Australian National University
KIT Karlsruher Institut für Technologie
SAARLAND UNIVERSITY

ARTIFICIAL INTELLIGENCE IN HEALTH CARE

Shirley Gregor, Australian National University

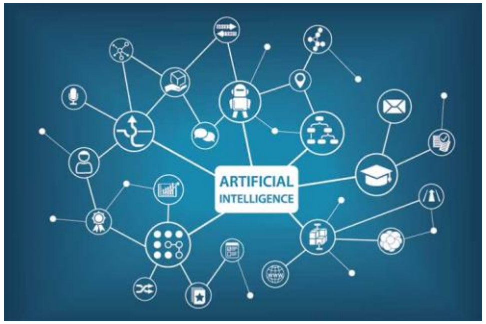
Alexander Maedche, Karlsruhe Institute of Technology

Stefan Morana, Saarland University

20 October 2020

Introduction

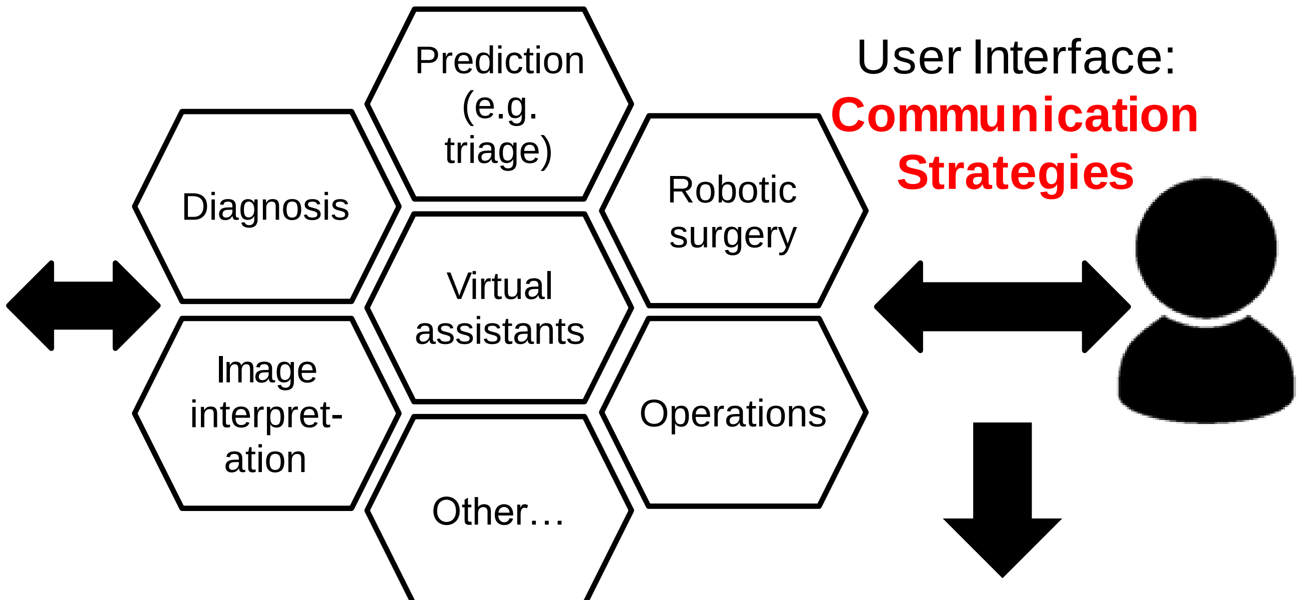
- Artificial Intelligence (AI) dates from 1950s
- Rapid growth in “AI Spring” since around 2010 with better technologies, very large amounts of data and more widespread use of technology
- Healthcare a very important application area
- Older systems more rule-based and algorithmic
  - Many Clinical Decision Support Systems effective
  - Transparency – can often explain reasoning
- Newer systems use Machine Learning
  - Machine learns itself from data
  - Concerns about accuracy, bias, privacy, lack of transparency and trustworthiness
  - Need for Explainable AI (XAI)



