

USING TECH/AI ETHICS: *WHY* ARE WE DOING THIS? (WHY DO TECHNOLOGY COMPANIES WANT ETHICAL INPUT?)

1. Provide (human language) metrics to determine AI tool's success
+ Enhance technical design
2. Shape company culture
3. Build user trust
4. Risk mitigation and legal compliance
→ Evaluations (Audits): How do you do them?

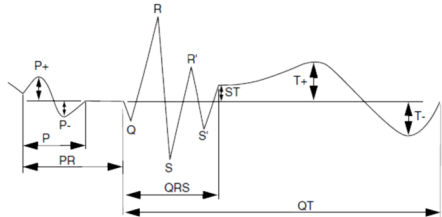
USING TECH/AI ETHICS: *WHY* ARE WE DOING THIS?

1. Provide metrics to determine success

Numerical accuracy only a proxy for success/failure of a tool

Cardio case: Privileging of high sensitivity accuracy led to numerical success (all patients with problems were alerted), but a failed tool...

Fixed by evaluating in terms of **patient autonomy** (ethics) instead of purely **numerical correctness** (mathematics)

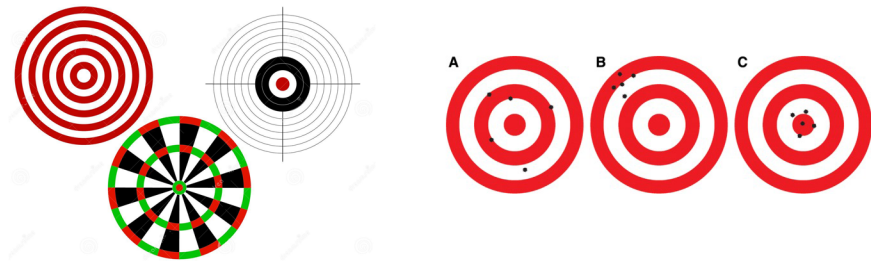


USING TECH/AI ETHICS: *WHY* ARE WE DOING THIS?

- + Enhance technical design: Understanding ethics empowers designing for success on the right target

Primary question:

Not how accurate/precise is the AI, but where does the AI aim?
Not how close to the target, but which target?



What *counts* as success? Ethics (which values?) → Math (which accuracy?)

USING TECH/AI ETHICS: *Why* ARE WE DOING THIS?



**Privacy.
That's iPhone.**

2. Shape company culture (+ individual...)

Work satisfaction increases with knowledge of outcomes apple.com/privacy

People like to know what they are doing with their lives

3. Build user trust

Users sensing fundamental rights protected more likely to trust and interact with the device

USING TECH/AI ETHICS: *WHY* ARE WE DOING THIS?

4. Risk mitigation and legal compliance

Because innovation is outstripping legal regulation, incorporating of ethics provides justification for decisions about the design and use of AI models, while also clarifying accountability.

Evaluations (Audits) increasingly required by law

OVERVIEW

USING ETHICS: WHY ARE WE DOING THIS?

1. Provide human language/metrics to measure AI success
+ Enhance technical design
2. Shape company culture
3. Build user trust
4. Risk mitigation and legal compliance

USING ETHICS: EVALUATIONS (AUDITS)

USING ETHICS: EVALUATIONS (AUDITS)

What are they?

- Cast technology in ethical/humanist terms

 - As opposed to scientific, economic, legal (iPhone example)

- Evaluate (ensure) how well the technology serves fundamental human rights

USING ETHICS: EVALUATIONS (AUDITS)

HOW TO DO THEM?

1. Social/Technical Scenario
2. Evaluate
3. Report/Recommend

USING ETHICS: EVALUATIONS (AUDITS): SOCIAL/TECHNICAL SCENARIO

What do you need to know to know what we are talking about?

- What does the AI/technology do?
- Who does it do it for?
- What counts as success?
- ...?

Gather stakeholders

Translation

Meeting/Zoom

Draft and commentary (Google doc)

USING ETHICS: EVALUATIONS (AUDITS): EVALUATE

Working group(s) with Principles

- EC Ethics Guidelines/Other...

Top down (Dignity: status of traditional music repurposed: At what point is the original author of the music no longer the author...? (*Whose music ethically versus legally?*)

Bottom up (Autonomy: new capacities to interact with music for everyone, for hearing impaired...)

