Redacting Documents and Records

An Coimisiún um Chosaint Sonraí
Data Protection Commission

Introduction

Redaction is the process of concealing information while leaving intact the rest of the document or record containing it. This is commonly completed by 'blocking out' the material to be obscured but, as will be discussed below, other approaches may be more appropriate, convenient or effective.

There are many situations in which it is necessary or advisable to limit the contents of documents or files that are to be published or otherwise disclosed. A document with uncontroversial contents might also contain confidential or proprietary information that should not be disclosed to the intended recipient or the public generally. Rules governing official secrecy, legal privilege or professional confidentiality can restrict disclosure of information either generally or to specific classes of person.

In the context of data protection, the most common situation in which redaction must be considered is when responding to a subject access request (or 'SAR') under Article 15 of the GDPR. (For more information on SARs, see the DPC's web page and links here.) This note will therefore focus on redaction in that context.

Subject Access Requests

In general terms, Article 15 of the GDPR provides that an individual (a 'data subject') whose personal data is being processed by or on behalf of a data controller may request and obtain a copy of that personal data. If the data subject makes his or her request by electronic means (such as email, text message, or through a web portal), the controller must provide the copies "in a commonly used electronic form" unless the data subject requests otherwise.

Article 15(4) adds an important qualification: the data subject's right to a copy of his or her data "shall not adversely affect the rights and freedoms of others." When considering whether data should be released in line with the requirements under Article 15(4) the data controller must take care when preparing documents to be released to the requesting data subject to avoid improper disclosure of any other individual's personal data. It is common for documents and records – whether in electronic or paper form – to name or refer to more than one identifiable living person, even if only tangentially. It may therefore be necessary to redact documents and records to ensure

that a response to a SAR does not accidentally disclose another individual's personal data.

General Principles

The following general principles should be observed whether the document or record being redacted is in electronic or paper form.

1. **Work on a copy:** This should not need to be said but is often overlooked. By definition, redaction involves the alteration of records and the concealing data that is presumably needed in its original form for other purposes. Working with a copy will also make it easier to correct any mistakes you might make as you work.

2. **Know what you are dealing with:** It is essential to understand the purpose, structure and terminology used in a document or record before you can be confident that you will correctly identify all personal data in it. For many types of document, this will be straightforward and involve simply reading through it. However, a person who is not familiar with professional terminology may find it difficult to accurately identify where individuals are referred to in a complicated report. Similarly, a database may contain multiple data points associated with a particular person, but not all may be directly linked to the individual's name. A document may contain an appendix or an index that is overlooked. The person responsible for preparing a response to a SAR must be sufficiently familiar with the material being examined to recognise data identifying the data subject and, equally importantly, data identifying other individuals.

3. **Keep records:** A checklist of redactions to be performed will help to prevent mistakes or omissions. Each redacted version of a document will be made for a particular purpose, so it is important to keep a clear description of what was redacted (if convenient, a copy of the redacted document itself), as well as a record of why it was redacted and its date of creation and transmission. This can also help to resolve any disputes concerning the redactions.

Paper Documents

There are several techniques for redacting information in a copy of a hard-copy document:

- **Cutting out:** It may be possible to use scissors or scalpel-style knife to cut out parts of a document containing information that is not to be disclosed. This will not always be easy or convenient, but is undoubtedly effective at permanently removing information.

- **Cover with tape:** Stationers and providers of office supplies commonly provide adhesive tape that is coloured – typically black or white – and completely opaque. The tape comes in a range of widths, allowing text of different sizes to be completely covered.
- **Block out by hand:** Do not simply scribble over text with a ball-point pen, as the text underneath will often remain wholly or partially legible. White correction fluid or a black marking pen of suitable width can be used to cover text that is not intended to be disclosed. The document should be photocopied or scanned for checking as the redacted text may still be legible when the document is held up to light.

Once all information to be withheld is completely removed or covered, you should prepare a further photocopy or scan of the redacted document for provision to the data subject.

Electronic Documents and Records

As with hard-copy documents, the first task is to identify and locate all information to be redacted. A checklist will help in large or complicated redaction tasks.

Use Search functions with caution: Search functions are helpful for identifying specific words or names in documents, but they have important limitations. These functions may not scan all parts of a document, such as tables of contents or references. They cannot scan text information that is stored in image format (e.g. a photograph of a certificate or document.) They can find only what they are told to look for, so it is important to bear in mind that names can be misspelled and people can be referred to in different ways. A document could refer to one person in a variety of ways – as 'Seán', 'Mr Murphy', 'the Appellant's neighbour', 'my grandfather' or even just by initials. This highlights the importance of being familiar with a document before deciding whether and how to redact it.

Concealed information: Files produced by word-processing applications, email clients, spreadsheets, presentation programs or databases typically include considerably more information than appears on screen. It is important to be aware of the types of information that can be concealed and how to find them:

- **Hidden content:** It is common for applications such as spreadsheets, databases and word processing applications to include a facility for concealing part of the contents of files. Spreadsheets and databases may contain concealed columns, tables or worksheets, while document files may contain revision histories, editing mark-ups, comments and other types of information that may not be immediately apparent.

- **File properties** can be recorded and stored with a file by the application that created it or by the operating system. They can contain information such as the name of the person who created or edited the file, dates of creation and revision history.

- **Metadata:** Like file properties, some types of file can contained detailed information generated by the application that created them, by operating systems or other sources. For example, email metadata can show the time and date of creation, the sender's email and IP address, recipients' addresses and the route taken to them. Some mobile phones and cameras add metadata to files including the username of the person who took a photo, the date and time, the location – and even the lens and camera settings.

Consider working from print-outs: Using paper print-outs of electronic files or documents can sometimes make it easier to review and cross-check for material that should not be disclosed.

Once you have identified all content that should not be disclosed, you should be ready to proceed with redaction.

Remember to work on a copy: Again, this cannot be over-emphasised. Do not take risks with data that may be needed for other purposes.

Do not rely on highlighting and/or changing the colour of text to redact information: While doing so may render sections of the document unreadable on screen or when printed, the text content will not be affected, as it can be seen by copying and pasting it as plain text into a text editor. Even if the altered material is exported or virtually 'printed' to a different format such as a PDF document, the text may still be legible or recoverable.

Replace text to be redacted: The simplest way to redact text is to select and replace it with symbols or words such as **[REDACTED]** or **[XXXXXXX]**. This will prevent the replaced text being disclosed, while giving a clear indication of where the redaction was performed.

Exporting to plain text or CSV format: Plain text and comma-separated values (CSV) are the simplest formats in which electronic records are processed:

- **Plain text** contains only text and number values, with minimal formatting content such as tab stops and line breaks. Font effects such as bolding, italics, colours or different typefaces or sizes are also removed.

- **CSV** displays values from spreadsheets or databases and does not include formulas or cross-references. The contents of each row in a spreadsheet or table are given in plain text, with a comma separating each column's value.

Exporting a document to be redacted into one of these formats has the advantage of disclosing or stripping out hidden content such as concealed tables, cross-references or various types of meta-data. This can in turn allow for easier redaction and greater confidence that all relevant material has been reviewed.