Data 6, Summer 2023 **Weichert and Ferrari**

Data Privacy

Discussion 4 *Summer 2023*

Nowadays, data are everywhere. Google knows our search history, Apple Watches track your exercise and heart rate, and BeReal knows our location. But is this ubiquity of our personal data a good thing?

In this discussion, we will discuss the implications of the world of 'big data', and the importance of embedding privacy into our work as data scientists.

1 My Data is Safe, Right?

1. To start off, think about the apps you use every day on your phone, tablet or computer. What might these apps know about you? Are you confident that the apps keep your data private? Do you restrict any access to your location, media, or search history?

   Next, read this article by a student in *The Hoya*, Georgetown University's student newspaper. In the commentary, Srishti Khemka explains how BeReal, the popular social media app that encourages users to post candid photos of themselves during random times of the day, may know too much about your daily life.

2. Do you agree with Khemka? Do you use BeReal and if so, are you considered about how much it knows about you? What about the data that TikTok has?

3. How would you define 'data privacy'? What does (or should) *private* mean in an age of cellphones, BeReal, and GPS monitors?

4. Do you think it is important for apps like BeReal or TikTok to keep your data private? Do you think these apps or companies are doing a good job at keeping personal data private?

5. Do you think data scientists have an obligation to keep personal information private? Is there a trade-off between privacy and being able to gain insights from data that might be used to help people?

2 Data and the Government

What privacy means to you and what privacy means to the government are likely two different things. When is it 'reasonable' for the government to intrude on your privacy or obtain your personal data without your consent?

The US Supreme Court has long dealt with issues of privacy. Key cases have dealt with personal privacy involving marital relationships (Griswold v. Connecticut), unreasonable searches and seizures (Mapp v. Ohio), wiretapping (Katz v. US), and more. In 2018, the Court decided another important privacy case, this time involving cellphone GPS data.

1. Read the summary of the case, Carpenter v. US. What was at issue in the case and what did the Supreme Court rule. Do you agree or disagree with the Court's conclusion and why?

2. Which justices voted in favor of ruling against the "warrantless acquisition of [the defendant's] cell-site records"? How do you think the Supreme Court would rule on the case if it were brought before them today?

3. What are some implications of the Carpenter decision? Does this court case prevent the government from ever obtaining your cell-site location information or other personal data?

4. A newer controversy over data privacy is brewing in the United Kingdom, where Apple is threatening to block access to iMessage due to a UK law preventing Apple from using end-to-end encryption for its messaging service. Read this Guardian article about the controversy. In your opinion, is there a difference between someone's location data and their private text messages? Should different privacy rules apply?

5. Should the government be the one to decide when data should remain private and when it can be shared with others? If not the government, who should decide? Does this have implications for us as data scientists?