# Ethische aspecten

Bij het ontwikkelen van het systeem is er nagedacht over ethische implicaties. In dit hoofdstuk worden de belangrijkste aandachtpunten besproken.

## Transparantie en uitlegbaarheid

Een belangrijk ethisch aandachtspunt is de transparantie van het systeem. De pipeline maakt gebruik van veel verschillende modellen, zoals objectdetectie, Human Action Recognition en LLM. Hierdoor kan het lastig zijn om te achterhalen op basis van welke informatie uit de video het LLM zijn antwoord heeft gegenereerd. Om de transparantie zoveel mogelijk te waarborgen, worden alle uitkomsten van de modellen opgeslagen, samen met de bijbehorende tijdstempels in de video. Daarnaast worden ook alle tussenuitkomsten van het LLM bewaard. Op deze manier kan zo goed mogelijk worden gecontroleerd hoe het model tot zijn antwoord is gekomen.

## Privacy en datagebruik

Hoewel er geen persoonsgegevens worden gebruikt tijdens het trainen van de modellen zelf, de data die binnen het project wordt gebruikt bestaat uitsluitend uit technische video’s en foto’s, is er toch nagedacht over de mogelijke privacy risico’s. Er kan bijvoorbeeld uit de interactie met het systeem informatie worden afgeleid over het kennisniveau van een monteur, wat in theorie zou kunnen leiden tot monitoring. Om dit te voorkomen, is besloten dat het systeem uitsluitend bedoeld is ter ondersteuning van het leerproces.

## Menselijke controle

Een ander ethisch aandachtspunt is het ontbreken van controle op het antwoord van het LLM. De monteurs kunnen de output van het systeem gebruiken ter ondersteuning, maar zijn niet verplicht om deze op te volgen. Omdat er geen controle van de modeloutput plaatsvindt, ligt de verantwoordelijkheid bij de eindgebruiker om de gegenereerde informatie kritisch te beoordelen. Het is daarom belangrijk dat de gebruiker zich hiervan bewust is.

## Toegankelijkheid en inclusiviteit

Een ander ethisch risico is dat niet alle gebruikers in gelijke mate toegang hebben tot de voordelen van het systeem. Bijvoorbeeld, monteurs met minder digitale vaardigheden of beperkte taalvaardigheden kunnen moeite hebben met het gebruiken van het systeem en of het interpreteren van antwoorden. Dit kan ertoe leiden dat de kloof tussen gebruikers juist groter wordt.

## Risico op function creep

Er bestaat een risico dat het systeem in de toekomst wordt ingezet voor andere doeleinden dan oorspronkelijk bedoeld, zoals het beoordelen van de prestaties van werknemers. Om het vertrouwen van de gebruikers te behouden is het van belang dat dit niet gebeurt.