medConfidential Data Sharing Code of Practice Response

We expect the update to the Data Sharing Code of Practice to reflect a strong reading of the GDPR and Data Protection Act 2018, the central intent of both being to increase protections for individuals in an increasingly hostile data environment where malicious, capricious and outright unlawful exploitation of personal information is endemic.

We start this submission with a consideration of principles that should run through the new Code, and end with some more narrow specific points.

All data sharing must be just, safe, and transparent, and special category data sharing must be *consensual,* safe, and transparent.

Consensual (special category personal data)

The need for consent in the sharing of special category data may take different forms, not necessarily always individuals' prior and explicit informed consent – as for example in the case of 'implied consent' for medical treatment, or explicit statutory bases for sharing such as in public health emergencies, or statutorily-defined processes to enable data sharing on a case-by-case basis, that may respect active *dissent*.[1] It is acknowledged that in situations such as those involving fraud, the democratic process has required that non-consensual counter-fraud measures exist, for very obvious reasons.

medConfidential would strongly urge a discussion with the National Data Guardian for Health and Social Care about the definition and limits of sharing and re-use of clinical data, and the duty of confidentiality within the NHS and by NHS clinicians. While the details are likely beyond the scope of the Code, there must be a reference in the new data sharing Code to the NHS context, and an NDG-owned document which covers it.

It should not be possible for a naïve individual to read the ICO's Code of Practice and not realise that special category data may be subject to specific rules applying to the originating organisations. While this should not need stating, situations arising from such readings of previous Codes show it is.

[1] A particular example being the recommendation by the Confidentiality Advisory Group at HRA for 'Section 251 support' to be given for the sharing of NHS patients' data for research or non-research purposes.

Just (non-special category personal data)

A lynch mob may be deemed ‘community justice’ by those within it, but it is not justice in the broadly understood meaning of the term. Groups of professionals may apply rules to themselves that are obscure or counterintuitive to outsiders, though they make sense within a particular framework or practice. To avoid being perceived as another form of ‘parochial justice’, the principles of the Data Protection Act and GDPR must offer justice in real terms, addressing the real world.

All data sharing must be, and must be seen to be, just to a fully informed and *partial* observer – just both in the sense of “morally right, lawful and fair” to the individual concerned, and to the wider community. This would include being just for those who have objections to their personal data being shared for ‘secondary uses’.

Any Data Protection Impact Assessment (DPIA) must be available to all those who use that service, whose personal data may potentially be shared – i.e. DPIAs should be done in advance, most likely be held in a register, and must be kept up-to-date.

For data controllers and processors in the public sector, the Code should *require* that all data sharing agreements and processes are placed in a published Data Sharing Register – and that Register potentially also be included in a Register of Data Sharing Registers.[2] This is already required by the Code of Practice for data sharing under the Digital Economy Act 2017, and DCMS/GDS are currently finalising that Register for publication. *[It is due out any week now, but is not currently public. We strongly urge you to talk with* [redacted] *or* [redacted] *at DCMS for the current details; someone in ICO should know more than DCMS has shared with us...]*

The ICO should consider maintaining a list of Data Sharing Registers, irrespective of their legislative provenance. If the ICO does not, medConfidential expects to do so – along similar lines to our work to improve NHS Digital’s Data Release Register and consent choices, which can be found at TheySoldItAnyway.com. It would be far better were the regulator to evidence bad behaviour and rule-breaking, but we remain content to provide the necessary transparency while institutions and organisations continue breaking laws, legal agreements and public promises to service users.

One special case for registers and citizen-focussed transparency is data sharing around fraud – where it must be permissible for the standard text about data sharing for fraud purposes to be published, rather than it only being shown to those who are suspected of fraud. This is effectively the current practice, that should be copied into the wider context of registers.

---

[2] Such a Register of Registers should include those covered under the different Codes already in operation – such as in the Digital Economy Act 2017.

Safe

Data sharing is an inherently risky activity; two copies of data held or processed in different places are at greater risk of loss or compromise than a single copy in one place.

During the Royal Free / DeepMind debacle,[3] DeepMind argued that its copy of 1.6 million NHS patients' data was perfectly safe because 'we run gmail'. While that assertion might carry some weight with a naïve observer, it has turned out to be less than honest,[4] and the existence of a copy held by DeepMind inarguably presents a greater risk than *not having that copy there at all*. Even if risks are minimised, they still exist. This point must be explicitly stated in the Code, if only to avoid the 'tech-bros from Shoreditch' and others from believing they are infallible.

Where a data controller wishes to argue that data is not identifiable, additional steps for such scenarios must be covered in the Code. Without an update to the Anonymisation Code being available when the new Data Sharing Code is published, the latter must make reference to the GDPR's expansion of the definition of identifiers.