AI in Health Care

AI Technologies

- Natural Language Processing
- Computer Vision
- Machine Learning
- Knowledge Representation & Reasoning
- Robotics
- other



AI-based Systems

User Interface:
Communication Strategies

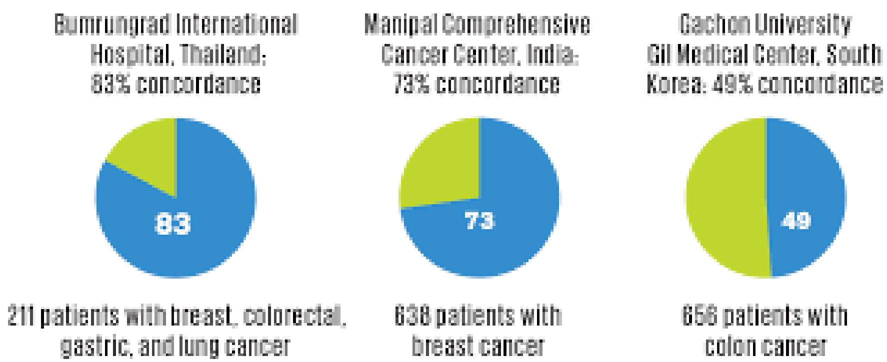
Outcome Goals
(affective, behavioural, cognitive)

Russell & Norvig (2016)

Sung et al (2020); Wang et al (2019)

Concerns with Machine Learning

Watson for Oncology "oversold" [1]



Discrimination By Artificial Intelligence In A Commercial Electronic Health Record [2]

- Commercial provider - predicts "no-shows" based on EHR
- Potential for explicit discrimination - ethnicity, financial class, religion, and body mass index

Source: [1] Strickland (2019) Jacklevic (2017); [2] Murray et al (2020)

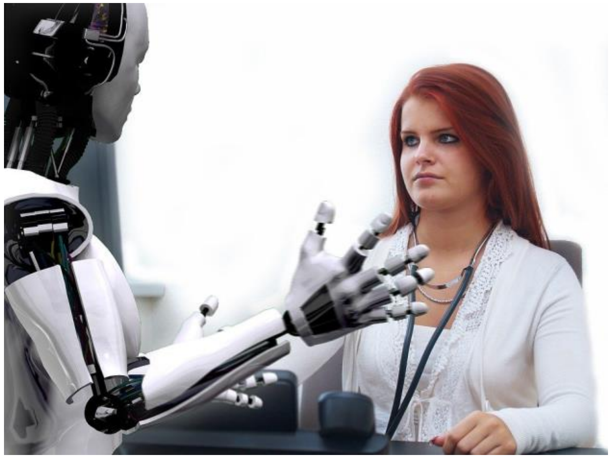
Our focus

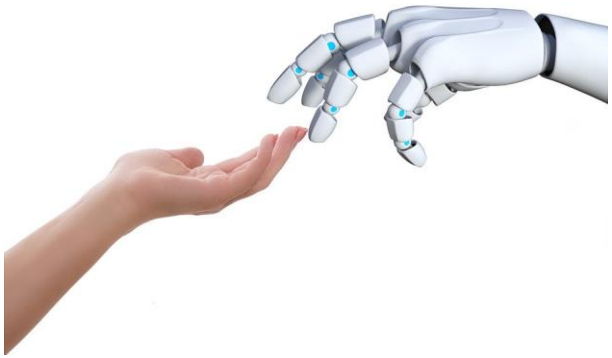
The design of the interface for AI influences trustworthiness and effectiveness – but can also be manipulative

- We propose a framework for **communication strategies** to achieve desired goals and aid assimilation
- New ODiSAI theoretical framework (Gregor, Maedche, Morana 2020)
  - strategies adapted from Habermas's (1984) *Theory of Communicative Action*.
  - based on theory and prior research
  - considers performance, transparency, & ethical concerns about manipulation/influencing
  - When are explanations needed?
- Note: In addition to interface design, good practice should be followed overall in design, implementation and governance and ethics guidelines followed, leading to *built-in integrity* (e.g. see OAI 2020)

