
TABLE ANALYSING CLAIMANT'S EXAMPLES OF SECONDARY DATA AND NON-PROTECTED CONTENT

[Defendants'] Introductory comments

This is the Defendants' response to the table that was provided to the Court by the Claimant during the course of argument on Day 1 of the hearing.

Where the Claimant cites examples from the various Codes of Practice, the Defendants accept in principle that such examples could (depending on the precise facts) constitute secondary data (or, as the case may be, non-'protected material' or non-'protected data') under the 2016 Act.

In relation to secondary data (as defined in s.137 of the Act), the question of whether information satisfies the statutory definitions of either (i) 'systems data' or (ii) 'identifying data' meeting the requirements of s. 137(5)(a)-(c)[1] will depend on the facts. In relation to a number of the categories identified by the Claimant, there will often be a factual question of whether identifying information can be logically separated from the remainder of the communication and if it were so separated, would not reveal anything of what might reasonably be considered to be the meaning (if any) of the communication. If such separation is not possible, the information would not constitute the kind of identifying data that falls within the definition of secondary data in s. 137(5) of the Act. Similar points apply to the definitions of 'protected material' (in the bulk equipment interference context: see s. 193(9) of the Act) and 'protected data' (in the BPD context: see s. 203 of the Act) save that, in relation to those definitions, it will also be necessary to consider whether the information is "*private information*", which "*includes information relating to a person's private or family life*", another fact-sensitive question.

It is also important to remember that information corresponding to the examples given in the Claimant's table could only ever be obtained under a bulk interception warrant or a bulk equipment interference warrant, or retained under a bulk personal dataset warrant, following approval **of the warrant** by a Judicial Commissioner **(save in urgent cases in relation to bulk equipment interference and bulk personal datasets)**, and could only subsequently be selected for examination to the extent that doing so is proportionate and necessary for an express 'operational purpose' under the warrant concerned **at the time of the selection for examination**.

Claimant's Response

There has been a degree of backtracking from the submissions made during the oral hearing. Further, the table below is inconsistent with the concessions made in the Defendants' note handed up on Friday 21 June 2019 during the hearing (the "21 June Note"), attached to this document. Detailed comments are provided below.

The Claimant has made some clarifications above in red text.

---

[1] Broadly speaking, this covers identifying data which can logically be separated from the remainder of the communication without revealing anything of what might reasonably be considered to be the meaning of the communication.

No	Example	Defendants' Response	Claimant's Response
1	"*contact 'mailto' addresses within a webpage*": Interception CoP ¶2.20 **[AB/3/24] (Identifying data)**	This is an example is taken from the Code. (Depending on the facts, a 'mailto' address on a website may alternatively be systems data.) If it is not systems data and if it is 'identifying data' and can be logically separated from the communication, it is likely to be secondary data.	**Agreed.**
2	"*location of a meeting in a calendar appointment*": Interception CoP ¶2.20 **[AB/3/24] (Identifying data)**	Taken from the Code: agreed, subject to the introductory comments above.	**Agreed.**
3	"*historical contacts from sources such as instant messenger applications or web forums*": Interception CoP ¶2.19 **[AB/3/24]**. Therefore: address books or contact lists containing e.g. telephone numbers, email addresses and user IDs **(Systems data)**	Taken from the Code: agreed, subject to the introductory comments above. Certain 'fields' in an electronic address book may have no role in enabling / facilitating (etc.) the functioning of the relevant communications service (cf. s. 263(4) of the Act, in which case information entered in such fields would not constitute systems data). It is not the case that all information in an address book or contact list would necessarily amount to systems data.	**It appears to be accepted in principle that information in address books or contact lists used by an email or messaging service to enable or facilitate the functioning of a service would constitute systems data. Address books and contacts are routinely used to enable or facilitate a service – for example, to store contact details that are then used for emails or calls, to convert real names into email addresses or usernames for the service, to autocomplete partial addresses, etc.** **Under s 263(4)(c), (d), and (e), any data that enables or facilitates any telecommunications service provided by means of a telecommunications system, a "relevant system" itself (that is, <u>any</u> system that holds communications or information: s 263(5)), or any service provided by means of a relevant system is "systems data".**
4	Full URL of websites visited beyond the 'first slash' **(Systems data)** (e.g. "*www.nhs.uk/<u>conditions/contraception</u>*")	Mostly agreed, subject to the introductory comments above and the following point. The example given here is a relatively simplistic example of a URL. A URL can contain a number of components as set out in paragraphs 2.60-2.67 of the communications data code of practice. This may include optional query parameters, identified by a '?', and	**"*Unstructured search data*" is not a term used in the Act or defined in any Code of Practice.** **If this term is intended to refer e.g. to a Google search term, it is content, in the ordinary and non-**

