- **Classifying text:** clinician notes & observations to identify common findings, approaches, mistakes, inefficiencies + recommendations for amending protocols
- **Data-driven AI:** distinguish between neuroimages, e.g. 'neuroimaging biomarkers' using MRI, to identify brain disorders
- **Data analysis models:** operating on established question-based results from diagnostic tools

cf. Rainey, Erden, & Resseguier, 2021

Making 'good decisions'

(Grote and Berens, 2020)

Good decisions in psychiatry

- More than being *right*…
- Lived experience
- Complex realities
- Can emerge discursively

‘Get Cape. Wear Cape. Fly’ By muffincopter

Three philosophical challenges

(Chin-Yee & Upshur, 2019, p. 237)

(1) Epistemological-ontological problem (theory-ladenness of big data / measurement)

(2) Epistemological-logical problem (algorithm limitations, (un)reliability, interpretability)

(3) Phenomenological problem (irreducibility of human experience to quantitative data)

(1) Epistemological-Ontological Problem (p. 238-43)

Algorithms + big data =	Big Data =	Measurement =
• **promise for:** diagnoses, prognoses, integrative analysis; • 'unbiased', 'theory-free'; more data = greater accuracy…; • *predictive abilities; from curing to prevention (Grote & Berens, 2020, p. 205);* • *address clinician flaws: cognitive bias/diagnostic errors (Grote & Berens, 2020, p. 205).*	• **theory-laden;** selected; epistemic interests; • logical positivism; neglects causal reasoning; • black-boxes obscure causal relations (inputs / outputs) • *possibility of accuracy at expense of opacity (Grote & Berens, 2020, p. 205);* • *values of transparency & evidence vs reliability & efficacy (Grote and Berens, 2020, p. 209).*	• **theory-laden;** semantic understandings; selected tools; • constructs ontologies; quantitative / abstract vs family resemblance • can mislead: false concreteness; misdiagnosis; decontextualised 'at hand' data with errors / omissions.

(2) Epistemological-Logical Problem (p. 244-47)

Limits of logic and inference

- complex realities: phenomenal, biological, social, psychological, historical
- dynamic complex systems: past empirical success ≠ future success
- programs remain incomplete; underdetermine complexity in models; can't account for rare or idiosyncratic factors
- CACE principle: 'Changing Anything Changes Everything'
- question of who assumes responsibility
- *distributed / collective responsibility + legal system? (Grote & Berens, 2020, p. 209)*

(3) Phenomenological Problem (p. 248-52)

Phenomenology / lived experience / being-in-the-world:

- patients experience; health-care research; clinical reasoning / judgment / interpretation
- data = (dynamic) meaning: can be missed / context can exacerbate epistemic injustice
- key content may be implied, may emerge through dialogue / interpretation
- deep learning cannot easily adapt or cope with structural change; big data does not change this

Dialogue is essential. Can AI be part of that dialogue?

EPISTEMIC AUTHORITY (OF CLINICIANS)

Current ML algorithms challenge epistemic authority of clinicians (Grote & Berens, 2020, p. 207)

Image: Trikle Trade, *Stories Of People Helping Others*

Epistemic models

Clinician:

- tests hypotheses; full certitude unattainable; variety of information gathering activities & treatments; takes calculated risks
- collaborative medical diagnoses (patients, peers); (time) constraints

Machines:

- rule-based algorithms: encoding expert knowledge
- ML algorithms: extracting patterns or structure in examples (p. 206)
- complexity impacts transparency; management of risks
- exert normative force: evidential standards
- algorithms as additional source of evidence?

Foundations for clinical decision making

(Grote & Berens, 2020; Chin-Yee & Upshur, 2019)

solitary reasoning: self-evaluation; cognitive bias; overconfidence; limits to knowledge; unreflective value judgments (disease definitions; treatment rationale)

collaborative reasoning: spotting errors; (equally) competent peers disagree on a proposition; explanation / justification for reasons, decisions, action; accountability for judgments

stalemate: 'equal weight view' (different views diminish confidence) vs. 'steadfast view' (own beliefs given epistemic privilege)

physicians and phronesis *(practical wisdom): flexible, interpretive capacity: context + best action (Chin-Yee & Upshur, 2019)*

ML + clinician decision making: dialogical engagement (p. 209)

- **ML increases opacity / uncertainty**: (1) corporate secrecy, (2) technical illiteracy, (3) complex (unintelligible) representation
- **Risks**: different training / reasoning; epistemic vices (dogmatism, gullibility); paternalistic decision-making / undermining shared decision-making; human confidence ≠ mathematical confidence (ML typically overconfident)

Two key questions

What does it mean to make good decisions?