ODiSAI Strategies – Giving Advice

1. Instrumental
   – *Take my advice because it works.*
2. Strategic
   – *Take my advice (it works for me and maybe for you).*
3. Expressive
   – *Take my advice because I am well-intentioned and sincere.*
4. Normative
   – *Take my advice because it is the right thing to do/others in your community do it.*
5. Communicative Action (ideal)
   – *Take my advice because it works, I am sincere, it is right, and I have given well-founded agreed-upon reasons (explanations) to show it is justified.*





1. Instrumental Strategy

Use to achieve task outcomes of efficiency, accuracy and effectiveness. Include capabilities congruent with accepted principles of human-computer interaction.

- The instrumental strategy applies across a wide range of applications and AI technologies.
- The user should find the system easy-to-use and advice should be comprehensible (e.g. terms/language should be understandable, terminological explanations can be used - as in all strategies)
- Can suit when concerns about cognitive load of user if explanations given

Case: Allegheny Family Screening Tool

- Response to perceived lack of duty of care in child protection
- Tool carefully developed in consultation with stakeholders and transparent processes
- Accuracy evaluated in independent studies & tool updated
- Care to avoid racial bias
- Uses data mining to search for patterns
- Predictive analytics using LASSO method. Algorithm available.

- Gives a screening score from 1 to 20 for a child on a call
- Screeners choose (sometimes) whether an investigation needed
- Screeners originally felt threatened but now find valuable
- No explanation given of key factors in individual case
- Report found more accurate than alternatives, so unethical not to use

Sources: Center for Public Impact (2018); Allegheny County (2020) https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx

- The **screening score** is from 1 to 20 (with a subset of referrals being mandatory screen ins)
- The **higher the score, the higher the chance of the future event** (e.g., abuse, placement, re-referral) according to the data

Image source: https://www.muckrock.com/news/archives/2019/jul/10/algorithms-family-screening-Pennsylvania/

2. Strategic Strategy

Use to achieve biased or manipulated decision making, based on theories of persuasion and cognitive bias. Include capabilities such as user profiling, persuasion techniques, leverage of human cognitive biases, and, on occasion, deception.

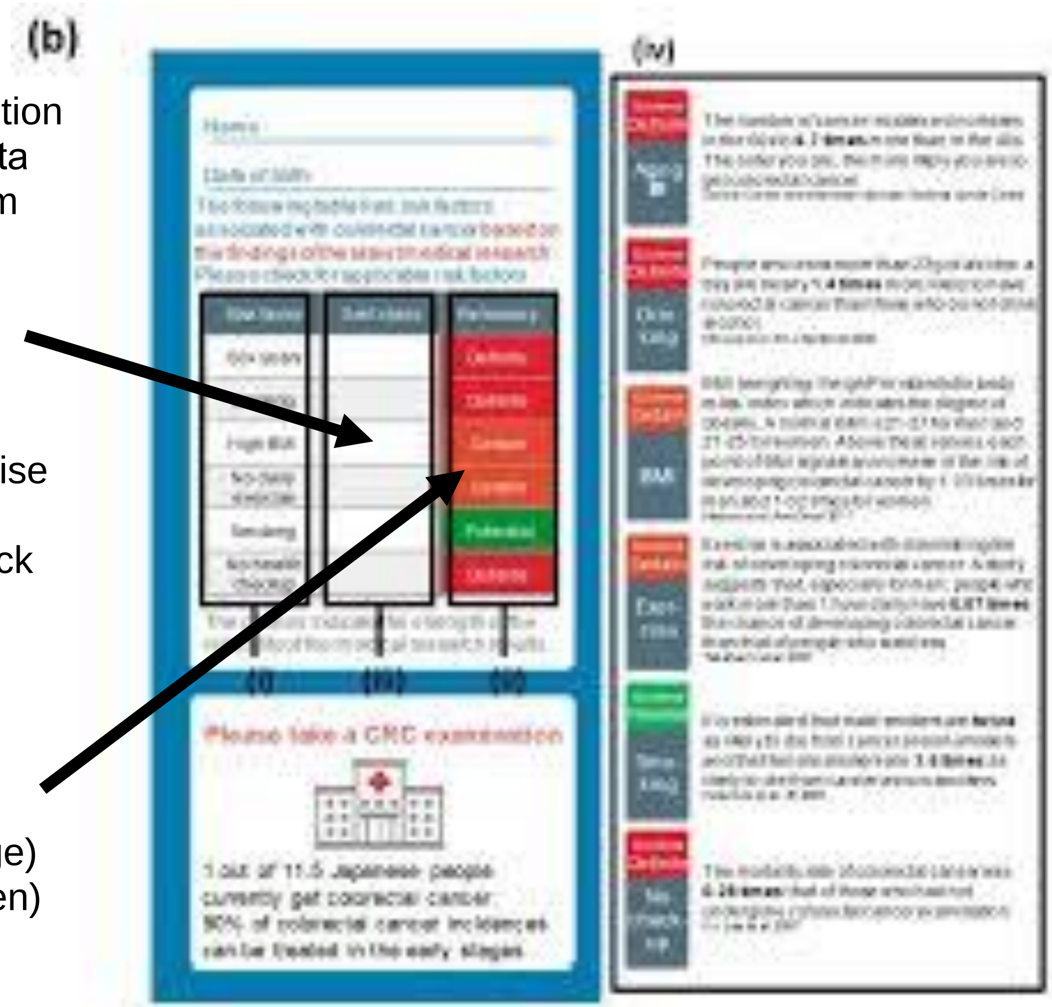
- Persuasion can occur through non-cognitive means eg nudges.
- Much work in this area for recommendation agents.
- Deception not recommended. Should be aware it can happen.

