Fostering Responsible AI Innovation: Insights and Recommendations from InterLabs' Approach

Furkan Güzelant 21901515 Computer Science

Ahmet Şahin 21902702 Computer Science

Tolga Yılmaz Ardor 21901918 Electrical and Electronics Engineering

Uğur Sesli 21902853 Electrical and Electronics Engineering

Helin Arya Gündoğdu 22003524 Industrial Engineering

1. Introduction

Artificial Intelligence (AI) technologies have gained significant momentum in recent years, finding diverse and beneficial applications in fields such as healthcare, production, education, and transportation. These technologies encompass various utilities, including autonomous vehicles, content recommendation tools, medical diagnosis models, and chatbots. However, the development and utilization of AI technologies have also raised concerns in recent years [1]. People have expressed apprehensions about how the use of AI technologies can lead to issues such as discrimination, bias, and concerns related to privacy, security, and safety. After the first pedestrian was killed by an autonomous car in Arizona, concerns in society increased even further [2]. As AI technologies continue to evolve, they introduce new concerns and ethical considerations, sparking controversy regarding their benefits and drawbacks, and how to effectively manage these technologies.

Computer vision, one of the subfields of AI, empowers computers and systems to extract meaningful insights from digital images, videos, and other visual inputs. The significance of computer vision is evident in real-world scenarios, showcasing its importance in various domains such as business, entertainment, transportation, healthcare, and daily routines. We specifically address mass surveillance technologies that employ computer vision, primarily for security and safety purposes. However, concerns arise regarding the potential misuse of video surveillance technologies, leading to potential abuses and invasions of privacy [3]. As the development and usage of the technology involves the collection and processing of human data, the technology imposes a risk of potential misuse.

It is evident that companies involved in the development of AI technologies, especially those related to mass surveillance and powered by computer vision, need to act responsibly to prevent the unintended consequences of their technology. To shed light on this issue and gain insights into the innovation process within the AI industry, we have examined the methods and strategies employed by InterLabs, a Research and Development (R&D) company specializing in advanced computer vision technologies. For the methodology, we conducted three interviews with three engineers from InterLabs to gain insight into their approaches to innovation processes in the context of the social consequences of their technology. For our analysis, we will adopt Responsible Innovation, an STS approach facilitating interaction among various societal actors and incorporating values into the design and usage of the technology, as the theoretical framework for our case study. This study highlights InterLabs' approach to addressing social risks in technological innovation, particularly in the areas of privacy and equality. While commendable, the company's commitment to responsible innovation could be strengthened through a more robust assessment of technologies from social and ethical perspectives. Recommendations include collaboration with experts in social sciences and law, and

increased dialogue with customers and users, emphasizing the broader significance of integrating responsible innovation practices in the AI industry.

2. Theory

Introduction and Definition of Responsible Innovation

In this paper, the principles of Responsible Research and Innovation (RRI) were used to address ethical problems and concerns between society and InterLabs, one of the AI companies. RRI can be described below:

> "A higher-level responsibility or meta-responsibility that aims to shape, maintain, develop, coordinate and align existing and novel research and innovation-related processes, actors and responsibilities with a view to ensuring desirable and acceptable research outcomes"[4]

The definition made by B. C. Stahl emphasizes that the societal effects of the product after its release should be acceptable and responsible. With that view of RRI, InterLabs's applications will be examined by questioning what will be the effect of that product on society, what degree of responsiveness companies will have while facing discrimination in that area/ bias of the AI datasets, whether users will become more comfortable or their more risks about their privacy than before? These are included questions on the topics of RRI, therefore in this paper, it is a suitable STS approach to analyze the findings.

Key Principles of Responsible Innovation

Responsible innovation has six main titles: stakeholder engagement, broadening current assessments, placing values at the center stage, experimenting for responsiveness, monitoring RRI progress, and aiming for shared value. Stakeholder engagement is a strategic process where companies involve individuals and groups affected by their actions. Early engagement, crucial in Responsible Research and Innovation (RRI), ensures transparency and aligns innovations with societal needs. Broadening the current assessment suggests companies can expand recent evaluations by leveraging existing analyses of innovation's ethical and social impacts instead of initiating new assessments from scratch [5].

Placing values at the center stage is crucial for bridging the RRI-industry gap. Emphasizing values like sustainability, safety, and integrity as a common language simplifies RRI in the industry.Experimenting for responsiveness is crucial for companies to build trust, address societal issues, and reduce uncertainty in technology. Monitoring and measuring RRI performance can give insight into its benefits and assist companies in understanding its value. Aiming for shared value suggests creating shared value among consumers, NGOs, and companies rather than trying to gain trust [5].

Responsibility in AI Innovation and Research

Responsible Innovation involves a proactive approach to the development and deployment of new technologies, including anticipating potential risks, addressing societal concerns, and ensuring that the benefits of innovation are distributed equitably. In the context of artificial intelligence (AI), RI involves the ethical and thoughtful development, deployment, and use of AI technologies. Given the profound impact AI can have on individuals, communities, and societies, responsible innovation in AI is crucial for ensuring positive outcomes and minimizing potential risks. As AI technologies can cause ethical implications such as privacy, discrimination, bias, and safety, they require responsible actions taken by actors playing a role in the innovation such as engineers, scientists, and managers in the AI industry. However, it is observed that the approaches used by the AI industry are often insufficient due to a limited understanding of the societal implications and the instrumental view of engineering [2]. We will analyze the approaches and strategies of InterLabs, an R&D company developing AI technologies, in their AI research and innovations using responsible innovation as our theoretical approach.

3. Background Research

3.1. Technology

In recent years, the field of computer science has witnessed a rapid and remarkable evolution in the development of artificial intelligence (AI). The term AI is commonly defined as “a system’s ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation.” Since the 1950s, AI was formally recognized as an academic discipline but remained relatively obscure and of limited practical interest. Today, with the emergence of Big Data and advancements in computing power, AI has become integrated into the business landscape and public discourse [6].

The increasing computational power of computers, driven by advances in hardware technologies, has enabled them to store larger volumes of data and perform computations more rapidly. This has facilitated the creation of more sophisticated AI models capable of processing greater amounts of data, leading to an accelerated pace of AI development. These progressions gave rise to the emergence of a concept known as deep learning, referring to "computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction.” Deep learning techniques have facilitated progress in various domains, including visual object recognition and object detection, and brought success to the field called computer vision, which involves machines extracting meaningful insights from digital images, videos, and visual inputs, enabling them to take actions or provide recommendations based on the interpreted information [7, 8]. The availability of extensive image datasets like ImageNet in the

2010s played a crucial role in laying the groundwork for deep learning models capable of recognizing and detecting objects in images.

In today's digital world, enormous amounts of visual data pass through visual devices such as cameras and phones [8]. This tremendous amount and range of data provide a trainable environment for various applications [8, 9]. However, only some are utilized in real-world applications, even if extensive research is ongoing [8]. Therefore, computer vision applications still have promising potential for future applications [9]. The most established straightforward computer vision applications include object detection, classification, recognition, and detection. The primary reason for the significance of such applications lies in the fact that AI applications operate faster than individuals, providing higher accuracy in results [10]. Moreover, these applications serve as the foundation for future technological advancements.

There are numerous complex computer vision applications, including self-driving cars with lane detection and traffic sign recognition, as well as action detection and safety monitoring [8, 11, 12]. Action detection applications analyze videos and observe the actions of objects in that frame using its pre-trained algorithm [11]. These actions could be falling, fighting, running, or any common activities that the model encountered during training [11]. Similarly, safety monitoring applications detect unsafe behavior and help users take action in those situations [12]. Both applications are vital for public safety, as they share the common goal of reducing concerns and enhancing overall well-being. The accuracy of sample applications is still lower than its potential due to the limited availability of datasets to train models [12]. However, the future of applications is promising as the availability and diversity of datasets increase [13].

In this paper, we focus on a technology called video analytics that utilizes computer vision algorithms. Typically, video analytics collectively leverage advanced computer vision (CV) and artificial intelligence (AI) to solve the four-W problem. That is to identify (Who) has done something (What) at a specific place (Where) at some time (When). Over the last decade, there has been significant growth in video surveillance and video analytics, driven by practical needs and a diverse range of applications, particularly in the field of public safety. Currently, with large-scale video surveillance, more and more intelligent cameras are used in public safety systems [14].

3.2. Values

The ethical implications of human data and privacy are becoming more prominent as artificial intelligence (AI) develops. InterLabs, a company that specializes in image processing and artificial intelligence technologies, provides an insightful case study with its applications in retail and kindergarten environments.

