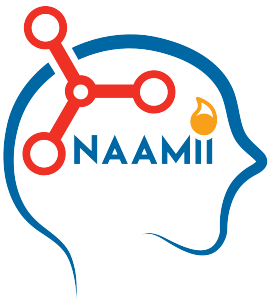


AI Ethics Survey in Nepal

Shreyasha Paudel, Aatiz Ghimire

This project is generously funded by the educational grant from UNESCO Asia and Pacific Regional Bureau

Contents

List of Tables

List of Figures

Acronyms

- **AI:** Artificial Intelligence
- **LMIC:** Low and Middle Income Countries
- **ML:** Machine Learning
- **MOOC:** Massive Open Online Course
- **OECD:** Organisation for Economic Co-operation and Development
- **STEM:** Science, Technology, Engineering, and Management
- **UGC:** University Grant Commission

Acknowledgements

This report is prepared as part of the research project 'A survey of AI Ethics among stakeholders in Nepal' funded generously by a grant from UNESCO Asia and Pacific Bureau.

We would also like to thank Mr. Pesal Rai who contributed to the project as a volunteer. He assisted with survey dissemination and data analysis. This project would also have not been possible without the guidance from Dr. Bishesh Khanal and assistance from NAAMII team. Lastly, we would also like to thank the survey respondents, focus group discussion participants, and workshop participants who contributed their time and knowledge to help with our research.

Chapter 1

Introduction

The main goal of this project is to do a landscape mapping of understanding and approaches to Artificial Intelligence (AI) and AI Ethics in Nepal among three key stakeholders: Technology Students, Technology Professionals, and Policymakers.

There is an immense potential of AI to transform the economic situation of countries like Nepal [1]. However, there is also a big risk that the AI-driven development dramatically broadens the gap of the existing digital divide. In the context of Nepal, the development of AI is scaling rapidly as evidenced by a growing number of AI communities, AI companies, and AI courses [1, 2]. However, such development is happening in a vacuum without government policies, regulations or even reliable sources of information that track the application and use of AI. In such a context, it is likely that AI application and education would also be dominated by few privileged groups in urban areas similar to other technical fields [3]. In addition, most people equate automation with progress and there is a lack of active ecosystem to keep technology practice accountable. As a result, there is a risk of AI development being unequal and exacerbating existing inequalities in the country.

To avoid such a risk, there is a need to contextualize responsible AI practices within the economic and cultural context of Nepal as well as to initiate conversations among different stakeholders to understand each other's concerns and challenges. This project aims to fill this gap through a survey among three key stakeholders: students, professionals, and policymakers. In this report, we publish findings from the survey to create an informational resource on understandings and current practices in Nepal, identify gaps in understanding of AI ethics among surveyed stakeholders, and use these insights to recommend future

actions to strengthen responsible AI practices in Nepal.

The rest of the report is organized as follows. We first present a short overview of existing AI ethics research and policies in the global context as well as the information we could find about state of AI in Nepal. We follow that with a brief description of our research design and data analysis for this survey. Finally, we present results for each stakeholder group - students, professionals, and policymakers. We conclude with a short discussion of the limitation of the project followed by recommendations.

Chapter 2

Background

2.1 Overview of AI and AI Ethics

Artificial Intelligence (AI) can be defined as the ability of man-made agents to successfully perform tasks that are generally associated with human intelligence. While AI research is often associated in popular media as trying to replicate human cognitive function, much of AI research and product development today focuses on statistical AI to solve specific narrowly defined problems [4]. With advancement in machine-learning and deep-learning, AI is capable of solving many tasks such as image classification, optical character recognition, speech transcription to an accuracy comparable to human performance [5]. Such applications have been deployed in many consumer products [5]. However, these algorithms are statistical models built using large amount of training data and thus are limited in their adaptability and generalizability. As the integration and use of such AI models has increased in many criticial decision-making areas, so has the the possibility of harms. Some of these harms have already materialized as evidenced in gender and racial bias [6, 7], social media manipulation [8], and many others documented by AI Incident database [9]. Some of the main concerns highlighted by AI ethics researchers include loss of privacy and increased surveillance due to the need for large amounts of data, algorithmic bias, lack of explainability of AI models, labor exploitation through automation, and unethical application of AI such as its use for spreading misinformation [7, 10].

AI ethics has emerged as a response to counter the growing concerns regarding AI's impacts. At a very high level, AI ethics is a framework for formalizing the impacts and

implications of AI. While AI ethics is sometimes used to philosophize about responsibility and morality of robots, we consider ethical AI as a way to ensure human autonomy and human welfare and prevent algorithmic harms in this work [11]. In this context, AI ethics is a discipline in a very nascent stage and has been influenced by a variety of predecessors such as law, philosophy, engineering ethics, and science and technology studies.

As expected, a major influence in AI ethics comes from Western moral philosophy. Researchers have explored the philosophical principles of utilitarianism, human rights, and virtue ethics as frameworks an principles to theorize impact of AI [4]. More recently, researchers have also explored Ubuntu [12], relational ethics [13], and Buddhist ethics [14] as frameworks to decolonize AI. Outside of philosophy, academics and activists have connected the socio-technical nature of AI with similar issues seen in the past technologies using frameworks from science and technology studies and medical ethics [15]. Simlarly, scholars have looked at issues of AI fairness, bias, and discrimination using human rights and existing legal frameworks [16]. Lastly, practitioners have also drawn from enginering ethics and existing codes in environmental, civil, and mechanical engineering that discuss and debate social and environmental impact of engineering projects to do the same for AI and software [4, 17].

2.2 State of AI Ethics Globally

Over the past few years, tech corporations, governments, civil society groups, and multi-stakeholder organizations have published dozens of high-level AI ethics principles [17, 18]. The main goals for these principles are to maximize benefits from AI systems and minimize potential harms. These principles have major topics in common such as transparency, justice, fairness, and equality, non-maleficence, responsibility, privacy, human dignity, freedom, and autonomy. [4,17,18]. Some early attempts at international governance include the OECD AI Principles [19], the G20 Human-centred AI principles [20], and UNESCO's AI Ethics guidelines [21]. In addition, various nations have defined their own principles in their national AI strategies such as the Pan-Canadian AI strategy, the Singapore national AI strategy, Colombia's regulatory system for AI, and various others [17, 18, 22].

Along with guidelines from international organizations and national frameworks,

multinational technological companies such as Microsoft [23], Google [24], Facebook [25], and IBM [26] have also released their own AI ethics principles. While these principles are also based on similar high level principles as the previous guidelines, they are also designed with a goal of guiding AI developers and engineers within these companies. These principles also serve as a way for these companies to set AI development agenda and build public trust in their products. Our review of these guides showed that the principles closely adhere to OECD principles, but Google has not explicitly stated accountability while IBM hasn't explictly mentioned privacy. Critics decry these efforts from private companies as "their way of lobbying and marketing for self-governance" and avoiding accountability. Besides AI companies, there are a number of non-profit civic organizations such as Montreal AI Ethics Institute, Parternship in AI, and Ada Lovelace Institute and research organizations such as AI Now Institute and Data and Society have been developing AI ethics guidelines and advocating for responsible AI ethics practices. Similarly activist organizations such as Algorithmic Justice League and Algorithm Watch are publicizing existing harms via algorithmic audits and advocating for more accountability. Recently, media publications such as The Markup and Rest of World have also been founded with an objective of reporting on technology and AI related issues. However, most of these organizations are based in North America or Europe. Our research could not find a similar ecosystem in the Global South.

Despite a growing body of work and focus on global principles that prioritize human well-beings, AI ethics as a discipline is still limited in its representation of values and perspectives from different regions, entities, and communities [17]. As the above list shows, more economically developed countries are shaping the guidelines and values. Low and Middle Income countries from Asia, Africa, Latin America, and the Caribbean are severely underrepresented. Another area of concern is how the debate around responsible AI largely focuses on how to improve these algorithmic systems and rarely question the incentive mechanisms and structural factors that guide technological product development [17]. Lastly, there is a need for operationalizing these principles into practice and ensuring accountability [27]. This can only be done by having non-ambiguous definitions of principles that is applicable to local contexts. More work and debates are necessary for this to happen.

2.3 State of AI in Nepal

Even though AI companies and communities are rapidly growing in Nepal, there is a lack of reliable information about these communities and companies. As far as we are aware, neither the government nor research universities have collected and shared this information. We found a few crowd-sourced list of AI companies collected by individuals which list nearly thirty-six companies working in the field of AI [2, 28]. Most of these companies are outsourcing companies working for international clients. Very few of them have local products for Nepal such as chatbots, recommendation engines, or health AI application. Besides companies, there are a few active ML/AI communities who run in-person and virtual networking and community events such as AI for Development, AI/ML community Nepal, DN: AI Developers Nepal, and more. However, many of these groups appear to be short-lived. There is also a lack of centralized information source about these communities, their members, and their impact.

At the University level, both computing and non-computing students are enthusiasts about Artificial Intelligence and its technologies. Based on preliminary research and conversation with students, most computing related students complete an AI related project as their final year projects. In addition, many students self-learn AI skills on their own through boot camps, MOOCs, special schools including NAAMII's annual winter school, AI fellowships and more. Many students publish their final projects and independent projects as public github projects and in blogs. However, it is difficult to get a comprehensive information about scale and impact of these projects since there has not been a central database or systematic study about these projects and their impacts.

Nepal is also lacking in terms of AI related policies. While the Government of Nepal has prioritized AI and information technology in its Science, Technology and Innovation Policy and Digital Nepal Framework, these policies lack practical details for implementation [29, 30]. The use of AI is governed by the Electronic Transaction Act which ensured digital privacy and security and use of online algorithms [31]. However, this act released in 2008 has no provisions specific to the use of AI or automated systems. Research ethics for academic research in AI is governed by the UGC Policy regarding Research Misconduct, and the ethical issues for health related application falls under National Ethical Guidelines for Health Research [32, 33]. However, these policies usually

only cover specific research projects. It is also unclear whether AI practitioners are aware of these policies as there is a lack of enforcement. More general policy framework and enforcement mechanism guiding all data collection and use for AI projects is needed.

This work attempts to bridge the gap between state of AI in Nepal with the global AI Ethics literature. Despite progress made globally, our research shows that there is a lack of information, policy, and robust AI ethics ecosystem in Nepal.

Chapter 3

Research Design

3.1 Research objectives

This project is an attempt to start conversations about AI Ethics and Responsible AI alongside other AI related developments. As a survey research project, the main objectives of this study are as follows:

- Create an informational resource on understandings and current practices in Nepal.
- Identify shared concerns and conflicting opinions among surveyed stakeholders.
- Identify gaps and limitation in current understanding among surveyed stakeholders.
- Use insights to spark discussions.
- In the long run, help contextualize responsible AI practices for Nepali society.

