Notes- Why Data Privacy Matters

Aditya Pal

Data Privacy is not about hiding stuff. It is about

1. Freedom of choice
2. Non-discrimination/equal opportunity
3. Transparency and accountability

Challenges of Regulations

1. What little control users have over their data is solely dependent on the UI revealing and correctly executing that control.
2. Some regulations have enforced UI improvements that have enabled certain users, for example, to erase their data from platforms. Such restrictions are difficult to apply because they ultimately depend on our capacity to check the accuracy of software.
3. Worse, users can remove their data by following the UI instructions, but the knowledge retrieved from this data may have already been incorporated into the models as a whole – a change that cannot be reversed.

What We Can Do

1. Inquire whether providing your data is essential.
2. read the terms and conditions to understand what is done with your data
3. clear cookies and consider what activity patterns describe you (PII)
4. Consider what data you leak across applications - strive to compartmentalize, making it impossible to connect data from several sources.
5. Think on the choices you've made.

Differential Privacy

Differential privacy (DP) is a mechanism for publicly sharing information about a dataset by describing the patterns of groups within the dataset while keeping information about individuals inside the dataset private. The theory behind differential privacy is that if the effect of making a single random alteration in the database is tiny enough, the query result cannot be used to infer anything about any specific individual, and therefore ensures privacy.

Privacy Tools for the Average User

Classifiers can be fooled by generating adversarial examples