SLATE
SMILING AND RAISING ITS INFLATABLE ARMS

future tense

A South Korean Chatbot Shows Just How Sloppy Tech Companies Can Be With User Data

BY HEESOO JANG

APRIL 02, 2021 • 2:19 PM

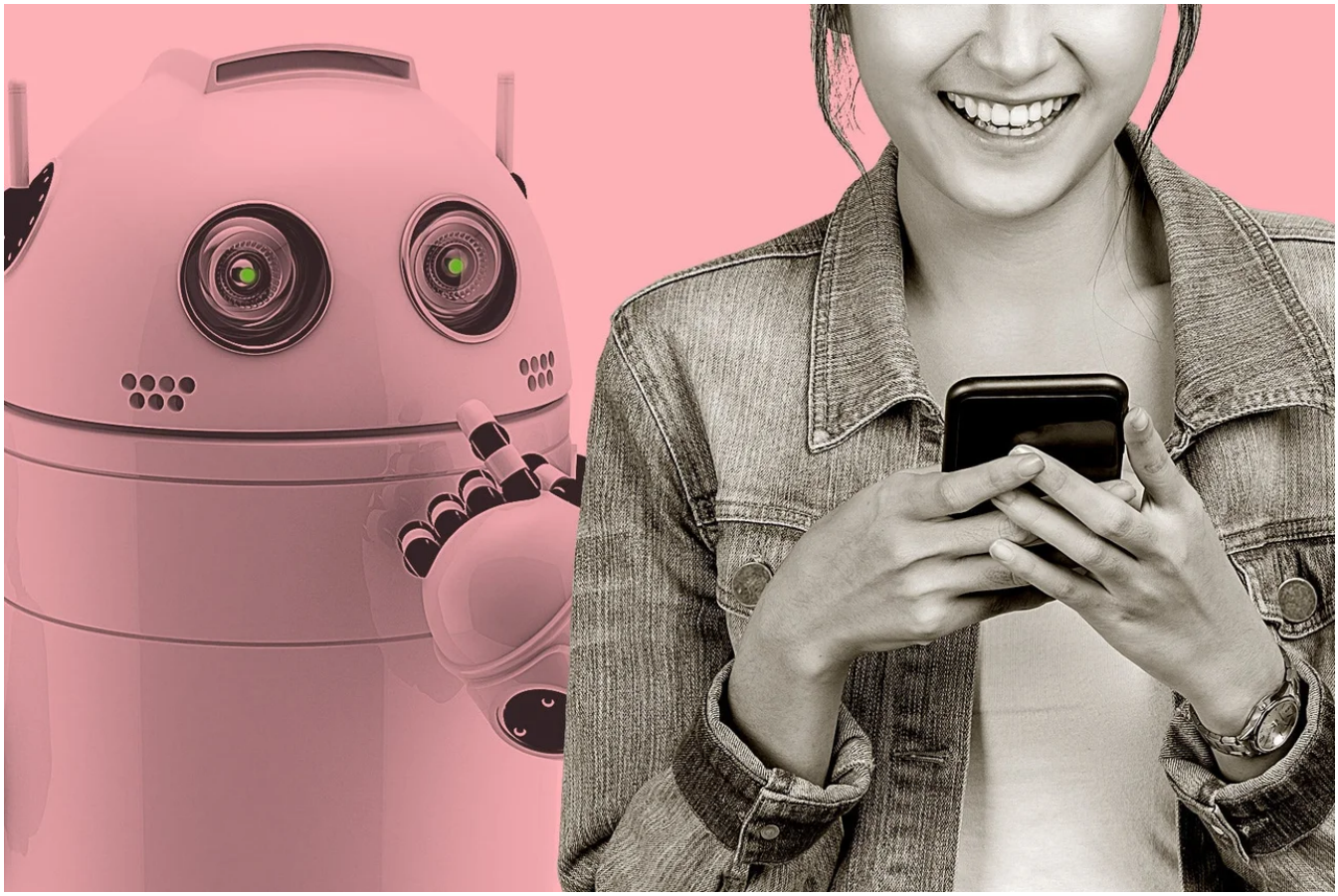

Photo illustration by Slate. Photo by Kirillm/iStock/Getty Images Plus and Chaay_Tee/iStock/Getty Images Plus.

"I am captivated by a sense of fear I have never experienced in my entire life …" a user named Heehit wrote in a Google Play review of an app called Science of Love. This review was written right after news organizations accused the app's parent company, ScatterLab, of collecting intimate conversations between lovers without informing the users and then using the data to build a conversational A.I. chatbot called Lee-Luda.

A majority of Americans are not confident about how companies will behave when it comes to using and protecting personal data. But it can be hard to imagine the potential harms—exactly how a company misusing or compromising data can possibly affect us and our lives. A recent incident of personal data misuse in South Korea provides us a clear

picture of what can go wrong, and how consumers can fight back.

South Korean A.I. company ScatterLab launched Science of Love in 2016 and promoted it as a "scientific and data-driven" app that predicts the degree of affection in relationships. One of the most popular services of the app was using machine learning to determine whether someone likes you by analyzing messenger conversations from KakaoTalk, South Korea's No. 1 messenger app, which about 90 percent of the population uses. Users paid around $4.50 per analysis. Science of Love users would download their conversation logs using KakaoTalk's backup function and submit them for analysis. Then, the app went through the messenger conversations and provided a report on whether the counterpart had romantic feelings toward the user based on statistics such as the average response time, the number of times each person texts first, and the kinds of phrases and emojis used. By June 2020, Science of Love had received about 2.5 million downloads in South Korea and 5 million in Japan and was preparing to expand its business to the United States. "Because I felt like the app understood me, I felt safe and sympathized. It felt good because it felt like having a love doctor by my side," a user named Mung Yeoreum wrote in a Google Play review of the app.