3.2 Survey Design and Dissemination

Three sets of survey for three groups of stakeholders – students, IT professionals, and policymakers was designed. The student survey was open to all the students pursuing Bachelors degree and above in Nepal. Engineering, computing, and business students who were taking classes in AI or data science were specifically targeted. We defined the professionals as individuals who were working in a computing related job either in Nepal or building a product for Nepal if they were working abroad. For this project, policymakers

were defined as individuals who were responsible for setting technology related policy for the governments, were working at a policy research group, or worked in a high level decision-making capacity at academia. The respondents are described in more detail in section 4.1,section 5.1, and section 6.1

Each of the survey covered demographic information in the first section, multiple choice and open-ended questions to gauge the understanding of the participants about AI ethics related topics in the second section, and optional questions to judge interest in AI Ethics at the end. The survey for the students and professionals also had an additional section with questions about AI ethics related topics in the curriculum and AI ethics process in the workplace respectively. Based on our preliminary analysis and experience from prior work, the survey for policymakers was designed to be much shorter and less technical than the other two groups. The surveys are shown in section 7.3

All the surveys were released online and shared through social media between December and February 2022. We also conducted targeted outreach for each group. For the students, we sent emails to major engineering and computing institutions in Nepal. The survey was also shared to the student participants at a 10 day AI school organized by NAAMII. Similarly, for disseminating the professional survey the team sent targeted email and messages to individuals within professional networks of the team as well as leading AI and data related companies in Nepal along with a social media campaign. The distribution for policymakers was more targeted and conducted via emails, phone calls and messaging on online platforms.

3.3 Data Analysis

The survey data was analyzed both qualitatively and quantitatively. Prior to analysis, any outlier data entry (i.e. the respondents who did not fit into the desired criteria as described in section 3.2) were removed along with any identifying information such as e-mail and contact info for the respondents. The quantitative data analysis for all categorical, numerical, and yes-no questions were done using Python. The open-ended and text based questions were analyzed through a thematic coding. The responses were also analyzed based on different demographic categories such as gender, education, work-role. The

demographic categories used for analysis was chosen based on whether the responses were diverse and whether the sample for the different categories were sufficiently representative.

Chapter 4

Survey Results: Students

4.1 Survey Participants

The survey was filled out by 217 students throughout Nepal. Of them, 82 % were male and 18 % were female. Though the survey had an option for choosing non-binary or refusing to disclose one's gender, none of the respondents chose to do so. The gender ratio of respondents reflects the gender ratio of enrolment in engineering campuses in Nepal [3]. Similarly, the majority of respondents were pursuing their undergraduate degree. About 61 % were pursuing an engineering degree and 37 % were pursuing a non-engineering but computing related degree. As shown in Figure 4.1, most of the respondents were from Kathmandu valley. In terms of their caste and ethnicity, majority were upper caste with about 52 % identifying as Brahmins and 11 % as Chhetris. Of the remaining, 14 % identified as Janjatis, 11 % as Madhesis, 9 % as Newars and 3 % as others. The other category included religious minorities such as Muslim as well as ones who refused to disclose their caste. None of the respondents identified themselves as Dalits. The overall demographic distribution of the survey respondents is shown in Figure 4.1.

4.2 Overall impression about AI Ethics

The students unequivocally considered ethics and AI ethics as an important topic with 85% agreeing that it is a topic of concern. Their main concerns included the need to keep AI systems safe as their applications increase in society along with the need to

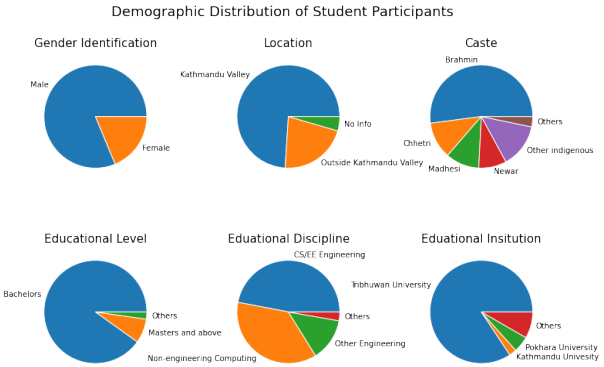


Figure 4.1: Demographic distribution of student respondents

respect human dignity, privacy risks and a lack of regulation. However, when asked if AI ethics consideration in low and middle income developing countries like Nepal would be different than more technologically advanced countries, only about 48% said yes while 26% were unsure if this is the case. The most cited reasons for such a difference in AI ethics considerations included lack of AI skills, lack of infrastructure and a lack of AI policy and good governance in LMIC countries.

The survey also asked students to rate a range of statements about AI and its impact. Figure 4.2 shows the results. This graph shows that the students were in general optimistic about the need and impact of AI in Nepal and less aware of the potential harms and technical limitations. While students were aware of data bias and discrimination as potential risks, they tended to think of these as technical problems which can be solved by larger datasets, better evaluation benchmarks, and more advanced technologies. Majority of respondents did not think of underlying societal factors that may cause AI and training data to be biased.

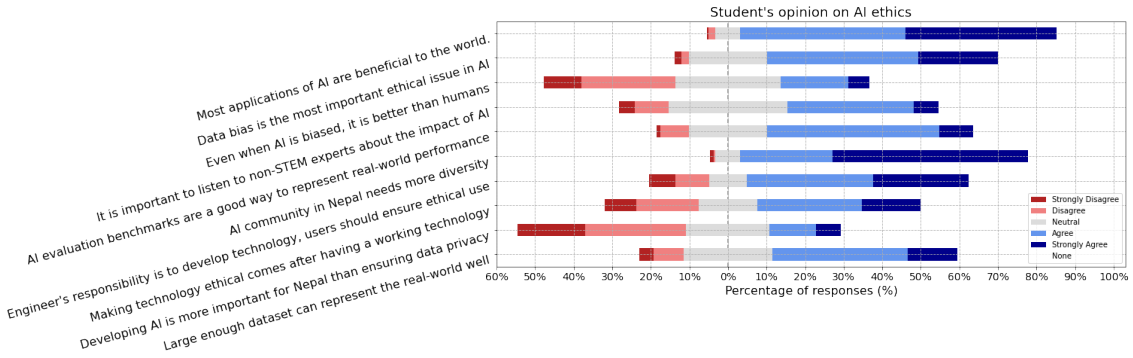


Figure 4.2: Students' response on whether they agree with various statements regarding AI and AI ethics.

4.3 Ensuring AI Ethics in Nepal

When asked about risks of AI to Nepal, most students were concerned about a potential for labor disruption, potential for AI misuse and harm along with risks to data privacy and security. This was also reflected when asked to choose the most important topic for AI Ethics in Nepal as shown in Figure 4.3where 70% of the respondents chose Privacy followed by 62% choosing explainability and 58% choosing fairness and bias related issues. Despite the concerns for potential for labor disruption in their qualitative answers, only 27% thought labor was an important topic for AI Ethics.

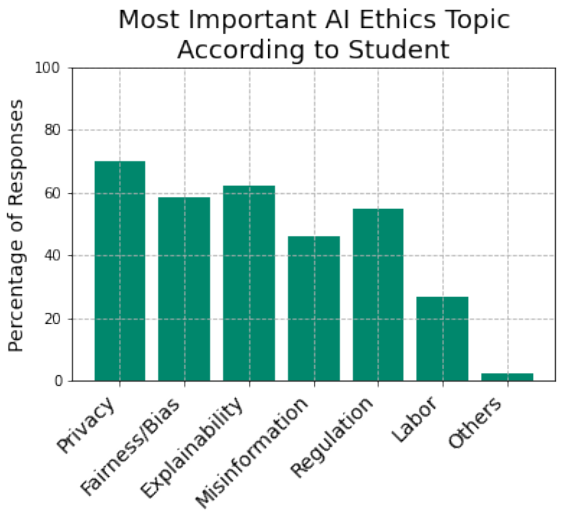


Figure 4.3: Students' response to the most important topics for AI ethics in Nepal

The students were also asked to select their top three choices among different ways for ensuring ethical AI in Nepal.As shown in Figure 4.4, there was quite a variation among the responses with a majority choosing regulations and audits by AI developers.The students were more unanimous on their choices for who is most responsible for ensuring ethical AI as shown in Figure 4.5. Almost 80% chose AI engineers and developers, followed by more than 50% choosing government and AI companies.

4.4 AI Ethics in Education

The students were asked about their courses related to AI and AI ethics. 34 % of the respondents said they were taking formal courses in AI. Of them 48 % said they had a formal ethics curriculum and 26 % said they had AI ethics curriculum.

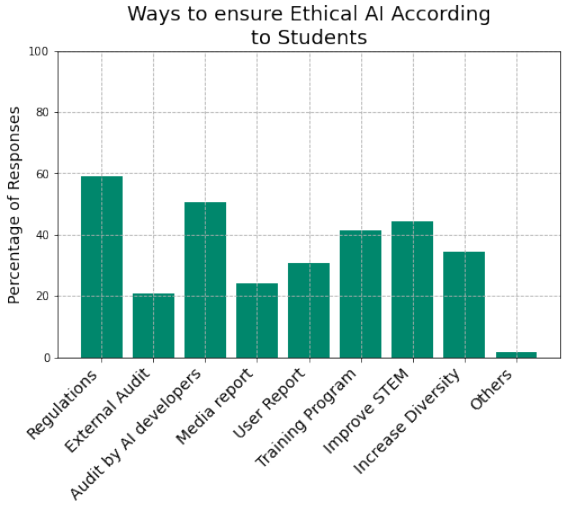


Figure 4.4: Students' response to ways for ensuring ethical AI in Nepal

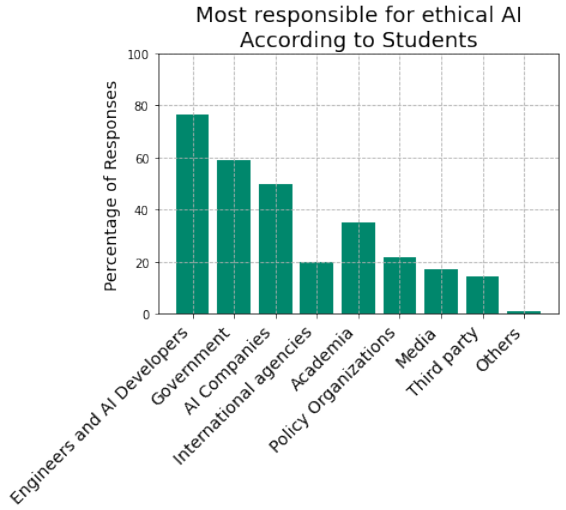


Figure 4.5: Students' response to who is most responsible for ensuring ethical AI in Nepal

The survey also asked whether the respondents considered risks, limitations and real-world impact in their academic and personal projects. 44 % said they considered risks in all course related technical projects, 58 % said they consider real-world impact in course-related technical projects, and 41 % said they considered real world impact specifically in AI related projects.

The survey also asked the students to list the sources they use to get information about AI and AI ethics. Most students received this information from blog posts, social media (especially Twitter), research papers. Similarly, when asked to list researchers who they follow for information about AI Ethics - the most common answers were Elon Musk, Andrew Ng, Lex Friedman, and Timnit Gebru. Of the researchers listed, only one Dr.

Gebru worked specifically in the area of AI ethics. Similarly, all the researchers worked and built AI products catered to North America and other developed countries. This indicates that the students may not be getting information related to use and risks of AI in Nepali context.

4.5 Demographic Differences

4.5.1 Difference in response based on gender

Both male and female students had similar attitudes and opinions about AI ethics in Nepal and there was no significant difference in response based on gender for most of the questions in the survey. This is evident from the results of the opinion poll shown in Figure 4.6.