InterLabs' AI technologies for retail concentrate on image processing to provide insightful data on product engagement, gender distribution, and customer behavior. Processing this kind of data has several phases within the data processing: acquiring the data, interpreting the data, transforming the data, and communicating the result. [15] These data are used in artificial intelligence technologies and they are applied in artificial intelligence algorithms to analyze. According to Michael Segalla and Dominique Rouziès' article *The Ethics of Managing People's Data*, these data-driven strategies highlight how crucial it is to use consumer data responsibly and transparently [16]. As discussed in that article, data-driven approaches emphasize the importance of using consumer information transparently and responsibly. Because of this, the company's values are in line with ethical principles of transparency and responsibility when dealing with personally sensitive data. Consent, privacy, and responsibility are critical ethical considerations because AI applications are becoming more and more dependent on human data as the privacy of people becomes more critical.

Concerns about privacy become more pressing in InterLabs' kindergarten app, which allows parents to remotely monitor their children. The potential risks associated with this technology is apparent, such as the possibility of privacy violations and misidentification. This aligns with the moral precepts presented in Danny Greefhorst's book, *The Human Data Processor*, which dives into the complex nature of managing human data and emphasizes the necessity of strong security measures for protecting privacy [17].

The intersection of artificial intelligence and privacy is a dynamic environment in which legal frameworks are struggling to keep up. The way InterLabs interacts with data protection authorities is indicative of its dedication to complying with current laws, including the KVKK Authority, and actively obtaining parents' informed consent. The changing nature of legal frameworks emphasizes the importance of ongoing ethical reflection and adaptation [15].

Issues with facial recognition technology—specifically, how different phenotypes respond to it differently—align with more general concerns about bias and discrimination in AI applications. This observation corresponds to the findings presented in the article *Human Data: What It Is and Why It Matters* [18]. The article emphasizes the importance of combating biases in AI models, particularly those related to race, in order to ensure fairness and reduce discriminatory outcomes. is drawing [18].

At last, the ethical considerations that InterLabs considers when employing AI technologies highlight the delicate balance that must be struck between technological innovation and the protection of human privacy [19]. The significance of taking proactive measures, maintaining engagement with stakeholders, and following changing ethical standards in the rapidly evolving field of AI technology is highlighted by this case study.

There are a lot of ethical and privacy concerns regarding AI supported mass surveillance systems. Artificial intelligence is growing as a disruptive technology, and it is expected to become more common in the next few years. As big data and autonomous systems grow more prevalent, the legal and political examination is increasing [20]. In order to prevent inconveniencies regarding these concerns according to parent's demands, governments or the ministry of education need to take action for some additional regulations and laws. Therefore, AI companies must be responsible for research and development, and manufacture of their mass surveillance product and other products.[21]

In addition to these ethical and privacy concerns, the developments in AI raises concerns even more such as AI can extract and interpret far more useful information from raw data sources, such as automating the recognition and categorization of human faces in surveillance videos. Most privacy problems will rise as more data is collected [20,22].

The increasing complexities of AI-powered surveillance systems necessitate a thorough and varied strategy to address rising ethical challenges. Apart from privacy concerns, the central ethical debate includes autonomy, openness, accountability, and societal ramifications. To negotiate these complicated ethical boundaries, the evolving AI landscape necessitates increased responsibility from all stakeholders, including AI developers, regulatory organizations, lawmakers, and educators.[20,21]

To sum up, this development of AI-powered mass surveillance highlights the crucial significance of ethical frameworks and strong policies. These frameworks must be rigorously built to be compatible with growing technical capabilities, societal values, and individual rights. The creation and implementation of ethically acceptable AI surveillance systems can avoid potential risks while maximizing the technology's beneficial potential by fostering collaborative efforts between governing authorities, industrial participants, and ethical experts.

3.3. Values By Design

"Values by Design" in computer vision and video analytics technologies refers to the intentional incorporation of ethical and moral considerations into the design, development, and deployment of these technologies. To incorporate values by design, AI researchers and practitioners must integrate moral, societal, and legal values into the design of AI systems. The responsible development of AI entails the ability to extract and depict human values, transform these values into technical specifications, address moral dilemmas and preferences, and assess systems based on their impact on human wellbeing. Traditionally, the development of AI algorithms has prioritized enhancing performance, resulting in opaque systems. Shifting the focus towards prioritizing human values in AI systems requires researchers and developers to transition from optimizing for performance to emphasizing

transparency. This shift is anticipated to spur the development of innovative techniques and applications in the field [22].

The AI technologies involving children users are also relevant to our study due to InterLabs' application for kindergarten children. AI's increasing use in monitoring children, propelled by advancements in machine learning and deep learning, has raised concerns. While surveillance enhances security to some extent, it also brings risks such as compromising children's privacy, safety, and introducing biases. In educational settings, surveillance may hinder children's freedom to take risks and express themselves openly. Therefore, researchers suggest that AI should be designed with a human-centric approach, prioritizing the rights and best interests of child users in all aspects, especially those impacting education. Both legal and ethical responsibilities exist for states and private entities, including tech companies, to uphold and protect children's rights in the development and provision of AI technologies, with adherence to the Children's Rights by Design (CRbD) standard [23].

During data processing, the biometric data is transformed and stored in various ways in various contexts. Another issue with AI in computer visioning technologies is from time to time raised whether or not the person to whom the information relates can still be identified, directly or indirectly by a third hand [24]. That is another serious concern of privacy of users and to minimize these concerns, data should be protected and regulated properly. Hence, risks that arise from the ability of an adversary to attack a biometric data profiling system, as well as risks related to the dangerous functionality of such a system, which can be due to implementation problems are all included in the issues and controversies of such systems [25]. Common issues of computer vision technologies are mostly privacy and discrimination concerns which are in general caused by the inability to process and respond to data effectively from an ethical and technological perspective. That situation can be shifted by placing the values at the center of innovations.

Making designs in accordance with public values forms a basis for mutual understanding and a common ground between the firm and the users [26]. Hence, beside the privacy by design, values by design is a crucial element of making responsible innovation. Discussing values at the very beginning of the design stage also puts conflicts and disagreements on the table [26]. Therefore developers can know what are the concerns of the users such as privacy concerns and can react to them while designing by the values.

4. Method

In order to gain insight into the field of computer vision technologies, artificial intelligence, and its ethical dilemmas, we have obtained the opportunity to interview important individuals from InterLabs. InterLabs was chosen due to their capability and competence in the AI field. Moreover, one of our group members is currently working for that company, and we could arrange interviews with qualified engineers thanks to him. After selecting our theory as responsible innovation, we conducted three interviews with senior engineers from InterLabs, and related questions about our topics were asked.

We contacted InterLabs and arranged our first interview with a senior software engineer. He is specialized in the field of computer vision, with a particular emphasis on deep learning model optimization. In the interview, we asked several questions about the ethical perspective and the STS question for their kindergarten project and successfully completed the interview.

For the second interview, we contacted another senior software engineer from InterLabs; he specializes in the fields of image processing, such as object recognition and face recognition systems. Despite the questions from the first interview, we delve deeper into the ethical concerns to better analyze the kindergarten project.

After successfully finishing our second interview, we contacted a senior AI engineer for the third interview. He is also specialized in the field, with a particular emphasis on person re-identification and multi-camera multi-target tracking. After completing the third interview as well, we generated the necessary transcripts for the interviews. By using QDA Miner Lite, we conduct a qualitative analysis of our interview transcripts to use in the findings and conclusion.

In order to evaluate each situation correctly and continue our research, the necessary sources were researched, including required help, progress paths, and ideas, and the interviews were examined in detail.

5. Findings

We conducted three interviews with three engineers from InterLabs, an R&D company that mainly works in artificial intelligence technologies. We have categorized our findings into the following four categories.

5.1. Technology

Table 1: Findings about technology

Images and AI	“We, at InterLabs, focus on artificial intelligence technologies. Specifically, our specialization lies in the field of **image processing**, with a particular emphasis on video. In image processing, we deal with both **image and video processing**.” [Senior Software Engineer].
	“We primarily work with videos, such as live video feeds, security **camera footage, streaming videos, or video files**.” [Senior Software Engineer].
Hype	It might sound a bit like "Big Brother is watching you," but it can go towards technology in a very **utopian sense**. But realistically, these possibilities could potentially **provide us with many things**. [Senior AI Engineer]
Algorithms	“We apply artificial intelligence algorithms to analyze these videos. Some of the solutions we develop include **object recognition, human recognition, facial recognition, motion detection, pose detection, scene change detection, anomaly detection**, and even **color recognition** for both objects and humans.” [Senior Software Engineer].
Technological Accuracy	“The **accuracy and performance of facial recognition vary based on the camera's location and the area of pixels the face covers within the camera frame**. The more significant the area the face occupies in the camera, the more accurate the recognition” [Senior Software Engineer].
Applications	“In the context of image processing, for example, **retail stores** use our solutions to monitor human traffic. They **track daily footfall, gender distribution, age distribution, and the location of customers within the store**.” [Senior Software Engineer].
	“We have another product in development, not yet on the market, which is geared towards **kindergarten centers**. In daycare centers, parents often feel the need to **monitor their children and stay informed about their well-being**.” [Senior Software Engineer].
	“For example, there is a kindergarten project where parents can monitor their own children within that system, which will protect the **privacy of other children and keep track of what their own children do during the day**” [Senior Software Engineer].

