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

Data Privacy Solutions

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?

   Solution: Mobile and computer applications may know a variety of things about you, depending on what permissions you've given them. Think about what applications may track your location (even if you're not currently using the app), what apps have access to your contacts, or what some services know about your interests based on the content you like or what videos you watch.

   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?

   Solution: Even if BeReal doesn't have access to your location data, could the app infer that from the pictures you take or the friends you've added on BeReal? What information about you might BeReal be able to generate based off of when or where you post, and what might TikTok 'know' based off of the types of videos you like?

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

Solution: Privacy in the more traditional sense refers to preventing unauthorized third parties from being able to obtain or 'see' information about you, the user. This could include preventing a company from seeing your GPS or account password, or preventing a government from obtaining your internet history, for example. However, sometimes it is not enough to simply hide some of your personal data if apps and algorithms are able to 'learn' information about you even if you don't explicitly share it with the application. TikTok provides a good example: within just a few minutes of creating a TikTok account and browsing/liking videos, the app's algorithm builds an understanding of you as a user, which may include information about your likes, dislikes, geographic location, etc. It can *infer* this information without having you input it yourself.

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?

Solution: Your answer to this question may depend on what you consider to be 'private' and the extent to which you are ok with companies knowing certain information about you.

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?

Solution: As people charging with working with and learning from data, data scientists have an ethical obligation to obey any privacy or confidentiality agreements, and to protect the privacy of user data. However, the goal of data privacy may sometimes stand in conflict with a desire to 'learn' from user data. If data scientists don't have access to some personal data, they may not be able to build an accurate picture of what is going on in the real world. This may hamper the ability of policymakers to make targeted interventions based on the most in-need populations.

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?

   Solution: At issue in Carpenter v. US was whether the government (in this case, the FBI) could legally obtain the cell-site location information (CLSI) data of a suspect without a search warrant. The Court's majority ruled that this was a violation of the defendant's 4th Amendment rights protecting him against unreasonable searches and seizures. In its opinion, the majority recognized that technology like CLSI has made it easier than ever to gather wide-reaching data about people, and that new standards should be applied to the obtaining of this type of data.

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?

   Solution: The 2018 Carpenter decision was a 5-4 decision in which Chief Justice John Roberts joined with four traditionally liberal justices (Breyer, Kagan, Ginsburg, and Sotomayor). As of 2022, only 3 of these justices remain on the Court, and only one (Breyer) has been replaced with an ideologically-similar successor (Jackson). One could speculate that if this case were re-decided today, a conservative majority would overrule this decision.

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?

   Solution: One of the biggest legal implications of the Carpenter decision is that new technologies may necessitate new legal frameworks to decide questions of privacy. However, this ruling does not prevent the government from ever obtaining your CSLI data, but does require them to obtain a search warrant to do so.

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?

Solution: There's no right or wrong answer here. Some consider location data to be less 'private' than private text messages, and argue that the latter should be more protected than the former.

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?

Solution: These are normative questions that are up to you to think about.