CANON Canadian Anonymization Network

Spectrum of identifiability

A risk-based spectrum approach to defining identifiability allows for a broad range of innovative uses of information, while accounting for, and mitigating, a reasonable amount of residual risk. To that end, we propose the following spectrum of identifiability with 3 specific states of information:

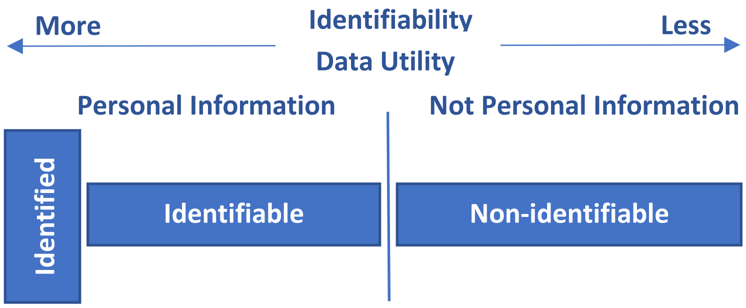

Figure 1. States of Information

DEFINITIONS – Spectrum of Identifiability

Identified information: Information which, by itself, directly identifies an individual.

Identifiable information: Information for which there is a serious possibility in the circumstances that it could be associated with an identifiable individual.

Non-identifiable information: Information for which there is no serious possibility in the circumstances that it could be associated with an identifiable individual.

The phrase "**in the circumstances**" involves consideration of contextual factors that include: likelihood that an actor will attempt to identify the information; what other datasets are accessible to an adversary which might be combined with the information; the environment in which the information will be used or into which it will be released; and, any controls associated with the information.

Where information ultimately lies on the spectrum of identifiability will be influenced by two factors:

- the innate identifiability of the information (e.g., some information may be non-identifiable from the moment of collection); and,
- the controls applied to further reduce identifiability.

Notes

- Rendering information non-identifiable may be achievable through multiple different techniques, including aggregation, data transformations, data synthesis, homomorphic encryption, and others. Any privacy-enhancing technology or process that creates non-identifiable information should be considered equally valid.
- Information to which the above techniques have been applied may be identifiable or non-identifiable, depending on the circumstances.
- We recommend against the use of the terms "anonymous information" and "anonymization," because the term has been used in the past to convey a range of meanings, and in some cases to mean information for which there is no possibility of identification.