	"As a team, we are working on image processing using AI and we are developing a product on **video analytics**. This product **monitors humans and objects and we can also identify objects and humans** according to their features" [Senior AI Engineer]
	"There are also a lot of fields that you can use it on. Mainly, they are **security and civil life**. About security, it is hard to understand at first but in civil life, which we focus on, for example in retail, you can **see how many people visit your shop or how many times an individual comes**. You can see what products are being used. You can also use that system in a factory. You can follow the accidents of your workers etc." [Senior AI Engineer].

The primary focus of InterLabs is on image processing, particularly in the realm of video analysis. One of the software engineers described their technology focus as:

> *"We work mainly on image processing. We work on many models in image processing, for instance, we are working on a system where all of them can be used commonly in many image processing areas such as object recognition, face recognition, and age-gender." [Senior Software Engineer].*

The main technology developed by InterLabs is called video analytics, which involves the real-time processing of video to convert it into meaningful and intelligent data. Although the technologies are considered controversial and thought to pose risks by the public, the AI engineer seemed to be optimistic about the development of technologies, as the AI engineer stated:

> *"It might sound a bit like "Big Brother is watching you," but it can go towards technology in a very utopian sense. But realistically, these possibilities could potentially provide us with many things." [Senior AI Engineer]*

The company develops these technologies using computer vision algorithms. They build deep learning models that are capable of many different tasks. One of the software engineers mentioned these algorithms and their capabilities:

> *"We apply artificial intelligence algorithms to analyze these videos. Some of the solutions we develop include object recognition, human recognition, facial recognition, motion detection, pose detection, scene change detection, anomaly detection, and even color recognition for both objects and humans." [Senior Software Engineer].*

The engineers also put their finger on how the accuracy of their algorithm depends on some environmental factors. Based on these factors, the AI model that uses the computer vision algorithm can produce inaccurate results. One of the software engineers gave an example of this issue about the technology:

> *“The accuracy and performance of facial recognition vary based on the camera's location and the area of pixels the face covers within the camera frame. The more significant the area the face occupies in the camera, the more accurate the recognition” [Senior Software Engineer].*

InterLabs has many diverse applications for its technology. Two engineers from InterLabs mentioned two main applications of the technologies they developed. One notable application of InterLabs solutions is in retail stores, where their systems are employed to monitor human traffic. This information proves invaluable for optimizing store layouts and improving the overall shopping experience. This application is described by one of the software engineers:

> *“In the context of image processing, for example, retail stores use our solutions to monitor human traffic. They track daily footfall, gender distribution, age distribution, and the location of customers within the store.”*

Another application of InterLabs is a cutting-edge product currently in development that targets kindergarten centers. It allows parents to get information about their children while the children are in kindergarten. One of the software engineers explained the features of the application:

> *“For example, there is a kindergarten project where parents can monitor their own children within that system, which will protect the privacy of other children and keep track of what their own children do during the day” [Senior Software Engineer].*

5.2. Stakeholders

Table 2: Findings about stakeholders

Parents	"For example, the **customers here are actually the kindergarten children**, but the **parents will be the ones providing the demand**" [Senior Software Engineer].
	"Therefore, we obtain written consent from every parent. They are informed that the images will be processed. We must obtain **permission from each parent of the child to reprocess images**." [Senior Software Engineer].
Children	"There is a daycare application there as a sub-project, so our **customers are the kindergarten children**." [Senior Software Engineer].
Kindergarten Centers	"Additionally, some legal documents need to be signed by the **nursery authorities**." [Senior Software Engineer].
Law and Authorities	"In this context, since it falls under the jurisdiction of the data protection law (KVKK), **one of the stakeholders is the KVKK Authority**." [Senior Software Engineer].
User Feedback	"Most parents want to watch and keep an eye on their children rather than remain uninformed. **They acknowledge some level of risk, but they are willing to accept it**" [Senior Software Engineer].
	"However, some parents exhibit a high level of sensitivity to this topic. Their concern primarily revolves around technology in general rather than our specific product. A few parents feel they are being observed, and we have encountered cases **where they refuse to have their children monitored using such technologies**." [Senior Software Engineer].
	"Of course, since we are not in the market right now, **we do not receive direct feedback from stakeholders.** At first, we will be able to focus on these parts more in the demo stages. We can only act accordingly if someone comes to our own mind or says that there may be such and such problems within the company, **but it will be the demo version that can give us real feedback from real customers.**" [Senior Software Engineer].

	“At first, we will be able to focus on these parts more in the demo stages. We can only act accordingly if someone comes to our own mind or says that there may be such and such problems within the company, but **it will be the demo version that can give us real feedback from real customers**.” [Senior Software Engineer].

As explained earlier, InterLabs has diverse applications utilized in various fields, including kindergarten centers and retail stores. Consequently, the company has a large customer base and user population, as articulated by the AI engineer:

> *“Our main users are in a very wide range, however, we generally focus on the civil side of it instead of the security side. Here, the main user is the one who already uses a camera(campuses, airports, malls, shops)“.*

The interviewees mainly told us about the stakeholders involved in the kindergarten project. One of the stakeholders of the project is the children in the kindergarten, who are monitored by their parents. One of the software engineers identified children as the customers:

> *“There is a daycare application there as a sub-project, so our customers are the kindergarten children.”*

Another stakeholder of the project is parents, who are the users of the application. One of the software engineers explained the role of parents in the project as:

> *“For example, the customers here are actually the kindergarten children, but the parents will be the ones providing the demand”*

In addition, we noted the interaction of the company with their users and customers. The engineers explained about the requirement of the consent for processing private data from the parents:

> *“Therefore, we obtain written consent from every parent. They are informed that the images will be processed. We must obtain permission from each parent of the child to reprocess images.”*

Other stakeholders involved in the development process of the application are identified as kindergarten centers and KVKK authorities. One of the software engineers explained how they are engaged with these authorities:

“Additionally, some legal documents need to be signed by the nursery authorities.”

“In this context, since it falls under the jurisdiction of the data protection law (KVKK), one of the stakeholders is the KVKK Authority.”

The engineers at InterLabs also put their finger on the feedback obtained from their users in the kindergarten project. One of the software engineers described the opinions of the users, that is parents, regarding the application developed by the company:

“Most parents want to watch and keep an eye on their children rather than remain uninformed. They acknowledge some level of risk, but they are willing to accept it. However, some parents exhibit a high level of sensitivity to this topic. Their concern primarily revolves around technology in general rather than our specific product. A few parents feel they are being observed, and we have encountered cases where they refuse to have their children monitored using such technologies.”

We asked the engineers at InterLabs about their procedures for receiving user feedback for the assessments of their technology in their other applications. While some user engagement activities are observed in the kindergarten project, one of the software engineers noted that these activities are not consistently followed, showing the project being at the development stage as the reason:

“Of course, since we are not in the market right now, we do not receive direct feedback from stakeholders. At first, we will be able to focus on these parts more in the demo stages. We can only act accordingly if someone comes to our own mind or says that there may be such and such problems within the company, but it will be the demo version that can give us real feedback from real customers.” [Senior Software Engineer].

5.3. Solutions

Table 3: Findings about solutions

Privacy enhancing	"With our video image processing solutions, which include facial recognition, we can ensure that **only the parent's child is visible** on the feed, and **other children are blurred out**. This preserves the privacy of other children while allowing parents to see their own children." [Senior Software Engineer]
	"As a result of our research, we have formulated the following plans to further protect privacy. **Instead of directly transferring images**, even if they are blurred, we propose **representing each child as a unique color-coded icon** on a sketch." [Senior Software Engineer]
Bias	"We are working on **separate models for different races** where they will also have datasets."[Senior Software Engineer]
Stakeholder requests	"We also aim to address parental queries such as '**What was my child doing at that time?**' or '**What actions was my child engaged in**?' To achieve this, we've developed an application capable of detecting a person's actions based on their skeletal structure, utilizing 17 key points on their body. This technology can **identify activities like running, falling, eating, and sitting**." [Senior Software Engineer].

