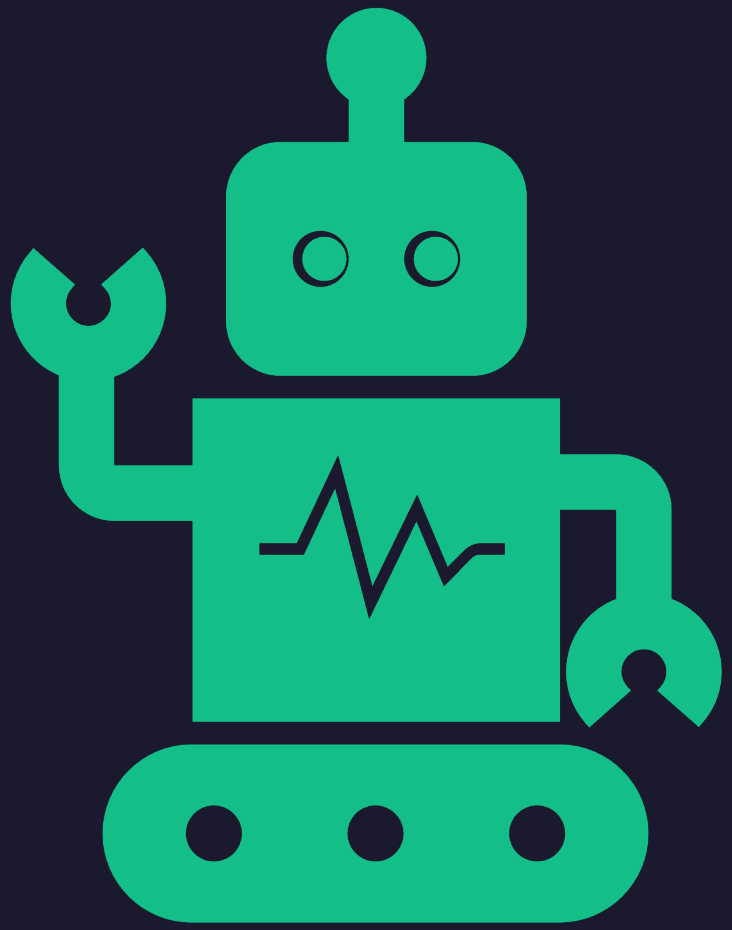
Ethically Aligned Design: An empirical evaluation of the RESOLVEDD-strategy in Software and Systems development context

Authors - Ville Vakkuri, Kai-Kristian Kemell, Pekka Abrahamsson.

Faculty of Information Technology, University of Jyvaaskyla¨ Finland.
Year - 2019

Presented By – Raman Srivastava



Fact

- **One of the most interesting technology facts about artificial intelligence is that by 2045, AI is expected to completely surpass human intelligence.**

Agenda

Abstract

About Study (Background)

Goal Of study

RESOLVEDD strategy

Ethically Aligned Design

ART Principles

Finding of 5 PEC

My view

END (Thank You)

Abstract

The widely spread presence of AI/AS systems forces us to analyze more profoundly under what type of ethical norms, rules and regulations AI systems should operate, and what kind of ethical standards should designers and developers hold when building these systems.

As software engineers, developers are constantly making decisions when building systems. In doing so, they build their own values into the systems, which end up reflecting their views. It is known that developers are not well-informed and aware of ethics.

Combined with the current lack of tools to support ethical AI development, this results in a situation where developers do not have the necessary means to tackle potential ethical issues, or even recognize them during development

About Study (Background)

The study was conducted in an Information Systems (IS) course at the University of Jyväskylä. Bachelor level students were introduced to the RESOLVEDD strategy as a part of the system design and development methods.

In the course, the students were given the task of developing a concept and prototype of a futuristic innovation that could be possible in the near future.

The projects were carried out as a group work in five groups of 4-5 students. Choosing from a list, the students had to decide which technology they would want to utilize as part of their solution. For example, the students could make solutions that utilized Augmented Reality (AR), AI while keeping mentioned strategies in mind.