- explanation / justification + reasons, decisions, action
- accountability for judgments

Can AI / ML be a valuable participant in clinical dialogue?

- contribution to solitary / collaborative reasoning
- impact of different reasoning / confidence etc.

Psychiatry

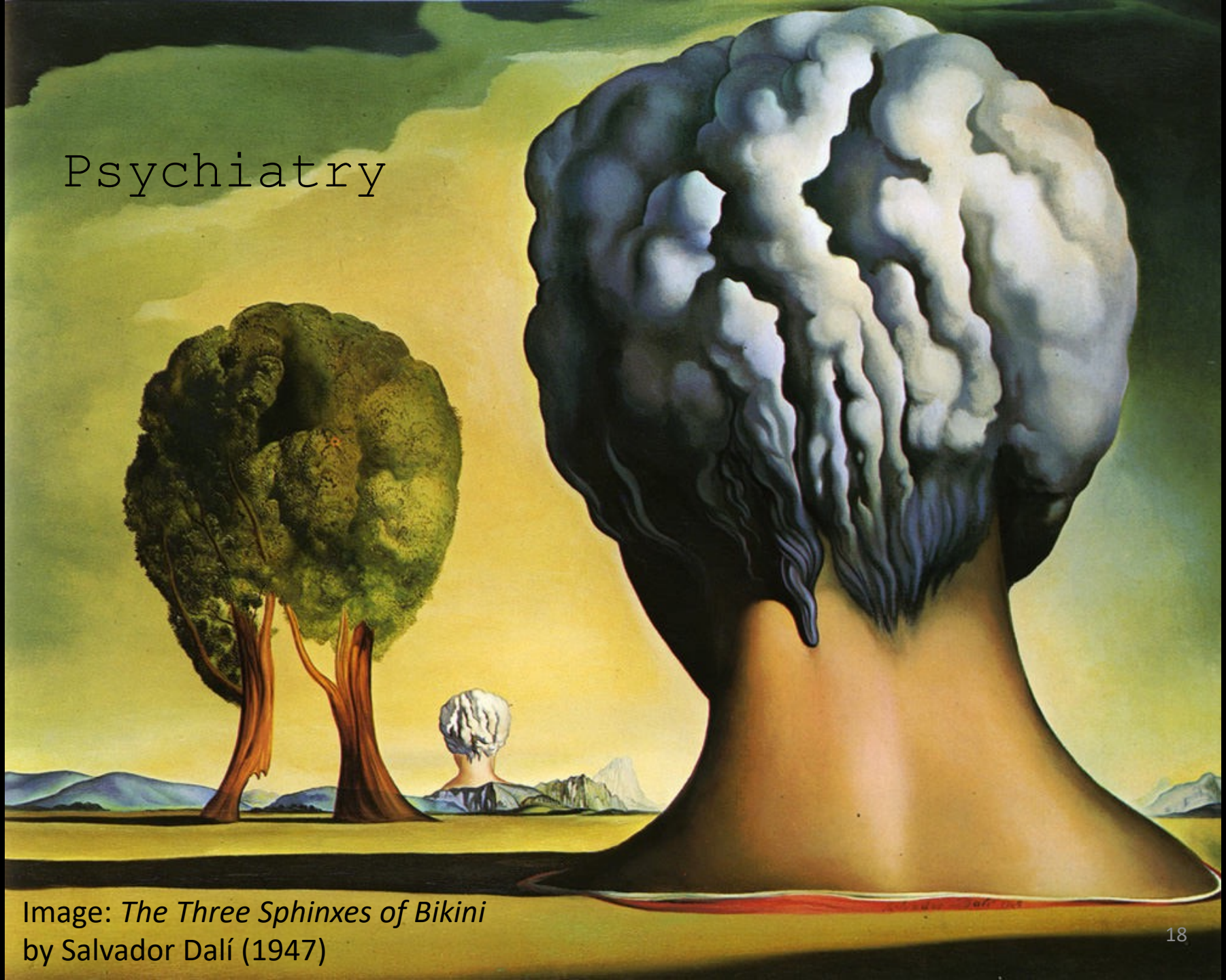


Image: *The Three Sphinxes of Bikini*
by Salvador Dalí (1947)

Definition

CLASSIFICATION
CATEGORISATION

Image URL

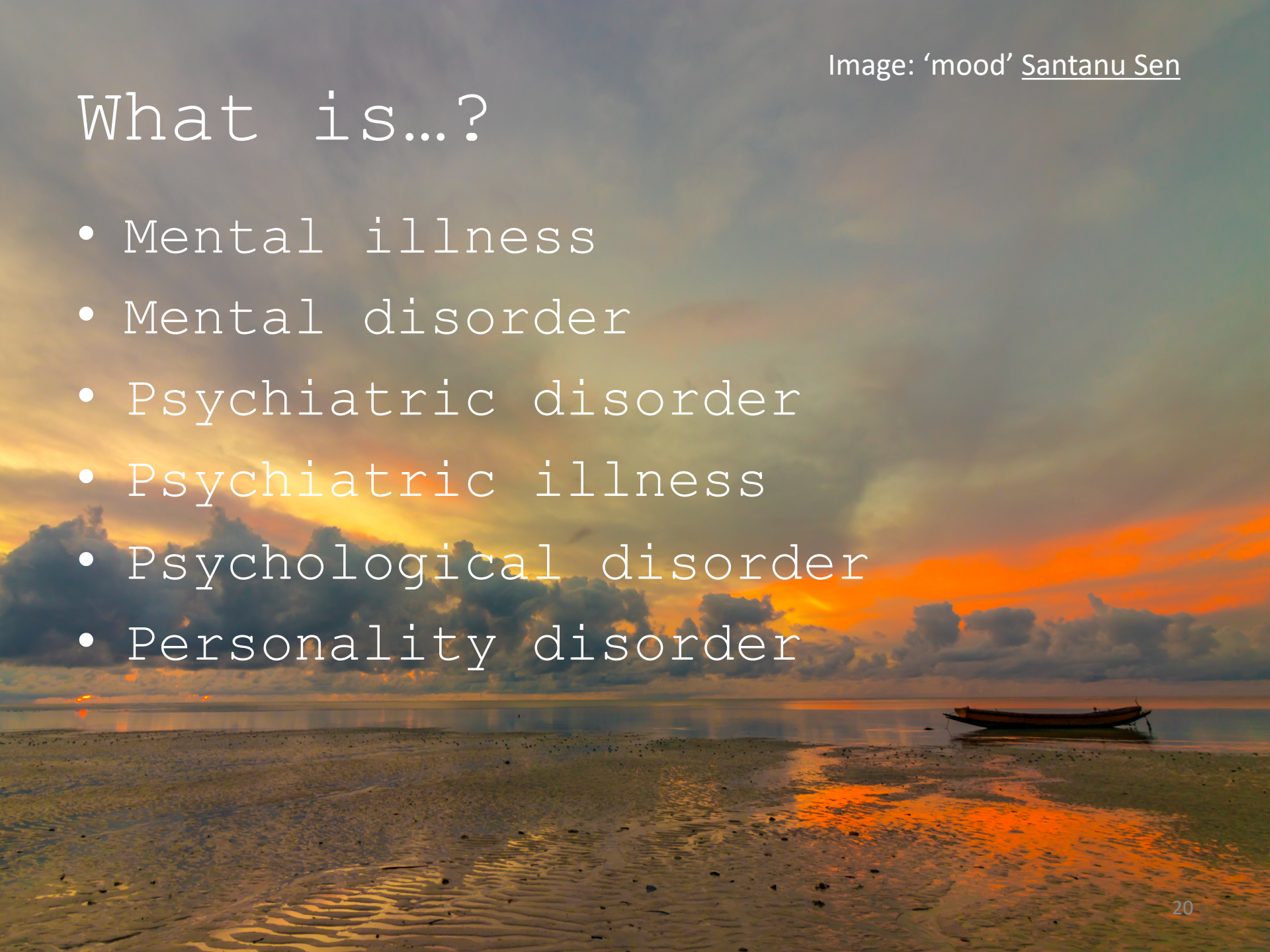


Image: ‘mood’ Santanu Sen

What is…?

