A Call to Technologists

Aaditeshwar Seth
IIT Delhi and Gram Vaani, India
aseth@cse.iitd.ac.in, aseth@gramvaani.org

Abstract—As technologists, we all want to ensure that responsible outcomes arise from our innovations. Through a chain of reasoning, we describe several limitations faced by technologists in achieving this, ranging from gaps in the design methods in use currently, common approaches to spot problems and find solutions, the organizational structures within which technologists operate, and the wider political economy of technology itself. We suggest several measures to address these challenges and conclude with a call to technologists to have a clarity of purpose as a tool to navigate these challenges.

Keywords—computer science, ethics, responsibility, design, deployment, technology workers, power, social development, inequality, political economy, communication, participatory media

I. ETHICAL GUARDRAILS FOR TECHNOLOGY

As the designers and managers of computer based systems, for all the wonderful positive examples of our innovations used to make the world a better place, we also have a plethora of negative examples of misuse or (unintended) undesirable harms arising from these technologies. The argument that technology in itself is neutral is often used to evade responsibility or distribute responsibility between design and deployment, developers and managers, researchers and industrialists, but is this the best that we can do? How can we ensure that responsible outcomes arise from all of our hard labour in designing and managing computer based systems? Afterall, not many of us would say that we are here to do harm.

Medical ethics has evolved a rich set of protocols to ensure minimal unintended harm from new innovations. It ultimately relies on doing extensive evaluations before new innovations are permitted to be released in the world. Innovations in computer science and engineering however do not go through such a check. A complex ecosystem of finance, research, entrepreneurship and even politics has supported the rapid growth and scaling of new technologies around the world, and the Internet has provided an unprecedented reach and speed for the diffusion of these developments. Why are ethics-based review protocols missing in our ecosystem? One explanation could be that our existence as biological beings has been known for a long time, and therefore medical ethics evolved to test and rigorously evaluate any artefacts that affect our biology. However, our existence as informational beings is only now being discussed, of how information affects our beliefs, shapes our actions, and not just individually but at a societal scale [1]. Hence only now are we seeing discussions about the need to develop ethical protocols for computer systems that channel information, and sometimes even produce information, for our consumption.

Now that we do realize this, how can we begin to build ethical guardrails for our innovations? Possibly the awareness among designers and managers of information systems to question the impact of the technological artefacts created by us, or processes governing the management and use of these artefacts, has not been higher ever before. Concerns about privacy, fairness, safety, etc of information systems seem to be actively discussed and debated at various fora these days. The current methods being developed though to build the required ethical guardrails, are not sufficient, in our opinion. The dominant approach that has emerged in our field is to incorporate ethics into the design of the artefacts [2, 3], but we argue that this method is limited in its potential. The rationale for ethics by design probably emerges from how it is done in medical ethics, and even regulatory authorities favour this route, based on the assumption that innovations designed upfront with principles like privacy or fairness encoded in them, will not go wrong. We argue that regulatory compliances to prove adherence by design, may certainly be a necessary step forward, but it is not sufficient.

There are many reasons for this. We have discussed elsewhere in detail based on our decade long experience of running Gram Vaani, a technology based social enterprise to support participatory media platforms in rural areas for less-literate populations, that even beyond careful initial design, concerns can arise at the socio-technological interface when people begin to use and adapt the technologies to their needs [4]. Aspects like who is included or excluded from access to the technologies, shaping of appropriate usage norms of the technologies, the line between flexibilities and constraints for appropriation of the technologies, ensuring social impact, etc, needed careful management of the platforms. This management was done by evolving a rich set of processes, such as collecting feedback about the use of the platforms, development of editorial policies, and whether or not to allocate financial resources towards meeting the social impact mission of our enterprise. Design changes in the technology were of course subsequently needed to implement these processes effectively, thus making design an evolving process itself, but our point is that just like the approach to incorporate ethics into the design of innovations it is also required to incorporate ethics into the processes of managing the innovations. We argue that ethics can be that common denominator which can provide the necessary guardrails to evolve management processes that can ensure responsible outcomes from technology, beyond what might be ensured through incorporating ethics into the initial design of the technology [5]. Undesirable outcomes can arise due to broken management processes to notice and react to problems
unforeseen during the design, and we discuss later many circumstances when this might happen.