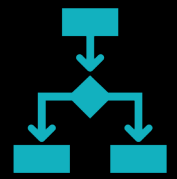
Goal of Study

Goal Of study

- **The goal of this study is to better understand how the introduction of an ethical tool would affects developers' ethical consideration in the design process and how the RESOLVEDD-strategy works in the given context.**
- **Use of artificial intelligence (AI) in human contexts calls for ethical considerations for the design and development of AI-based systems. And this Can be done using RESOLVEDD Strategy.**

Strategy One

What is RESOLVEDD Strategy



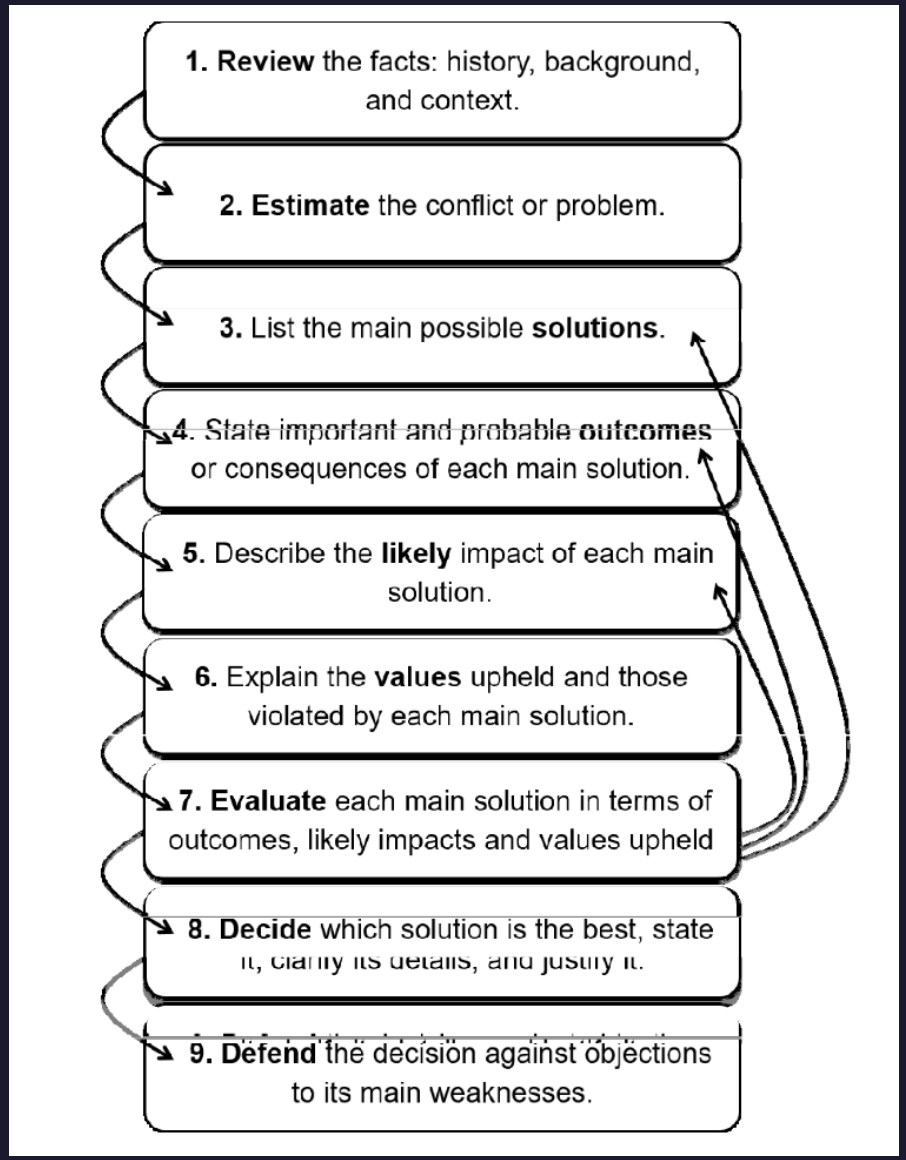
The RESOLVEDD Decision-Making Strategy Is A Nine-Step Method For Coming To A Decision About Hard Ethical Cases

RESOLVEDD strategy is intended to support its users in considering ethical issues and tackling them through their own set of values or through an ethics theory of their choice.

By using the method, one is able to justify and explain the decision-making process leading up to whatever actions were ultimately taken.

