



Fontys

> FOR SOCIETY

BUILDING VALUE- BASED TECHNOLOGY TOGETHER

INAUGURAL SPEECH
BART WERNAART
20 JANUARY 2022

All I've got is a red guitar, three chords and the truth.

This Bob Dylan song is, in my view, a perfect metaphor for moral design and its challenges. Imagine that the guitar resembles new technology. The three chords are possible moral programming solutions that can be used to 'play' this new technology. And the truth is an unreachable moral state that we desperately seek but is hard to find when we morally design new technology. Ladies and gentlemen, welcome to my inaugural speech. It is an honor to introduce our new professorship of moral design strategy.

1.1 A red guitar

Let's start with the guitar. The acoustic guitar isn't very loud. It suits a campfire song but has no effect when played on a large stage in front of a crowd of dancing and cheering fans. When we use an amplifier to turn the acoustic guitar into an electric guitar, it will amplify each chord. The beautiful sounds, but also the dissonant and uncomfortable chords you play. The amplifier is a technological solution for a problem. It is also another metaphor for what technology

is: an amplifier that can be used in good ways and bad ways. It is also important to establish that technology is never without value. The designer of the amplifier has a motive for designing new products and an agenda for creating new technologies. For instance, it is difficult to play flamenco guitar on a heavy metal amplifier. In other words: the motive of the designer plays a role in how technology can work. And this motive is mostly far from neutral. Amplifiers can be designed to sound as loud as possible, as warm as possible, or as trashy as possible. Also, the mere existence of the amplifier can change the music scene beyond the intention of the designer of the amplifier. Take the rotating Leslie speaker used in many Beatles recordings. This usage was never foreseen by the creators of the equipment, and profoundly changed the musical style of many music groups. Such new recording techniques also led to occasional social resistance. For instance, the famous 'backmasking' – or playing songs backwards – was invented by the same Beatles, but also led to sounds

that some would classify as 'satanic'. The values embodied in the new technology did not always match society's values.

So, if we want to discuss the moral design of technology, we should reflect on what technology amplifies, and how we can align the intended values of the designers with social values.

1.1.1. Technology amplifies

Let's first discuss the amplifying role of technology. When we look at the world, there are various important societal changes we need to consider when we assess the amplifying role of technology. The most important are privatization and globalization.

I'll explain this further. For various reasons, governments encourage privatization. The idea is simple: private undertakings are more capable of offering social services, since they are able to work more efficiently compared to governmental institutions. Also, a business will come up with more creative and innovative solutions due to the pressure of competition. The promise is that in this

privatized world, the end-user will have more choice for the best price, resulting in a better and financially healthier welfare system. Whatever the effect or success of privatization is, it undeniably leads to a change in moral perception on both the side of the supplier and the end user. A government is simply led by different values compared to a business organization. Where a government primarily acts in support of public interest, a business is led by private interest. Where previous public services— such as healthcare, public transport, or energy supply — are now provided by private organizations, we have to ask ourselves how the public interest is served, and how public values are protected by these private organizations. It becomes even more complicated when complex technologies are used by businesses to realize public values, such as the use of predictive technologies in health care and health care insurances. When complex technology is used, the citizen will find it harder to democratically control governments. For instance, it appeared to be difficult for citizens to review the use of advanced surveillance systems for city safety in Roermond. Furthermore, consumers will not easily use their buying power to make sure businesses behave ethically when they do not know exactly how the involved technology works. For instance, consumers are usually more than willing to give up some of their privacy in exchange for a social media profile, or access to a webshop, mostly due to a lack of knowledge in how the related algorithms work. All this leads to a knowledge and power gap between the individual and the institution, amplifying the effects of distrust and the feeling of dependence, where individual choice no longer has a significant or relevant influence.

Now, let's discuss globalization. Especially economic and technological globalization has a fading effect on borders between nations. This enables us to do business on a worldwide scale, which leads to more intense competition, larger markets and larger businesses. One of the most important pitfalls of such a system is 'regulatory competition'. This means that nations that have the most flexible, or fewer regulations regarding

for example data protection, sustainability or labor laws, will attract most business, including tech-businesses. That also means that the laws of these flexible nation states will not always offer meaningful protection against malpractices among tech-companies. If you consider social media platforms and other internet-based companies, technology easily crosses borders. In the end, this regulatory competition may be at the cost of important human values.