InterLabs displays proactive behavior by figuring out and resolving potential problems. They are committed to seeking the protection of privacy. One of the software engineers described their approach as:

> *"With our video image processing solutions, which include facial recognition, we can ensure that only the parent's child is visible on the feed, and other children are blurred out. This preserves the privacy of other children while allowing parents to see their own children."[Senior Software Engineer]*

Recognizing that blurring faces may not be sufficient, InterLabs has taken extra precautions. The same software engineer detailed it as:

> *"As a result of our research, we have formulated the following plans to further protect privacy. Instead of directly transferring images, even if they are blurred, we propose*

representing each child as a unique color-coded icon on a sketch." [Senior Software Engineer]

Additionally, InterLabs has taken precautionary measures to mitigate side effects related to race bias in video processing. One of the software engineers described their precautions as:

> *"We are working on separate models for different races where they will also have datasets."[Senior Software Engineer]*

Moreover, InterLabs is aware of the importance of stakeholder demands in shaping its solutions. Some stakeholders expressed a need for more detailed explanations of their child's activities beyond the colored icon in the sketch. One of the software engineers described their aim to fulfill this need as:

> *"We also aim to address parental queries such as 'What was my child doing at that time?' or 'What actions was my child engaged in?' To achieve this, we've developed an application capable of detecting a person's actions based on their skeletal structure, utilizing 17 key points on their body. This technology can identify activities like running, falling, eating, and sitting." [Senior Software Engineer].*

5.4. Values

Table 4: Findings about values

Transparency	"We obtain written consent from every parent. They are informed that the images will be processed. We must obtain permission from each parent of the child to reprocess images... One of the stakeholders is the parents." [Senior Software Engineer]
	"For example, by covering other children, blurring them, or preventing access, but of course, we wonder if they can see them here or if they can remove the blur or monitor the children from outside when there is an external intervention in the system, these are security concerns." [Senior Software Engineer]
Bias and Discrimination	"The reality in the market and academia is that AI systems have bias... facial recognition systems, for example, cannot work very discriminatively on black people. That is, they cannot work very well. In fact, it works better with white people." [Senior Software Engineer]

	"I think you mean our models. We are working on separate models for different races where they will also have datasets, but we have not encountered much of a problem so far." [Senior Software Engineer]
Privacy	“If such a situation were to occur, where someone else's child is recognized as their own, it would constitute a breach of privacy." [Senior Software Engineer].
	"In our kindergarten application, due to violation of personal privacy, there is a prohibition to visualize kids." [Senior AI Engineer]
	"The aim is to ensure the privacy of other people's children while someone is monitoring their own child." [Senior Software Engineer]

In the ever-evolving landscape of artificial intelligence (AI), some ethical considerations play a crucial role in shaping the responsible development and implementation of technologies. Transparency is one of the critical components of ethical development in the context of artificial intelligence (AI). As the senior software of InterLabs engineer put it,

> *"We obtain written consent from every parent. They are informed that the images will be processed. We must obtain permission from each parent of the child to reprocess images."*

This commitment attempts to make parents aware of the image-processing processes that affect their children and to involve them as stakeholders. But maintaining transparency is difficult, especially when it comes to privacy concerns. Senior software engineer brings up important issues:

> *"By covering other children, blurring them, or preventing access, but of course, we wonder if they can see them here or if they can remove the blur or monitor the children from outside when there is an external intervention in the system, these are security concerns."*

This indicates that possible outside interventions that might put at risk the privacy of the system were carefully considered. In order to create a fair and inclusive technology, bias in AI systems must be addressed too. The senior software engineer acknowledges this fact, saying,

> *"The reality in the market and academia is that AI systems have bias... facial recognition systems, for example, cannot work very discriminatively on black people. That is, they cannot work very well. In fact, it works better with white people."*

This open acknowledgment draws attention to the ongoing problems with bias in face recognition applications. In response to potential bias issues, Senior software engineer proposes a proactive approach:

> *"I think you mean our models. We are working on separate models for different races where they will also have datasets, but we have not encountered much of a problem so far."*

In an effort to reduce bias and ensure equitable representation for all races, this proactive approach highlights initiatives to develop diverse and inclusive models. Concerns about privacy take the center stage when it comes to applications involving child monitoring. A serious privacy breach could happen, as the senior software engineer points out:

> *"If such a situation were to occur, where someone else's child is recognized as their own, it would constitute a breach of privacy."*

This acknowledgment of the possibility of misidentification highlights how crucial accuracy is in AI applications, particularly in sensitive circumstances involving children. Therefore, privacy protection becomes the most important factor in a kindergarten application:

> *"In our kindergarten application, due to violation of personal privacy, there is a prohibition to visualize kids."*

This restriction shows a genuine effort to find an equilibrium between giving parents useful information and maintaining the privacy rights of all the children involved.

These insights into transparency, bias&discrimination, and privacy provide a comprehensive picture of the challenges and considerations involved in AI development. Recognizing these complexities and actively working to address them brings the industry closer to developing technologies that are not only advanced but also ethical and considerate of diverse user needs and privacy concerns.

6. Analysis and Conclusions

Placing Value at the Center Stage

One of the key lessons for RI is placing value at the center stage. This approach involves adopting certain values such as transparency, safety, and privacy. With this approach, the companies can demonstrate the values they strive for and back up their decisions in the case of value tensions [27]. In the domain of AI-driven mass surveillance technologies, key considerations for stakeholders include issues of privacy infringement, individual autonomy, power dynamics, civil liberties, discrimination, and the ethical use of technology. The paramount objectives in assessing progress in mass surveillance technologies involve upholding ethical principles, respecting human rights, and prioritizing the well-being of individuals and communities [19,28-30]. The companies developing such technologies need to be transparent about their innovation processes and they should prioritize the social and ethical values in their innovation process.

In our case study of InterLabs, we identify privacy as a central value for their application for kindergarten centers. The application constitutes a solution for parents who want to monitor their children at a day center. However, the application was concerning about the privacy of other children. The engineers noticed this concern in the design process and took proactive action to minimize the potential risks that the technology could bring. As a solution, they developed their facial recognition model to show only the children of the parent user and blur the faces of other children. This solution can be related to the concept of privacy by design. We see that InterLabs engineers inscribed privacy into the design of the application at the early stage of innovation. We can argue that it can make a good impression for InterLabs on customers, therefore, it could benefit both the company and users. In addition, we can argue that the company follows a macro-level approach in its innovation process as it foresees the possible risks and takes action accordingly, in the early stage of the innovation process, so that it can prevent those risks from being realized.

Furthermore, as discussed in the background research, It is crucial to have a clear understanding of data collection methods, particularly when dealing with sensitive information concerning children, as illustrated earlier. Ensuring that data usage aligns with explicit consent is becoming more critical to fostering ethical innovation and safeguarding the privacy of participants, specifically children [15-20]. In this regard, the engineers at InterLabs seem aware of the privacy risks that human data holds and consent requirements. As one of the software engineers mentioned, in the development of the kindergarten application they requested the consent of parents for processing the images of their children. However, in the interview with the AI engineer, it can be noted that there is an ambiguity about the legality of collecting images online. The company collects such data to train its AI models. It can be

argued that the company needs to ensure that it collects data legally. It should also have a system in place to protect the privacy of users and ensure that their images are used only for training purposes and not for any other purpose. To ensure the legality of the data collection process, the company can work with some experts in this field like lawyers and law academicians.

Another value we observed in the innovation process at InterLabs is equality. As the background research we conducted demonstrated, AI models have the potential to contribute to the discrimination of certain social groups, such as black people [31]. It is also mentioned that the Chinese system that detects Uyghur people from their faces. One of the software engineers shared information about measures taken to prevent discrimination in their AI models. As an illustration, they employ a varied dataset comprising individuals from diverse races to ensure equal assessments by the AI models. The acknowledgment of the broader implications of AI on discrimination and privacy aligns with a commitment to the well-being of individuals and communities, a core principle in responsible innovation [19, 28-30]. This is not only a moral imperative but also a strategic advantage, fostering trust among users, clients, and the wider public.

Stakeholder Engagement

Within the framework of responsible AI development, InterLabs, a company specializing in artificial intelligence (AI) technologies (especially image processing), operates on the difficult path of engagement with stakeholders. The company's commitment to ethical considerations is evident in every step it takes to address stakeholder concerns and potential risks associated with innovations.

When it comes to tackling ethical and privacy issues associated with AI technologies, stakeholder interaction should be considered. The establishment of assessment procedures is notable, especially in the context of the kindergarten application. Here, parents are the main stakeholders, and their worries about their children's privacy are carefully addressed. At the same time, authorization is crucial, and the company seeks written consent from parents in compliance with data protection laws [Interview 1].

