

Dear ACCC,

For my public submission, I note the following from your preliminary report:

[De-identification of 'personal information', P187.](#)

I would ask that the ACCC reconsider its de-identification example using "hashing".

The example given I would suggest is actually the creation of a "unique identifier" using a hash function, rather than de-identification.

The creation of a unique identifier is confirmed through your description of the use of this "deidentified data", whereby two parties confirm that they are talking about the same person, but not by name.

This is a danger with using hash functions, in that while the original string is no longer identified by name, hash functions are designed to create a statistically unique number from a string of characters.

To find out the original credit card number from a hashed value, one can simply build a data base of known credit card values and hashes. This would take around 3.6 days to create a lookup table of hashes to credit card values for all Visa Cards, around the same for Mastercard.

Source: <https://jimshaver.net/2015/02/01/doing-the-math-on-hashing-credit-card-numbers/>

As such, I would recommend that the ACCC consider the work of the Data61 unit of the CSIRO in regards to de-identification.

<https://data61.csiro.au/en/Our-Work/Safety-and-Security/Privacy-Preservation/De-identification-Decision-Making-Framework>

Yours Sincerely,

Brad Ellis