- Mental illness
- Mental disorder
- Psychiatric disorder
- Psychiatric illness
- Psychological disorder
- Personality disorder

Image: ‘mood’ Santanu Sen

What is…?

- Mental illness
- Mental disorder
- Psychiatric disorder
- Psychiatric illness
- Psychological disorder
- Personality disorder = as above?
- Brain disorder = neurological? (e.g. epilepsy, Parkinson’s disease?)

Section II: Diagnostic Criteria and Codes

Neurodevelopmental Disorders

Intellectual Disabilities
Intellectual Disability (Intellectual Developmental Disorder)
Global Developmental Delay
Unspecified Intellectual Disability (Intellectual Developmental Disorder)

Communication Disorders
Language Disorder
Speech Sound Disorder (previously Phonological Disorder)
Childhood-Onset Fluency Disorder (Stuttering)
Social (Pragmatic) Communication Disorder
Unspecified Communication Disorder

Autism Spectrum Disorder
Autism Spectrum Disorder

Attention-Deficit/Hyperactivity Disorder
Attention-Deficit/Hyperactivity Disorder
Other Specified Attention-Deficit/Hyperactivity Disorder
Unspecified Attention-Deficit/Hyperactivity Disorder

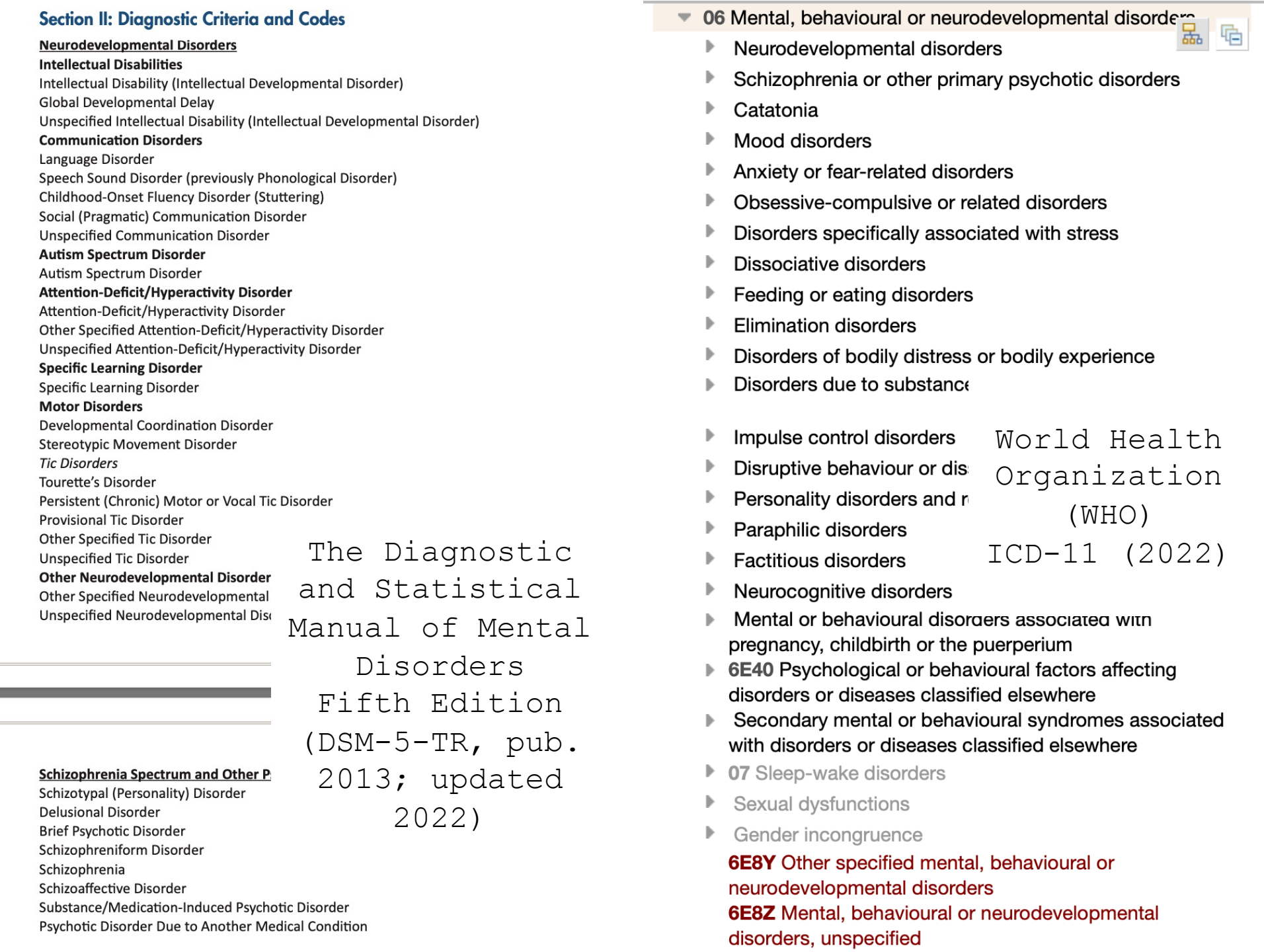
Specific Learning Disorder
Specific Learning Disorder