Another reason why ethics by design is not sufficient is the current context in which information systems have deeply permeated our lives already. These systems are being used by millions and billions by people, and are embedded in a complex global web of finance and politics, that it is daunting to even conceive an eventuality in which they will be redesigned or replaced. We may be seeing the classic Collingridge dilemma of failing to control technology until it is too late. Even corrective regulatory mechanisms that might evolve like compliance protocols to validate ethics in the design, or even the management processes, could be corrupted by powerful vested interests. Paltry fines even in proven incidences of misplaced responsibilities seem to point towards what could become an unfortunate reality [6]. Further, an approach of regulatory compliances tends to lead to an attitude of outsourced morality, rather than take a pro-active approach of owning responsibility of the outcomes from our actions.

Given these limitations, what prevents us as designers and managers of these innovations, to exercise some discretion ourselves in our work? It was heartening for us to hear during a session on fairness in AI at a prominent data sciences conference, from the developers working at some of the largest technology companies, that they are in a powerful position to shape what their companies do with technology built from their labour [7]. Whether this is an illusion or indeed true that white collar workers can influence their employers and shareholders in technology governance, is an open question to which we will probably find answers in due course. For now, we want to draw attention to several perceptible gaps that seem to exist in our own understanding as designers and managers of information systems, in how we can exercise some discretion ourselves.

II. SPOTTING PROBLEMS

An important gap is often a lack of clarity about the goals that our innovations should achieve through their deployment. An example is information feeds in content recommendation systems, where there is an inherent ambiguity of whether to optimize the feeds for business metrics such as time spent by the users [8], or normative user experience metrics such as diversity in content recommendations [9]. Metrics such as time spent are known to lead to filter bubbles and prioritization of information that is sensational, which can lead to good business but causes social harm, whereas metrics such as diversity may lead to less user engagement but would demonstrate a commitment towards pluralism and its potential in shaping social norms. With an ambiguity about the objectives of our innovations and what are they meant for, even as well-intentioned technologists we can get lost in what should we be challenging in the first place: The objectives themselves, or a lack of clearly articulated objectives, or problems in the methods to meet the objectives. How can we achieve this clarity?

Shown in Figure 1 is a three-layered framework where we suggest that a common ethical system needs to provide underlying guidelines in defining the objectives, design, and management processes, of any technological system [5]. The ethical system can provide clarity in framing what might be good objectives to pursue, and consequently in guiding the design of the system to meet the objectives, followed by processes to manage different aspects of the socio-technological interface during the deployment of the system. The ethics of the design itself, at the middle layer, can be considered on at least three fronts. First, the user interface may have persuasive elements that can nudge users towards certain actions, raising questions about informed choice and appropriate design that should be settled based on the underlying ethical system. Second, the data and algorithms would raise questions about privacy, biases in the data, the definitions incorporated for algorithmic fairness, etc, which similarly need to be clarified based on the ethical system. Third, the system design would present choices such as whether to build systems that have centralized decision making elements Vs a decentralized design, assisted access Vs private access, aimed at collective interactions Vs individualized interactions, and other such design choices that can shape the relationships between the direct and indirect users of the technological system, and with the technology itself. In summary, we suggest that in any information system, each of these three elements of the design, along with the objectives of the information system, and the management processes, should be guided by a common underlying ethical system that is followed by the designers and managers of the information system. We have discussed this framework elsewhere in detail and also used it to run simple consistency checks on a few information systems (Facebook, Aadhaar – the biometric based unique identification system in India, and Mobile Vaani – a voice based participatory media platform for less literate users) to evaluate if their stated objectives, and observed design and management processes, are internally consistent with one another in terms of the ethical values they demonstrate [5]. For example, in the case of Aadhaar, the choice of biometrics as an authentication method seems to have been inspired by ethical values of inclusion of even less-literate users, but the handling of false negatives arising due to technological or other problems was inadequate and led to denials of welfare benefits to the needy, revealing that different ethical values were operating during the management of the Aadhaar platform. Examining innovations through an ethical lens can therefore provide a systematic approach to technologists to overcome ambiguities of where to place their attention. The objectives, design, and management processes can all be questioned based on a clearly articulated common ethical system, and deviations from it can be identified and addressed.