Nudging towards health check

- City of Hachioji, Tokyo
- Application of machine learning and **nudge theory**
- Data obtained from designated periodical health examinations, digitalized medical insurance receipts, and medical examination records for colorectal cancer
- Deduce segments for whom the examination was recommended
- Some messages sent with personalized risk factors

- Uptake rate for colorectal cancer examinations significantly increased
- Are there privacy concerns?

Misawa et al (2020)



Customized condition had recipient’s data from medical exam filled in for 6 risk factors;

- Age
- Drinking
- High BMI
- No daily exercise
- Smoking
- No health check up

Risk level:

- Definite (red)
- Certain (orange)
- Potential (green)

3. Expressive Strategy*

Use to achieve affective outcomes such as emotional trust, based on theories such as CASA. Include capabilities such as the adoption of human-like characteristics (e.g. a persona, social presence) and use of appropriate social cues.

- CASA is Computers are Social Actors Theory (Nass et al 1994)
- This strategy can promote trust by adopting a benevolent, sincere persona.
- The trust could be mistaken as the AI could be insincere
- An avatar is not necessary. A smart watch can exhibit a caring persona by messaging, so the user feels attached to it.
- This strategy can have some aspects of the strategic strategy.

(*In German, *bewährungsverpflichungen,* parole-obligation)

Digital Avatar
Sophie
(iDAvatars Inc)

Para the harp seal robot (Battey 2016)

Robot Caregiver (Riek 2017)