RESOLVEDD

- 1) Review the background, context (All Facts)
- 2) Estimate the conflict or problem
- 3) List all the main possible Solutions
- 4) Define important and Probable Outcomes or consequences of each main solution
- 5) Describe the likely impact of each main solution
- 6) Explain the values upheld and those violated by each main solution
- 7) Evaluate each solution in terms of Outcomes, its impact and Values upheld
- 8) Decide which solution is the best, state it, clarify its details , and justify it.
- 9) Defend the decision Against objections to its main weaknesses.



Strategy Two

Ethically Aligned Design's

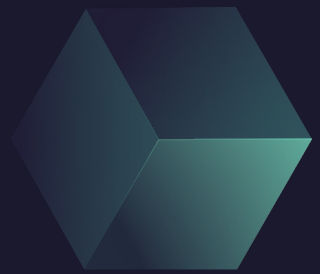
Ethically Aligned Design refers to the involvement of ethical consideration in practice while decision-making and designing AI and autonomous systems and technologies.

These are some AI-specific values proposed, to better incorporate human values into the design process of AI systems.

We must therefore use ART principle's (Accountability, Responsibility, Transparency) frameworks to guide, ethical development of AI systems

ART principles (Accountability, Responsibility, Transparency)

Goal Of ART



- **Together, the ART principles provide a framework for individuals and organizations to make ethical decisions and behave in a responsible and accountable manner. By following these principles, individuals and organizations can help to promote trust, integrity, and positive social impact.**
- **The ART constructs have a central role in determining design protocols that take into consideration the designer, the product, and the end-users.**

Commitment

Commitment involves making a promise to take specific actions or deliver certain outcomes.

Commitment is an important aspect of the ART framework, as it involves following through on the principles of accountability, responsibility, and transparency.