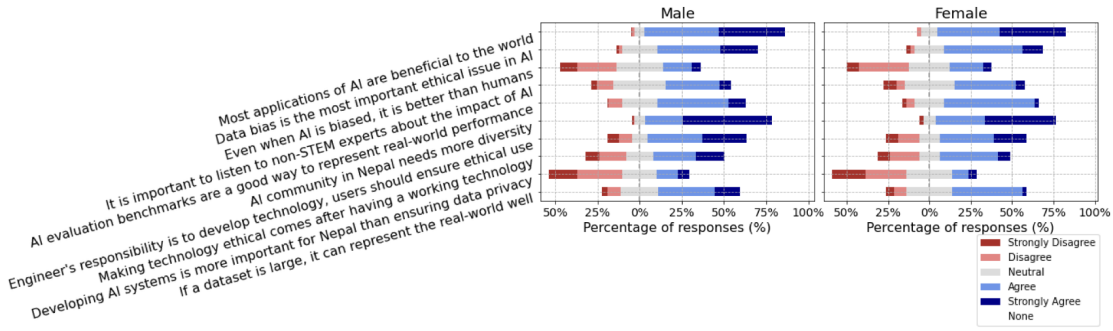


Figure 4.6: Plot comparing male and female students' opinions on various statements about AI and AI ethics

However, women were more likely to place a higher priority on ethical considerations. When asked to rate the importance of ethical issues in their career decisions, the average rating for women was 1.45 compared to 2.09 for men. More men were likely to say these considerations were not important at all compared to women. Similarly, 43% of the women said they considered real-world implications during their AI project while only 32% of

men said so. There was no difference when they were asked about considering real-world implications due to course requirements.

4.5.2 Difference in response based on degree

We compared the sets of responses between undergraduates (Bachelors) and graduate level students (Masters and above). Graduate level students were more aware of technical limitations of AI and potential harms as evidenced by Figure 4.7. Graduate students were also more aware of AI ethics issues being different based on local contexts. 68% of master's students said AI ethics consideration for Nepal and other low and middle income countries would be different compared to only 47% of undergraduate students.

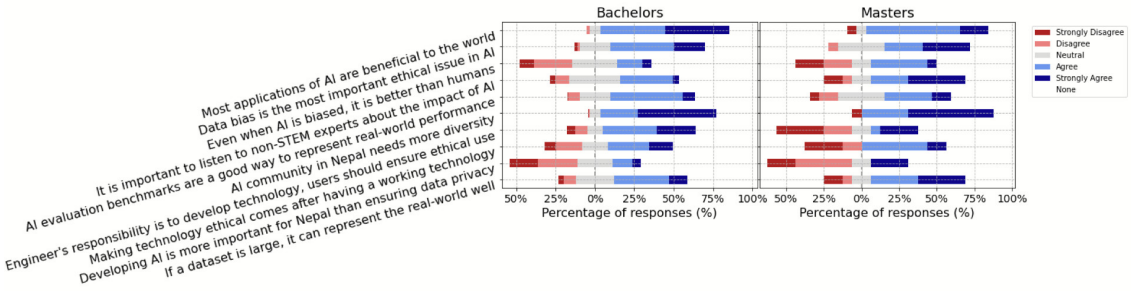


Figure 4.7: Plot comparing undergraduate and graduate students' opinion on various statements about AI and AI ethics

In terms of the curriculum, more Masters students have formal courses in AI but fewer of them mentioned having ethics or ethics related curriculum. On the other hand bachelors students were unsure about about whether they had AI ethics related courses in their curriculum with 32 % saying maybe. The percentage of response is shown in Table 4.1

Table 4.1: Responses to questions on formal AI and AI ethics curriculum

Questions	Undergraduate (%)			Graduate (%)		
	Yes	No	Maybe	Yes	No	Maybe
Are you taking formal courses in AI ?	33.5	61	0	50	43.75	0
Do you have an ethics course in your curriculum?	32.5	60	0	43.75	50	0
Do you have an AI ethics course in your curriculum?	16	46.5	31	18.75	75	0

Surprisingly, Masters level students were less likely to consider risks and real-world impact of their projects. While 42 % of Bachelors students said they considered real-world implications of their AI projects, only 21 % of Masters students said yes. The questions and the percentage response is shown in Table 4.2

Table 4.2: Responses to questions about considerations of real-world impact

Questions	Undergraduate (%)			Graduate (%)		
	Yes	No	Maybe	Yes	No	Maybe
For technical projects required for your degree, are you required to consider risks and limitations?	42	5.5	43	68.75	12.5	12.5
For technical projects, do you have to consider real-world implications, including potential harms of the project?	58.5	11.5	19	56.25	31.25	0
Have you considered real-world implications in your AI projects (school or personal) ?	42	48	0	31.25	62.5	0

4.5.3 Difference in response based on the type of degree

We also compared the responses based on the degree that the students were pursuing. Namely, we compared the responses between engineering students and students who were pursuing a computing degree outside of an engineering degree. Table 4.3 and Table 4.4 show their responses regarding whether they have formal AI and AI ethics curriculum as well as whether they are required to consider limitations and real-world implications when doing projects. As the tables show, we did not find a significant difference between the two groups. This indicates that both engineering and non-engineering curriculum is very similar in their (lack of) coverage of AI ethics related topics. We also did not observe significant difference in their attitudes and responses about general questions of AI.While the survey also had responses from students who were pursuing other degrees besides these two, the sample size was too small to be considered in further analysis.

Table 4.3: Responses to questions on formal AI and AI ethics curriculum

Questions	Engineering (%)			Non-eng computing (%)		
	Yes	No	Maybe	Yes	No	Maybe
Are you taking formal courses in AI ?	34.35	60.30	0	36.25	58.75	0
Do you have an ethics course in your curriculum?	28.24	64.88	0	38.75	53.75	0
Do you have an AI ethics course in your curriculum?	15.26	44.27	34.35	17.5	55	21.25

Table 4.4: Responses to questions about considerations of real-world impact

Questions	Engineering (%)			Non-eng computing(%)		
	Yes	No	Maybe	Yes	No	Maybe
For technical projects required for your degree, are you required to consider risks and limitations?	39.69	6.10	44.27	51.25	6.25	35
For technical projects, do you have to consider real-world implications, including potential harms of the project?	57.25	9.16	21.37	60.0	18.75	12.5
Have you considered real-world implications in your AI projects (school or personal) ?	41.22	48.85	0	41.25	50.0	0

4.5.4 Difference in response based on other demographic categories

During our study, we also considered differences in response based on location, caste, and educational institution but did not find any significant difference to be reported. This might be due to the small sample size and over-representation of upper-caste respondents and those from Kathmandu valley.

Chapter 5

Survey Results: Professionals

5.1 Survey Participants

The survey was filled out by 50 professionals. As shown in Figure 5.1, the respondents were primarily male, from Kathmandu valley and upper caste. Though the survey had an option for choosing non-binary or refusing to disclose one's gender, very few chose to do so. Of the respondents, 78% identified as male and 18% identified as female. The gender imbalance among respondents is likely due to both the existing gender imbalance among tech professionals in the country and the lower likelihood for women to fill out online surveys. Similarly, 72 % or the respondents were from Kathmandu valley, 10 % outside the valley within Nepal, 6 % outside Nepal, and the remaining refused to disclose their location. In terms of caste and ethnicity, 36 % chose Brahmin, 24 % identified as Adivasi or Janajati, 12 % as Chhetris, 10% and Madhesis, 6 % as Newars and none as Dalits.

68 % of the respondents had their Bachelor's degree and 30 % had Masters or above. Among the respondents with a Bachelor's degree, most people have a background in computer or electronics engineering. Among Masters, MBAs were the most common with few respondents with MSc. in Data Science, GIS or computing. The participants were quite well distributed in terms of their work-role - 26 % of the respondents worked in research or academic environments, 20 % had a leadership role, and the majority (44 %) had an engineering or developer's role.

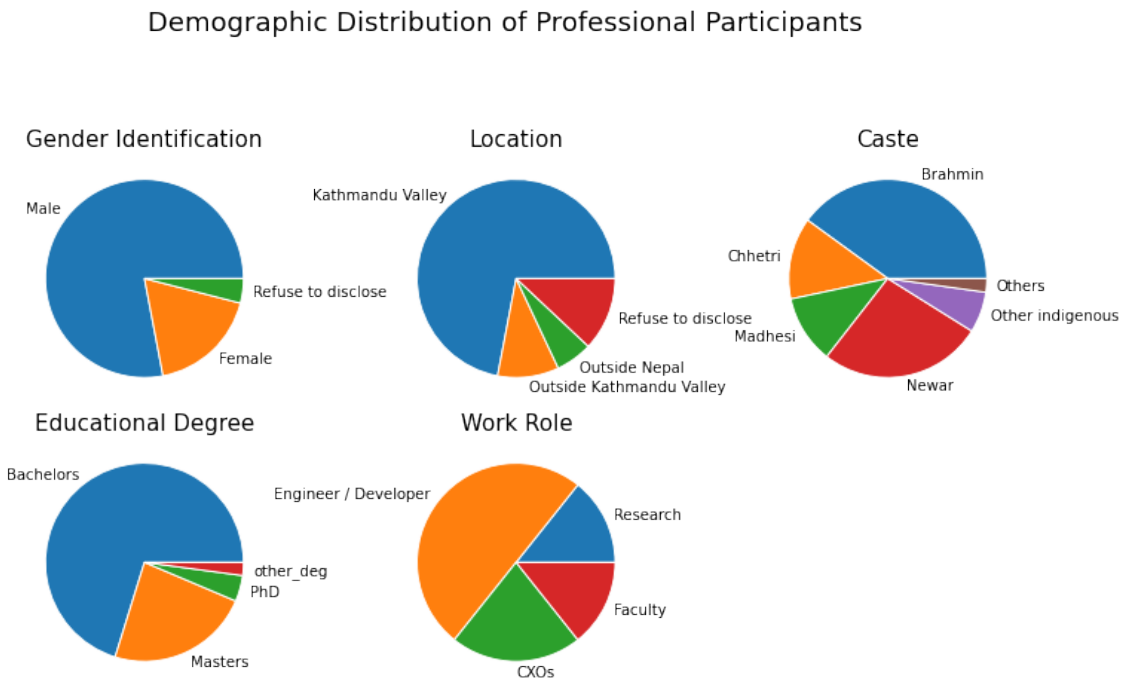


Figure 5.1: Demographic distribution of professional respondents

5.2 Overall Impressions about AI Ethics

The professionals considered AI Ethics as an important topic with 90% agreeing that it is a topic of concern. Compared to students, professionals rated ethical concerns as being less important to their career considerations with an average rating of 2.42. Another difference between professionals and student group lay in where they get their information about AI ethics. When asked to list AI ethics researchers that they follow, about 68% said they do not follow research about AI ethics. Of the remaining, most listed AI ethics researchers or journalists such as Timnit Gebru, Karen Hao, and Kate Crawford.

Figure 5.2 shows the opinion poll on various statements about AI and AI ethics in Nepal. The professionals are more aware of harms and technical limitations of AI as evidenced by the majority disagreeing on the statement that biased AI are still better than humans and large enough datasets can represent the real world well. The professionals also agree with the need for diversity in AI community and the need for non-STEM experts to have a say in matters of AI ethics.