4. Normative Strategy

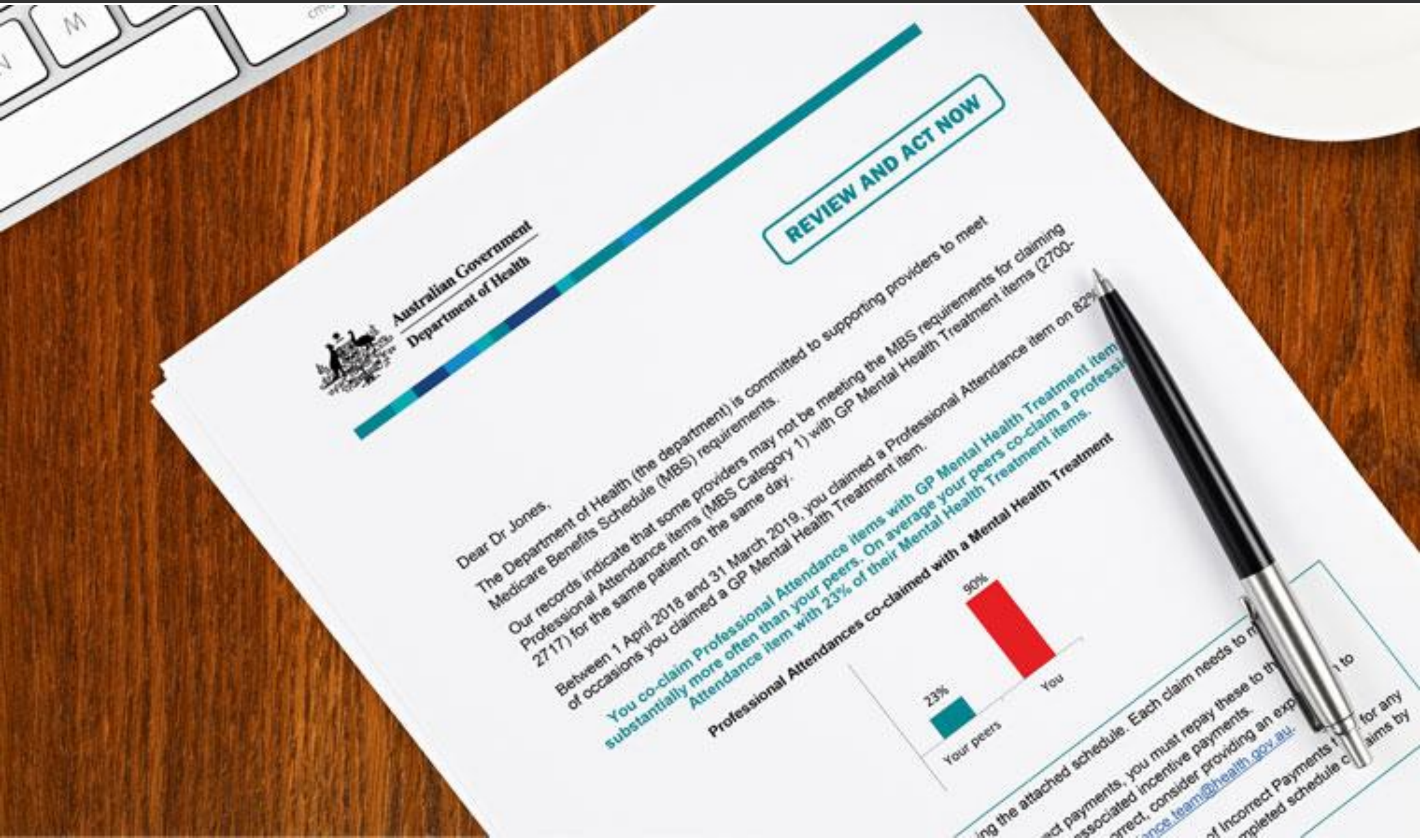
Use to achieve outcomes of compliance with, or enforcement of norms, based on role and social norms theory. Include capabilities such as norm description, norm desirability indicators, behaviour monitoring, critiquing, censoring, guidance, feedback and encouragement.

- There is the implication that one “should” or “ought to” take the action advised.
- Justification can make reference to well-accepted practice, expected community behaviour.
- This strategy could overlap with the strategic strategy (i.e. nudging)

Case: Department of Health Compliance Audits

- Aust. Dept Health monitors compliance by medical practitioners eg on amounts claimed under the Medicare Benefit schedule
- Uses information from analytics to identify targets
- Sends letters to targets asking them to verify compliance or voluntarily refund over-payments
- Depending on response, issues a *Notice to Produce*
- Currently subject to audit by Australian National Audit Office
- Australian Medical Association submission to ANAO says:
  – Not just about compliance, trying to change behaviour
  – Blunt use of metrics such as 80th percentile of users overlooks individual practice circumstances
  – For opoid over-prescription, letters seen to be ill-targeted and threatening and data questionable
  – In any compliance activity, the messaging that is used should be “sense tested” by clinicians

Sources: Australian Medical Assoc (AMA) 2018, 2019; Dept Health (2020)



Source: https://www1.racgp.org.au/newsgp/professional/these-gps-worked-with-australia-s-most-vulnerable

5. Communicative Action Strategy

Use to achieve outcomes of improved performance, learning and increased trust, based on theory of argumentation, cognitive learning theory and trust theory. Include capabilities such as explanations and conversational style exchanges.

