Data Flow Diagram illustrating the OpenPseudonymiser process to enable sharing and linkage of two discrete data sources.

Version 2.0.2

Project X

Data Source A

OpenPseudonymiser Software used by data controller:
Combine NHS number with project specific PKI encrypted salt to create Pseudo1
Optionally rounding dates & validating NHSNumber

OpenPseudonymiser Software used to combine Pseudo1 with a second salt known only to linking organisation 1 to create Pseudo2. Pseudo1 data not retained.

Researcher 1

Gives an anonymised subset of the data for Research Projects/3rd Parties

Organisation Boundary

Project Boundary

Organisation 1 is able to link the data from the two different sources A and B using the Pseudonymised key (the digest). Pseudo1 is a one way hash (SHA2-256) so linking organisations are unable to get back to the identifiable elements.

Project Y also needs data from Source B and can match this to data from Source C. However, organisation 2 is unable to use the data from Source A since it was Pseudonymised using a project specific key.

Data Linking Organisation 1
User Links sources A and B using Pseudo 2

Research Projects

Organisation 2

Data Source B

OpenPseudonymiser Software used by data controller:
Combine NHS number with project specific PKI encrypted salt to create Pseudo1
Optionally rounding dates & validating NHSNumber

OpenPseudonymiser Software used to combine Pseudo1 with a second salt known only to linking organisation 2 to create Pseudo2. Pseudo1 data not retained.

Data Linking Organisation 2
Links sources B and C using Pseudo 2

Researcher 2

Organisation Boundary

Project Y

Data Source C

OpenPseudonymiser Software used by data controller:
Combine NHS number with project specific PKI encrypted salt to create Pseudo1
Optionally rounding dates & validating NHSNumber

OpenPseudonymiser Software used to combine Pseudo1 with a second salt known only to linking organisation 2 to create Pseudo2. Pseudo1 data not retained.

Key:
- Identifiable data
- Pseudo1 data
- Pseudo2 data
- Anonymised data

Open Pseudonymiser by Julia Hippisley-Cox, University of Nottingham is licensed under a Creative Commons Attribution-ShareAlike 2.0 UK, England & Wales License. See www.openpseudonymiser.org for more.