Safe Disclosure of Information Guidance - Removing Personal Data from Information Requests & Datasets

This guidance supplements the Trust's Data Protection Confidentiality and Disclosure Policy; all SFH staff should be familiar with the provisions of the Policy and should not take this guidance in isolation.

It highlights key points for staff to be aware of when disclosing information derived from personal data either as part of a subject access, freedom of information or environmental information requests.

Redacting personal data from the information requested means that some information can be released without breaching Data Protection principles. Redaction can also be used to remove information which is out of the scope of a request.

Tools such as spreadsheets and documents allow data to be put into a more understandable form to educate, inform or convey information hidden within the data. Such summarisation methods can also be an important anonymisation technique and provide a privacy protecting mechanism to allow information to be published without the risk of identification and are commonly used to prepare information in response to a Freedom of Information (FOI) request, submission of mandatory datasets and other reporting purposes.

Example	Solution
Hidden data Data being disclosed in error can occur when data is not immediately visible on the screen but elsewhere within the file. This can be due to a range of design choices or the rendering of certain formatting styles. For example, when setting up a template a user might have chosen to 'hide' certain data by setting the font colour to be the same as the background (e.g. white on white or black on black in word or excel). Whilst hiding data in this manner prevents personal data being disclosed on a printed copy of the file, it will still remain within the source file leaving personal data at risk of accidental disclosure if the electronic version is distributed. Another example of where data might be hidden from obvious view is when it is placed in the fringes of a file where it is not expected to be found. As an example, Microsoft Excel 2007 and upwards support up to 16,384 columns and 1,048,576 rows of data. A user might place data outside of the normal visible area with the aim of hiding it from being displayed on a standard sized monitor. Hiding personal data in this manner is not good practice, is an ineffective way of removing or masking personal data for the purposes of redaction and therefore should be avoided.	**Best Practice** A more appropriate practice would be to control access to the file containing the personal data. An alternate view of the data can be created for those individuals who do not need access to the identifying elements. If you are trying to control printed versions of the file or the data must exist within the file a more appropriate practice would be to set a print area. Ensuring that worksheet have comprehensive and descriptive documentation which indicate the location and content of all data can also reduce the risk that data is 'lost' within a particular file.
Hidden rows and columns This is a common method of 'hiding' data within a spreadsheet. Selecting 'Unhide' from the appropriate submenu in a series of clicks will return the data to full view and once the column has been unhidden personal data contained within the column is accessible to anyone who has access to the file. Hidden rows and columns can be identified from the fact that the (column and row) headings do not flow in a consecutive order. In some versions of Microsoft	**Best Practice** Export the data into a simple text format such as CSV. CSV is a format where only the visible text is exported; and columns in spreadsheets are separated with a comma. The CSV file format does not support complex features such as hidden data fields, formulae, type formatting (e.g. bold and italic) or comment boxes. If a cell contained a formula then just the formula result would be exported. The exported data can be

Excel, hidden data can also be identified using the Document Inspector function which is accessed via the Check for Issues button.	manually validated by opening the CSV file and inspecting the data.
Pivot tables A pivot table can be used to summarise a large set of data. This can create an automatic summary of the underlying data. As with hidden data fields, despite the fact that the underlying data is not immediately visible on the screen it can still be accessed. A double - click on the pivot table can signal to the software to automatically extract the data used to calculate the clicked data and display this in a new worksheet. Even if the worksheet containing the original data is deleted from the workbook or if the pivot table is copied into a new workbook, the underlying data may be copied across with it, making the data accessible to other users.	**Best Practice** Export the worksheet containing the pivot table to CSV. The exported pivot table can be validated by opening the CSV file and manually inspecting the data. Copy the pivot table and paste only the values to a new workbook. This is sometimes referred to as a 'paste special' operation. The copied data can also be checked by double - clicking cells within the copied pivot table to ensure there is no link back to the source data. However, exporting the data to CSV would provide a greater assurance and provide greater compatibility with other software packages.
Charts As with pivot tables, charts can also contain an embedded copy of the source data; a further risk could arise when a chart is embedded into a document or presentation as the embedded chart could also contain a copy of the source data. If the chart is embedded within a Microsoft Word document then a copy of the underlying data is also copied across and embedded within the document. Simply double - clicking on the chart and selecting the data worksheet can reveal the underlying data.	**Best practice** - Copy the chart and paste as an image file (e.g. jpg or png) into the destination file. - Create the chart using a summarised version of the data (e.g. a pivot table with source data removed). Providing a copy of the anonymised chart source data would therefore not be a disclosure of personal data.
Functions Functions such as LOOKUP and VLOOKUP also create and store a cache of the source data that can be exposed through careful manipulation of the function.	**Best Practice** Export the VLOOKUP values to CSV. The exported values can be validated by opening the CSV file and manually inspecting the data.
Ineffective redaction The purpose of redaction is to irreversibly remove the exempt information from the redacted copy of the information. Care should be taken to protect against deleting data from the original file. The use of the black highlighter tool in Microsoft Word to add a black box around text marked for redaction may be sufficient if the document is disclosed as a printed copy (or printed, scanned and emailed), this would prevent disclosure of the desired information. However it is important to recognise that the information still exists underneath the black box in the original electronic file. If this file was retained this may not be clear to future readers and may lead to inadvertent disclosure if sent electronically. Using the 'Save as PDF' or 'Print as PDF' function is unlikely to provide effective redaction because the PDF file format can support formatting marks such as the highlighter. Also, copying the highlighted text and pasting to a text editor will reveal the content because the formatting will not be copied across.	**Best Practice** - If you are using a highlighter tool to mark text for someone else to redact, do not use a black highlighter. A different colour (e.g. yellow) will clearly indicate which text requires redaction yet also show that the original text remains. - For permanent redaction use specific redaction software.
Photography and video For the redaction of personal data from still images and video (e.g. CCTV footage,	**Best Practice** - Redact information within an individual image through the use of a simple

