is Ai ethical

AISLA ELSi Framework

Introduction

isITethical⊛?

The workshop was a chance to make space, anticipate, notice, and start to address some ethical tensions in the AISLA project. We aim to move beyond the concept of ethics as a tick box exercise, and instead under-stand ethics as a contextual, creative, participatory, and iterative process of doings and critical reflections. Working with ethics in this way helps to remove it from the isolated bubble of an ethics office, and find routes toward ethical conduct.

To do this, we co-create a framework of Ethics through Design, following three principles

Principle of Art Thinking
Using creative practices and art interventions, including performance, film, dance, play, poetry, painting, making, to nurture spaces of movement, and open other ways of knowing ethical issues

Principle of Engagement
Value is dependent upon the possibilities of encounter, with humans and more-than-humans. We resort to nurturing engagement as a process of mutual loss and gain.

Principle of Anticipatory Ethics
Values are a constellation of material, discursive, contextual, and iterative practices. We engage with dynamics that make visible and accessible value and the meaning of value for actors involved with a technological innovation, building on ideas of contextual and participatory ethics.

This is what we co-created:
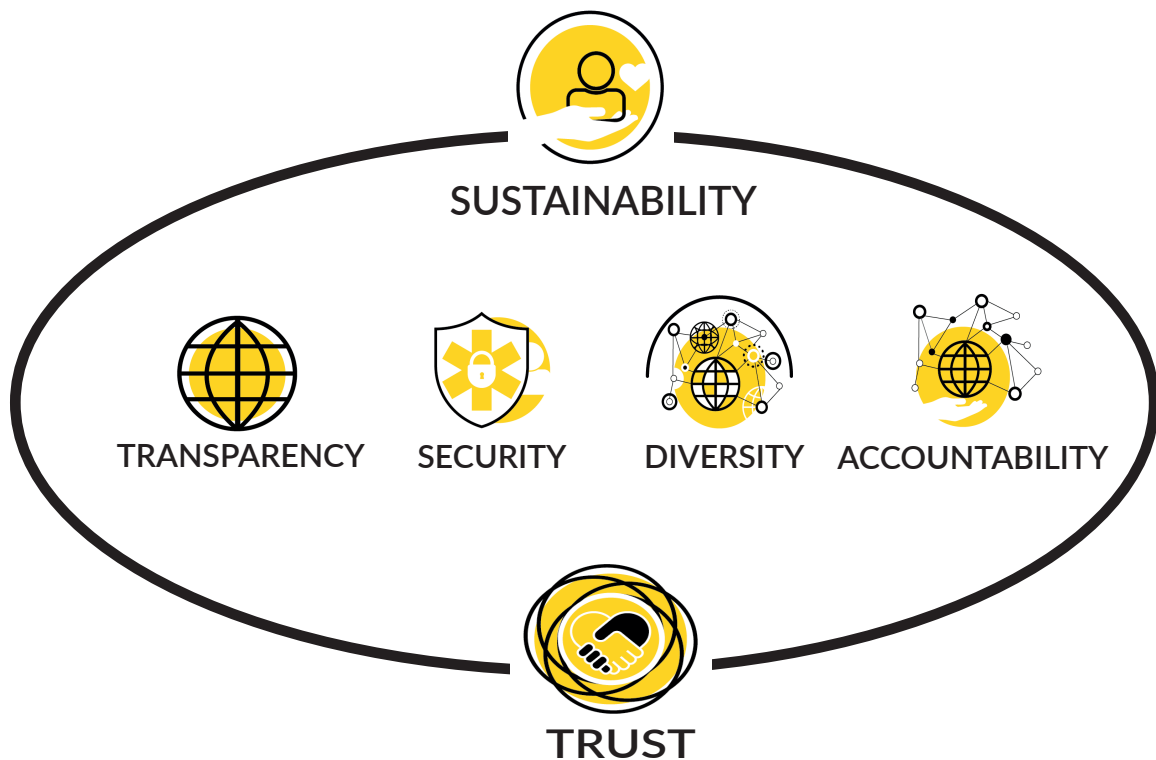
SUSTAINABILITY