Furthermore, acknowledging potential risks that could lead to violations of privacy, such as face recognition mistakes, shows the company's commitment to ongoing improvement and risk minimization [Interview 1]. This is in line with the industry's recognized understanding of the complex obstacles associated with investing in AI [15].

Five Ps of Ethical Data Processing provides a useful perspective for understanding InterLabs' emphasis on provenance and purpose in data utilization.

Similar to academic Institutional Review Boards (IRBs), incorporating such ethical components is a strategic step for responsible AI deployment [15].

A further approach to innovation that could improve stakeholder engagement is the investigation of new privacy-preserving technologies. The concept of face filters, which is now being explored, is a step in this direction [Interview 3]. InterLabs is currently developing a solution that addresses stakeholders' concerns about data privacy by allowing users to safeguard their privacy by masking their faces from the system.

The emphasis on source and purpose in data processing, is compatible with InterLabs' commitment to match data use with original consent and ethical considerations.

Increasing the level of transparency in AI applications is another possible innovation path. One example of this attempt is the suggestion to depict individuals as color-coded symbols on a drawing rather than using images [Interview 1]. This strategy not only eliminates the possibility of inaccuracies in facial recognition, but also offers complete privacy protection by guaranteeing that fixes are issued as soon as feasible, even in the event of errors.

As demonstrated by its evaluation processes and proactive handling of privacy concerns, InterLabs' approach to stakeholder involvement is based on a dedication to the development of ethical AI. The company is portraying itself as a responsible actor in the field of artificial intelligence (AI) by actively including stakeholders in decision-making and considering new approaches such as face filters and alternative data representations.

Aiming for shared value

Companies pursuing shared value must balance technological advances with social needs and ethical considerations.[27] As we discussed in placing value at the center stage, InterLabs aims to address privacy issues and align with societal needs in kindergarten centers by blurring images. In other words, InterLabs aimed to improve the security of their application through video analytics. This shows their commitment to shared values. InterLabs' dynamic attitude towards obtaining explicit consent from parents is another important example of shared values. This approach involves parents in decision-making processes, creates mutual understanding, and ensures transparency in the use of artificial intelligence. This mutual effort contributes to the ethical improvement of technology. This approach is strongly related to stakeholder engagement, which is important to build trust and maintain transparency during the development of InterLab’s application.[27]

An AI engineer at InterLabs also mentioned the legal and ethical challenges in collecting data. When considering the law, the use of online images poses a subtle challenge due to the lack of specific laws for this practice. This vagueness leads to ongoing discussions among the AI community, highlighting the need for responsible and fair data use. InterLabs demonstrates its commitment to increasing mutual trust and fostering shared values in the technological landscape by openly addressing the legal vagueness and ethical considerations surrounding collecting data.[27]

Monitoring RRI Performance

RRI performance in InterLabs can be evaluated in various aspects such as social impact awareness, sustainability, risk management and safety. In all those topics, the company has general consideration and awareness about how they can fulfill requirements and become more responsible. Before deciding on risk management strategies, first social impacts should be understood and the question “What might be the outcome of those societal impacts?” should be asked so that risk and concerns about the technology they are developing become clear.

In InterLabs, a kindergarten application is developed on AI computer visioning, therefore some risks about that technology that company also pondered over are the representation of people of color, safety of data collected of preschool children from the cameras and privacy of other children while monitoring one. There are some solutions designed by the company for these concerns, for example, processing children's data and sharing them after blurring the faces of others, or removing the use of images completely and showing them only as figures such as dots on the map. Regarding the representation of people from different ethnic backgrounds, InterLabs aims to reduce the effect of big data bias by using different datasets for users with marginal characteristics and they try this algorithm to work on the models created for them. Although the company has produced solutions for discrimination and privacy concerns and is still producing different solutions, these risks in computer monitoring technology have not been minimized yet.

Since the kindergarten application has not been released yet, it is not yet possible to fully observe the social effects of some of the issues in this example; instead, some effects can only be considered by making predictions. In this regard, there is also the possibility that the social impacts that the company has thought about and designed the application accordingly are not comprehensive enough, and this is a factor that affects the company's responsiveness in RRI perspective. While monitoring company’s performance on RRI methodology, Based on what was said in the interviews about risks, security issues and future developments, it actually shows that the company is quite aware of responsible innovation and is responsible by proactively maintaining its responsiveness.

Experimenting for Responsiveness

To achieve a higher level of responsiveness the key feature is the adaptation ability to evolve systems or approaches to better meet expectations of stakeholders in an innovative and responsible way. In order to control the responsiveness of companies there should be some indicators, for example user-centered design, co-creation, ethical considerations, organizational culture, iterative prototyping and knowledge sharing. At InterLabs, many of these indicators are considerations of the team that are developing and designing the applications. For example, in terms of user-centered design, who the users are, what kind of application is needed and demanded, and what the users' values are, were taken into consideration as the most basic elements in all steps from the idea stage of the kindergarten application to the stage when the software was completed. Thus, the application and design are developed around users' demands, needs and concerns.

Another work InterLabs has done to enhance its responsiveness is creating a prototype of the application and testing it with different considerations. For example, in applications requiring biomonitoring, they measured the accuracy and reliability of the output by including images of black people in the prototype of the application in order to represent people of different races more accurately and to process the collected data in a bias-free manner.In this case, they tried to comply with ethical issues by trying to eliminate discrimination and also prepared their application for possible future scenarios by testing application prototypes. It can be inferred that their proactivity and inclusive attitude build on the company's responsiveness.

Broadening Current Assessment

From the standpoint of broadening the current assessment, we propose possible solutions based on the cumulative data from the interviews and background research to ensure that InterLabs promotes their kindergarten project in terms of responsible innovation practices for their kindergarten monitoring. To widen the present assessment, they may develop an internal committee or board responsible for assessing and approving data collection and usage rules, as well as assuring compliance with ethical norms and regulatory criteria. Second, to ensure continuing responsible innovation, undertake frequent audits and evaluations to analyze any biases, accuracy, and privacy threats. Third, give personnel participating in video processing substantial ethical training to enhance awareness of potential biases, privacy problems, and the necessity of responsible innovation. Fourth, one of the most important components in promoting their project is transparency. They might improve openness by establishing clear communication channels with parents and providing transparent information on how their children's data is collected, processed, and used inside the kindergarten monitoring system. Create thorough

and transparent methods for collecting, storing, using, and disposing of video data from kindergarten centers. A hypothetical example of such communication is taking a group of parents and showing them the inside procedures behind the software with engineers. The fifth possible solution for promoting responsible innovation research and company benefit could be to evaluate the impact of AI algorithms on diverse demographics on a regular basis, ensuring fairness and inclusivity in facial recognition and object detection across different races, genders, and age groups. Because software is prone to error and may be biased against certain ethnic minorities. As a result, regular AI assessments may broaden the existing assessment. Finally, they must respond to and resolve complaints and problems on a frequent basis. As a result, they could put in place a system to resolve parental concerns or grievances about data privacy and utilization.

To summarize, the recommended solutions for InterLabs' kindergarten project are centered on strengthening responsible innovation methods. Establishing an internal board for ethical compliance, performing frequent audits, and offering intensive ethical training to people are among the recommendations. Improving openness through open lines of contact with parents and different demographic assessments of AI algorithms are also critical. Establishing a system to handle parental concerns about data privacy ensures responsibility. These solutions aim to support responsible innovation by establishing ethical standards and constructing a solid framework for the kindergarten project's childcare monitoring system's success.

Conclusion

In conclusion, during the analysis of our interviews we observed that overall, the engineers at InterLabs are aware of the possible risks regarding social dimensions such as privacy breach and discrimination. They follow the macro-level approach during the innovation process, that is, they foresee the possible risks and social implications in the early stages of the technological innovation process and inscribe social values such as privacy and equality. We examined how these values are incorporated into the design in particular, in the kindergarten project and facial recognition models implemented by the company. Furthermore, we also analyzed the principles of the company in the data collection and processing processes. We observed that the engineers take account of the privacy concerns regarding human data and they employ the necessary actions such as asking for consent from the parents in the kindergarten application development. However, we also identified some approaches of the company that could be improved in terms of responsibility for innovation. Mainly, we observed that the company lacks a reliable assessment of their technologies in social and ethical perspectives. The company can work with experts in social sciences and law to gain insight into how their practices comply with the RI approach. In addition, the company can increase the dialogue with possible customers and users to learn about their concerns and values.

Through this project, we have gained insights into the principles employed by an AI company and the impact of social and ethical considerations on the technological innovation process, utilizing responsible innovation as our theoretical framework.

7. Appendix

7.1. Interview with a Software Engineer at InterLabs - 1

Furkan Güzelant: First, can you please give us an overview of the technologies your company is developing? What kind of technologies are you working on?