Motor Disorders
Developmental Coordination Disorder
Stereotypic Movement Disorder
Tic Disorders
Tourette's Disorder
Persistent (Chronic) Motor or Vocal Tic Disorder
Provisional Tic Disorder
Other Specified Tic Disorder
Unspecified Tic Disorder

Other Neurodevelopmental Disorder
Other Specified Neurodevelopmental
Unspecified Neurodevelopmental Dis

The Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5-TR, pub. 2013; updated 2022)

Schizophrenia Spectrum and Other P
Schizotypal (Personality) Disorder
Delusional Disorder
Brief Psychotic Disorder
Schizophreniform Disorder
Schizophrenia
Schizoaffective Disorder
Substance/Medication-Induced Psychotic Disorder
Psychotic Disorder Due to Another Medical Condition



World Health Organization (WHO) ICD-11 (2022)

Main criticisms of DSM-5:

- 'an unhealthy influence of the pharmaceutical industry on the revision process'
- 'an increasing tendency to "medicalise" patterns of behaviour and mood that are not considered to be particularly extreme'

NHS analysis

Do theories in psychology stand up? e.g. “social priming”

World J Psychiatry. 2014 Dec 22; 4(4): 133–140.
Published online 2014 Dec 22. doi: 10.5498/wjp.v4.i4.133
PMCID: PMC4274585
PMID: 25540728

Racial disparities in psychotic disorder diagnosis: A review of empirical literature

Robert C Schwartz and David M Blankenship

Professional Psychology: Research and Practice
1994, Vol. 25, No. 1, 55-61

0735-7028/94/$3.00

Sex Bias in the Diagnosis of Borderline Personality Disorder and Posttraumatic Stress Disorder

Dana Becker and Sharon Lamb

AI in psychiatry

Can improve accuracy, consistency, efficiency, effectiveness

Analysis of large data sets to recognise key phenotypes (machine learning / neural networks)

Objective systems can help avoid clinician bias

Prediction capacity from theory + data

Fill gaps: (1) specialists (2) knowledge or uncertainty (3) in cases of novelty (4) second opinions

cf. Erden, Hummerstone, & Rainey, 2021

AI and…

- Meaningful language use / understanding
- Emotions/desires
- Autonomous intentionality
- Qualia (what it is *to be* something)

MultiSense

SimSensei

SimSensei nods at appropriate points in the dialogue.