TRANSPARENCY   SECURITY   DIVERSITY   ACCOUNTABILITY

TRUST

We want to create a TRANSPARENT, SECURE, SUSTAINABLE, and DIVERSE AI system informed by empirical research for enhancing language learning that is SUSTAINABLE AND TRUSTWORTHY

AISLA ELSI Framework

We want to create a TRANSPARENT, SECURE, SUSTAINABLE, and DIVERSE AI system informed by empirical research for enhancing language learning that is SUSTAINABLE AND TRUSTWORTHY

Our project two core project values are:

SUSTAINABILITY

o Our interests cannot be harming other humans, species or environment.

o Proactively weigh risks and benefits avoiding the distribution of risks among the most vulnerable and the distribution of benefits only amongst the better off.

o Support user and participant competence in engaging with research: How users and participants directly benefit from their research engagement? (avoid data extrationism)

Related ELSi Guidance :

- Community Distributed Design
- Sustainable AI
- Responsive Governance

.
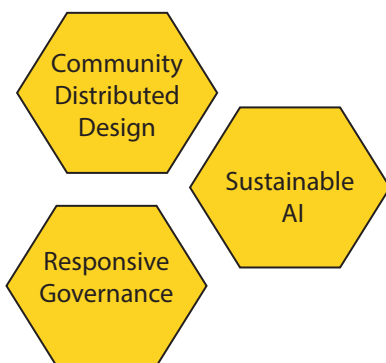
Reference:

'The entire issue of sustainable development centres around inter- and intragenerational equity anchored essentially on three-dimensional distinct but interconnected pillars, namely the environment, economy, and society' (in Wynsberghe, 2021, p. 215)

Resources:

Resources: van Wynsberghe, A. (2021) Sustainable AI: AI for sustainability and the sustainability of AI. AI Ethics 1, 213–218, available here

TRUST

o Consider how trust in AISLA system will be scaffolded

o Consult others when there are uncertainties

o Identify positive expectations and enable them to be regularly met.

o Configure in the project plan systems, spaces, methods and channels for reflection or feedback, as well as mechanisms of response

Trust Covers:

RESPONSE-ABILITY

o Consider how actions impact those engaged in AISLA system and beyond. Integrate research capacity and technical ability to respond to changes.
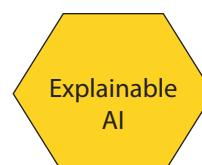
Actions:

Explaining how the App works
What it can and can't do
What the role of the App is
What the purpose is
What the constraints are

Carry out user training
Explain the system and the technology
Integrate explainability in the website

Related ELSi Guidance :

- Explainable AI

TRANSPARENCY

o Explainable AI makes its inner logics and functions accessible

o Find technological and social mechanisms to share the inner workings of systems being created with users and those being served by its use.

o Help users understand the inner logics and functions of AISLA including the classification systems, taxonomy, access controls, etc

Actions:
Open Scholarship and Open Science:
Decide what to make available to all and what to keep private.

Lobby for, and integrate in Project Outputs:
Awareness and request of information from suppliers of technology about their security

Related ELSi Guidance :

Explainable AI

Research Assumptions

Responsive Governance

Transparency Covers:

ADAPTIVE SYSTEM
o Our system technical capacity to respond to individual user's need and behaviours.

RELATIONALITY
o Ai that visibilises relations of non-renewable resources, labour, and data it requires to work

NON-DISCRIMINATION
o Reflect on our position, biases, and privileges, and that of data subjects and users of AISLA

FAIRNESS
o Be aware of inequalities - social, financial, racial - between data subjects and users

SOVEREIGNTY
o Beyond tecnical success, ensure benefits for the community, integrate capacities for everyone to have control over their data usage (data footprint).

Reference:

The principle of transparency requires that any information and
communication relating to the processing of those personal data be easily accessible
and easy to understand, and that clear and plain language be used.

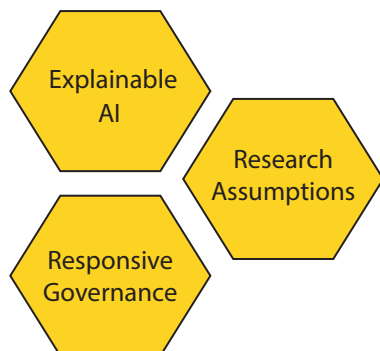EU Working Party, Guidelines on Transparency under Regulation 2016/679, p. 6.