photos, X-ray), obscuring information from a single image can be a straightforward task as most operating systems include a simple image editing tool that can be used to blur or cover part of the image. It is however important to ensure that the redacted image is exported to a simple 'un – layered' format to ensure that the redactions are permanent. It is also worth considering whether information that you have not redacted may still result in someone being identified. It can be more complex to obscure information from video in part due to the larger volume of data and this is likely to require the use of specialist software to achieve this effectively.	image editing tool included with most operating systems or using a specialist redaction software tool. - X-ray images can be anonymised by radiology by contacting the radiology department.

Checklist

The following checklist highlights a number of things to consider when disclosing certain data types that may contain personal data. It is good practice to keep a record of all transformations or redactions you make and to retain the original records used.

File type	Considerations
Spreadsheet e.g. xls(x)	- Are you sure you know where all the data is? - Are there hidden columns? - Are there hidden rows? - Are there hidden work sheets? - Do pivot tables contain linked data? - Do charts contain linked data? - Are there formula included which link to external files? - Is there any meta-data that should be removed? - Is the file size larger than you might expect for the volume of data being disclosed?
Word processor e.g. doc(x)	- Are there any comments within the document that should be removed? - Does the document contain a version history? - Do pivot tables contain linked data? - Do charts contain linked data? - Is there any meta-data that should be removed? - Does the document title or filename contain any personal data (e.g. Letter to John Smith)? - Has a header or footer been automatically added to a print-out?
Presentation e.g. ppt(x)	- Are there any presenter notes which should be removed? - Do pivot tables contain linked data? - Do charts contain linked data? - Is there any meta-data that should be removed?
PDF	- Are there any comments which should be removed? - Are all redactions effectively applied? - Is there any meta-data that should be removed?
Email	- Is there data within any attachments that also needs to be redacted? - Is there any meta-data that should be removed?
Image and video e.g. jpg, avi	- Is there attached meta-data data (In particular, photographs taken with smart phones and tablets can include the GPS coordinates of where the image was taken)?

	- Is there personal data that needs to be obscured (e.g. faces of third party individuals?)

Meta - data

Electronic files rarely contain just the information entered by the author or just what is displayed on the screen. So called meta-data or 'data about data' is embedded within the file and can include information such as previous authors, changes made to previous versions, comments or annotations, location etc.

- Photographs taken with smartphones and tablets can contain the GPS coordinates of where the image was taken, time and date or information about the type of device used.
- Emails contain information about the sender and recipient as well as routing information about how the message was delivered. Releasing the original electronic version of an email may also disclose any attachments. If you need to disclose an email without meta-data you can disclose a printed version of the message (or print – as – PDF version).

Microsoft Office software suite including Word, Excel and PowerPoint can embed information such as the author, comments and version history automatically into files. The Document Inspector mentioned previously can highlight certain file properties in Microsoft Office files including comments, annotations and version history.

Additional guidance

https://ico.org.uk/media/for-organisations/documents/how-to-disclose-information-safely-removing-personal-data-from-information-requests-and-datasets/1432979/how-to-disclose-information-safely.pdf

Version	Changes	Author	Date
1.0	New Document		July 2016