- In many situations this is the “ideal” strategy

Case: Decision Support for Rehabilitation

- Adherence to guidelines in clinical practice is often low
- Decision support system for cardiac rehabilitation (CARDSS) developed (Goud 2009)
- Used in 40 Dutch cardiac rehabilitation outpatient clinics
- Advice rationale (explanations) provided

- Evaluation 2016 (Verheul 2016)
- Improved guideline adherence by increasing user's understanding
- Overcame inertia of prior practice
- Reduced guideline complexity eg in calculations and data interpretation
- Increased patient's willingness to adopt recommendations when these were shared
- Systematic reviews show success of many similar systems (e.g .Moja et al 2014)

Source: https://www.intechopen.com/books/efficient-decision-support-systems-practice-and-challenges-in-biomedical-related-domain/guideline-based-decision-support-systems-for-prevention-and-management-of-chronic-diseases

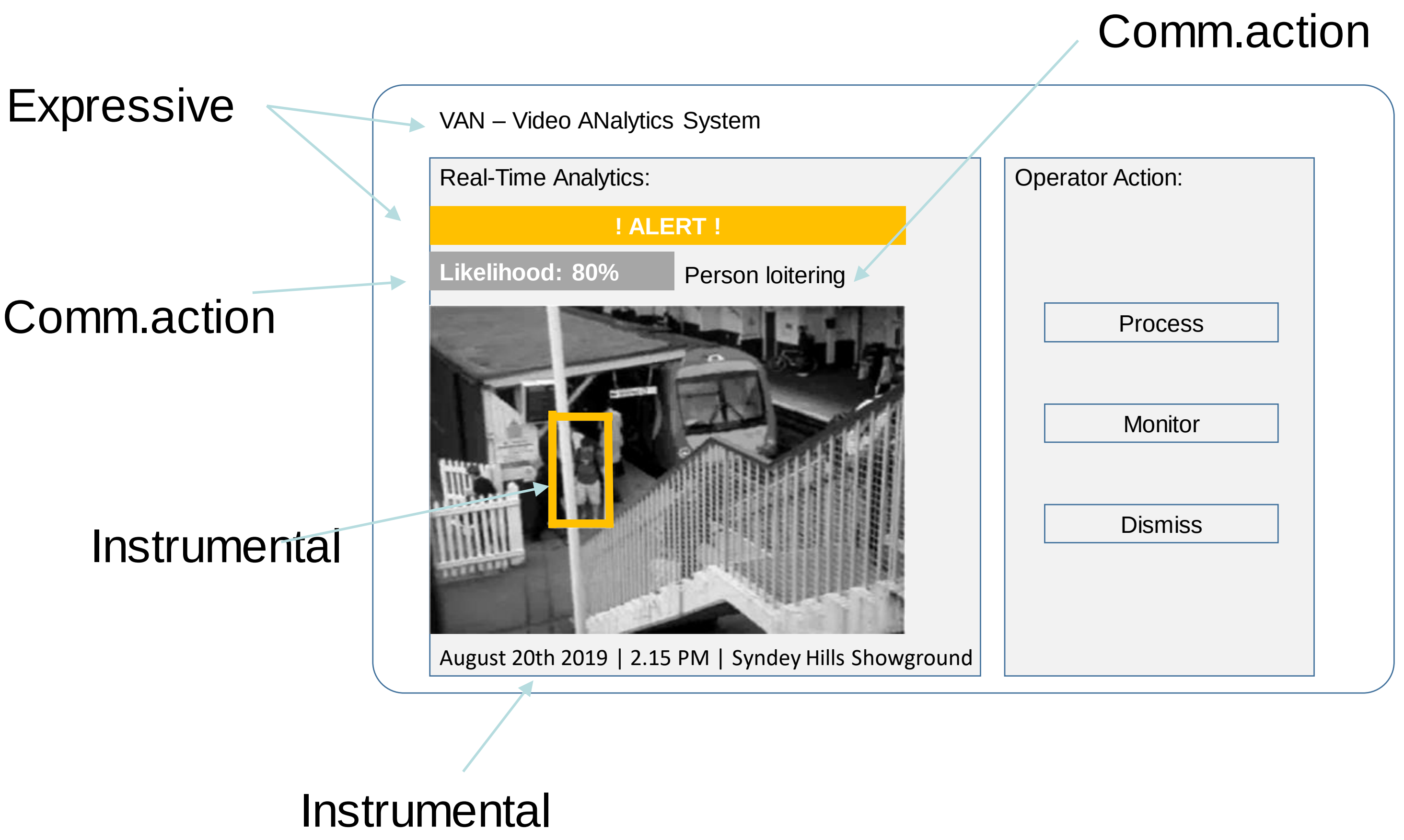
Cases with Machine Learning?