Available here

Resources:

Elements of AI Online Course, University of Helsinki, https://www.elementsofai.com/

SECURITY

o Take the time to make a security plan

o Ensure all measures are in place to protect security of researchers, users and stakeholders, including data subjects

o Ongoing assessment that your research does not infringe security

o Security needs to balance the right to privacy, civil liberties, and vulnerability to hacking

Actions:
Develop data storage strategy

Hire people to be responsible for system/data security.

Related ELSi Guidance :

Data Practices

GDPR

Security Covers:

ADAPTABILITY
o Our methods can withstand need for change, exception, and improvisation as collaborations shift over time

PRIVACY
o We are taking the all the measures we can to protect data collection and processing from unauthorised access
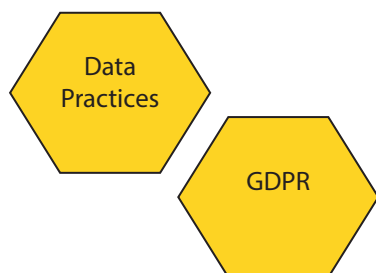
DATA PROTECTION
o We secure platforms for data storage and processing and provide mechanisms for data subjects to exercise control over their data

ANONYMITY
o We consider how anonymity might support trust or distrust

Reference:

Data is 'processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisa-tional measures' (Hildebrant, 2020, p. 150)

Resources:

European University Institute (2019) 'Good Data Protection Practice in Research', available here:

Hildebrandt, M. (2020). 5.5. 'Data Protection Law' in Law for computer scientists and other folk. Oxford University Press.

With SUSTAINABILITY and TRUST at the core, these are the value that inform AISLA framework of practices

ACCOUNTABILITY

o Capacity to stay answerable for our choices, actions and expectations of our role.

o Capacity of algorithms, programming, and languages to account for their affordances in an intelligible way

o Integration in both our research project and ASILA system of spaces and capacities to respond to the unknown and unexpected.

Accountability Covers:

SELF-DETERMINATION

o Ensuring that users are not dominated by the agent, but they can decide how they use it.

Actions:
Explain how data will be stored and used

Related ELSi Guidance :

Autonomy

Reference:

A data controller should be accountable for complying with measures which give effect to the principles of individual partici- pation - confirmation of what data is stored, access to this data or reasons for denial of access, right to request data deletion (Hildebrant, 2020. pp. 134-5)

Resources:

Doshi-Velez, F. et al., (2017). 'Accountability of AI under the law: The role of explanation'. Available here

Hildebrandt, M. (2020). 5.5. 'Data Protection Law' Law for computer scientists and other folk. Oxford University Press.

DIVERSITY

o Identify the need of the user, rather the impose alien visions or perspectives.

o Integrate methods to allow differences to be visible in research and innovation (gender, nationality, race, abilities, socio economic backgrounds, geographies)

o Proactively avoid reproduce stereotypes and cultural assumptions.

o Be aware of assimilation. Beyond inclusion, be open to other modes of knowledge.

Actions:

More diverse content, .e.g.
Different varieties of English; idioms.

Related ELSi Guidance :

Community Distributed Design

Research Assumptions

Reference:

'Embrace Pluralism. Data feminism insists that the most complete knowledge comes from synthesizing multiple perspectives, with priority given to local, Indigenous, and experiential ways of knowing.' (D'Ignazio & Klein, Data Feminism, Chapter 5)

Resources:

D'Ignazio, C. and Klein, L.F., 2020. Data feminism. Mit Press.
https://cyberfeminismindex.com/
Indigenous Protocols for AI Position Paper
Available here

isITethical Ai ELSi Guidance

AUTONOMY
Consider different meanings and levels of autonomy

When do humans intervene? Map relations and dependencies. Who is accountable? When? How?

COMMUNITY DISTRIBUTED DESIGN
Beyond User Centered Design. This approach includes: centering the most vulnerable members of the community where our systems will operate.

