

Dilemma of Privacy vs. Ethics for a Social Robot



Pepper greets a visitor for the second time in a mall and says, “You remember, yesterday you bought this bracelet”, then shows the picture on the screen. The intention was to connect better with the person, but what if today the guy is with some lady who was not offered that bracelet?

Remember the Microsoft AI chatbot ([Tay](#), aiming at learning based on conversational understanding), which became in less than 24 hours a “racist” and that the company had to take down? People can teach robots “bad things”, and a robot can behave “unethically”. The Innovation department, which focuses also on learning from interaction, is

proactively trying to handle such situations. The first part of the [video](#) (using our interaction-based learning framework) shows how easily a simple interaction can lead the robot to learn potentially unethical behaviors from the person, on the name of teaching privacy. The second part of the video (mostly scripted) is a set of different possible responses from the robot. Innovation is planning for a survey-based user study, using this video with the aim to find the unknown answer to “how” the robot should behave in this dilemma of Privacy vs. Ethics and to guide our developers to program the robots to minimize such “risks”.