Whether selling snake oil or hiding behind well-defined marketing language, most (if not all) commercial 'de-identification' solutions engage only with such data items as they are designed to engage, which is not the full list of identifiers listed in GDPR. This being the case, if such technologies are to be used in data sharing, for all data fields in a dataset that are to be shared, the Code should require the DPIA to clearly state which data fields are subject to which de-identification or 'anonymisation' techniques, and which fields are not subject to any protection techniques at all. This is necessary for any reader of the DPIA to be able to understand what protection has been applied and where, and, more importantly, where protection has not been applied..

For <u>all</u> special category data, the use of a 'safe setting' should be required for any individual-level data shared where identifiers remain intact and the recipient organisation is not already processing data about that individual.

Transparent

The current Code was finalised in early 2011. GDPR notwithstanding, it is safe to say that the understanding and expectations of the public have evolved significantly since that time – as have commercial and government technology, and government's approach to publication of information.

The principles of ethics, data protection, GDPR, and simple good governance require that a data subject understands how data about them will be used. For public bodies, acting as monopoly providers of social safety net services, there is no real choice available to citizens; they either take the service, or they don't. When it comes to paying tax, there is similarly no

---

[3] https://medconfidential.org/whats-the-story/health-data-ai-and-google-deepmind/
[4] https://www.theverge.com/2018/7/2/17527972/gmail-app-developers-full-email-access

choice. This adds obligations to Government that are not present in competitive markets with genuine consumer choice.

Point 10 of the Government's Technology Code of Practice requires "making sure users of transactional services have access to data held about them – the service should clearly communicate how data will be used".[5] This was written to include Data Sharing Agreements; the obligation should be on every data controller that shares data to ensure the data subject can see how their data has been used.

For public bodies that offer a login mechanism – which are predominantly used for transactional services (as the Technology CoP reflects) – that must mean showing the data subject how their data has been shared, which includes linking to the relevant documents.

The publication of Data Sharing Agreements, and a record of bulk data sharing, allows members of the public to examine where their data has gone, and on what basis.

NHS Digital, for example, publishes a Data Release Register which covers both every new agreement and every dissemination of a new bulk data file.[6] Public Health England, by comparison, simply publishes new agreements, to varying levels of detail and accuracy. Given past failures to maintain the PHE code (which can go 8 months between "quarterly" updates) and irregularities in other publications as well, wherever a register is stated as being published "on a regular basis", if a scheduled update is significantly missed, the next update to the register should explain the reason for non-publication of the previous one.

The Data Sharing Code, especially for special category data, must require that all government and public sector data sharing agreements appear in a Register of Data Sharing Agreements, and a separate Register of Bulk Disseminations – since not all agreements may result in a dissemination record. (API access on a 'lookup' basis would not be a bulk dissemination, but would appear only in the records of individual data subjects.)

To ensure such registers do not have gaping loopholes, all data sharing agreements and contracts must require that any onward sharing is only with the permission and agreement (which must also appear in a register) of the original data controller. The processes for granting such permissions may be different to those of the initial sharing.

Fraud and malfeasance aside, **the new Code must require that all bulk sharing is listed in a Register (both agreements and issuances), and that individual acts of data sharing are shown to the data subject – via a digital service if possible.** For instances of fraud and malfeasance, the Register can cover the existing statutory obligations to publish documents, but not highlight any particular project as affecting any particular data subject.

Specific points relating to DPA 2018 and the current Code

[5] https://www.gov.uk/guidance/make-better-use-of-data

[6] A single project may have an agreement covering 3 years (one agreement), with monthly data updates per year disseminated as part of that agreement (36 disseminations).

Democratic engagement

Given the broad carve-out for "democratic engagement" in section 8(e) of the 2018 Act, the Code must be explicit in the obligations for satisfying that exclusion. We would expect that, if data is *shared* for the purpose of "democratic engagement", then such data sharing must have been stated and explained to the data subject at source. If a data processor or controller cannot justify that citizens knew their data would be used for that purpose, the carve-out cannot apply if other obligations under DPA18 / GDPR have not been demonstrably met.

Machine learning / AI

Given past breaches of Data Protection law, the Code should probably make clear that just because a data processor (or controller) wishes to use "AI", machine learning, deep learning, blockchain, or whichever new technology emerges next, the law still applies. We would oppose removal of the section explaining public bodies', and public bodies' suppliers', obligations with regard to Human Rights as well.

Sharing data for the purpose of processing via AI must follow the same rules as sharing for the purpose of processing by anyone or anything else.

Template forms and good practice

Given the likelihood that any examples and templates included in the Code will be used as the model for what many organisations will do in practice, we would hope considerably more thought and effort would be put into the design patterns provided than is evident in those included in the previous Code (pages 44 and 45 in particular).

We would hope also that stronger, more explicit links to good practice are made in any 'checklists' provided. (Experience suggests these are used by some as a form of 'tick box' compliance check.) While they are mentioned in the 'Governance' section of the current Code, for example, Privacy Impact Assessments are not explicitly mentioned or linked to at the relevant points in the checklists in section 15.