ADF COMPANION DOCUMENT G: UKAN Anonymisation with Differential Privacy

What is Differential Privacy?

Differential Privacy guarantees a limit to the amount of information, specific to any individual, that is revealed by an analysis. Differential Privacy is not an algorithm itself, but a property of an algorithm. There are a variety of differentially private algorithms for use in situations such as releasing aggregate statistics, creating synthetic data, and training machine learning models.

These algorithms typically rely on introducing a small, controlled amount of noise to the analysis. This noise is calibrated to be of the same magnitude as an individual's possible contribution to the release, effectively masking the contribution of any one individual. One important distinction is whether noise is added to data as it is collected (known as local, or distributed Differential Privacy), or is added when analysis is run on centralised data (global).

The amount of information revealed is parameterised by a variable epsilon (ϵ), which also controls noise added. Lower epsilon means less can be learned, with more noise added. Higher epsilon means more can be learned, with less noise added.

Differential Privacy is a mathematical statement[1]. Let X be an arbitrary data domain, and let T be an arbitrary output domain. A randomised algorithm mapping a dataset to output $A : X^n \to T$ satisfies ε-Differential Privacy if for all dataset pairs $x, y \in X^n$ where x and y differ by a single entry, and for all subsets S of the output domain T

$$\Pr[A(x) \in S] \leq e^{\epsilon} \Pr[A(y) \in S],$$

where the probability is over the randomness of the algorithm A. At most, the probability of any one outcome can change by a factor of e^{ϵ}.

The benefits of Differential Privacy

The Differential Privacy approach gives a provable mathematical guarantee in terms of how much could be learned about a specific individual from the released data. Fundamentally, all analysis published from a dataset limits the domain of possible underlying values in the data, and this is the root cause of attacks that reverse engineer confidential data from published outputs. Differential Privacy addresses the root cause of attacks - disclosure of information specific to individuals - rather than specific attack methodologies. This gives attack-method-agnostic protection.

[1] Note that this is a formulation of "pure" Differential Privacy. Multiple variants exist varying by what they guarantee about the similarity of outputs.

This formal framework gives a quantity epsilon that can be related simultaneously to both disclosure (information revealed about any individual) and utility (distortion of the data due to added noise) impact on the output, facilitating reasoning about this inherent trade-off. Disclosure risk in terms of epsilon can be understood across multiple differentially private algorithms (or repetitions of one algorithm) run on the same data, a property known as composition. In this case, for i algorithms the overall epsilon is $\epsilon = \sum \epsilon_i$.

The Differential Privacy guarantee is robust. It cannot be weakened regardless of what is done to the release, allowing it to be used without fear of additional disclosure risk. Nor is it weakened by revealing any parameters of the protection itself, including algorithm details and epsilon values. This can allow those using the output to take noise perturbation into account, potentially enhancing utility, and can reduce reliance on assessments of the data environment.

Current limitations of Differential Privacy

Setting epsilon determines the maximum amount of individual-specific information an analysis could reveal, but how much is appropriate? Setting epsilon also determines the noise added to the data, and the effect on utility must also be understood. A systematic method for setting epsilon has yet to be agreed upon, either in theory or practice.

Differential Privacy allows for tracking of total information revealed by multiple analyses. The overall epsilon across these analyses is a sum of the individual epsilons, and places a bound on the total information revealed. This allows the setting of a total epsilon budget of an acceptable amount of information to reveal, whereby a target overall epsilon can be split across multiple analyses. Whilst the idea of a fixed privacy budget restricts the amount of analysis that can be performed on a dataset, it is important to note, this privacy budget is not unique to Differential Privacy: information is revealed *whenever useful analysis is produced from a dataset*. Differential Privacy simply exposes this reality.

The relationship between Differential Privacy and anonymisation

Under European law, the property of a dataset of being ‘anonymous’ relates to the likelihood that an individual in that dataset could be identified. When identification is considered sufficiently unlikely, taking into account all objective factors and the means reasonably likely to be used, the data is out of the scope of European data protection law.

Differential Privacy does not provide a measurement of the risk of re-identification, instead it offers a way of limiting the maximum amount of information a release reveals about an individual. To understand anonymisation in terms of Differential Privacy one must consider how this information relates to the risk individuals could be identified.

A differentially private release will therefore not necessarily be anonymous. Whether it is anonymous will depend on the value of epsilon selected and the specifics of the data situation. There is a correlation between noise added and the likelihood of re-identification. The less noise, the more an attacker can change their confidence that an individual had a given property, which constitutes re-identification by inference under the GDPR. But whilst an important factor, noise is not the only factor that needs considering when evaluating if data is anonymous, it will also depend on the wider data situation.

Sometimes Differential Privacy is characterised as the 'strongest' protection for statistical releases. It is strong in the sense that it provides a mathematically robust guarantee that released information is limited. This limit, however, is chosen by the user and will not necessarily always afford a desired level of protection, such as anonymity.

Many existing concepts of anonymisation are based on protections against known attacks. As such there is a risk that as compute power and data science advance they will become vulnerable. Differential Privacy, as a formal guarantee, is 'stronger' than existing methods because the guarantee it offers is not at risk of being broken.

Using Differential Privacy

Differential Privacy, whilst rapidly developing, is still nascent outside of academia and as such can be challenging to use well. However, there are a number of open source toolkits and resources available, such as Harvard's OpenDP project[2].

When evaluating whether or not to use Differential Privacy, consider:

- Whether the possibility for attack can be prevented by environmental controls such as sharing analyses within a controlled environment with trusted parties.
- Whether you are confident that the data is not vulnerable to attack.
- Whether a formal guarantee of disclosure risk is required in order to access and perform analysis on the data, such as guaranteeing individuals have a form of plausible deniability.

As Differential Privacy is a fast evolving research field that is beginning to see large scale real world applications[3], the likely trend is that it will become easier to use in a wider range of scenarios over time. Differential Privacy is a highly promising field, but can be challenging to use effectively, as such it is worth engaging with an expert early in your project and ensuring that they are able to assist not just in delivering a differentially private output, but in managing the privacy budget to ensure the output is both safe and useful, which can be more challenging.

If you want to learn more about Differential Privacy, there are a range of good resources[4][5] that can provide more information.

[2] https://privacytools.seas.harvard.edu/opendp

[3] https://www.census.gov/about/policies/privacy/statistical_safeguards/disclosure-avoidance-2020-census.html

[4] https://gss.civilservice.gov.uk/wp-content/uploads/2018/12/12-12-18_FINAL_Privitar_Kobbi_Nissim_article.pdf

[5] https://privacytools.seas.harvard.edu/publications/differential-privacy-primer-non-technical-audience-preliminary-version