Meeting/Zoom

Consolidate in terms of principles

Draft and commentary (Google doc)

USING ETHICS: EVALUATIONS (AUDITS): EVALUATE

TWO OVERLAPPING APPROACHES

Direct Principles/Values (Example: *EC Ethics Guidelines Trustworthy AI*)

- Is human autonomy served or limited by LinkedIn?
- Does the ethics of dignity allow Woebot lie to patient to help cure depression?
- Can a society in a pandemic demand to know everywhere you have been in the name of social wellbeing?

Tensions/Tradeoffs

- Autonomy *versus* Dignity (Woebot: mental health)
- Privacy *versus* Fairness (exAID)
 - Jess Whittlestone, Rune Nyrop, Anna Alexandrova, and Stephen Cave. 2019. The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society DOI:<https://doi.org/10.1145/3306618.3314289>

USING ETHICS: EVALUATIONS (AUDITS): EVALUATE

EXAMPLE GOOGLE DOC...

USING ETHICS: EVALUATIONS (AUDITS): REPORT/RECOMMEND

Convert Google doc into (academic) narrative

Major areas:

- Autonomy (Cardisio case)

- Privacy (Covid Case)

- Fairness + Explainability (exAID case)



Document question: Academic *or* Public policy *or* private company...

USING ETHICS: EVALUATIONS (AUDITS): REPORT/RECOMMEND

Convert Google doc into academic narrative

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

1

An Ethics Framework for the Internet of Musical Things

James Brusseau and Luca Turchet, IEEE Senior Member

Abstract—The Internet of Musical Things (IoMusT) is an emerging field of academic and industrial research that extends the Internet of Things to the musical domain. Scarce research has been conducted on the field's ethical aspects, and to fill this gap we propose a framework for the ethical design and evaluation of IoMusT technologies and applications. Besides being ethically rigorous, the framework seeks to be accessible for information engineers, musicians, and the wider circle of participants in the IoMusT. The purpose is to facilitate and quicken the process of ethically designing and evaluating work at the intersection of network-based technology and musical creativity. Finally, we exemplify the framework by applying it to an IoMusT experimental performance. Beyond facilitating the ethical evaluation of technologically enhanced music, the framework also advances work in contemporary AI ethics in two ways. First, by introducing the principles of creativity and decentralization as critical to ethically exploring musical creativity. Second, by organizing the principles of AI ethics under a human-centric logic.

Index Terms— Internet of Musical Things, ethics, ethics evaluation, design framework

I. INTRODUCTION

THE Internet of Things (IoT) refers to the network of physical objects augmented with sensing, actuation, computing and communications capabilities [2]. This paradigm is constantly evolving and is impacting several

engineering and art research in both academic and industrial contexts. It is positioned at the intersection of the Internet of Things [2], New Interfaces for Musical Expression [3], Ubiquitous Music [4], human-computer interaction [5], artificial intelligence [6], technology-mediated audience participation [7], and Semantic Web [56].

As the general IoT, the IoMusT is based on the central pillars of heterogeneous devices, communication infrastructures, applications and services, as well as ecosystems of humans and machines [54, 55]. From a computer science perspective, the IoMusT refers to the networks of computing devices embedded in physical objects (musical things) that produce and receive musical content [56]. Musical things, such as smart musical instruments [8] or musical haptic wearables [9], are connected by an infrastructure that enables multidirectional communication, both locally and remotely. The digital and physical domain connections are enabled by information and communication technologies, which foster novel musical applications and services.

The ecosystems associated with the IoMusT include interoperable devices and services linking musicians and audiences, and to support musician-musician, audience-musician, and audience-audience interactions. The IoMusT, consequently, enables and supports a wide range of novel human musical experiences including: 1) augmented and immersive concert experiences; 2) audience participation; 3) remote education; 4) music learning; and 5) smart studio

USING ETHICS: EVALUATIONS (AUDITS): REPORT/RECOMMEND

Convert Google doc into academic / public policy narrative

272

IEEE TRANSACTIONS ON TECHNOLOGY AND SOCIETY, VOL. 3, NO. 4, DECEMBER 2022

Assessing Trustworthy AI in Times of COVID-19: Deep Learning for Predicting a Multiregional Score Conveying the Degree of Lung Compromise in COVID-19 Patients