- Operational deep machine learning system with explanations not yet located
- Technologies exist (see Gilpin 2018):
  - Decision tree re-construction
  - LIME
  - *What-if* open source tool
  - Counterfactual explanations
  - Vendor tools eg DataRobot
  - “debugging” research – user changes the system
- Still concerns:
  - Whether users want these explanations?
  - Are the ML explanations good enough?
  - ML explanations might allow attacks and fraud

6. Overarching Orchestration Strategy

Orchestrate the use of strategies in combination. The communicative action strategy rather than the instrumental is preferable in many situations. Consider other strategies depending on context.

- Can be across organization (e.g. Australian Taxation Office)
- Can be within system
- Our research proof-of-concept:
  - Intelligent Video Surveillance at Sydney Trains to detect suspicious behavior preceding suicide attempts (ARC Linkage Project)
  - Interface developed in collaborative design workshop
    - (Gregor Maedche Morana 2019)



Note: nothing for Strategic or Normative

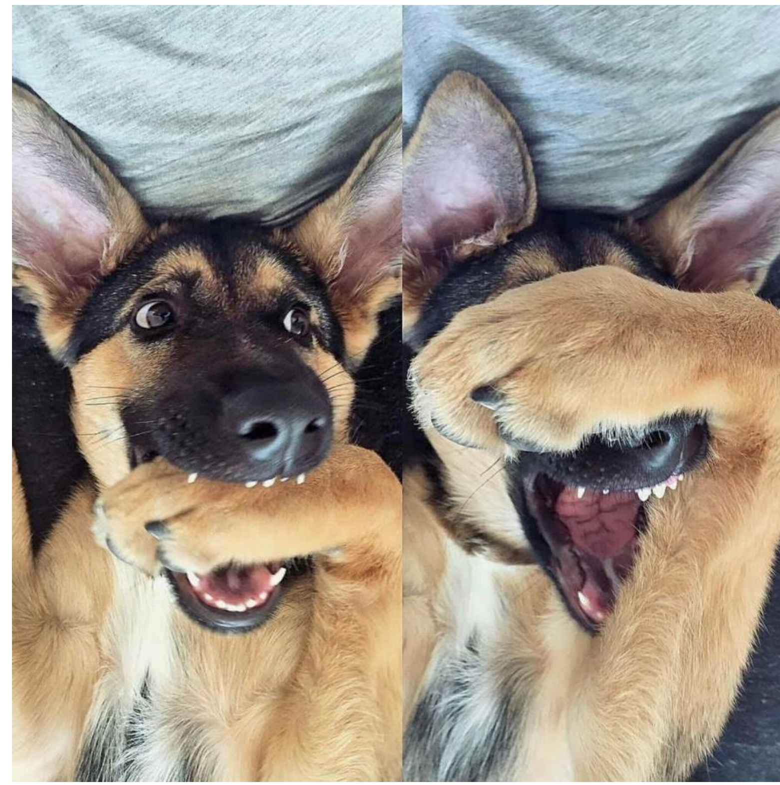
Discussion

- Have shown how design principles derived from Habermas and research on AI can be applied to healthcare case studies for communicating with AI
- Gives an integrated view not previously provided
- Questions:
  - Is the application of the design theory to the case studies evidence of credibility of design principles?
  - Would professionals in healthcare find it useful?