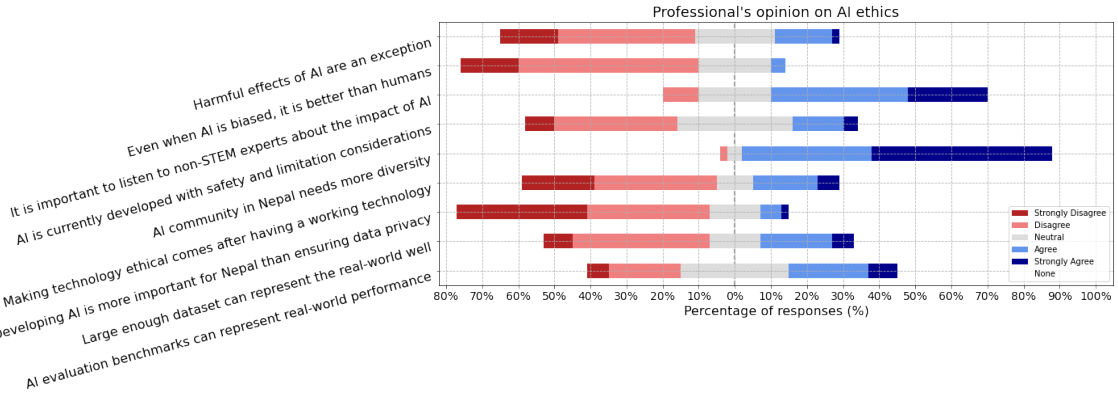


Figure 5.2: Professionals' response on whether they agree with various statements regarding AI and AI ethics.

5.3 Ensuring AI Ethics in Nepal

Among the professionals, the major risks identified for Nepal included data privacy violation, lack of AI literacy, lack of technical accuracy, and the potential for misuse.

Similar to the students, the professionals also considered privacy, fairness and explainability as the most important AI ethics related topics in Nepal as shown in Figure 5.3. This was reflected in the opinion questions as well. 66% of respondents agreed or strongly agreed that data bias was the most important ethical issue in Nepal. Similarly 66% of respondents disagreed that biased AI is better than humans unlike the responses from students where majority thought that even when biased, AI is better than humans.

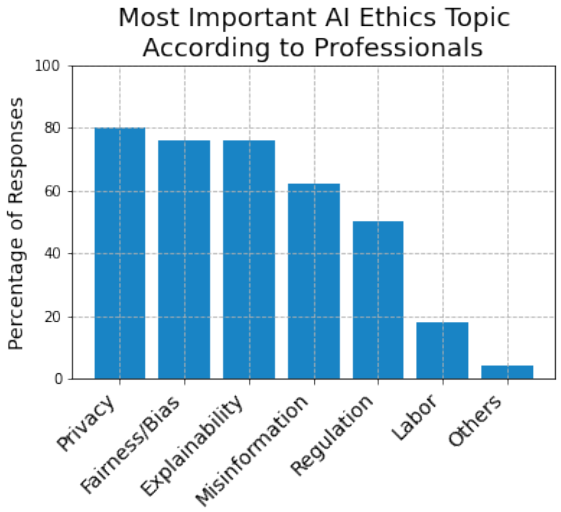


Figure 5.3: Demographic distribution of professional respondents

The practitioners thought engineers and AI developers were the most responsible for en-

suring ethical AI, with government and AI companies sharing some of the responsibilities. The percentage who chose these three groups are at 70%, 66%, and 66% respectively as shown in the figure. However, when asked whether they considered themselves personally responsible for considering bias in their AI models, only 34% said yes. Most respondents said managers or team leads were responsible for these tasks. Similarly, most respondents did not know about tools related to ensuring AI ethics or did not think them necessary for their jobs. As shown in the figure, the respondents were also divided in their opinion when asked to agree or disagree with the statement that 'The primary responsibility of engineers is to develop technology, ensuring it is used ethical is the job of the users.'

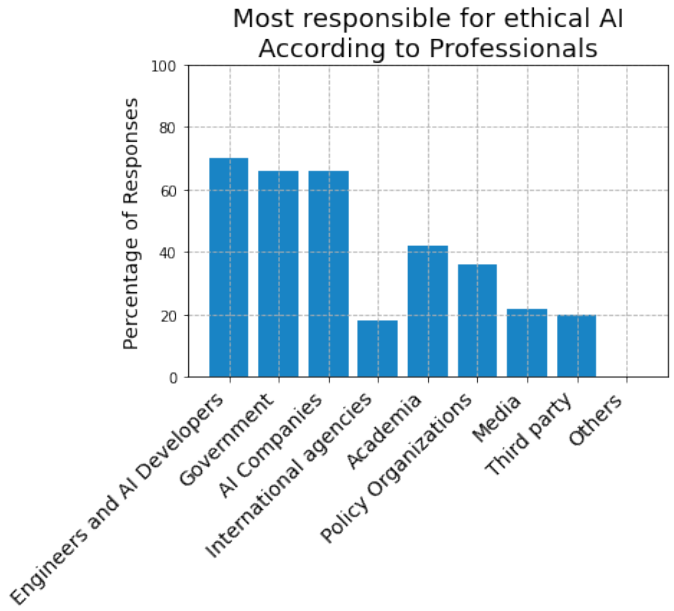


Figure 5.4: Professionals' response to who is most responsible for ensuring ethical AI in Nepal

The respondents also thought that AI is not currently developed with safety and limitations in mind. 76% chose regulations as one of the most effective ways to ensure ethical AI, while 48% chose external audits, and 40% chose internal audits. 54% of participants also thought there should be ethical AI related training programs while 34% wanted increased diversity.

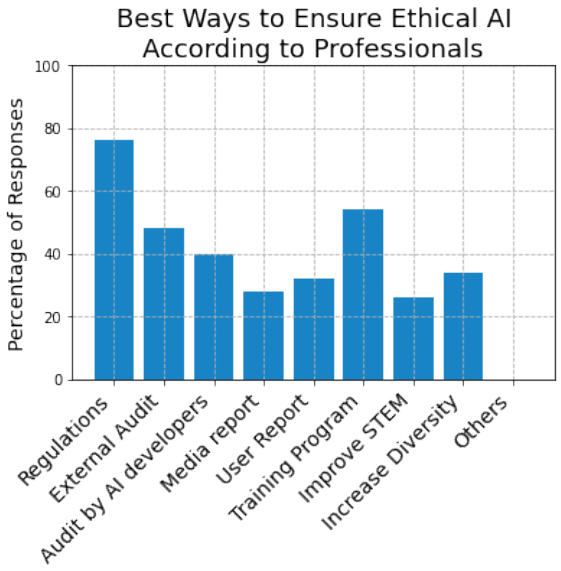
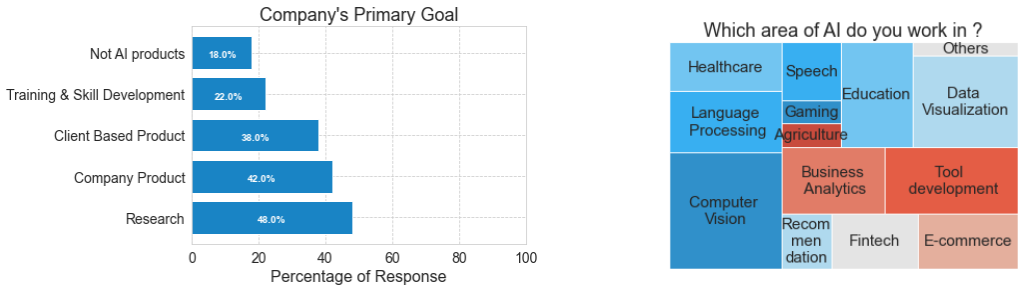


Figure 5.5: Professionals' response to different ways that could be most effective for ensuring ethical AI in Nepal

5.4 AI Ethics in practice

The survey for professionals had a section with questions regarding their professional practices including the area of AI, data source, tools they used, their responsibilities regarding risk estimation, and ethical practices at work. Of the respondents, 18 % of the respondents said they did not develop AI products while most of the remaining worked in AI research or developing AI products for their company. The most common area of work was computer vision followed by data visualization, tool development, and business analytics. Other significant areas included language processing, fintech, and recommendation systems. Multiple participants chose multiple areas which indicates that the same companies and engineers may be working in very different application domains within AI. The overall distribution of participants among different application is shown in Figure 5.6b.

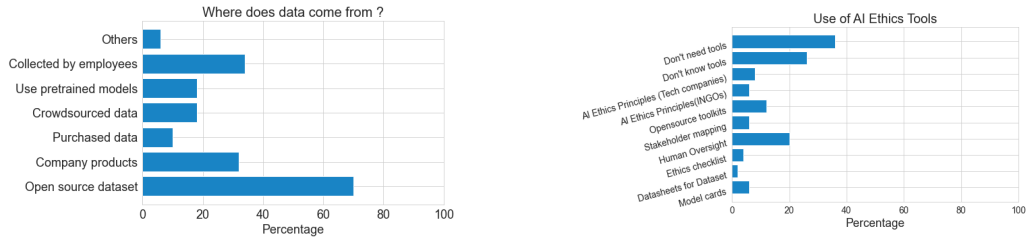
To understand how AI ethics translates to practice, our survey asked the professionals questions about where they got their datasets, ethical issues during their work, and any tools they may use while developing AI. In terms of data source, about 70 % used open source datasets while 34 % used dataset collected by employees and 32 % used data collected from company products as shown in Figure 5.7a. Despite the majority of respondents concern that privacy is the most important ethical issue, most participants agreed or strongly agreed



(a) (b)

Figure 5.6: Professionals' response to primary goal and areas of application of AI in their workplaces

that the data they used were collected with consent. When asked about what tools were available for them to build ethical models, most participants chose that they did not know about the tools or did not think they needed AI ethics specific tools. Among the participants who chose a tool, most chose human oversight. The overall distribution of different ethics tools is shown in Figure 5.7b. The response to this question particularly indicates a severe lack of awareness of different AI ethics tools among the professionals.



(a) (b)

Figure 5.7: Data source and AI Ethics tools used in their company

We also tried to measure how the participants felt about ethical practices at their workplace as shown in Figure 5.8. Most participants agreed or felt neutral about the positive impact of AI products they were working on. The respondents mostly had neutral feelings about ethical issues at their companies when asked about ethical dilemmas during projects, adequate time to ensure unbiased datasets and models, delaying a project due to ethical problems, and using risk analysis as PR tools. However, the respondents strongly agreed that the data they used was collected with consent and that the AI products they are building are good for everyone. While positive, this is in contradiction with the previous response about majority using open-source datasets which indicates that the respondents may not have access to information about data collection processes.