So, with the amplifying capabilities of technology in mind, we have to conclude that not only the public sector plays an important role; the business world does too. And business ethics should be considered in light of techno-moral challenges. That is why this research group operates in both the public, and the private domain. Especially in our Brainport region, where technology development is our trademark, a necessity. We also need to build another bridge, namely between the economic domain and that of technology. I have noticed that ethics is high on the agenda in both domains. However, being an agenda item isn't enough: we need to act. That is why our professorship is part of Fontys School of Business and Communication, but is also composed of researchers from our school of ICT. Remarkably, the domains use different jargon. In managerial language we speak of concepts like triple bottom line, corporate social responsibility and external cost-accounting. In techno-ethics we speak of internal and external values of technology, or perhaps intended, embodied and realized values of technology. What both domains have in common is that until now, the ethical discourse has had an 'inward' focus. In the economic domain, the language is managerial and there is a focus on internal company processes. In the technological domain, the focus is on the internal values of technology, and not so much on societal values. It is our ambition to build a bridge between the domains combining the best of two worlds, and internalize societal values when we innovate technology.

1.1.2. The designer's choice

Let's zoom in on the designer's choice. In the metaphor of the guitar, this is about the intention of the designer when designing the guitar. In general, there is value judgment in the designer's choice. One way or another. For instance, many social media platforms are designed to optimize advertisements. Not to function as a source for news gathering, or not even a platform to meet friends. It is meant to earn money by displaying as many and relevant

advertisements as possible. By showing these advertisements, the platform is designed to let the user spend as much time on the platform as possible.

There are at least two important things to consider when we discuss the intention of the technology designer. First, does the intention align with the society's expectations? In case of social media, it becomes clearer by the day that the expectations of society do not match the designer's intentions. The spreading of fake news is amplified by the algorithmic systems of social media platforms. Not because the designers of these platforms want to spread fake news, but because they want to optimize advertising effects. These 'social bubbles' help that process. To address this issue, we work closely with the Fontys Readership Journalism

and Responsible Innovation, chaired by Danielle Arets.

And second, does the moral character of the designer matter when technology becomes value-laden? In many cases, technology is designed by highly educated white males, a very homogenous group. The characteristics of the white male resonates in the technology they create. A simple example: a recent comparative study across 189 face recognition algorithms revealed that the technologies involved were considerably less accurate in the case of women of color. The reason? They were mainly tested amongst those who created the technologies.

Not only features such as gender and color resonate in the technology we create. Our moral character also seems to resonate. For instance, when Microsoft launched their self-learning chatbot, it was trained in the toxic environment of twitter, which is dominated by angry males (more than 70 per cent). In no time, the chatbot issued sexist and racist messages, including holocaust denial. What we learn from this is that personal features and context of the designer matter. Let me explain.

Empirical ethics teaches us that three major features influence an individual's moral character and decision making. First, these are individual features, such as demographics. Gender, for instance, is also an important feature. Carol Gilligan already revealed in the 80s that men and women have different ethics. Considering that the vast majority of ethicists has always been male, the female voice in ethics urgently needs to be incorporated in how we organize ethics. Secondly, situational factors also play a role in how we 'conduct' ethics. These can be work related. For instance, the reward system within a company can greatly determine how employees behave morally. The biggest corporate scandals in history involve a reward system or work climate that encourages immoral behavior. Think of the reward system that led to the Wells Fargo Accounting Scandal, or the often criticized toxic and 'cult-like' working environment of Facebook. What these cases have in common is that the organizational context created an atmosphere that led people to behave unethically. A third factor is the intensity of the moral issue itself. When moral consequences are not near to the designer, or designers feels they have limited influence on the moral issue, the eagerness to take a moral responsibility will decrease. This is something that typically happens in case of long product chains or processes in which many 'hands' are involved. Take bad labor conditions in supplying countries in complex product chains: these effects are not near to the designer or the end-producer of technology. Or think of complex algorithmic systems used by governments to monitor fraudulent behavior amongst citizens. Each individual actor –including the programmer of the involved technologies- plays a very small role, and it is not always easy or convenient to feel morally responsible for that small role.

When technology designers represent a homogeneous group, we risk that the moral characteristics that resonate in technology are homogenous as well. As a consequence, it will be a challenge to align the designer's values with societal expectations. This could in part be solved by making sure that technology designer teams are less homogenous. So, diversity is

not a left-winged 'woke'-hobby. It is a necessity to create useful and meaningful products. In our research we will make use of methods that measure moral impact of new technology. An example is the Technology Impact Cycle Tool, created by various Fontys researchers. One of the key-components of this method is to make sure that designers are aware of their own composition, and how this relates to our society. Another example is our Moral Intensity Dashboard for Designers, that helps designers to allocate moral responsibility to those whom it concerns, and not those who happen to design.