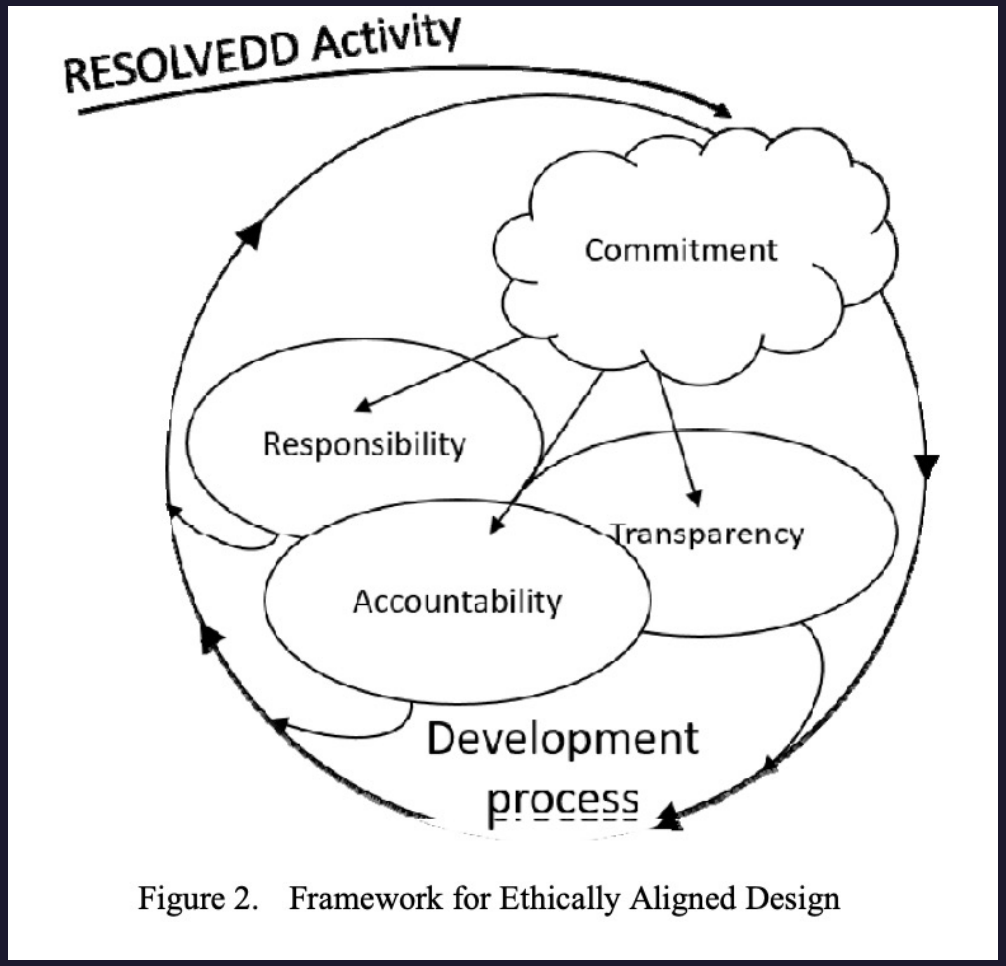


Figure 2. Framework for Ethically Aligned Design

Accountability

Accountability is referred to the explanation and justification of one's decisions and one's actions to the relevant stakeholders.

This principle emphasizes the importance of taking responsibility for one's actions and decisions. It involves being aware of the potential consequences of one's actions and taking steps to address any negative outcomes.

Responsibility

Responsibility in the context of this study connects the designer to the outside world, to others as stakeholders. In order to be responsible, one must make weigh their own actions and to consciously evaluate their choices. E.g., one very simple way of considering responsibility would be to ask oneself "would I be fine with using my own system?".

This principle focuses on the idea that individuals and organizations have a responsibility to act in a way that promotes the well-being of society and the environment. It involves considering the impact of one's decisions and actions on others and taking steps to minimize any negative effects

Transparency

In the ART model, transparency, focus on the transparency of the algorithms and data used, as well as their provenance and their dynamics. Systems need to be transparent so that the reasons behind unwanted results can be understood

Transparency has been considered to be crucial for the ethical design and use of AI/AS since it provides a simple and objective way of understanding what an AI/AS is doing and why.

This principle emphasizes the importance of being open and honest in one's communications and decision-making. It involves providing clear and accurate information about one's actions and decisions and being willing to answer questions and address concerns from others.

Findings In Study

Findings of 5 PEC (Primary Empirical Conclusions (PEC)

Commitment to Ethically Aligned Design - When considering commitment to EAD, it is important to understand what the true concerns of the developers are. In this case, the teams were more concerned about the usefulness and viability of their product than its ethical aspects.

Transparency in design - When the RESOLVEDD strategy is followed step-by-step a paper trail is born where each decisions made, and the respective justification can be found. In this case, this produced transparency in the design process, but it does not promote transparency at the product layer.

Findings of 5 PEC (Primary Empirical Conclusions (PEC) Continued

Accountability in design - The question of accountability divided the teams. It was not clear to the teams who can be held accountable for the design. Teams defended their position (not being accountable) by arguing that the systems are only concepts and prototypes.

Responsibility in design - The mere presence of an ethical tool has an effect on ethical consideration creating more responsibility among teams.

Findings

- Findings have shown, initially, that while it is possible to introduce EAD by force, but results will not sustain over time. The RESOLVEDD strategy needs to be adjusted in practice. One important adjustment done by case teams was the introduction of group discussions as the primary means to do EAD in practice.

