Administrative Data Research Network

An ESRC Data Investment

'Personal information' in relation to de-identified data

This guidance note looks at the concept of 'personal information' in relation to de-identified data, drawing on recent court case judgements in this area.

In the 'abortion statistics' case (*Department of Health (DH) v. Information Commissioner (IC),* Case Ref EA/2008/0074)[1] a request to the DH for a full statistical breakdown on the number of late term abortions carried out on prescribed grounds was declined on the grounds that publication would allow for public identification. The DH based its case on the risk that the relevant patients and/or doctors might be identified by those sufficiently motivated to do so. The IC had concluded that numerical information was not personal data and therefore did not fall within the Freedom of Information Act 2000[2] exemptions. On appeal, the Information Tribunal concluded that the disputed information did constitute personal information in the hands of the DH for the purposes of the Data Protection Act; but it was not satisfied that the FOIA was effective, as there was insufficient risk of identification. By failing to disclose the information, the DH had breached the FOIA.

The DH then appealed to the High Court. In the judgement to the 'abortion statistics' appeal (*Department of Health v IC* [2011] EWHC 1430 (Admin)), **Cranston J** made several important points on the concept and extent of 'personal data', especially in relation to the grey area of statistics or other anonymous data which is derived from other data which is more clearly personal data. The judgment referred in detail to the leading House of Lords decision on this subject, the CSA case[3] which concerned a request for childhood leukaemia statistics in Scotland.

In the 'abortion statistics' appeal, Cranston J agreed with the Information Commissioner that the statistics were not personal data and held that that the Tribunal had been entitled to find that the proposed disclosure was of fully anonymised data, which would not identify or (in combination with other information) lead to the identification of any of the individuals concerned.

The Information Commissioner had argued that one route to that conclusion was to adopt the approach of Baroness Hale in the CSA case: anonymised statistics remain personal data and therefore subject to the protection of the DPA in the hands of the data controller (who possesses the underlying data from which individuals could be identified) but not in the hands of the general public (who do not)[4]. Cranston J, however, rejected that route in the 'abortion statistics' appeal, as he found that it was rather the reasoning of Lord Hope in CSA which had attracted the majority's support.

[1] See our guide to the Freedom of Information Act for researchers
[2] S.40 FOIA (personal information exemption) and s.44 FOIA (disclosure prohibited by an enactment)
[3] Common Services Agency v Scottish Information Commissioner [2008] UKHL 47, [2008] 1 WLR 1550
[4] This approach commended itself to the Upper Tribunal in the case of All Parliamentary Group on Extraordinary Rendition v Information Commissioner [2011] UKUT 153 AAC

Following Lord Hope's approach, the definition of personal data under s.1 DPA provides for two means of identification: either from the data itself (which does not apply in the case of anonymous statistics) or "*from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller*". Lord Hope's approach to such situations is to ask: does the 'other information' (if provided to a hypothetical member of the public) add **anything to the statistics** which would enable them to identify the underlying individuals? If the answer is no, the statistics are not personal data. If identification can be achieved from the 'other information' **in isolation** (rather than when added to the statistics) then the statistics themselves are truly anonymous, and are not personal data. The statistics in the 'abortion statistics appeal' failed Lord Hope's test, and were thus not personal data.

Mr Justice Cranston went on to find in the 'abortion statistics' appeal:

- that in the CSA case, Lord Hope had recognised that the CSA itself could always identify the children involved from the original information which it held, but
- it did not follow that statistics derived from that data, if disclosed in a fully anonymised form would still be personal data. This is a point on which, he found, all members of the House of Lords in the CSA case demonstrated 'a shared understanding'
- the House of Lords decision to refer the issue back to the SIC in that case was to determine whether barnardisation[5] could prevent the public from identifying any of the children involved (not to determine whether barnardisation would transform the data into data that would no longer be personal data in the hands of CSA, as that process was not capable of doing so, since the CSA would still hold data identifying the children concerned)
- the Tribunal in the 'abortion statistics' case had thus made an error of law. It should have found that disclosure of fully anonymised abortion statistics to the public did not involve a disclosure of personal data, even though the DH, the data controller, could still identify each women involved
- any other conclusion would be 'divorced from reality'. The DH's argument that the data remained personal data, even if disclosed to the public in fully anonymised form, because the data controller could identify those involved, would lead to the conclusion that to reveal that 100,000 women had an abortion in a particular year would be to disclose personal data about every one of them, which "is not a sensible result" and would seriously inhibit the publication of medical statistics.

The DH decided not to appeal the High Court ruling on the definition of personal data in the 'abortion statistics' appeal and it disclosed the disputed statistics on the number of late abortions carried out.

The current **ICO guidance** on this issue is that data is truly anonymised if, on the balance of probabilities, individuals cannot be identified by cross-referencing the 'anonymised' data with information or knowledge already available to the public[6]. Truly anonymised data is not 'personal data' within the meaning of the DPA and is therefore not caught by the Act – and such data can therefore be disclosed without reference to the DPA. The ICO does not accept that where a public authority holds information to

[5] a method of disclosure control that involves randomly adding or subtracting 1 to or from some cells in the table

[6] With a table of statistics it would still be necessary to consider whether, for example, a low cell count revealing that only one person in a small geographical area had a particular illness would allow people from that area to use their local knowledge to identify that person. Similarly with a report into the treatment of a patient it would be necessary to consider whether information gleaned from any news reports of the allegations would enable someone to identify who the references in the report related to.

identify living individuals from the anonymised data, this turns the anonymised data into personal data. The ICO's position is drawn from and supported by the House of Lords' judgment in the CSA case and the High Court decision in the 'abortion statistics' appeal. However, if individuals can be identified by cross-referencing the anonymised data with other information that was available, then the information is personal data. Whether it is possible to identify individuals from the anonymised data is a question of fact based on the circumstances of the specific case. Where the data controller that produces the anonymised data is in possession of other information that allows individuals to be identified, the data that relates to those individuals would be their personal data and the DPA would apply.

However, if the data controller sends the anonymised data to a third party which is not able to identify any individuals from the information in its possession then the DPA will not apply to the disclosure or the third party.

Please note that any advice provided by the Administrative Data Research Network is for information only and not legal advice. If you are unsure about any aspect of administrative data research, please consult your legal department or the relevant data holding organisation.

For further reference, see:

- Panopticon blog on information law
- Common Services Agency v Scottish Information Commissioner [2008] UKHL 47 Department of Health v (1) Information Commissioner (2) Pro Life Alliance EA/2008/0074
- Department of Health v Information Commissioner[2011] EWHC 1430 (Admin)