Interviewee: We, at InterLabs, focus on artificial intelligence technologies. Specifically, our specialization lies in the field of image processing, with a particular emphasis on video. In image processing, we deal with both image and video processing. We primarily work with videos, such as live video feeds, security camera footage, streaming videos, or video files. We apply artificial intelligence algorithms to analyze these videos. Some of the solutions we develop include object recognition, human recognition, facial recognition, motion detection, pose detection, scene change detection, anomaly detection, and even color recognition for both objects and humans. Essentially, we can perform various analyses on both objects and humans, such as identifying the color of clothing on a person or the color of a car.

Furkan Güzelant: Thank you. The second question pertains to your applications and users. What are your main applications, and who uses them? Could you elaborate on this?

Interviewee: We serve several different areas with our technology. In the context of image processing, for example, retail stores use our solutions to monitor human traffic. They track daily footfall, gender distribution, age distribution, and the location of customers within the store. We have another product in development, not yet on the market, which is geared towards kindergarten centers. In kindergarten centers, parents often feel the need to monitor their children and stay informed about their well-being. However, kindergarten centers have multiple children, not just the parent's own child. Privacy is a significant concern here. With our video image processing solutions, which include facial recognition, we can ensure that only the parent's child is visible on the feed, and other children are blurred out. This preserves the privacy of other children while allowing parents to see their own children. However, as I mentioned, this product is still in the development stage.

Tolga Yılmaz Ardor: I have a question as well. You mentioned the technologies you're developing and the blurring solution. What are the risks associated with these technologies?

Interviewee: There are indeed some risks involved in this field. The accuracy and performance of facial recognition vary based on the camera's location and the area of pixels the face covers within the camera frame. The more significant the area the face occupies in the camera, the more accurate the recognition. As a result, there is

a rare possibility of incorrect matching. In other words, there is a chance that someone else's child might be recognized as the parent's child. The accuracy is also influenced by the camera's resolution, lighting conditions, and contrast. These are the risks associated with the technology. The reason this product is still in development is to observe and minimize these risks. If such a situation were to occur, where someone else's child is recognized as their own, it would constitute a breach of privacy. I can say that this is a consideration.

Uğur Sesli: May I ask a question as well? Do you have routine evaluation procedures to assess users and stakeholders? Do you gather feedback and conduct user assessments?

Interviewee: In this context, since it falls under the jurisdiction of the data protection law (KVKK), one of the stakeholders is the KVKK Authority. This is because it involves privacy concerns and facial recognition, which is considered biometric data and is a protected and regulated form of data. Processing it requires consent. Therefore, we obtain written consent from every parent. They are informed that the images will be processed. We must obtain permission from each parent of the child to reprocess images. The images cannot be shared until a blur effect is applied. Only after that can they be shared with the parents, and they are shared exclusively with the parents. One of the stakeholders is the parents. Additionally, some legal documents need to be signed by the nursery authorities. We ask parents which part of their children they want to be shown. This carries some risk since if the system makes a mistake and selects the wrong child's face, how will the parents react? We are currently conducting investigations on this matter.

Furkan Güzelant: Have any parents objected to or complained about this technology?

Interviewee: Most parents want to watch and keep an eye on their children rather than remain uninformed. They acknowledge some level of risk, but they are willing to accept it. The majority of parents share this perspective. However, some parents exhibit a high level of sensitivity to this topic. Their concern primarily revolves around technology in general rather than our specific product. A few parents feel they are being observed, and we have encountered cases where they refuse to have their children monitored using such technologies. Nevertheless, such instances are rare, and most parents are willing to embrace the associated risks.

Uğur Sesli: What are your thoughts on the future of this technology, and what plans do you have for developing its transparency and enhancing stakeholder perspectives?

Interviewee: As a result of our research, we have formulated the following plans to further protect privacy. Instead of directly transferring images, even if they are

blurred, we propose representing each child as a unique color-coded icon on a sketch. This approach ensures that even if our application mistakenly selects the wrong child, any correction can be made in less than 20 seconds, providing 100% privacy protection. However, this alone may not suffice. We also aim to address parental queries such as 'What was my child doing at that time?' or 'What actions was my child engaged in?' To achieve this, we've developed an application capable of detecting a person's actions based on their skeletal structure, utilizing 17 key points on their body. This technology can identify activities like running, falling, eating, and sitting. By overlaying action-based text, such as 'Performing this action here and now on the map' on the sketch, we can convey information nearly as effectively as an image while safeguarding privacy in case of system errors. We are currently working on implementing these enhancements.

Furkan Güzelant: Do you have any concerns about bias or discrimination in your AI application?
Interviewee: The reality in the market and academia is that AI systems have bias. There is a situation like this: facial recognition systems, for example, cannot work very discriminatively on black people. That is, they cannot work very well. For instance, it works better with white people. In fact, the performance of facial recognition systems varies according to race. I'm unsure if it goes into many details or discrimination. There is such a situation. Apart from that, it is close to technology in general. It is not something we are involved in, but there is a situation like this as the world agenda. For example, the Chinese have made a system in the Uyghur region. That is, they can tell from a person's face whether he is Uyghur. The world criticizes this, and there is news about this. Some can use technology in the wrong way. Let's just say that this technology can be used for bad purposes worldwide.

7.2. Interview with a Software Engineer at InterLabs - 2

Furkan Güzelant: First of all, we would like you to talk about your technologies as a company, in general, in which areas do you work?

Interviewee: We work mainly on image processing, we work on many models in image processing, for instance we are working on a system where all of them can be used in a common way in many image processing areas such as object recognition, face recognition, age-gender. Briefly like this.

Uğur Sesli : What are your applications and information? Can you give examples for your applications, who are the main users?

Interviewee: There is a kindergarten application there as a sub-project, so our customers are the kindergarten children,but are actually parents in a way. For

example, there is a kindergarten project where parents can monitor their own children within that system, which will protect the privacy of other children and keep track of what their own children do during the day. For example, the customers here are actually the kindergarten children, but the parents will be the ones providing the demand. Other projects can branch out a little more and focus on security, where more or less any corporate structure can be involved.

Uğur Sesli :Finally, can you explain things like what are the risks in this business regarding values? Do you have concerns about privacy? Is transparency important to you or are there biases in this technology? Things like that.

Interviewee: Of course, we want to ensure this security in the daycare application. The aim is to ensure the privacy of other people's children while someone is monitoring their own child. For example, by covering other children, blurring them or preventing access, but of course, we wonder if they can see them here or if they can remove the blur or monitor the children from outside when there is an external intervention in the system, these are security concerns. In a different way, if we expect it to work in closed systems, we do not expect much security problems, but of course, if the institution receiving the data does something illegal within itself, there may be a problem.

Uğur Sesli : Are you subjected to discrimination? Do you have such a concern because of race or skin color?

Interviewee: I think you mean our models. We are working on separate models for different races where they will also have datasets, but we have not encountered much of a problem so far. For example, we tried his experiments on basketball player Lebron James and we did not encounter such a problem yet, but of course, they were detailed photographs focused on the face, we do not know how they will be done with real-time cameras, but I don't think there will be much problem when the model is special for them.

Tolga Yılmaz Ardor: I want to ask a question too. Do you have a process that investigates these stakeholders' requests or concerns and produces solutions accordingly?

Interviewee: There is no such team at the moment. If we think of anything that comes to mind, we discuss it among ourselves.

Tolga Yılmaz Ardor: So, you are constantly taking into account the opinions of these stakeholders, right?

Interviewee: Of course, since we are not in the market right now, we do not receive direct feedback from stakeholders. At first, we will be able to focus on these parts

more in the demo stages. We can only act accordingly if someone comes to our own mind or says that there may be such and such problems within the company, but it will be the demo version that can give us real feedback from real customers.

Uğur Sesli : So, what do you think about the future of this technology? Simply put, what are your future plans? Where do you think this technology is heading, how can it be advanced?

Interviewee: Have you watched the series Person of Interest? At the end of the day, although not ours, similar systems can evolve here, that is, when all systems are connected to each other, as in Person of Interest, when e-mails, texts, sounds, images are all connected to each other. It is actually difficult, but it does not seem impossible when you look at it.

7.3. Interview with an AI Engineer at InterLabs

Furkan: We are curious about your technology. What are you developing and what kind of fields are you working on?

Interviewee: For that question, first, I want to mention that as a team instead of company. As a team, we are working on image processing using AI and we are developing a product on video analitics. This product monitors humans and objects and we can also identify objects and humans according to their features. You can also structure a data when it doesn't have that quality. You can also search it. There are also a lot of field that you can use it on. Mainly, they are security and civil life. About security, it is hard to understand first but in civil life, which we focus on, for example in a retail, you can see how many people visit your shop or how many times an individual comes. You can see what products are being used. You can also use that system on a factory. You can follow the accidents of your workers etc.. I will explain it further.

