Corporate policy

Anonymisation and Pseudonymisation Standard

Issue sheet

<table>
<thead>
<tr>
<th>Document reference</th>
<th>NHSBSADPN005a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document location</td>
<td>S:\BSA\IGM\Mng IG\Developing Policy and Strategy\Develop or Review DPA Policy\Current and Final</td>
</tr>
<tr>
<td>Title</td>
<td>NHS Business Services Authority Anonymisation and Pseudonymisation Standard</td>
</tr>
<tr>
<td>Author</td>
<td>Gordon Wanless</td>
</tr>
<tr>
<td>Issued to</td>
<td>All NHSBSA staff</td>
</tr>
<tr>
<td>Reason issued</td>
<td>For information / action</td>
</tr>
<tr>
<td>Last reviewed</td>
<td>15 November 2017</td>
</tr>
<tr>
<td>Date of Equality Assessment</td>
<td>N/A</td>
</tr>
<tr>
<td>Date of Fraud Review</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Revision details

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Amended by</th>
<th>Approved by</th>
<th>Details of amendments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial release</td>
<td>15.11.2017</td>
<td>CG</td>
<td>APF</td>
<td>Restructure wording from previous Data Protection and Freedom of Information Polices to an ISMS Standard</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. Anonymisation and Pseudonymisation Standard

1.1 This standard supports and acts as a measure of compliance with both the Freedom of Information Policy and the Data Protection and Confidentiality Policy.

1.2 This stand will comply with the legal and professional obligations set out in the:

ICO Anonymisation Code of Practice

1.3 The following Anonymisation standard will be applied:

- Remove all direct identifiers such as Name, Address, DOB, NHS Number National Insurance Number, Full postcode;
- Assess the risk of identification from indirect identifiers.
- Could the Individuals concerned identify themselves from that data
- Could the NHSBSA reverse the anonymisation process to identify the individuals based on information we hold?
- Consider this as a disclosure to anyone and assess the risks of re-identification
- What harm that could be caused to the individuals if the data were to be re-identified.
- Assess the current information that could be linked to this information to identify the individual
- What technology available that could be used to re-identify the information
- Consider who may have access to linking or identifying information and access if their motivation would make it likely they would expend considerable time and resource to re-identification.
- Assess how a re-identification test could be achieved and how often this should be repeated as new data sets may become available that could enable the anonymisation to be reversed.
- Be aware that unauthorised re-identification is a criminal offense.

- Reduce the risk of re-identification to that which is low/remote by Consider which of the following is most effective:
  - Aggregate totals less than five, use pseudonyms,
  - Rounding of totals
  - use age rather than date of birth,
  - only use the first part of a postcode,
  - provide range banding rather than exact data,
  - mask parts of the data.
Swap data attributes;
- change the sort sequence to prevent inference,
- provide only a sample of the population
  - Keep a summary of the techniques used prior to any disclosure

- consider the impact of the above on the usability of the anonymised data.
- How would you monitor if re-identification has occurred.

1.4 The following Pseudonymisation standard will be applied:

- Pseudonymisation techniques should be used wherever possible to complying with the principle of privacy by default and design. This ensures access to identifiable personal data is on a need to know basis.

- Pseudonymised data is still personal data that has been made more secure through pseudonymisation. Encryption can be considered to be a form of pseudonymisation. Consequently:
  - Only display the pseudonymised data items that are required.
  - Apply the same security measures as for personal data
  - Ensure a data sharing agreement and privacy impact assessment have been completed prior to sharing with a trusted third party

Carefully consider the feasibility of Pseudonymisation and consider factors such as:
- the impact on performance of the pseudonymisation process as this carries a processing overhead.
- Confirm with the software supplier if the information technology employed supports pseudonymised data

The Pseudonymisation algorithm must be:

- Consistent across the data sets it will be used in from an agreed starting point.
- Strong enough to prevent reverse engineering during the lifetime of the data
- Any change in the algorithm needs to be risk assessed and justified
- Destructive and non-reversible when re-identification is not required

The algorithm used must have strong management controls in place including:
- The Pseudonymisation keys must be stored in a database encrypted to NHS Encryption Standards.
- Access must be restricted to a minimal number of authorised staff
• There must be clear agreed rules for each data set as to when reversal of the pseudonymisation process is justified.
• Have a clear process as to how the algorithm would be changed if the keys were compromised.
• Where a key is shared outside the NHSBSA then there must be a clear agreement as to responsibilities to ensure this standard is adhered to.
• A destructive Pseudonymisation technique should be used when disclosing to third parties.