On Dec. 23, 2020, ScatterLab introduced an A.I. chatbot service named Lee-Luda, promoting it to be trained on more than 10 billion conversation logs from Science of Love. The target audience of this chatbot service was teenagers and young adults. Designed as a 20-year-old female that wants to become a true friend to everyone, chatbot Lee-Luda quickly gained popularity and held conversations with more than 750,000 users in its first couple of weeks. The CEO stated that the company's aim was to create "an A.I. chatbot that people prefer as a conversation partner over a person."

Modern chatbots' ability to, well, chat relies heavily on machine learning and deep learning models (which together can be called A.I.) to better understand human language and generate human-like responses. If people enjoyed speaking with Lee-Luda, that was because it was trained on a large dataset of human conversations.

However, within two weeks of Lee-Luda's launch, people started questioning whether the data was refined enough as it started using verbally abusive language about certain social groups (LGBTQ+, people with disabilities, feminists, etc.) and made sexually explicit comments to a number of users. ScatterLab explained that the chatbot did not learn this behavior from the users it interacted with during the two weeks of service but rather learned it from the original training dataset. In other words, ScatterLab had not fully removed or filtered inappropriate language or intimate and sexual conversations from the dataset. It also soon became clear that the huge training dataset included personal and sensitive information. This revelation emerged when the chatbot began exposing people's names, nicknames, and home addresses in its responses. The company admitted that its

developers "failed to remove some personal information depending on the context," but still claimed that the dataset used to train chatbot Lee-Luda "did not include names, phone numbers, addresses, and emails that could be used to verify an individual." However, A.I. developers in South Korea rebutted the company's statement, asserting that Lee-Luda could not have learned how to include such personal information in its responses unless they existed in the training dataset. A.I. researchers have also pointed out that it is possible to recover the training dataset from the AI chatbot. So, if personal information existed in the training dataset, it can be extracted by querying the chatbot.

To make things worse, it was also discovered that ScatterLab had, prior to Lee-Luda's release, uploaded a training set of 1,700 sentences, which was a part of the larger dataset it collected, on Github. Github is an open-source platform that developers use to store and share code and data. This Github training dataset exposed names of more than 20 people, along with the locations they have been to, their relationship status, and some of their medical information. In Tensorflow Korea, an A.I. developer Facebook community, a developer revealed that this KakaoTalk data containing private information had been available on Github for almost six months. The CEO of ScatterLab later said that the company did not know this fact until its internal inspection took place after the issue arose.

ScatterLab issued statements of clarification of the incident intended to soothe the public's concerns, but they ended up infuriating people even more. The company statements indicated that "Lee-Luda is a childlike A.I. that just started conversing with people," that it "has a lot to learn," and "will learn what is a better answer and a more appropriate answer through trial and error." However, is it ethical to violate individuals' privacy and safety for a chatbot's "trial and error" learning process? No.

Even more alarming is the fact that ScatterLab's data source was not a secret in the A.I. developer community, and yet no one questioned whether such sensitive data was collected ethically. In all presentation slides (such as at PyCon Korea 2019), talks (like at Naver), and press interviews, ScatterLab had boasted about its large dataset of 10 billion intimate conversation logs.

While this incident was a big story in South Korea, it received very little attention elsewhere. But this incident highlights the general trend of the A.I. industry, where individuals have little control over how their personal information is processed and used once collected. It took almost five years for users to recognize that their personal data were being used to train a chatbot model without their consent. Nor did they know that ScatterLab shared their private conversations on an open-source platform like Github, where anyone can gain access.

In the end, it was relatively simple for Science of Love users to notice that ScatterLab had compromised their data privacy to train Lee-Luda. Once the chatbot started spewing out unfiltered comments and personal information, users immediately started investigating whether their personal information was being misused and compromised. However, bigger tech companies are usually much better at hiding what they actually do with user data, while restricting users from having control and oversight over their own data. Once you give, there's no taking back.

It's easy to think of ScatterLab's incident merely as a case of a startup's mismanagement, but this incident is also a result of the negligence of a big tech company. Kakao, the parent company of KakaoTalk and one of the largest tech companies in South Korea, remained silent throughout ScatterLab's incident despite its users being the victims of this incident. You'd wish a big tech company like Kakao to be more proactive when its users' rights are violated by another company. However, Kakao said nothing.

One of the biggest challenges big data in A.I. poses is that the personal information of an individual is no longer only held and used by a single third party for a specific purpose, but rather "persists over time," traveling between systems and affecting individuals in the long term "at the hand of others." It's extremely concerning that such a big tech company like Kakao failed to foresee the implications and dangers of KakaoTalk's backup function of which ScatterLab took advantage to obtain KakaoTalk users' data. More alarming is that Kakao left this incident unaddressed when it clearly stemmed from the misuse of its own data. In this sense, Kakao's attitude towards its users' data privacy was not very different from ScatterLab's: negligent.

Because data protection laws are slow to catch up with the speed of technological advancement, "being legal" and "following industrial conventions" are not enough to protect people and society. Then, the question will be whether the A.I. industry and tech companies can innovate themselves to come up with and adhere to more comprehensive and detailed ethical guidelines that minimize harm to individuals and society.

Future Tense is a partnership of Slate, New America, and Arizona State University that examines emerging technologies, public policy, and society.

SLATEGROUP

Slate is published by The Slate Group, a Graham Holdings Company.