Article URL

"a future AI…"

Technical: data, analysis, prediction, e.g. efficient diagnosis/treatment-planning using data-driven taxonomies of mental illness

Relational: conversation, support, understanding, e.g. via individualised, non-judgemental AI

Image: ‘Depression’ by mattwalker69

What would an AI need to know in order to distinguish between clinical depression, and bereavement?

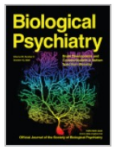
Article | Open Access | Published: 05 May 2022

Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder

Jonathan T. Megerian, Sangeeta Dey, Raun D. Melmed, Daniel L. Coury, Marc Lerner, Christopher J. Nicholls, Kristin Sohl, Rambod Rouhbakhsh, Anandhi Narasimhan, Jonathan Romain, Sailaja Golla, Safiullah Shareef, Andrey Ostrovsky, Jennifer Shannon, Colleen Kraft, Stuart Liu-Mayo, Halim Abbas, Diana E. Gal-Szabo, Dennis P. Wall & Sharief Taraman

ELSEVIER

Biological Psychiatry
Volume 92, Issue 8, 15 October 2022, Pages 643-653



his article

Archival Report

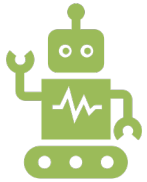
Robust, Generalizable, and Interpretable Artificial Intelligence–Derived Brain Fingerprints of Autism and Social Communication Symptom Severity

Kaustubh Supekar [a 1], Srikanth Ryali [a 1], Rui Yuan [a], Devinder Kumar [a], Carlo de los Angeles [a], Vinod Menon [a b c]

AI weaknesses?

AI lacks nuance and flexibility

AI struggles with discursive, contextual analysis



AI requires simplified (simplistic) classifications

‘This ability of human intelligence to draw on **‘common sense’** when needed, and **uncommon sense** when especially needed (i.e. that of specialists with their unique training data), means that humans will almost certainly remain critical for managing and interacting with even relatively simple psychiatric cases for the foreseeable future.’ (Christian Brown et al., 2021, p. 132)

Core principles: AI for psychiatry

1. AI lacks capacities crucial to psychiatry: nuance / flexible thinking about mental states (beliefs, desires, intentions, preferences, needs), plus context / experience
2. AI for psychiatry should not undermine necessary relational aspects of care, especially where technological fixes seem to offer respite for resource heavy fields
3. Statistical data / analysis include inferences, while models of psychiatry are located within (shifting) normative frameworks. AI should not cement simplistic classifications or exacerbate harmful biases (design, data), and care should be taken in the selection of theories of mind, brains, and human behaviour. AI systems need to be flexible enough to adapt as these theories likewise adapt, or to change tack where theories are discredited or papers retracted
4. Brain data needs particular scrutiny given potential to bypass self-reporting / interpersonal, discursive methods
5. AI should be sufficiently transparent, with methods, processes, (brain) data sets, including for training, open to critique

References

- Becker, D., & Lamb, S. (1994). Sex bias in the diagnosis of borderline personality disorder and posttraumatic stress disorder. *Professional Psychology: Research and Practice*, 25: 1, 55.
- Duncan, B. L. and Reese, R. J. (2012). Empirically Supported Treatments, Evidence-Based Treatments, and Evidence-Based Practice. In G. Stricker & T. Widiger (Eds.), *Handbook of psychology: Clinical psychology* (2nd ed., pp. 977–1023). Hoboken, NJ: Wiley.
- Erden, Y. J., Hummerstone, H. & Rainey, S. (2021). Automating autism assessment: what AI can bring to the diagnostic process. 'Philosophy Thematic Issue', *Journal of Evaluation in Clinical Practice*, 27: 3, 485-490.
- Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data (ed. Tom Tregenza). *PLoS ONE* 4: 5, e5738.
- Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine* 2: 8, e124.
- Kirmayer, L., Crafa, D. (2014). What kind of science for psychiatry? *Frontiers in Human Neuroscience*, 8(435), 1-12.
- Larsson, S. & Heintz, F. (2020). Transparency in artificial intelligence. Internet Policy Review, 9: 2.
- Open Science Collaboration. (2015). Estimating the Reproducibility of Psychological Science. *Science* 349: 6251, aac4716.
- Rainey, S. & Erden, Y.J. (2020). Correcting the Brain? The Convergence of Neuroscience, Neurotechnology, Psychiatry, and Artificial Intelligence. *Sci Eng Ethics* 26**,** 2439–2454.
- Rainey, S. (2022). Datafied brains and digital twins: Lessons from industry, caution for psychiatry. *Philosophy, Psychiatry, & Psychology*, 29: 1, 29-42.
- Regier, D. A., Kuhl, E. A., & Kupfer, D. J. (2013). The DSM-5: Classification and criteria changes. *World psychiatry: official journal of the World Psychiatric Association (WPA)*, 12(2), 92-8.
- Ritchie, S. (2020). *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science*. London: The Bodley Head.
- Schwartz, R. C., & Blankenship, D. M. (2014). Racial disparities in psychotic disorder diagnosis: A review of empirical literature. *World journal of Psychiatry*, 4: 4, 133.
- Straw, I. (2020). The automation of bias in medical artificial intelligence (ai): decoding the past to create a better future. *Artificial Intelligence in Medicine*, *110*, 101965–101965.