Himanshi Allahabadi, Julia Amann, Isabelle Balot[✉], Andrea Beretta, Charles Binkley[✉], Jonas Bozenhard[✉],
Frédéric Bruneault, James Brusseau[✉], Sema Candemir, Luca Alessandro Cappellini[✉],
Subrata Chakraborty[✉], *Senior Member, IEEE*, Nicoleta Cherciu, Christina Cociancig, Megan Coffee[✉], Irene Ek,
Leonardo Espinosa-Leal, Davide Farina, Geneviève Fieux-Castagnet, Thomas Frauenfelder[✉], Alessio Gallucci,
Guya Giuliani, Adam Golda[✉], Irmhild van Halem, Elisabeth Hildt[✉], Sune Holm, Georgios Kararigas[✉],
Sebastien A. Krier, Ulrich Kühne, Francesca Lizzi[✉], Vince I. Madai, Aniek F. Markus[✉], Serg Masis[✉],
Emilie Wiinblad Mathez, Francesco Mureddu, Emanuele Neri, Walter Osika, Matiss Ozols[✉], Cecilia Panigutti,
Brendan Parent, Francesca Pratesi[✉], Pedro A. Moreno-Sánchez, Giovanni Sartor, Mattia Savardi[✉],
Alberto Signoroni[✉], Hanna-Maria Sormunen[✉], Andy Spezzatti, Adarsh Srivastava[✉], Annette F. Stephansen,
Lau Bee Theng[✉], *Senior Member, IEEE*, Jesmin Jahan Tithi, Jarno Tuominen[✉], Steven Umbrello[✉],
Filippo Vaccher, Dennis Vetter[✉], Magnus Westerlund, Renee Wurth, and Roberto V. Zicari[✉]

Abstract—This article's main contributions are twofold: 1) to demonstrate how to apply the general European Union's High-Level Expert Group's (EU HLEG) guidelines for trustworthy AI in practice for the domain of healthcare and 2) to investigate the research question of what does "trustworthy AI" mean at the time of the COVID-19 pandemic. To this end, we present the results of a *post-hoc* self-assessment to evaluate the trustworthiness of an AI system for predicting a multiregional score

pandemic. The AI system aims to help radiologists to estimate and communicate the severity of damage in a patient's lung from Chest X-rays. It has been experimentally deployed in the radiology department of the *ASST Spedali Civili clinic* in Brescia, Italy, since December 2020 during pandemic time. The methodology we have applied for our *post-hoc* assessment, called Z-Inspection[®], uses sociotechnical scenarios to identify ethical, technical, and domain-specific issues in the use of the AI system in the context

USING ETHICS: EVALUATIONS (AUDITS): REPORT/RECOMMEND

Convert Google doc into private company narrative...

(Ethics and/versus law...)

USING ETHICS: EVALUATIONS (AUDITS)

QUICK EXAMPLE CASE (MAYA.AI, LOW STAKES)

1. Social/Technical Scenario
2. Evaluate
3. Report/Recommend



USING ETHICS: EVALUATIONS (AUDITS)

QUICK EXAMPLE CASE (MAYA.AI, LOW STAKES)

Optimize For:

Autonomy/Dignity/Privacy/
Fairness/Equity/Social wellbeing
Explainability/Safety/Performance

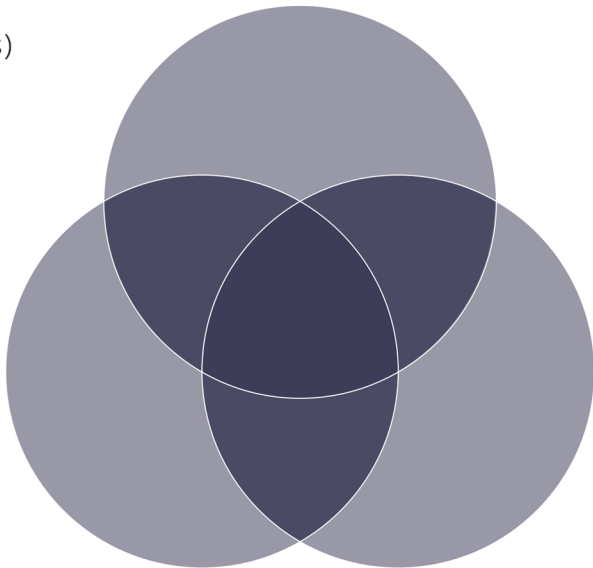
What does this mean?

For whom?

How can you do it?

What if there is no "sweet spot"?

*(Does maximizing user autonomy
subtract from privacy, dignity?)*



OVERVIEW

USING ETHICS: WHY ARE WE DOING THIS?

1. Provide human language/metrics to measure AI success
+ Enhance technical design
2. Shape company culture
3. Build user trust
4. Risk mitigation and legal compliance

USING AI ETHICS: EVALUATIONS (AUDITS)

What are they?

How do you do them?

1. Social/Technical Scenario
2. Evaluate
3. Report/Recommend

Quick case