My View

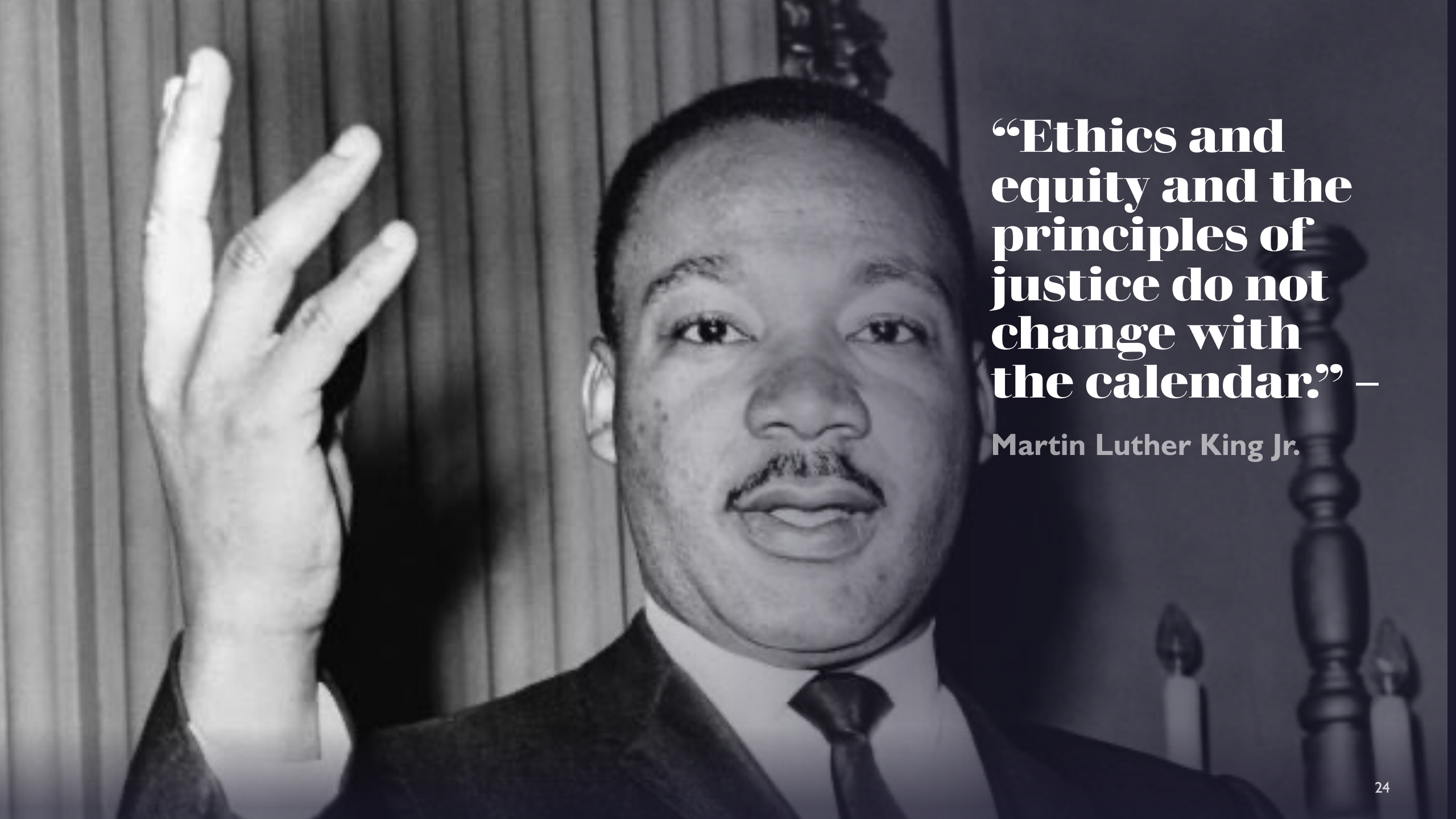
For strategy to work, Direction's , expectation, end result , should be clear from starting.

Implementing accountability mechanisms, such as auditing and review processes, can help ensure that AI systems are being developed in an ethical and responsible manner

Conducting ethical impact assessments can help identify potential ethical issues that may arise from AI development and help develop strategies to mitigate or address those issues.

Providing training and education on ethical considerations in AI development can help increase awareness and understanding of potential ethical issues

Last, Always remember RESOLVEDD And ART while Developing AI Systems.


“Ethics and equity and the principles of justice do not change with the calendar.” –
Martin Luther King Jr.

References

C. Allen, W. Wallach and I. Smit, "Why Machine Ethics?" IEEE Intelligent Systems, vol. 21, (4), pp. 12-17, 2006 doi: 10.1109/MIS.2006.83.

P. Abrahamsson and N. Iivari, "Commitment in software process improvement - in search of the process," Proceedings of the 35th HICSS, pp. 3239-3248, 2002. doi: 10.1109/HICSS.2002.994403.

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition. IEEE. 2019.

Ethically Aligned Design: An empirical evaluation of the RESOLVEDD-strategy in Software and Systems development context - Authors - Ville Vakkuri, Kai-Kristian Kemell, Pekka Abrahamsson , Faculty of Information Technology, University of Jyväskylä Jyväskylä, Finland. Year - 2019

Thank You For Your Patience

Feel Free To ASK any Questions

END OF PRESENTATION