Another issue that I have already briefly addressed is that technology developers have a natural interest in the internal values of technology, such as technological enthusiasm, effectiveness, reliability, robustness, maintainability, compatibility, quality and rationality. In other words: technology should work well and you should be able to sell it. As mentioned before, there is an inward focus when technology is designed. These values are not always the same as external values that have societal significance, such as safety, health, well-being, sustainability, justice, democracy and inclusiveness. The challenge is to include external values in the internal design of technology. So, apart from more diverse designer teams, we need to make sure that the voice of those who are affected by the new technology is heard at the designer's table. This is why we build moral labs. Our moral labs offer society a seat at the designer table by enabling individuals to choose their own preferred ethical settings in possible technology scenarios.

1.2 Three chords

We discussed the red guitar and its amplifier. Now it is time to reflect on the chords that are suitable to be played on that guitar. As discussed, the moral intention of a designer is never neutral. This leads to the question whether the moral intentions of a designer can be embodied in the technology that is designed. In other words: is it possible to program ethics in machines? We can determine that machines do not work the same way as humans do. This also means that human ethics is not the same as machine ethics. For example, programming a self-driving vehicle entails many ethical questions. When confronted with a crash, how can a designer program the car to respond to that crash? What would you consider to be more ethical? Hitting the older man that will die anyway within a couple of years, or hitting a younger man that has his whole life ahead of him? Is moral programming like this Thomas the Tank-Engine puzzle? There are four ways of doing it right, but are those four ways mutually exclusive? Or can we combine ethical viewpoints

and still create meaningful programming solutions? Or, to use the guitar example, can we play one chord on the guitar at a time, or can we combine chords? One thing is certain: the more complicated technology gets, the higher the risk of failed value alignment. The intention of the designer gets lost in technological translation, and the outcome of the programming works counterproductive.

Another example I'd like to use is the paperclip maximizer, coined by Nick Bostrom. The question is: what if we ask Artificial Intelligence to produce as many paperclips as possible, without a safety net? It will probably end up destroying all human life on earth and transform this planet into one big paperclip factory. Each human being will understand the context of this request: 'produce as many paperclips as possible' and recognize that paperclips only have meaning when they can be put to good use by humans, and that it is not allowable to destroy life in the pursuit of paperclip production. But AI does not see that, unless you instruct it precisely. So we can ask ourselves

the following question: is technology morally charged? Or is technology just a tool, depending on those who design or use it, and therefore morally neutral? The answer to that has been disputed since Plato and Aristotle.

However, nowadays ethicists mostly agree that technology is morally charged, regardless the intention of the designer or user. The amps used by the Beatles, or social media or the self-driving car, transformed our modern society in ways that were well beyond the imagination of its original creators. For this reason, we all have to reject the so-called 'value neutrality thesis' and recognize that technology can have its own moral life.

It is also important to continually monitor whether the intended values for technology, and the embodied values in technology, are still in line with the moral expectations of those who are affected by this technology. If the red guitar is designed to play certain chords, we need to make sure these chords continue to sound great. This is why we build moral labs.



1.3 The truth

This brings me to the third part of the song: The truth. Does moral truth exist? Should we look for it in the first place? This is a question that philosophers have been asking since the existence of philosophy. Especially Enlightenment-thinkers assumed there was moral truth to be found. Utilitarianists claim that the greatest happiness for the greatest number is an all-inclusive rule that can be applied by all human beings for each decision. We call this universalist ethics. Great thinkers such as Immanuel Kant, with his categorical imperative, or John Rawls, with his Theory of Justice proposed moral fundamentals as non-negotiable principles. This is called moral absolutism. There is, however, more to how morality works. As we have seen, individuals do not base their ethics on such theories, but are led by different considerations. This means that in practice, moral truth -if found at all- is more an individual than a collective affair. We therefore need to gain a better understanding of the moral solutions that individuals seek in technological solutions. We need to involve those who are affected by new technology in the design process at an early stage; interact with basic prototypes and enable them to change the moral context to their personal liking. This is why we build moral labs. In a way, we need people to give us the building blocks of new technology. For example; imagine an HRM bot that can match perfect candidates with perfect jobs. What kind of data should be most important? Someone's CV? Someone's social (online) behavior or perhaps even someone's DNA? This was one of our cases in the moral lab installation that we used during the Dutch Design Week in 2019, in collaboration with