Furkan: As our second question. What are your applications on technology and who are your main users in those technologies?

Interviewee: Our main users are in a very wide range, however we generally focus on the civil side of it instead of the security side. Here, the main user is the one who already uses cameras(campuses, airports, malls, shops). This product improves cameras with AI in a systematic way and main feature is getting cameras smart. We can apply that to everyone who has a camera.

Furkan: Thank you. Next question is what are your values as a company? What are the risks that your applications includes? For example, investigating with camera

might be a privacy problem or transparency problem? Can you explain your concerns and customers' concerns. We would be happy to hear about that.

Interviewee: First of all, I want to give a general information which I think you should know too. Now, AI, especially image processing AI includes everything especially things about humans (like processing the face). As data, you need images so that you can build and educate an AI model using those images and use it as a result. Whole AI world has uncertainty about that procedure. For instance, last year, we made a work inside of the company about face recognitions' ethical side. How should be the license of it in an ideal scenario? At first, Law side of it, law departments etc. coming behind in technology. When you produce something, they argue whether that product is suitable or not after 5-10 years of production. They are always coming from behind. AI has the same problem too. About face recognition, there are no articles about that in LAW literature. There was an article in 2022, which made the business impossible, for instance when you take a photograph from Google, both Google and the site Google took the photo should allow you. So, it wasn't very beneficial. There is not a precise solution to that. For example, your photos in public are being used for face recognition too. They are available to get used in AI models. Even in our models, we can use any photo online. Moreover, it is understood as legal to use images from youtube (even if there is no precise law about it). Videos that you took from youtube or a news channel are seems as okey but only your personal pages like facebook or instagram seen as a concern. In addition to that, we are doing counter movements too. For example, In our video analytic application, we provide the individuals who are unknown, that system won't use in face recognition, we give an ID to that person. For instance, when you come to a shop, we give you an ID such as ID2. If you come a month later, we call you ID2 again but it is unknown that what are your personal information and shop doesn't have to share that info. We are trying to protect confidentiality in that way. Moreover, in our kindergarten application, due to violation of personal privacy, there is a prohibition to visualize kids. On the other hand, according to growing violence, it is asked to follow children too by their parents. As a result of this, we are trying to create a product that blurs other kids and around except that parent's child. Another example is, as a basic solution to that problem, we gave a project to Bilkent EE department last year students using networks that create face filters. When you process your image from the system, the system will not be able to recognize your face. This is for people to be able to upload their photos with the knowledge that their photo won't be used anywhere and they will feel safe. We are taking precautions against AI models. Hope that answer is enough for you.

Furkan: [00:11:56:03] Okay, You've answered the fourth question. I was going to ask about what kind of changes you've made. We mentioned that technologies change, and there are some unwanted aspects to technologies. Some changes are made in

response to that. I was going to ask about that. You've touched on it a bit. If anything else comes to your mind, you can elaborate on it.

Interviewee: [00:12:27:24] What do you mean by changes? What else was there, apart from what I've just explained?

Furkan: [00:12:33:01] It could be in terms of design or in terms of implementation. Changes of this kind.

Interviewee: [00:12:41:01] I guess it's not about ethics. Is it about technical changes this time?

Furkan: [00:12:44:13] It could be technical. I mean changes that address the concerns of users.

Interviewee: [00:12:52:23] Actually, I think it's a bit technological. For example, the video analytics side, artificial intelligence applications are advancing very specifically. You all see AI news constantly on LinkedIn, YouTube, Twitter, and so on. A new model comes out, and the processed version of that model in a video is shared, and you witness that the AI stack is constantly advancing as an industry every day. In fact, as students, you too. But when it comes to turning a product into something or meeting a real need in the field, it's not exactly like that. For example, right now, due to the increasing population and the crowded metropolitan life, security is a significant issue, and also learning statistics. So, in the V-tail side, statistics are becoming very meaningful, and in the video analytics side as well. Actually, the goal is this: to somehow bring the sophistication of AI to the user level. Anyway, in that field, these video management systems, which are already there in the industry, are advancing. We are developing a video analytics product a bit independently, so that it directly meets the needs of people, and when doing video, since the interface in those management systems is very outdated, in fact, in stores, campuses, etc., for people to use more comfortably. We are doing a new interface for it. We are doing it with new features, and if I need to explain it more technically, we are making it flexible. This means that you can use it with all cameras. You can place security cameras there, you can place the screen image on your phone on the screen of your computer, and with the right protocols, or you can place the camera on a drone. And we want to make it so that you can do this, which means this: if you can get it for three cameras, you can get it for ten cameras, a thousand cameras. These are options that you can find piece by piece in the industry. But right now, we actually have them all collected in one way. So, we are trying to do these things differently. I don't know if these are useful pieces of information.

Furkan: [00:16:13:03] Thank you. Let's move on to the next question. My next question is about secondary interests, for example, it could be users or customers.

Do you implement measurement procedures for them? Is there such a thing? Do you gather their opinions? How?

Interviewee: [00:16:42:23] Actually, we are currently an R&D company, and this product is only in the demo stage. It's not a product that has been launched in the market. Therefore, we are not yet receiving customer expectations because there is no customer directly using the product at this stage. So, we can't really measure it because we haven't received much feedback. But I can say that we did it very well before. At least at this stage, we are not adding something to the product just because we can do it; we are trying to put it according to a real need. So, I may not be able to give a complete answer to the question.

Ahmet: [00:17:33:00] Actually, because you make it according to the expectations of the customers, you have already answered the question. So, even if they haven't tried it directly, ultimately, you listen to their expectations and desires and create a product that suits them.

Interviewee: [00:17:46:11] What we hear is also available here. We also talk to the potential users of this product, people who might potentially use this product. We ask them if they would use it if there was such a feature. We make our own suggestions as well. There are also things that the other party says they need. Besides these discussions, we use a lot of things. Using the internet, YouTube, or Google, we really find out where it can be placed. Where it can be placed, what difference it makes when placed with which features, etc., according to their expectations. In fact, it becomes a system that is built according to their expectations. I can summarize it this way.

Furkan: [00:18:38:03] I understand, thank you very much. Let's move on to the last question. The last question is generally about the future of video analytics technologies. What do you see in the future, what will happen?

Interviewee: [00:18:58:17] In the future of video analytics, I see a few things. It might sound utopian, but I think this way. For example, right now, a video analytics might be watching you through an airport. But the airport's purpose is not to watch you; it's to get information about things related to the airport by watching everyone. Furkan Güzelant doesn't want to get information about what Ahmet Şahin is doing or will do there. Or it can monitor a store. The store's purpose is to gather statistics and increase sales. The purpose of a shopping mall may be to find the most visited store, the most used area, and make that store sell at a higher price, for example. But I also think that it will be like this. When we become widespread as individuals everywhere, these systems will allow us to get our own data; we can monitor ourselves. This is likely to add a lot to us because it's always visual. For example, imagine this: you know how many steps you've taken. You probably know it now. But this time, you'll know where exactly your posture was disrupted, and at what age it

was disrupted. Because the camera provides much more detailed information. For example, there is no data about what you wore on which day right now. But with this kind of system, it will be created. You can choose to pay in cash or have your name not written on the receipt; it doesn't matter because you are recognized by the system. It might sound a bit like "Big Brother is watching you," but it can go towards technology in a very utopian sense. But realistically, these possibilities could potentially provide us with many things. Realistically, it is expected that video analytics will greatly increase its market share right now. We think that we will do more testing of these. Anomaly detection, for example, will be much easier. When someone runs at an airport, for example. When they start running, for example. Because running at an airport is an abnormal condition. I'm not talking about running to catch a flight. Really. Like two people chasing each other. And these kinds of things within a second. As an alarm fee or... In a kindergarten, you watch. When a child falls, the parent will know within a second. Therefore, or think about what happens in a hospital. I think the information flow will speed up a lot because you provide all the information you can't see right now within three minutes, five minutes, ten minutes, or an hour. Think of it as a single-channel alarm system. Alarms are coming from everywhere. Like, an earthquake happens, we learn about it instantly. Like that, we will be able to learn and monitor everything, and we will be able to get a much faster information. Because now you provide all the information you didn't see right away, within three minutes, five minutes, ten minutes, an hour, or even later.

Furkan: [00:24:22:00] Thank you very much. My questions end here. Ahmet, do you have any question?

Ahmet: [00:24:27:19] Thank you very much. For the detailed information.

Interviewee: [00:24:30:11] You're welcome. Thank you very much.

Furkan: [00:24:31:14] Thank you.

Interviewee: [00:24:33:05] You're welcome. So, are we closing now?

