Application of Ethics in My Future Career

One thing I am particularly interested in is artificial intelligence and machine learning. Such forms of technology often require a great deal of data in order to build. That being said, if I'm ever in the position of choosing what data to gather and use, I'll need to make an important decision on what kind of data is ethical to use. This becomes an issue especially when dealing with human generated data. Obviously, if I have the express consent of a person to use their data in order to build models, then that is fine to use. And obviously if someone has expressly said that they do not want their data to be used, then it would be unethical to do so. But what about the gray area where I have neither been permitted or denied use? One such instance has to do with public places. While it is legal to film someone in a situation where there isn't a reasonable expectation of privacy, is it okay to do that? For instance, if I wanted to use the exterior security cameras of the building I worked at, or set up passive surveillance cameras in public areas in order to gather data on people to train some sort of model, is this okay? I would say, to be as ethical as possible, it is not okay. While there is nothing illegal about this activity, using the interactions of random people to train a model without their permission, especially when that model will be used to make a profit off of some product it is being used in, I would say is wrong. You're using data generated from people without their knowledge to make money without giving them a share. But what if it's not to make money, or if it's for some sort of noble cause? For instance, what if there is an area with a particularly high murder rate and you're trying to build a model that can detect when someone is in distress, in which case it will call the police? This is a bit trickier, especially since "justifications" for collecting and using data for the "noble" purpose I just mentioned can lead to a slippery slope that leads to areas that are definitely wrong. As they say, "the path to hell is paved with good intentions." In this case I would say that using data collected from people in a public place to create a system that would ultimately save lives is admissible. But there would need to be some sort of ethics committee set up before such actions were taken that clearly outlined what are "noble" reasons and when/where it is okay to use publicly acquired data in order to keep such actions from snowballing out of control. But even then there is still the problem of bias being introduced. This might be resolved at least partially by creating an ethics committee that was composed of a highly diverse group, but even then, in order to be there, they'll all have to have some common goal that will cause bias. I think it's possible to create a system that will keep a creation of such a model within ethical boundaries. But I won't always have a diverse group of individuals to go to about what is and isn't ethical. So before I dive into my career, I need to do two things. Create a list of both what is ethical in a general societal sense, and what I personally think is ethical. By doing this, I would have a predetermined ethical line in which I would not cross that would guide me as I encountered these problems. This will become especially important in the instance that I get involved in a project that I am passionate about. In such a case, without a predefined set of principles to follow, it would be easy to find ways to justify the means and do what I want in the moment. But with a predefined ethical line, I could keep myself in check and not be influenced by my feelings in the moment.