- Comments welcome

  Shirley.Gregor@anu.edu.au

Comments please?



References

- Allegheny County (2020). Allegheny Family Screening Tool. URL: https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx (accessed June 2020).
- AMA (2018) https://ama.com.au/ausmed/issues-be-aware-when-responding-compliance-audits
- Centre for Public Impact. https://www.centreforpublicimpact.org/how-to-make-ai-work-in-government-and-for-people/
- Battey, T. (2016) tombattey.com/design/case-study-paro/
- Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018, October). Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)* (pp. 80-89). IEEE.
- Goud, R. (2009). Computerized decision support to improve guideline implementation in cardiac rehabilitation: the CARDSS project. PhD thesis/R. Goud–Amsterdam.
- GovCC (2016) https://www.iqpc.com/media/1002724/66503.pdf
- Gregor, S. Maedche, A. and Morana, S. (2019). A Communicative Action Framework for Discourse Strategies for AI-based Systems: The MetTrains Application Case. *Proceedings of the Eighteenth Annual Pre-ICIS Workshop on HCI Research in MIS,* Munich, Germany, 15 Dec.
- Habermas, J. (1984) *The Theory of Communicative Action Reason and the Rationalization of Society*. Vol 1. Boston: Beacon Press,.
- iDAvatars Inc. https://biztimes.com/digital-avatar-sophie-brings-intelligence-compassion-to-interactions/
- Jacklevic, M (2017). https://www.healthnewsreview.org/2017/02/md-anderson-cancer-centers-ibm-watson-project-fails-journalism-related/
- Knight, W. (2017) The Dark Secret at the Heart of AI. *MIT Technology Review.*
- Misawa, D., Fukuyoshi, J., & Sengoku, S. (2020). Cancer Prevention Using Machine Learning, Nudge Theory and Social Impact Bond. *International journal of environmental research and public health*, *17*(3), 790.
- Moja, L., Kwag, K. H., Lytras, T., Bertizzolo, L., Brandt, L., Pecoraro, V., ... & Iorio, A. (2014). Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. *American journal of public health*, 104(12), e12-e22.
- Murray , S., Wachter, R. & Cucina , R. (2020). https://www.healthaffairs.org/do/10.1377/hblog20200128.626576/
- Nass, C; Steuer, J; and Tauber, E.R. Computers are social actors. In B. Adelson; S. Dumais; and J. Olson (eds.). *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94*. New York, New York, USA: ACM Press, 1994, p. pp. 72–78.
- OAI (Office for Artificial Intelligence). (2020). *A Guide to Using Artificial Intelligence in the Public Sector*. Government Digital Service. Version: S3BKR13R_V1.2 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/871178/A_guide_to_using_AI_in_the_public_sector__web_version_.pdf (accessed 1 July 2020).
- [OECD] Berryhill, J., Heang, K. K., Clogher, R., & McBride, K. (2019). Hello, World: Artificial intelligence and its use in the public sector. OECD.
- Riek, L. D. (2017). Healthcare robotics. *Communications of the ACM*, 60(11), 68-78.
- Russell, S. & Norvig, P. (2016). *Artificial Intelligence A Modern Arpproach* (3rd edn), Pearson.
- Strickland (2019) How IBM Watson Overpromised and Underdelivered on AI Health Care, *IEEE Expert.*
- Sung, J. J., Stewart, C. L., & Freedman, B. (2020). Artificial intelligence in health care: preparing for the fifth Industrial Revolution. *Medical Journal of Australia.*
- Verheul, M. M. (2016). *Building an infrastructure to improve cardiac rehabilitation: from guidelines to audit and feedback.*
- Wang, F., & Preininger, A. (2019). Ai in health: State of the art, challenges, and future directions. Yearbook of medical informatics, 28(1), 16.