Furkan: [00:24:37:18] Yes, I say goodbye.

Interviewee: [00:24:39:09] Goodbye, Thank you.

8. Credits

In this project necessary role distributions are as the following;

Furkan Güzelant (Editor and Interview Coordinator)

Furkan has a crucial role for the general organization of the report. His dedication for details and elaborations were crucial evidents for him to complete efficient editing process. Moreover, he was generally in charge of the groups' interview sessions. He made interview possible with coordinating them. He didn't just ensured the smooth process of the report, he also succeeded in smooth progress for interviews and he provided important insights that improved the quality of the report.

Tolga Yılmaz Ardor and Uğur Sesli (Content Researcher)

Tolga and Uğur performed a crucial role while conducting thorough research for the report and the interviews. Their effort on obtaining according datas and accurate sources enriched the reports quality. Their hard work on finding relevant articles and controlling their relevance with content played a crucial role for the credibility of our report.

Helin Arya Gündoğdu (Content Reviewer and Editor)

Arya contributed to the project by playing a very crucial role as completing missing parts and correcting mistakes in the report and content. Her attention and dedication on details and quality assurance was important to complete the final document in a proper way. She also conducted controls on headings and made sure the accuracy and coherence.

Ahmet Şahin (Creative Ideation Specialist)

Ahmet's dedication on showing a proactive approach for obtaining creative solutions, ideas and thoughts on each stage of the project helped to make the project original. His innovative ideas, solution skills contributed to the projects' final version. He generally conducted the brainstorms and suggested unique ideas on the content and titles.

However, each of the group members contributed in each speciality counted. Demonstrated individuals were just focused on that area furthermore. In addition, everyone exchanged ideas, contributed to the writing, and helped each other while writing the subheadings. Everyone tried to compensate for each other's shortcomings.

9. References

[1] Benjamins, Richard, Alberto Barbado, and Daniel Sierra. "Responsible AI by design in practice.", 2019.

[2] K. Joyce *et al.*, “Toward a sociology of Artificial Intelligence: A call for research on inequalities and structural change,” *Socius: Sociological Research for a Dynamic World*, vol. 7, p. 237802312199958, 2021. doi:10.1177/2378023121999581

[3] A. Senior *et al.*, “Enabling video privacy through computer vision,” *IEEE Security and Privacy Magazine*, vol. 3, no. 3, pp. 50–57, 2005. doi:10.1109/msp.2005.65

[4] B. C. Stahl, “Responsible research and innovation: The role of privacy in an emerging framework,” Science and Public Policy, vol. 40, no. 6, pp. 708–716, 2013.

[5] Lawlor, R. Engineering in Society Royal Academy of Engineering. “Statement of Ethical Principles”pgs. Excerpt: 20-22.

[6] M. Haenlein and A. Kaplan, “A brief history of artificial intelligence: On the past, present, and future of Artificial Intelligence,” California Management Review, vol. 61, no. 4, pp. 5–14, 2019. doi:10.1177/0008125619864925

[7] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature News, https://www.nature.com/articles/nature14539 (accessed Dec. 3, 2023).

[8] “What is Computer Vision?,” IBM, https://www.ibm.com/topics/computer-vision (accessed Dec. 3, 2023).

[9] S. Paneru and I. Jeelani, “Computer Vision applications in construction: Current State, opportunities & challenges,” *Automation in Construction*, vol. 132, p. 103940, 2021. doi:10.1016/j.autcon.2021.103940

[10] E. J. Topol, “High-performance medicine: The convergence of human and Artificial Intelligence,” *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2019. doi:10.1038/s41591-018-0300-7

[11] Y. Zhang, P. Tokmakov, M. Hebert, and C. Schmid, “A structured model for action detection,” *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. doi:10.1109/cvpr.2019.01021

[12] W. Fang *et al.*, “Computer vision applications in construction safety assurance,” *Automation in Construction*, vol. 110, p. 103013, 2020. doi:10.1016/j.autcon.2019.103013

[13] D. Bhatt *et al.*, “CNN variants for Computer Vision: History, architecture, application, challenges and future scope,” *Electronics*, vol. 10, no. 20, p. 2470, 2021. doi:10.3390/electronics10202470

[14] Q. Zhang, H. Sun, X. Wu and H. Zhong, "Edge Video Analytics for Public Safety: A Review," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1675-1696, Aug. 2019, doi:10.1109/JPROC.2019.2925910.

[15]A. Ballantyne, “Where is the human in the data? A guide to ethical data use,” GigaScience, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6047448/ (accessed Nov. 30, 2023).

[16] “The Ethics of Managing People’s Data.” *Harvard Business Review*, 13 June 2023, hbr.org/2023/07/the-ethics-of-managing-peoples-data.

[17] A. D. Greefhorst, “The Human Data Processor,” TDAN.com, https://tdan.com/the-human-data-processor/20296 (accessed Nov. 30, 2023).

[18] R. Garner, S. Deakins, and N. Gerrard, “Human data: What it is and why it is important?,” CitizenMe, https://www.citizenme.com/human-data/ (accessed Nov. 30, 2023).

[19] M. Wiesenthal, “The ethical implications of AI-based mass surveillance tools,” OPUS 4, https://repositorium.hs-ruhrwest.de/frontdoor/index/index/docId/770 (accessed Nov. 26, 2023).

[20]“The Science of Human Data Collection,” Q Analysts, https://qanalysts.com/the-science-of-human-data-collection/ (accessed Nov. 30, 2023).

[21] Clough, Ryan. “The Inevitability of AI Law & Policy:Preparing Government for the Era of Autonomous Machines.” *Public Knowledge*. Oct. 2018. https://publicknowledge.org/wp-content/uploads/2021/11/AI_Report.pdf

[21] Author links open overlay panelGenia Kostka a et al., “Under big brother’s watchful eye: Cross-country attitudes toward facial recognition technology,” Government Information Quarterly, https://www.sciencedirect.com/science/article/pii/S0740624X22000971 (accessed Dec. 12, 2023).

[22] Dignum, Virginia. “Responsible Artificial Intelligence: Designing Ai For Human Values.” *ITU Journal: ICT Discoveries*, Special Issue No. 1, Sept. 2017. https://www.itu.int/dms_pub/itu-s/opb/journal/S-JOURNAL-ICTS.V1I1-2017-1-PDF-E.pdf

[23] von Struensee, Susan. “Eye on Developments in Artificial Intelligence and Children's Rights: Artificial Intelligence in Education (AIEd), EdTech, Surveillance, and Harmful Content.” June 2021. Available at SSRN: https://ssrn.com/abstract=3882296 or http://dx.doi.org/10.2139/ssrn.3882296

[24] Kindt, E.J. The Risks Involved upon the Use of Biometric Data and Biometric Systems. In: Privacy and Data Protection Issues of Biometric Applications. Law, Governance and Technology Series, vol 12, (2013). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7522-0_4(accessed Nov. 27, 2023).

[25] Andronikou, V., Yannopoulos, A.,Varvarigou, T. “ Biometric Profiling: Opportunities and Risks.” In: Hildebrandt, M., Gutwirth, S. (eds) *Profiling the European Citizen*. Springer (2008)., Dordrecht. https://doi.org/10.1007/978-1-4020-6914-7_7 (accessed Nov. 27, 2023).

[26] Mirjam de Groot et al. “Visioning with the Public: Incorporating Public Values in Landscape Planning”, *European Planning Studies*, 2014, 1165-1181, DOI: 10.1080/09654313.2013.770218

[27] I. van de Poel et al., “Learning to do responsible innovation in industry: Six lessons,” Journal of Responsible Innovation, vol. 7, no. 3, pp. 697–707, 2020. doi:10.1080/23299460.2020.1791506

[28] Internet encyclopedia of philosophy, https://iep.utm.edu/surv-eth/ (accessed Nov. 26, 2023).

[29] S. S. U. of California et al., “Lawful device access without mass surveillance risk: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security,” *ACM Conferences,* https://dl.acm.org/doi/pdf/10.1145/3243734.3243758 (accessed Nov. 26, 2023).

[30] Overall, Jeffrey. "The Ethics of Mass Surveillance: An Anarchist, Objectivist, and Critical Theorist Perspective." Journal of Information Ethics 28, no. 2 (Fall, 2019): 34-50. https://www.proquest.com/scholarly-journals/ethics-mass-surveillance-anarchist-objectivist/docview/2317014392/se-2.

[31] Stephen Carradini et al. “The Potentials and Pitfalls of Computer Visioning and Machine Learning Methods for Communication Researchers”.Proceedings of the 41st ACM International Conference on Design of Communication. October 2023. https://dl.acm.org/doi/abs/10.1145/3615335.3623024#sec-ref (accessed Nov. 27, 2023)