DATA PRACTICES
This is about both precedence of data set and data gathering.
What is the data used for? How and why was it derived? How does it fit with other data it is being used with? How does the data match reality?

EXPLAINABILITY
If AI is to meet basic business-use, humans must comprehend its inner works. How are we ensuring AI is intrinsically explainable and human-compatible?

GDPR
What personal data is used and stored? When do we process personal data? How is it anonymised? Who can access the data? How long can they be stored?

SUSTAINABILITY
AI systems depend on extractionist practices: data extractions, carbon footprints, environmental impacts of digital infrastructure, and global extraction of labour. What measures can we include to off-set some of the risks?

RESEARCH ASSUMPTIONS
Beyond the technical possibility, is the AI desired and needed? By who? Who would rather not use it? How do we know?

RESPONSIVE GOVERNANCE
It refers both to our projects and the governance of AI systems. How are our projects responding to emerging needs and contexts? How are we ensuring sustainable governance of our AI systems?

isITethical ✳ Preliminary User Feedback

For this, two groups were interviewed 1) six 16 year old students across race, gender and level of proficiency and experience with second languages. 2) Two college-level language teachers

Students

What is an AI?
What is a Conversational Agent?
I would like to know how it works.

I'll use at home, not at school

How is it different from an app? (Duolingo)

Can the AI be a friend rather than a teacher?

We're tired of screens, can it be like a phone call?

Can I play and compete with my friends?
Can I meet and chat with my mates?

We want to learn slang, we want to speak like kids in real life.
Can I select the accent? (English, American, Indian, African, etc.?)

We know social media like Whatsapp are dangerous. Is this risky?
I would like the system to show me the risks.

Teachers

Is it to support class activities or for independent learning?
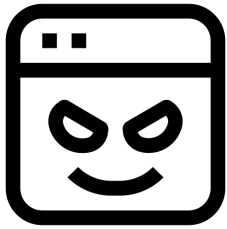
What about tone, humour, idioms?

How is it different from other digital tools?

We teach differently depending on the end:
grammatical (exams) or conversational (travel, etc.).

Individual or tailored feedback.
We rarely give the same feedback to two students.

How will the AI account for different types of learners?
(Visual, readers, listeners)

isITethical ✳ ELSi Scenarios

For circumspect reflection upon impact beyond the system; on people and planet

Abuser Online

A pedofile has found a way to hack into AISLA's servers and retrieve personal data about some of the school children using it. This includes ages, schools the children attend, and location information of where the children were when making certain interactions with AISLA. It is now very easy for this pedofile to track down the children in person.

I, stutter

Felipe is a 7th grade kid who moved to Germany from Chile with his family 6 months ago. He has struggled to make friends since starting in his new school, because Felipe as a speech impediment that makes him frequently stutter. Other children at the school have made fun of Felipe for this, and so he rarely speaks in his English class. The teacher advises Felipe to start using AISLA at home to practice his speech, but the conversational agent cannot understand Felipe and tells him he is making many mistakes. Felipe is losing confidence and feeling humiliated.
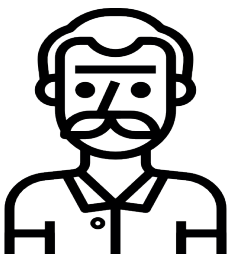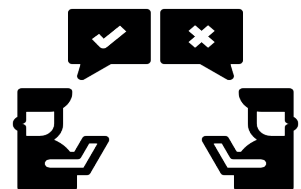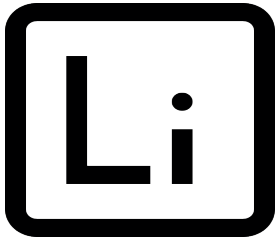
Teacher Teacher

A local school has just purchased an AISLA subscription, and is starting to implement it in class. Herr Jacobi has been teaching English at the school for 25 years, and has always had good relationships with his students. He believes firmly in the importance of physical textbooks and grading with pen and paper. Herr Jacobi is reluctant to start using AISLA, and when it's implemented school-wide, he doesn't know how to engage with its interface. Now, Herr Jacobi is adding to his already large workload trying to manage his students using AISLA, and is struggling to keep up.