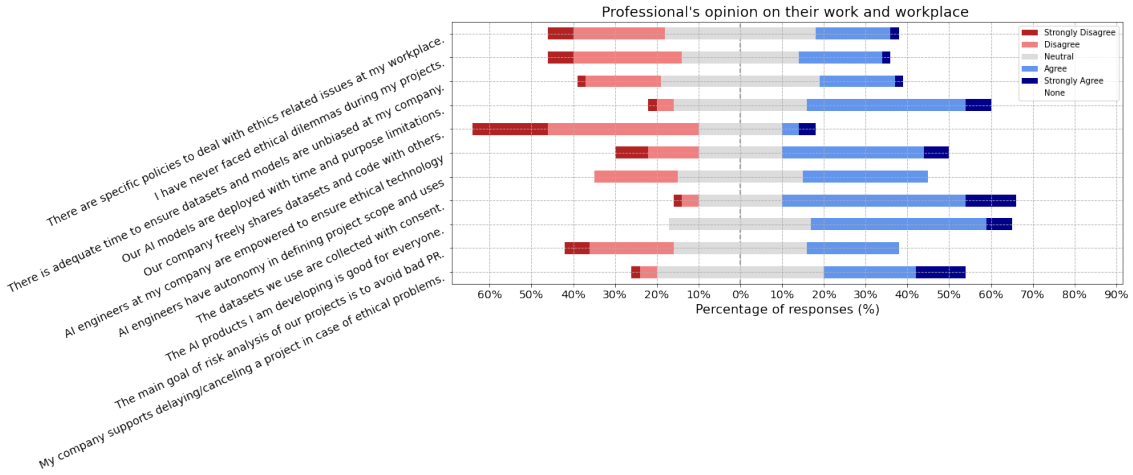


Figure 5.8: Professionals' response on whether they agree with various statements regarding their workplace practices

5.5 Demographic Differences in survey response

5.5.1 Based on gender

Among professionals, there was observable difference between male and female respondents in various questions. As shown in the Figure 5.9, women were more aware of harms and less enthusiastic about opportunities from AI. The difference is particularly stark in statements about tradeoffs between technology and privacy, and the ability of large datasets and evaluation benchmarks to be sufficiently representative. Women give more importance to ethical considerations as a part of developing AI technology. 44 % of women strongly disagreed that engineer's role is to build technology and ethical use is the role of users compared to 13 % of men. Similarly, more women disagreed that developing AI should come before making it ethical or ensuring privacy. This was also reflected on their rating for the importance of ethical considerations in career choices. The average rating for women was 1.78 while that for men was 2.38.

In terms of workplace practices, female respondents considered themselves more responsible for considering potential bias compared to male. However, there was no significant difference to their responses on whether they consider real-world impact of software projects as shown in Table 5.1. Similarly, the female respondents also tended to have stronger and more positive opinion about ethics related workplace practices compared to male.

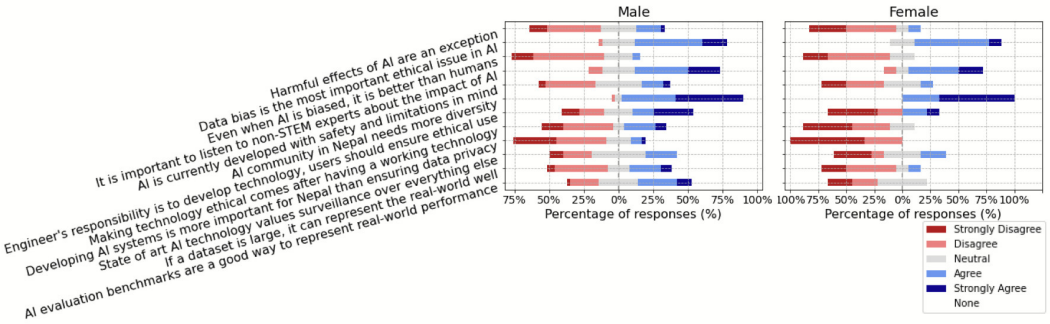


Figure 5.9: Plot comparing professionals' general opinions about AI and AI ethics based on their gender

Table 5.1: Professionals' responses to questions about considerations of real-world impact

Questions	Male (%)			Female(%)		
	Yes	No	Maybe	Yes	No	Maybe
Do you think AI Ethics is a topic of concern ?	92.31	7.69	0.0	100	0.0	0.0
Are you responsible for considering potential bias in your AI projects?	25.64	25.64	46.15	55.56	22.22	22.22
Do you consider real-world use and potential harms during of software projects?	53.84	7.69	33.33	55.56	0.0	44.44

5.5.2 Based on degree

We compared the responses to the professional's survey based on degree and degree type and did not observe any significant difference as evidenced the opinion poll in Figure 5.10. Comparing this observation to the difference in opinion among students based on degree level and noting that graduate students usually have a few years of experience, we think experience is more likely to shape opinions on AI ethics related topics than formal education.

In terms of responses to questions about consideration of real-world impact shown in Table 5.2, graduate respondents were more likely to consider themselves responsible for considering potential bias as well considering real-world implications of their project. It is unclear whether this is due to more awareness of graduate degree holders or the likelihood that these individuals probably are in a superior positions with more responsibility compared to undergraduates at their respective companies.

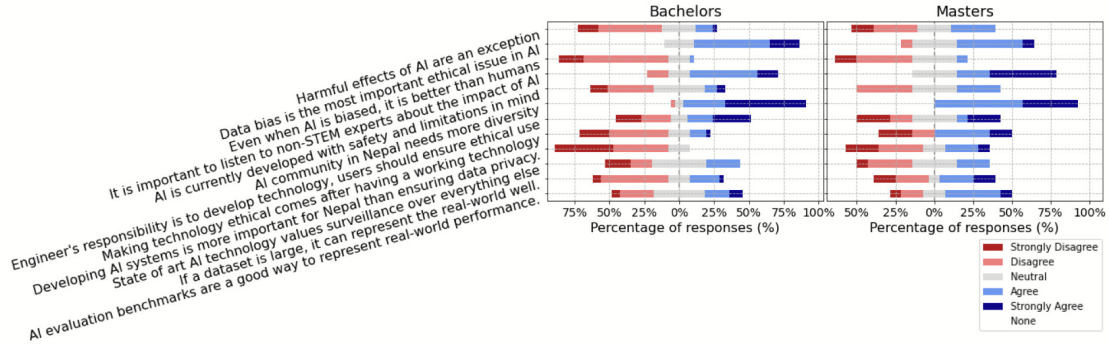


Figure 5.10: Plot comparing professionals' general opinions about AI and AI ethics based on their degree

Table 5.2: Professionals' responses to questions about considerations of real-world impact

Questions	Undergraduate (%)			Graduate(%)		
	Yes	No	Maybe	Yes	No	Maybe
Do you think AI Ethics is a topic of concern ?	93.93	6.06	0.0	92.86	7.14	0.0
Are you responsible for considering potential bias in your AI projects?	27.27	27.27	42.42	42.85	21.43	35.71
Do you consider real-world use and potential harms during of software projects?	51.51	9.09	33.33	64.28	0.0	35.71

5.5.3 Based on workrole

Lastly, we compared the responses based on the different professional roles and discovered significant differences on answers. Overall, the CXOs were more likely to consider real-world implications as well as considered themselves responsible for ensuring AI was ethical compared to academics and engineers as shown in Table 5.3.

There was also a difference between the different groups on the most important topics for AI ethics as well as who is responsible. While all the groups agreed on fairness, privacy, and explainability as topics of concern, CXOs were more likely to also consider regulation and labor as important topics compared to engineers and academics Figure 5.11

Interestingly, both CXOs and academics seemed to minimize their role in ensuring ethical AI as shown in Figure 5.12a. Only 33 % CXOs considered it to be the AI company's role to ensure ethical AI while 76 % of engineers and 75 % of academics thought AI companies should be responsible. Similarly, only 25 % of academics thought they were

Table 5.3: Professional's responses to questions about AI ethics based on their work-role

Questions	CXOs(%)			Engineers (%)			Academics(%)		
	Yes	No	Maybe	Yes	No	Maybe	Yes	No	Maybe
Do you think AI Ethics is a topic of concern ?	88.88	11.11	0.0	95.23	4.76	0.0	91.67	8.33	0.0
Are you responsible for considering potential bias in your AI projects?	33.33	33.33	33.33	28.57	23.81	42.85	57.64	17.36	33.33
Do you consider real-world use and potential harms during software projects?	88.88	11.11	0.0	38.09	4.76	52.38	58.33	0.0	33.33

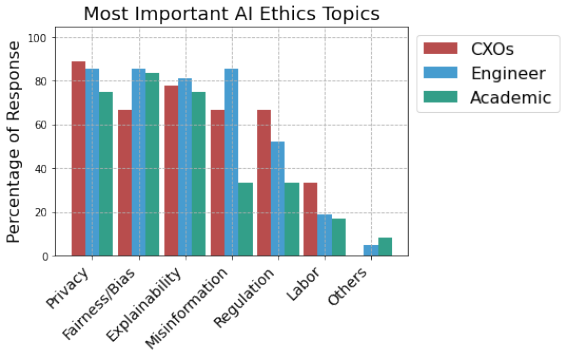
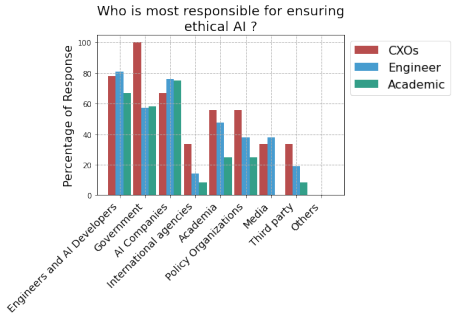


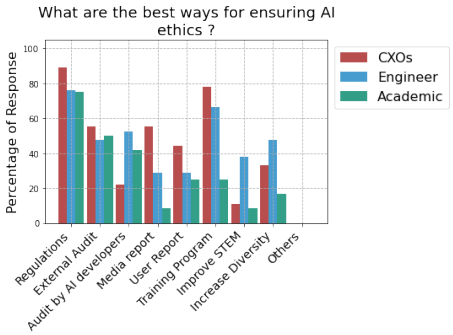
Figure 5.11: Plot comparing professionals' responses on most important topic based on their workrole

responsible while 48 % of engineers and 56 % of CXOs thought they were responsible. We also observed a difference in opinion in ways for ensuring ethical AI as shown in Figure 5.12b. While 52 % of engineers thought internal AI audits helped with ethical AI products, only 8 % of academics and 22 % of CXOs considered it useful. Unlike the other two groups, CXOs chose media reports and user reports as useful tools for ethical AI. These responses indicate that different roles will require different incentive mechanisms to ensure AI development processes are ethical.

It was also interesting to compare the opinions of CXOs and engineers on different



(a) Who is most responsible for ensuring ethical AI



(b) Best ways to ensure ethical AI

Figure 5.12: Comparing professionals' responses on who is most responsible for ensuring ethical AI and best ways to ensure ethical AI based on their work role.

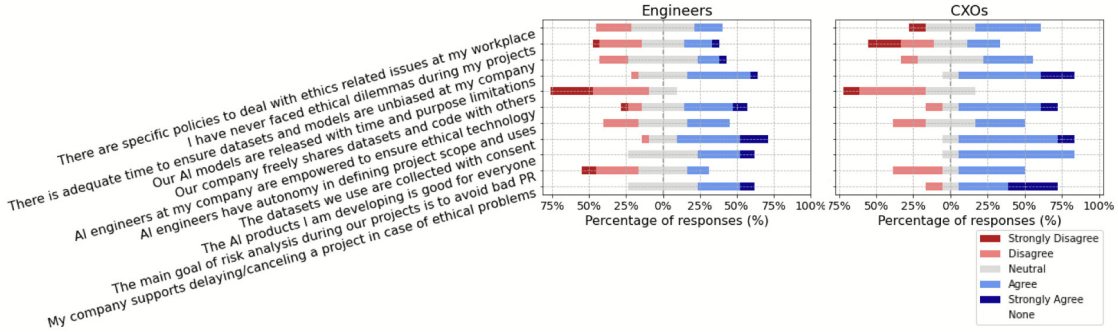


Figure 5.13: Plot comparing professionals' opinions about AI ethics practices in their workplace based on their workrole

aspects of AI ethics in practice. As expected, the CXOs had a more favorable opinion about the company work culture while the engineers were more neutral. This is shown in Figure 5.13

Chapter 6

Survey Results: Policymakers

6.1 Survey Participants

The survey was filled out by 11 policymakers. Of them, 4 respondents were female and the rest identified as male. Most of the respondents were in their 30s and 40s. The respondents included a few academics and the rest were employed within government or policy organizations. All the respondents had a graduate degree in a technical field. In terms of caste, 2 respondents identified as “Janajati” and the rest identified as Brahmins and Chhetriyas. Due to the small sample size of the respondent group, we have not presented the demographic distribution in a plot or used it to further analyze the differences in responses amongst different sub-groups.

