

Case Study 1:

<https://www.scu.edu/ethics/focus-areas/technology-ethics/resources/open-source-ai-to-release-or-not-to-release-the-gpt-2-synthetic-text-generator/>

Artificial Intelligence (AI) will change the world, and OpenAI knows this. As a company dedicated to the development of AI, they must be careful with the realities they create, since they can be used in a spectrum of ways on society, both positively and negatively. Upon the completion of the GPT2 Synthetic Text Generator model, the team was excited to announce it to the world. However, when they did so, there were some caveats. Instead of releasing the premier version of the model, the team opted to release a lesser trained version, preventing it from creating realistic, 'politically-charged, false texts that could be used to stir disorder. These decisions directly align with the ACM principle 1.2 to avoid harm. Omitting the 'intelligence' of the highly trained model allows people to explore the potential of the text generator without creating a potential for social unrest. To no surprise, maintaining this AI in privacy does more benefit than avoiding harm. ACM principle 1.3 of being honest and trustworthy is preserved through keeping this generator out of the hands of unethical users. Finally, upon release of the generator, the company was explicit about the potential negative uses of the product, following principle 2.5 (Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks). On their Github, they dedicate a large part of the description of the GPT-2 to its shortcomings and problems, opening the door for other researchers and engineers to play with the code, and find creative solutions.

Case Study 2:

<https://www.scu.edu/ethics/focus-areas/more/engineering-ethics/engineering-ethics-cases/copyright-concerns/>

As a new hire, there is a lot of resources that go into learning about the way a company operates, what its rules are, and how its culture behaves. At companies, especially large ones, employee training is an essential aspect of onboarding new roles. This is a fundamental issue that SDX Alliance should devote more resources to. Intellectual property (IP) is a highly important topic around the world, in nearly every industry, but arguably most importantly in software creations. By not making it an area of focus, SDX Alliance is making itself vulnerable to lawsuits and potential IP infringement, leaving principle ACM 1.5 off of its fundamental checklist. On the individual side, Ralph did not follow ACM principle 2.2, which outlines how developers should maintain high standards of professional competence, conduct, and ethical practice. Ralph did not display this given he was unaware of the consequences of the mistake. On the contrary, both the company and Ralph demonstrated a strong use of principle 2.1, achieving high quality in both the process and products of professional work, by applauding the improvements of the system after the code was added, and for using existing code, saving time and energy. Unfortunately, the code was copyrighted, but the idea of efficiency and quality is present. Finally, the four-month quality assurance process review checks principle 1.1 on the dot, ensuring the system contributes to society and human well-being, making sure no errors are present upon release that could be of detriment.