Bias

AISLA's AI has been extensively trained on data sets of English language syllabus, and speech patterns. However, the vast majority of these have been British and American English. When students with other accents and dialects of English try to speak to their conversational agents, it deems their English incorrect, even though they have not made any mistakes.

For circumspect reflection upon impact beyond the system; on people and planet

Extractionism

Li

El Salar, the world's largest flat surface, is located in southwest Bolivia at an altitude of 3,656 meters above sea level. It is a high plateau, covered by a few meters of salt crust which are exceptionally rich in lithium, containing 50% to 70% of the world's lithium reserves.The Salar, alongside the neighboring Atacama regions in Chile and Argentina, are major sites for lithium extraction.

This soft, silvery metal is known as 'grey gold.' Smartphone batteries, for example, usually have less than eight grams of this material.

According to the Aymara legends about the creation of Bolivia, the volcanic mountains of the Andean plateau were creations of tragedy. Long ago, when the volcanos were alive and roaming the plains freely, Tunupa - the only female volcano – gave birth to a baby. Stricken by jealousy, the male volcanos stole her baby and banished it to a distant location. The gods punished the volcanos by pinning them all to the Earth. Grieving for the child that she could no longer reach, Tunupa wept deeply. Her tears and breast milk combined to create a giant salt lake: Salar de Uyuni. "your smart-phone runs on the tears and breast milk of a volcano". This landscape is connected to everywhere on the planet via the phones in our pockets; linked to each of us by invisible threads of commerce, science, politics and power.
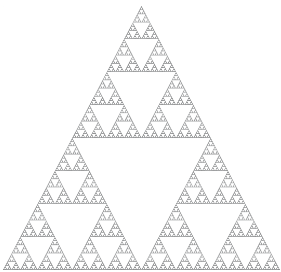
Supply Chain Complexities

This triangle of value extraction and production represents one of the basic elements of our map, from birth in a geological process, through life as a consumer AI product, and ultimately to death in an electronics dump. The triangles are not isolated, but interconnected. They form a cyclic flow in which the product of work is transformed into a resource, which is transformed into a product, which is transformed into a resource and so on. Each triangle represents one phase in the production process.

TIf we look at the production and exploitation system through a fractal visual structure, the smallest triangle would represent natural resources and means of labor, i.e. the miner as labor and ore as product. The next larger triangle encompasses the processing of metals, and the next would represent the process of manufacturing components and so on.

According to Amnesty International, during the excavation of cobalt, workers are paid the equivalent of one US dollar per day for working in conditions hazardous to life and health, and were often subjected to violence, extortion and intimidation. Amnesty has documented children as young as 7 working in the mines.

As a semiconductor chip manufacturer, Intel supplies Apple with processors. In order to do so, Intel has its own multi-tiered supply chain of more than 19,000 suppliers in over 100 countries providing direct materials for their production processes, tools and machines for their factories, and logistics and packaging services. That it took over four years for a leading technology company just to understand its own supply chain, reveals just how hard this process can be to grasp from the inside, let alone for external researchers, journalists and academics.

What is Next?

We opened a lot of questions during our time together, as well as unexpected avenues, and exciting opportunities.The ELSI framework is designed to be re-iterative, consultable at various points along the project process. And this is just the beginning. Some things to do:
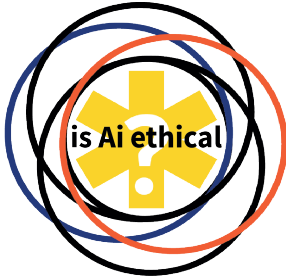
>Integrate spaces and activities in AISLA project for reflection

>Revisit AISLA values (use the isAiethical cards)

>Revisit ELSI Guidance

This framework was co-created by AISLA project and isITethical.
For more about us, please visit our website
https://www.isitethical.org

Or get in touch directly,

Dr Malé Luján Escalante
Senior Lecturer in CoDesign
& Knowledge Exchange
University of the Arts, London
m.lujanescalante@lcc.arts.ac.uk

Dr Luke Moffat
Research in Sociology,
Trustworthy Autonomous Systems Security Node
Lancaster University
l.moffat1@lancaster.ac.uk