Additional slides

The University Of Queensland, Queensland Brain Institute

Brain diseases and disorders

/ The Brain

QBI scientists are working to better understand brain function using a strategically balanced suite of scientific models and advanced technologies. Researchers are also studying brain function in the mouse, honeybee, fruit fly, frog, zebra fish and flatworm in an effort to address fundamental human neurological questions.

The Institute's neuroscientists are exploring the developmental causes of, and possible therapeutic treatments for, a wide range of mental and neurological diseases and disorders.

ADHD	Anxiety	Autism
Brain tumour	Concussion	Childhood brain disorders
Dementia	Depression	Motor Neurone Disease
Obsessive Compulsive Disorder	Schizophrenia	Stroke
Traumatic Brain Injury		

National Institutes of Health (Biological Sciences Curriculum Study, US, 2007) URL

‘There are many different **mental illnesses**, including depression, schizophrenia, attention deficit hyperactivity disorder (ADHD), autism, and obsessive-compulsive disorder’

To consider

- What is *rational* in decision making?
- Data quality + likelihood of appropriate outputs?
- How to differentiate, e.g. overlapping / co-morbid conditions?
- Scope for flexibility, e.g. new data/patterns/tasks?
- Scope for certainty in categorisation?

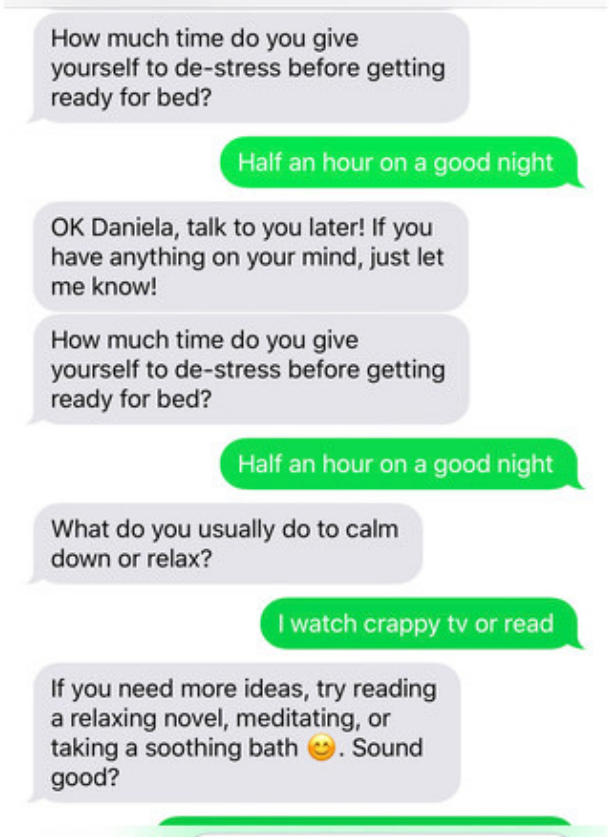
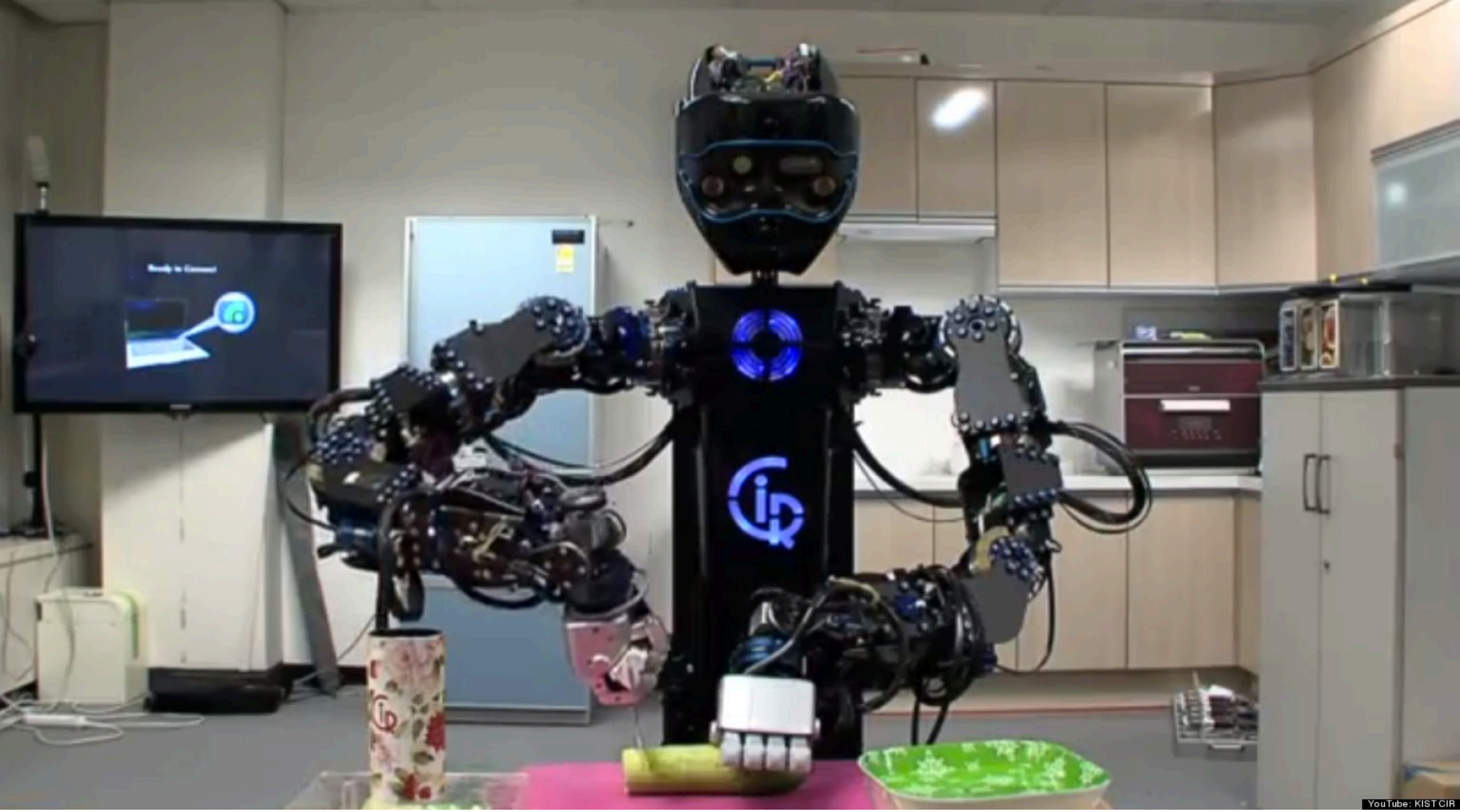


Image URL

PICTURE YOUR KITCHEN

Exactly as it was this morning…

Now consider how to make a cup of coffee



How to make a salad

THE FRAME PROBLEM

Understanding each other?

Gif URL

Similar to naïve physics?

Gif URL

Is understanding of others…

…intentional and directed?