Figure 1: Ethical underpinnings to information systems
III. Finding Solutions

What other factors can impede us in enforcing discretion in our actions, or in addressing inconsistencies in the ethics based framing of the information systems we design and manage? The next gap we feel arises from our dominant belief in a deterministic world, which likely arises from the positivist paradigm of science and technology in which most of us have been trained to operate. We tend to expect, much like a computer system, that specific inputs will lead to specific outputs. Dahlbom and Mathiassen discuss this passionately in their analysis of a romantic Vs mechanistic view of information systems, and argue that the outcomes of technological systems cannot be mechanistically determined [10]. Rather the outcomes are shaped by the context in which the technological systems are deployed. The culture, power relationships between the users, the agenda of users behind using the technology, etc, all influence the eventual outcomes and are equally shaped by the technology as well. It is probably the rarity of such a romantic view of technology among most of us that as technologists we often come up with purely technological solutions to solve problems that might be stemming from deeper societal issues, or we feel confident in designing systems that are flawless to begin with, or we ignore the inherent diversity of people and communities with a belief that we can model all this complexity in a machine. Facebook’s attempt at encoding communication norms into algorithmically driven community standards, or Uber’s attempts at formalizing relationships between drivers and users and cities, are examples of a dominant mechanistic view in approaching problems. We should not get carried away with our own technological prowess, and realize that there are limits to what technology can solve. How can we counter this limitation in our training so that we can come up with better solutions?

A more comprehensive education curriculum is certainly a no-brainer to address this problem that seems to have emerged from biases in how most of us have been trained [11]. Additionally, rich methods of participatory design have emerged over the years to help technologists embrace a romantic view of technology [12]. Methods of ethnographically informed user centred design, co-design, etc aim to understand the context in which the technology is deployed, and go on to understand the effects of the technology. We need to embrace such methods so that we can be better informed and effectively react to problems. The emphasis we have placed on management processes to govern the deployment also emerges from a similar realization.

Unfortunately however, in most cases these methods are applied only to help the designers and managers understand the users as consumers of their applications to increase usage. It is rare to come across cases where emergent undesirable outcomes in some contexts are noticed and addressed by them unless there are business or other advantages emerging from it for the organization. Extremely rare are cases where users can participate in the governance of the systems themselves, to control and shape the use of the systems in their respective contexts and achieve the outcomes that they desire, as opposed to helping designers and managers achieve their organizational objectives. The original Scandinavian tradition of participatory design was centred on such an ideal approach, to jointly determine with the users what objectives to achieve through the technology, and accordingly design it [13]. Mumford’s ETHICS (Effective Technical and Human Implementation of Computer Based System) framework even operationalized a similar approach [14]. Such methods become imperative in today’s platform dominated context where an unprecedented number of users engage on the same platform but are not involved in any governance of the platform. In this situation, what prevents us as technologists from giving more agency to the users of our innovations?

IV. Are Our Hands Tied?

This brings us to at least two additional formidable gaps that we face in imposing more discretion ourselves. The first challenge lies in the organizational structure and culture of the companies or government departments where we might be working [5]. Most large organizations are structured along functionally segregated roles that can prevent a free flow of information across all units. For example, engineers may not hear all that the ethnographic study teams have to say about problems that some users segments may be facing, and these problems may go unaddressed. Further, different teams and individuals may be operating under different ethical systems, and insufficient socialization among them can lead to inconsistencies in how they choose to respond to various observations. Teams that interact with users in person may be more empathetic in their response to user problems, whereas teams that are removed from direct user interactions may choose to prioritize other issues, or respond differently. For example, cases of defrauding less technically skilled users through fraudulent mobile money transactions may bring different responses depending upon different underlying ethical systems within which the team members may be operating. Some may want to run training workshops for the users, while others may want to improve the technology and find technical solutions to spotting fraud, while yet others may choose to do nothing and encourage the users to learn on their own, and some may want to create user associations that can democratically discuss and decide on the best route to choose.