6.2 Overall impression about AI Ethics

Similar to the other two groups, policymakers also unequivocally agreed that AI ethics is an important topic with 90 % calling it a topic of concern. Policymakers mentioned that ethics is an important framework to define harms and benefits for any application and an ethical framework for AI could be adapted from other existing frameworks. Their key concerns about AI included lack of awareness, lack of technical competence, and the potential of ”winner takes all dynamics” in AI.

When asked about their familiarity with AI on a scale from 1-5 the policymakers rated themselves an average of 3.27. 90 % of the respondents also expressed an interest to learn

more and requested focused events for them and availability of resources.

Figure 6.1 shows policymaker's opinions on various statement about AI and its impact. On the whole, the respondents were cautious about potential harms of AI and aware of the need for better policies in Nepal.They strongly agreed on the need for AI community in Nepal to be more diverse. Similarly, they disagreed with promoting AI development at the expense of privacy risks and ethical shortcomings.

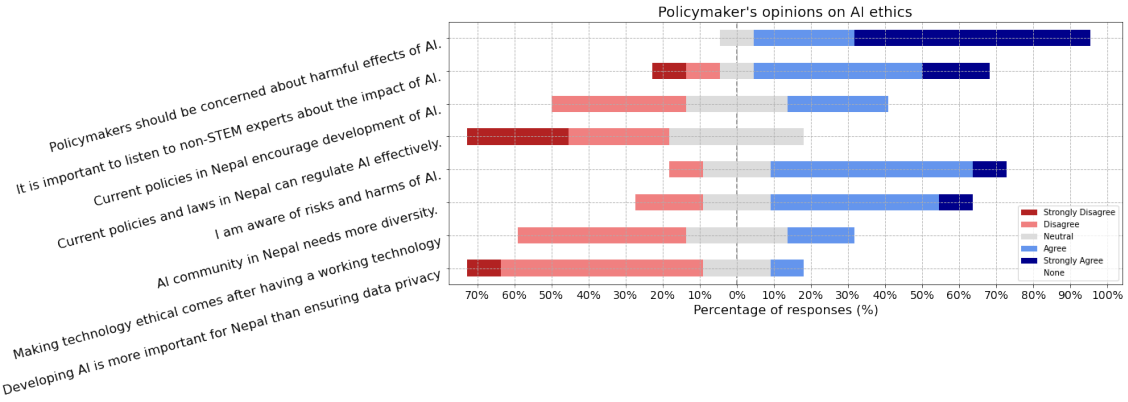


Figure 6.1: Policymakers' response to whether they agree with variou statements about AI and AI ethics policy in Nepal

6.3 Ensuring AI Ethics in Nepal

As expected, the policymakers did not have a rosy view about AI ethics in Nepal. 81% of the respondents said Nepal that does not have good policy currently. In terms of main risks to Nepal, they listed a lack of awareness, a lack of information security policy and infrastructure, a lack of technial competence, and the risk of ending up as consumers without a say in how AI is developed.

As Figure 6.2 shows, most policymaker respondents consider data privacy risk as the most important topic in Nepal. Besides privacy, they were equally concerned about digital divide, legal protection, lack of expertise, misinformation, and AI discrimination. Surprisingly, none of the respondents considered the potential for job loss and unemployment due to AI as their top 3 concerns.

As policy solutions, they suggested a need for more discussion, awareness, and advocacy. All of the policymakers in our survey considered the government as the most responsible for ensuring ethical practices for AI in Nepal as shown in Figure 6.3. Besides

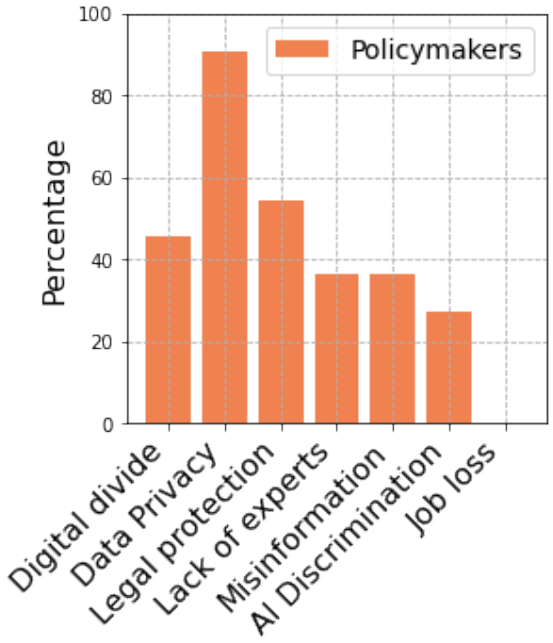


Figure 6.2: Policymakers' response to the most important topics for ethical AI in Nepal

government, they also considered engineers, university and policy organizations as important. However, the respondents in our survey did not think that AI companies, third party organizations, media had a significant role in ensuring ethical AI in Nepal.

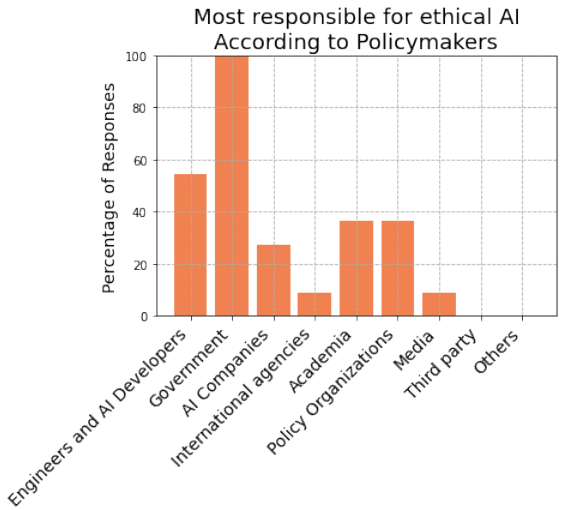


Figure 6.3: Policymakers' response to who is most responsible for ensuring ethical AI in Nepal

Chapter 7

Discussion and Future Recommendations

7.1 Gaps in AI Ethics understanding

Based on the findings described in Chapters 4, 5, and 6, we observed significant gaps in the understanding of AI and AI ethics in Nepal. As pointed out by most of our respondents across all three surveys, there is a lack of national policy and guideline related to Artificial Intelligence in Nepal. Existing laws are very limited in addressing many ethical problems that may arise from the rapidly growing AI industry.

Our analysis shows that the students are extremely positive about AI applications being good for the world and are not very aware of the harms and technical limitations of the systems that already exist today. This can be attributed partly to the current engineering curriculum which does not appear to prioritize ethics and real-world implications. In the absence of formal courses, the information ecosystem on AI is dominated by social media and blogs leading to filter bubbles. Most students appear to get information related to these topics from entrepreneurs and celebrity researchers, usually based in the Global North. Thus, our research shows that there is a lack of information ecosystem about responsible practices for AI contextualized for developing countries.

We also observed that our respondents mostly treated AI and ethics surrounding it as a technical problem ignoring the root causes that stem from pre-existing social inequalities in society. Students were unaware of technical limitations of large datasets and evaluation

benchmarks. Professionals were aware of these issues but not very reflective in their own practices. Despite raising issues of model bias and data privacy, our respondents were not concerned about lack of reliable training and evaluation datasets in Nepal. Most professionals used open source datasets in their practices. However, research has shown that open source data are biased in terms of geographical, racial, and gender representations, often embed stereotypes, and can cause AI models to be similarly biased [34, 35].

Lastly, there is a lack of accountability mechanism and ecosystem in Nepal. Our research showed that the participants considered themselves responsible in theory but did not really self-reflect on their own roles and responsibilities in practice. Most of our respondents also did not consider civic organizations, journalists, and researchers as responsible for playing an important role in ensuring ethical AI. This is in contradiction to rest of the world where external researchers, activists, and journalists have highlighted many significant issues with AI applications that have led to new laws, updated models, and even ban of certain AI applications [6, 36, 37] On the contrary, most of our respondents proposed protectionist and surveillance focused policies (such as banning biased algorithms, banning misinformation) as solutions without considering potential unintended consequences and complications that can arise in implementing such policy.

7.2 Limitations of this Project

Because of limited time and resources, the project was designed and conducted with a limited scope including targeting technology related audience for the stakeholders and disseminating an online and mostly qualitative surveys. However, truly understanding the impact of AI in society as AI ethics aims to do is a broad and inter-disciplinary task. In addition to technical experts, there should be lawyers, psychologists, economists, journalists, and social activists included in the conversation. By limiting the survey to technical experts, this work might have replicated the existing dominance of technical experts [17]. Follow up projects should pay special attention to include these other experts as well as AI users to value knowledge that comes from lived experience.

Even within the limited scope, the survey had a few limitations. One of the key limitation is the small sample size for the policymakers' survey. While the findings help us

understand some of the concerns and trends, the survey size is too small to make general claims. In addition, the respondents primarily were from Kathmandu valley and had a science and technology background which introduces an additional bias. We think that this is due to the online medium of the survey and COVID related constraints. We believe visiting different governmental and non-governmental organizations with policy experts with a paper survey and in-depth interviews would yield better results.

While the professional and student survey had larger samples, they were also primarily filled in by upper caste engineering students in Kathmandu valley. This is potentially both due to sampling bias caused by dissemination of survey via social media as well as inherent inequalities in Nepal on what types of students and professionals are likely to be exposed to AI and ML. Therefore, we should be aware that the findings and claims for this survey may not be perfectly generalized throughout the country. The respondent groups also point to a need for increased effort to improve diversity among students and professionals. Another limitation in analyzing the demographic distribution of the survey is the difficulty inherent in categorizing caste groups in Nepal. Due to widespread social stigma and discrimination, many Dalits and underprivileged caste groups choose to hide their caste identity [38] and they may have chosen to do so in the case of this survey as well. In addition, caste identity and ethnicity may intersect and different ethnic groups have their own caste-hierarchies which makes categorizing privileged caste groups and underprivileged caste groups more challenging.

Lastly, the quantitative nature of many survey questions is limited in its ability to capture many nuances and subjectivities that are inherent in any ethics based discussion. We also observed the respondents using a slightly different phrasing and repeating topics covered in the questionnaire itself as something that had not been covered in the survey. This indicates that the survey would have been improved by better phrasing or including definition as well as a potential lack of awareness among participants about common terminologies used in AI ethics literature.

7.3 Future Recommendations

Our survey showed that all three groups of stakeholders - students, professionals, and policymakers are concerned about AI Ethics in Nepal and are interested in learning more. While students are more optimistic and less aware of technical and harmful possibilities of AI, both professionals and policymakers had a basic understanding of global conversations happening around these topics. However, professionals did not always translate this awareness into their day to day work and were more likely to be disconnected from the impact of their own projects. Similarly, policymakers lacked knowledge about technical limitations of AI and were more likely to associate AI with general intelligence.