designer's collective We Are and the municipality of Eindhoven. We noticed that people have different thoughts about how candidates should be linked to jobs through automated processes. So, how do we translate these insights into design principles when people have different ideas about moral design? In our methods, we use so called 'augmented utilitarianism' as a way to explore functional moral settings in technology. This is complemented by translating moral sentiments into new technology that is attuned to a personal values dictionary. Now, we not only know what people want in terms of moral solutions and why, we also understand the conflicting values between the solutions. For instance, in the design of surveillance camera systems, it is important to understand why people find privacy important. The word privacy does not give us too much information. The motivation to aim for privacy reveals a bit more: do we find it important for personal security? For power over personal data? For realizing self-direction? Or perhaps a more universal principle to guarantee equality? And if privacy is important, how does it relate to public security? Or to the wellbeing of others in our society? In understanding all this, we are better equipped to create and explain moral choices in the design of our technology as well as be more transparent about, for instance, our business models. We are also better equipped to understand who should decide on the morality of technology. Is it the designer team? The end user? Or perhaps society as a whole through legislation? As a rule of thumb we could say that those who are affected the most by new technology should have the loudest voice at the design table. These could be individuals, but most of the time, this is a society-wide issue.



3. Our research group

I will not be doing this research alone. Moral design strategy is a wonderful research group composed of very talented and warm people. They are the backbone of our activities and our moral compass. Together, we will focus on three main moral design strategy challenges in the coming years. The first concerns moral design and democracy. We will develop methods and practices through which citizens can make their voice heard at the design table of new technology. The Moral Data City Hunt is an example of that. Our endeavors to morally design technological solutions to better deal with disinformation through social media networks is also an important part of this 'research pillar'. Together with our colleagues from Journalism and Responsible Innovation, we just received a research grant to explore how local media outlets can redesign their business models and align their online media with public values. The second challenge concerns applied human rights. What would happen when human rights are more than just legal concepts, and become design principles? What would governance, products and services look like? A new book will be published this year by Wageningen Academic in which this is further explored. We are also part of the LIME (Hogeschool Zuyd and Maastricht University) partnership where we use

moral labs to position moral product design in the health sector and recently explored how we can program moral drones with applied human rights as design principles. Last but not least, we will focus on moral strategy and leadership. To be able to keep up with the pace of constant innovation, we do not always need solutions. Instead, we need strategies that enable us to stay on course with moral design. For instance, our Karen project, where we collaborate with Ottawa University, is a good example of the use of a personal values dictionary to better understand memes, and how to respond to this when businesses (or other entities) get dragged into an online political discussion.

What a team! I'm happy to work with this ambitious group. We have so much plans and our enthusiasm continues to grow. Now, let's look ahead a bit. What would we like to see accomplished in four years' time? We want to make sure the red guitar is amplified so that everyone can hear and enjoy the chords they love. We want to become a leading research actor in translating morality to design perspectives. We want to guarantee a seat at the designer's table for society. We will do this by helping businesses and governments to implement strategies of moral leadership in technology. One thing is crucial in this: never lose track of the hopes, fears, and moral gut feeling of those who are affected by the

new technology you design or use. We want our initiatives to use design to stretch people's moral imagination and make them decide about future technologies becoming a standard part of innovation processes. We are currently planning to organize Moral City Hunts each year, involving multiple cities and multiple stakeholders in and outside the Netherlands. This way, we not only learn from the moral particularities of neighborhoods, but also compare them with twin cities. Our moral labs can be built in many contexts. As we speak, we are exploring possibilities to build such labs for a soccer stadium, quantum-technology applications in the financial sector, a city mobility app, health-care apps, delivery drones, social media platforms and datacenters. We look forward to it, and on a more personal level, I am grateful to be able to do that with a team of highly talented researchers, a warm and encouraging Fontys society, a network of inspiring actors, and academic brothers-in-arms all over the world.

All I've got is a red guitar, three chords and the truth. All I've got is a red guitar, ...the rest is up to you.

Bart Wernaart



> FOR SOCIETY

Bart Wernaart
Lector Moral Design Strategy
20-1-2022

FONTYS.NL