Therefore, the organizational structure, clarity in the organizational mandate, and cultures of assimilation between diverse teams, can be important determinants in getting heard
or having the required authority to impose responsible practices ourselves in the design and management of information systems. Clearly unless these aspects are not addressed, we may not have the required tools to impose meaningful discretion ourselves in governing our innovations. Mechanisms such as co-determination by employees through board representation [15], or tech-worker committees that can mandatorily draw attention and ensure redressal of various issues, may need to be imposed by law to address these gaps.

The next challenge we discuss lies in the broader political economy of technology itself [4]. Figure 2 shows the conflicts within which technologists are embedded. The nature of technology used in the information systems of today requires large amounts of capital, which in turn requires investment, and puts the company owners on a path to achieve rapid scale-up to meet investor expectations, as opposed to growing slowly with careful iterations applied on the design and management processes to ensure responsible outcomes. Next, social objectives and business objectives are often different from each other, and companies therefore naturally tend towards meeting business metrics often at the expense of causing or ignoring social issues that may emerge as a collateral outcome. Governments in theory do exclusively want to meet social objectives, and that too in a democratic manner, but many recent events have reinforced that tight interlocks exist between governments and companies, and shifts the priorities away from social objectives [16, 17]. In fact, right-wing governments like in India have effectively used information systems to mould public opinion not only for their own populist agenda but to also endorse a vision of bringing social change through technology while clouding the scenarios when it leads to harm. Further, governments have often chosen to use technology as a means of imposing greater control and coordination of the population with a high modernity assumption of bringing social good, but such initiatives often disempower the people and reinforce inequalities [18]. As technologists, we may be able to get better at spotting problems arising from our innovations, we may improve our own awareness to come up with effective solutions to handle these problems, we may acquire sufficient agency ourselves to enforce that our organizations pay attention to these problems, but how can we deal with an entire global system of capital and politics that seems to be too powerful to render all such efforts meaningless?

V. clarity of purpose

Let us not resign ourselves to disappointment though, and recall that all information systems designed by technologists like us have indeed been used to change many aspects of this system of capital and politics. The same innovations that may have been appropriated for other objectives or may have been mismanaged or inappropriately designed, can also be used as a force for good. We see hints all around us of viable pathways to steer systems towards this direction. Schumacher advocated many decades ago to design small technology which is appropriate to the local context and can be understood and controlled by the people, as opposed to large industrial systems that are instead commonly adopted because they may favour economies of scale [20]. We probably face the same question with the forms of information technologies in use today, and need to choose and support more appropriate forms of technology. Marx pointed out centuries ago the problems with private ownership of technology [21], Ostrom showed how people can cooperatively evolve methods to manage the commons [22], and open-source systems have demonstrated that rapid innovation can emerge through shared knowledge and resources. Such experiments that have happened over the years, and continue to unfold, can reveal suitable organizational structures within which to develop information systems going forward. Referring to the framework shown in Figure 1, similar to the neo-humanist paradigm advocated for information systems [19] we also suggest that power-based equality should be a universal value that information systems should uphold because it can counter many of the recurring problems shown in Figure 2 that arise in today’s global system of capital and politics [5]. Information systems have the potential to alter power relationships in society, and carefully designed and managed systems can alter these dynamics to empower the weak. Other universal values can be identified as well. It is these values that can help us discover meaningful objectives on how to solve some of the big problems we see around us of inequality, exploitation, intolerance, propaganda, environment, and so on.

We therefore cannot say that paths do not exist, they may just need to be discovered, and this indeed should be our purpose as technologists to ensure that technology becomes unanimously a force for good. In today’s socio-economic and environmental context, we can no longer remain naïve in our understanding of the world and forsake the responsibility to govern how our innovations get used. We need to learn to navigate the various complex variables mentioned above so that we can develop appropriate organizational structures, discover meaningful objectives to solve for, and design and manage the solutions in ethically consistent ways. Good intentions and a clarity of purpose can help us achieve this.

REFERENCES


Leveraging Web Data to Monitor Changes in Corporate-Government Interlocks in India. In Proc. COMPASS.