Based on these findings, we have following recommendations on AI related needs for Nepal on a national level:

- A national framework and accountability mechanism for ensuring responsible AI. This framework should be built with discussion and input from diverse stakeholders including AI users.
- Revise current engineering curriculum to introduce AI ethics related topics as well as encourage students to consider real-world impacts of their studies.
- Increased awareness about risks and limitations of AI across all levels of society.
- Increased diversity in AI community.
- Partnerships between academia, tech industry, and civic organizations. These could take the form of research collaborations, industry internships for students, consulting opportunities for professors, and interdisciplinary workshops and conferences.
- Inter-disciplinary AI collaborations including with non- STEM experts and digital rights activists.

We also see potential for follow-up work by NAAMII and other research organizations like it to fulfill some of the above listed needs. Potential future works are listed below:

- Publicize and contextualize these results for more awareness.
- Have follow-up survey projects for AI users and non-STEM experts.

- Have more qualitative studies to get a more in-depth understanding of the findings from this survey.
- Run training programs for students and professionals on AI Ethics toolkits and foundations.
- Organize community events and discussions on AI Ethics related topics.
- Build research collaborations with digital rights groups on data privacy and algorithmic discrimination in Nepal.

Appendix

Appendix 1: Survey for Students

Section 1 : Introduction

This is an online survey on the "Understanding of AI Ethics among Engineering Students in Nepal". This is a research project being conducted by NAAMII and is funded by UNESCO Asia and Pacific Regional Bureau for Education.

Participation:
You are invited to fill out this survey because you are a student in Nepal and interested in AI technology. Answering the survey should take around 20-30 minutes. Please understand that your participation in this survey is voluntary and you may, anytime, refuse to take part in this research. You may freely decline to answer particular questions that may be uncomfortable to you.

Confidentiality:
The answers to the survey will be stored by NAAMII with password-protected access. This information will not be shared outside of the team working on this project. We will not use any demographic information to personally identify you. The results of this survey will be shared in aggregate and anonymized formats. We've requested email address for raffle prizes and further contact. This email address will be stripped from the database before analysis of data.

Reward:
We will enter all entries with an email address into a raffle for 15 recharge cards. You will be able to choose the format for receiving the card (e-sewa, Ncell, NTC) on winning.

Contact:
Please reach out to NAAMII if you have any concerns or questions. We can be reached at ai-ethics@naamii.org.np

1.Please confirm that you have read and understood the information above, and agree to participate voluntarily in this survey.

- ○ Yes. Please take me to the survey.
- ○ No. Please take me to the end of this survey.

Section 2: Demographic Information

2.Please select your age group below.

○ Under 18 ○ 18-25 ○ 26-30 ○ 31-35 ○ 36 and above

3.Please select your gender.

○ Male ○ Female ○ Non-binary ○ Prefer not to say

4.What is your current institution?

5.What is your current degree and year?

6.What is your area of study?

7.Where are you currently located? Please mention city.
8.Which caste/ethnicity group do you identify with? Please choose all that apply. ☐ Brahmin ☐ Chhetri ☐ Madhesi ☐ Dalit ☐ Adivasi ☐ Others
Section 3 : AI Ethics Topics
9.Please define what AI means to you?
10.Do you think AI Ethics is a topic of concern for AI engineers? ◯ Yes ◯ No ◯ Maybe
11.Explain your answer to 10.
12.What are the most important topics for AI ethics? Choose all that apply ☐ Privacy ☐ Fairness/Bias ☐ Explainability / Transparency ☐ Misinformation / Disinformation ☐ Regulations ☐ Labor ☐ Others
13.Do you agree with papers, news articles, and blog posts that criticize AI? ◯ Yes. Most of the time. ◯ No. Most of them are biased or misunderstand AI. ◯ Sometimes. It really depends. ◯ I have never read such posts.
14.Which of the following sources have you used to get information about AI ethics online? Please choose all that apply. ☐ Research papers ☐ Newspaper articles ☐ Blog post ☐ Talks ☐ Podcasts ☐ Twitter ☐ Other social media ☐ I ignore anything related to AI ethics
15.Please list up to 5 researchers/writers that you follow regarding AI ethics topics. Separate researchers by a comma, semicolon, or an enter.
16.Do you think AI ethics consideration is different for developing (Low and Middle Income as defined by World Bank) countries like Nepal vs more developed countries? ◯ Yes ◯ No ◯ Maybe
17.Explain your answer to 16.
18.In your opinion, what is the biggest risk of AI for Nepal? Please explain your answer.

19.In your opinion, who is the most responsible for ensuring that AI technology is ethical? Choose top 3.

☐ Engineers and AI developers ☐ Government

☐ AI companies ☐ International agencies such as the UN

☐ Universities / Research Institutes ☐ Policy Organizations

☐ Media ☐ Third party agencies established to monitor AI

☐ Others

20.What are the best ways to ensure that AI is developed in an ethical manner? Choose top 3.

☐ Regulation ☐ Audits by external agencies

☐ Audits by AI developers ☐ Media reporting on AI ethics related topic

☐ Users reporting harmful impact from AI ☐ Training programs

☐ Improving STEM (Science, Technology, Engineering, and Math) in Nepal

☐ Increasing diversity in AI community

☐ Others

21.Please rank the following sentences (1 –Strongly agree, 5-Strongly disagree)

	1	2	3	4	5
a. Most applications of AI are beneficial to the world.					
b. Data bias is the most important ethical issue in AI.					
c. Even when AI is biased, it is better than humans.					
d. It is important to listen to non-STEM experts about the impact of AI.					
e. AI evaluation benchmarks are a good way to represent real-world performance.					
f. AI community in Nepal needs more diversity.					
g. The primary responsibility of engineers is to develop technology, ensuring it is used ethically is the job of users.					
h. The first step is to have a working technology. Making it fair and ethical comes after.					
i. Developing AI systems is more important for Nepali society than ensuring data privacy.					
j. If a dataset is large, it can represent the real-world well.					

22.Can you think of any other AI ethics related topic not covered by this survey?

Section 4: AI in curriculum

23.Are you taking formal courses in AI ?

◯ Yes ◯ No

24.Do you have an ethics course in your curriculum?

◯ Yes ◯ No

25.Do you have AI ethics course in your curriculum?

◯ Yes ◯ No ◯ Maybe

26.Have you covered any of these topics in your curriculum? Please choose all that apply

◯ Fairness/Bias ◯ Surveillance

Failure of AI systems
Privacy/Data Ownership
Data Ethics
Hidden labor in AI
Misinformation
Engineering Ethics

27.For technical projects required for your degree, are you required to consider risks and limitations?

◯ Yes ◯ No ◯ Depends on the course.

28.For technical projects, do you have to consider real-world implications, including potential harms of the project?

◯ Yes ◯ No ◯ Maybe

29.Have you considered real-world implications in your AI projects (school or personal) ?

◯ Yes ◯ No ◯ Maybe

30.How important would ethical considerations be in your future career plans? 1 - strongly important, 5- not important at all

◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

Section 5 : Conclusion

31.Would you be interested in joining events by NAAMII related to AI Ethics? Please check all that apply:

☐ Focus group discussions ☐ AI Ethics Reading group

☐ AI Ethics Workshop/Conference ☐ AI Ethics Seminar

32.Please provide your email address if you would like to be entered in a raffle for recharge card?

Appendix 2: Survey for Professionals

Section 1 : Introduction

This is an online survey on the "**Understanding of AI Ethics among Engineering Professionals in Nepal**". This is a research project being conducted by NAAMII and is funded by UNESCO Asia and Pacific Regional Bureau for Education.

Participation:
You are invited to fill out this survey because you are an **engineering professional in Nepal** developing AI related technology. Answering the survey should take around 30 minutes. Please understand that your participation in this survey is voluntary and you may, anytime, refuse to take part in this research. You may freely decline to answer particular questions that are uncomfortable to you.

Confidentiality:
The answers to the survey will be stored by NAAMII with password-protected access. This information will not be shared outside of the team working on this project. We will not use any demographic information to personally identify you. Your answers will not be connected to your organizations. The results of this survey will be shared and analyzed in **aggregate and anonymized formats**.

Contact:
Please reach out to NAAMII if you have any concerns or questions. We can be reached at **ai-ethics@naamii.org.np**

1.Please confirm that you have read and understood the information above, and agree to participate voluntarily in this survey.

- ○ Yes. Please take me to the survey.
- ○ No. Please take me to the end of this survey.

Section 2: Demographic Information

2.Please select your age group below.

○ Under 18 ○ 18-25 ○ 26-30 ○ 31-40 ○ 41-50 ○ 51 and above

3.Please select your gender.

○ Male ○ Female ○ Non-binary ○ Prefer not to say

4. What is your highest education level? If you are currently a student, mention the degree you are pursuing.

5.What is your current organization?

6.What is your official title ?

7.Where are you currently located? Please mention your current city.

8.Which caste/ethnicity group do you identify with? Please choose all that apply.

☐ Brahmin ☐ Chhetri ☐ Madhesi ☐ Dalit ☐ Janajati ☐ Adivasi

☐ Other _____

Section 3 : AI Ethics Topics

9. Do you think AI Ethics is a topic of concern for AI engineers?

◯ Yes ◯ No ◯ Maybe

10.Explain your answer to 9.

11.What are the most important topics for AI ethics? Choose all that apply

☐ Privacy ☐ Fairness/Bias ☐ Explainability / Transparency

☐ Misinformation / Disinformation ☐ Regulations ☐ Labor ☐ Others

12. Do you find that you agree with papers, news articles, and blog posts that criticize AI?

◯ Yes. Most of the time. ◯ No. Most of them are biased or misunderstand AI.

◯ Sometimes. It really depends. ◯ I have never read such posts.

13. Is there a researcher/expert that you follow for AI ethics related topics?

◯ Yes ◯ No

14.If yest to 13, please list them.

15. What is the biggest risk of AI for Nepal? Please explain your answer.

16.In your opinion, who is the most responsible for ensuring that AI technology is ethical? Choose top 3.

☐ Engineers and AI developers ☐ Government

☐ AI companies ☐ International agencies such as the UN

☐ Universities / Research Institutes ☐ Policy Organizations

☐ Media ☐ Third party agencies established to monitor AI

☐ Others

17.What are the best ways to ensure that AI is developed in an ethical manner? Choose top 3.

☐ Regulation ☐ Audits by external agencies ☐ Audits by AI developers ☐ Media reporting on AI ethics related topic ☐ Users reporting harmful impact from AI ☐ Training programs

☐ Improving STEM (Science, Technology, Engineering, and Math) in Nepal

☐ Increasing diversity in AI community

☐ Others

18.Please rank the following sentences (1 –Strongly agree, 5-Strongly disagree)

	1	2	3	4	5
a. Harmful effects of AI are an exception.					
b. Data bias is the most important ethical issue in AI.					
c. Even when AI is biased, it is better than humans.					
d. It is important to listen to non-STEM experts about the impact of AI.					
e. AI is currently developed with safety and its limitations in mind.					
f. AI community in Nepal needs more diversity.					
g. The primary responsibility of engineers is to develop technology, ensuring it is used ethically is the job of users.					
h. The first step is to have a working technology. Making it fair and ethical comes after.					
i. Developing AI systems is more important for Nepali society than ensuring data privacy.					
j. State of art AI technology values surveillance over everything else.					
k. If a dataset is large, it can represent the real-world well.					
l. AI evaluation benchmarks are a good way to represent real-world performance.					

Section 4: AI in Practice

19.What is the primary goal for AI projects at your company? Choose all that apply.

☐ Research ☐ Company product ☐ Client based projects

☐ Training and Skills development ☐ My company does not work on AI projects

20. Which area of AI/ML/Data Science do you currently work in? Select all that apply

☐ Computer Vision ☐ Language Processing ☐ AI for Healthcare

☐ Recommendation Systems ☐ Fintech ☐ E-Commerce

☐ Business Analytics ☐ AI/ML Tool development ☐ Agriculture

☐ Gaming ☐ AI for speech ☐ AI education

☐ Dashboards / Visualization ☐ Others (_____)

21.Do you consider real-world use and potential harms during of software projects?

◯ Yes ◯ No ◯ Maybe

22. Are you responsible for considering potential bias in your AI projects?

◯ Yes ◯ No ◯ Maybe

23. Who else is responsible for risk-analysis for AI projects at your company? Please provide as much detail as possible.

24. Which of the following AI ethics related tools have you used during your projects? Select all that apply.

☐ Model cards ☐ Data cards ☐ Ethics checklists

☐ Human oversight ☐ Stakeholder mapping ☐ Open source AI Fairness toolkits

☐ AI ethics principles from international organizations

☐ AI ethics principles from Google, Facebook, Microsoft, IBM

☐ There are tools for AI ethics? We would use it if we knew.

☐ No, we do not need such tools for our projects.

☐ Others _____

25. Where do you obtain datasets for the AI products at your workplace? Please select all that apply?

☐ Open-source datasets ☐ Datasets from company products

☐ Purchase data ☐ Crowdsourced data

☐ We use existing pre-trained models. We do not need additional datasets.

☐ Collect datasets by employees

□ Others _____

26.Please rank the following sentences (1 –Strongly agree, 5-Strongly disagree)					
	1	2	3	4	5
a. There are specific policies to deal with ethics related issues at my workplace. b. I have never faced ethical dilemmas during my projects. c. There is adequate time to ensure datasets and models are unbiased at my company. d. Our AI models are deployed with time and purpose limitations. e. Our company freely shares datasets and code with others. f. AI engineers at my company are empowered to ensure ethical technology. g. AI engineers have autonomy in defining project scope and uses. h. The datasets we use are collected with consent. i. The AI products I am developing is good for everyone. j. The main goal of risk analysis of our projects is to avoid bad PR. k. My company supports delaying/canceling a project in case of ethical problems.					

27. How important would ethical considerations be in your future career plans? 1 - strongly important, 5- not important at all

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

28. How can your company help you to develop ethical AI ?

29. How can governments and policymakers ensure AI in Nepal is ethical?

30. Do you work on AI projects outside your company role?

○ Yes ○ No ○ Maybe

31. If you answered yes to 30, can you please provide details about the project ? (Eg: school project, startup, hobby, etc)

32. If you answered yes to 30, do you consider real-world implications and ethics in those projects?

○ Yes ○ No ○ Maybe

Section 5 : Conclusion

33.Would you be interested in joining events by NAAMII related to AI Ethics? Please check all that apply:

☐ Focus group discussions ☐ AI Ethics Reading group

☐ AI Ethics Workshop/Conference ☐ AI Ethics Seminar

☐ Other _____

32. Please enter your email address if you want to be contacted about AI ethics events by NAAMII in the future?

Appendix 3: Survey for Policymakers

Section 1 : Introduction

This is an online survey on the "**Understanding of AI Ethics among Policymakers in Nepal**". This is a research project being conducted by NAAMII and is funded by UNESCO Asia and Pacific Regional Bureau for Education.

Participation:
You are invited to fill out this survey because you have experience with **Science,Technology, or IT related Policy** in Nepal. Answering the survey should take around 15 minutes. Please understand that your participation in this survey is voluntary and you may, anytime, refuse to take part in this research. You may freely decline to answer particular questions that are uncomfortable to you.

Contact:
Please reach out to NAAMII if you have any concerns or questions. We can be reached at **ai-ethics@naamii.org.np**

1.Please confirm that you have read and understood the information above, and agree to participate voluntarily in this survey.

- ○ Yes. Please take me to the survey.
- ○ No. Please take me to the end of this survey.

Section 2: Demographic Information

2.Please select your age group below.

○ Under 25 ○ 26-30 ○ 31-40 ○ 41-50 ○ 51 and above

3.Please select your gender.

○ Male ○ Female ○ Non-binary ○ Prefer not to say

4.What is your current position and organization ?

5.What is your education background ?

6.Which caste/ethnicity group do you identify with? Please choose all that apply.

☐ Brahmin ☐ Chhetri ☐ Madhesi ☐ Dalit ☐ Adivasi ☐ Others

Section 3 : AI Ethics Topics

7. Please rate your familiarity with Artificial Intelligence (AI) related topics. (5 -> very familiar, 1 -> not familiar)

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

8.Do you think Nepal currently has good policies for developing ethical AI ?

○ Yes ○ No

9.Explain your answer to 8.

10. Do you think AI Ethics is a topic of concern in Nepal ?

◯ Yes ◯ No ◯ Maybe

11. Please explain your answer to 10.

12.Please rank the following AI ethics topics in terms of their importance in the contetxt of Nepal (Most important -> Least important)

Data Privacy

AI Discrimination

Misinformation / Disinformation / Fake news

Legal protection

Labor (Job loss)

Digital divide

Lack of experts

13. What is the biggest risk of AI for Nepal?

14. In your opinion, who is the most responsible for ensuring that AI technology is ethical? Choose top 3.

☐ Engineers and AI developers ☐ Government

☐ AI companies ☐ International agencies such as the UN

☐ Universities / Research Institutes ☐ Policy Organizations

☐ Media ☐ Third party agencies established to monitor AI

☐ Others

15.Please rank the following sentences (1 –Strongly agree, 5-Strongly disagree)

	1	2	3	4	5
a. Policymakers should be concerned about harmful effects of AI.					
b. It is important to listen to non-STEM experts about the impact of AI.					
c. Current policies in Nepal encourage development of AI.					
d. Current policies and laws in Nepal can regulate AI effectively.					
e. I am aware of risks and harms of AI.					
f. AI community in Nepal needs more diversity.					
g. The first step is to have a working technology. Making it fair and ethical comes after.					
h. Developing AI systems is more important for Nepali society than ensuring data privacy.					

16. Do you have any other comments related to AI in Nepal ?
17. Would you be interested in learning more about AI and AI ethics ? ○ Yes ○ No
18. Please provide the best way of contacting you in future (email/phone number).

Bibliography

[1] Ashesh Timsina, "Artificial intelligence (ai) and it's impact in nepal," 2021.

[2] S Mahesh Acharya, "Top ai companies in nepal," 2021.

[3] Radha Sapkota, "Driven by purpose or passion: What do nepalese women have to say about their stem careers?," 2021.

[4] E. Kazim and A. S. Koshiyama, "A high-level overview of ai ethics," *Patterns*, vol. 2, no. 9, p. 100314, 2021.

[5] S. Du and C. Xie, "Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities," *Journal of Business Research*, vol. 129, pp. 961–974, 2021.

[6] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine bias," in *Ethics of Data and Analytics*, pp. 254–264, Auerbach Publications, 2016.

[7] S. M. West, M. Whittaker, and K. Crawford, "Discriminating systems," *AI Now*, 2019.

[8] F. González, Y. Yu, A. Figueroa, C. López, and C. Aragon, "Global reactions to the cambridge analytica scandal: A cross-language social media study," in *Companion Proceedings of the 2019 world wide web conference*, pp. 799–806, 2019.

[9] S. McGregor, "Preventing repeated real world ai failures by cataloging incidents: The ai incident database," *arXiv preprint arXiv:2011.08512*, 2020.

[10] D. Acemoglu, "Harms of ai," Working Paper 29247, National Bureau of Economic Research, September 2021.

[11] A. Birhane and J. van Dijk, "Robot rights? let's talk about human welfare instead," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pp. 207–213, 2020.

[12] S. Mhlambi, "From rationality to relationality: ubuntu as an ethical and human rights framework for artificial intelligence governance," *Carr Center for Human Rights Policy Discussion Paper Series*, vol. 9, 2020.

[13] A. Birhane, "Algorithmic injustice: a relational ethics approach," *Patterns*, vol. 2, no. 2, p. 100205, 2021.

[14] D. Duckworth, "A buddhist contribution to artificial intelligence," *Hualin International Journal of Buddhist Studies*, vol. 3, no. 2, pp. 27–37, 2020.

[15] R. Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, 2019.

[16] S. Barocas, M. Hardt, and A. Narayanan, "Fairness in machine learning," *Nips tutorial*, vol. 1, p. 2, 2017.

[17] M. Hickok, "Lessons learned from ai ethics principles for future actions," *AI and Ethics*, vol. 1, no. 1, pp. 41–47, 2021.

[18] J. Fjeld, N. Achten, H. Hilligoss, A. Nagy, and M. Srikumar, "Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for ai," *Berkman Klein Center Research Publication*, no. 2020-1, 2020.

[19] OECD, "Recommendation of the council on artificial intelligence," 2019.

[20] "G20 ai principles," 2019.

[21] UNESCO, "Toward a draft text of a recommendation on the ethics of artificial intelligence," tech. rep., UNESCO, 2020.

[22] F. Galindo L., K. Perset, "An overview of national ai strategies and policies," 2021.

[23] Microsoft, "Responsible ai."

[24] Google, "Artificial intelligence at google: Our principles."

[25] Facebook, "Facebook's five pillars of responsible ai," 2021.

[26] IBM Cloud Education, "Ai ethics," 2021.

[27] J. Morley, L. Kinsey, A. Elhalal, F. Garcia, M. Ziosi, and L. Floridi, "Operationalising ai ethics: barriers, enablers and next steps," *AI & SOCIETY*, pp. 1–13, 2021.

[28] amitness, "Ml companies in nepal," 2021.

[29] Government of Nepal Ministry of Education, Science, and Technology, "National science, technology and innovation policy, 2019," 2019.

[30] Government of Nepal Ministry of Communication and Information Technology, "2019 digital nepal framework unlocking nepal's growth potential," 2019.

[31] Government of Nepal, "The electronic transactions act, 2063 (2008)," 2008.

[32] The University Grant Commission, "The ugc policy regarding research misconduct, 2018," 2018.

[33] Nepal Health Research Council, "Final draft national ethical guidelines for health research in nepal, 2019," 2019.

[34] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," in *CVPR 2011*, pp. 1521–1528, 2011.

[35] A. Birhane and V. U. Prabhu, "Large image datasets: A pyrrhic win for computer vision?," in *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1536–1546, IEEE, 2021.

[36] M. Joseph Pisani, "Amazon bans police use of its face recognition for a year," 2020.

[37] BBC News, "Home office drops 'racist' algorithm from visa decisions," 2020.

[38] Nepali, Karna Bahadur, "Untouchability and caste based discrimination in nepal," *Year 20, No. 18, June, 2018 A Nepalese Journal of Participatory Development*, p. 49, 2018.