		fragments, identified by a '#', in the URL. These parameters contain data which helps to locate certain content but does not fit within a hierarchical path structure. Where there is unstructured search data as part of a URL string, this is considered to be and treated as content.	**statutory sense of the term. But the effect of s. 261(6)(b) IPA 2016 is that, if it is "systems data", it is deemed <u>not</u> to be "content" under the IPA, however revealing it may be of the meaning of the communication. "Systems data" is "secondary data", non-"protected material" and non-"protected data", even if it reveals meaning: ss 139, 193, 203.** **The Defendants appear to suggest that "optional query parameters" that help to locate certain content are not systems data. This is wrong: on the Defendants' own description, they enable or facilitate the provision of a telecommunications service provided by a telecommunications system (s 263(4)(c)) or a service provided by a relevant system (s 263(4)(e)). Similarly, it appears that the Defendants have in mind "unstructured search data" that contains and passes to a server the terms searched. That falls within the same definitions of "systems data".**
5	"*Photograph information - such as the time/date and location it was taken*": Interception CoP ¶2.20 **[AB/3/24]**; EI CoP ¶2.5 **[AB/3/22] (Identifying data)**	Agreed, subject to the introductory comments above.	**Agreed.**
6	"*information which describes the content held on a system, how that content got on a system and where it lives on a system*" **[GCHQ Compliance Guide: TB/5/153] (Systems data)**. Therefore: a file name, file type, creation date, modification date, and full directory structure for (including these details for all files on) a computer, mobile phone or other electronic device.	Agreed, subject to the introductory comments above.	**Agreed.**

7	"*a data field where the operation of the system depends on a valid value being entered in a data field for example, a passport number in a flight booking system that must be valid in order for the passenger to be able to fly*" **[GCHQ Compliance Guide TB/5/153; see also BPD CoP ¶4.40 [AB/3/23]] (Systems Data)**. Therefore: • Searches on Google or Facebook • All locations visited whilst using a navigation app • Everything searched for on a website, such as Amazon • Each swipe left or right on Grindr or Tinder (dating apps) • Login usernames and passwords • Encryption keys • All content entered into any website or apps that is processed as part of the service	The bullet-point examples given by the Claimant are not taken from the GCHQ Compliance Guide / BPD CoP. The treatment of these bullet-point examples under the Act is likely to be highly fact dependent. Data fields where the operation of the system depends on a structured valid value being entered, such as the Code examples, will constitute systems data, but it will be a question of fact whether the other examples given (or similar examples) meet that requirement, and where a new service is identified the data generated would need to be carefully considered and classified accordingly. Certain structured data may not enable or facilitate the functioning of any system and would therefore not be systems data. Unstructured data entered into a web form, such as searches on Google or Amazon, or free text or other unstructured boxes on webforms are considered to be and treated as content (although in certain cases some identifying data could be extracted from such content where the conditions set out above are met). Further, a password may well not be 'communicated' to a system (e.g. on Facebook) at all because it is automatically converted to a unique value in an Application before any communication takes place. Certain structured data may not enable or facilitate the functioning of any system.	**It is notable that the Defendants now appear unwilling to engage with the practical examples given.** **First, the Defendants conceded at §5 of the 21 June Note that "*text entered into Google would constitute systems data, since it enables the functioning of the search engine, and yet it might also (depending on the text entered) reveal the meaning of the communication*." This concession was correctly made. The text entered facilitates or enables the functioning of the search service, itself provided via a telecommunications system (the internet) and/or a relevant system (the computer or phone from which the search is made).** **The remainder of the examples all enable or facilitate the functioning of a telecommunications system, telecommunications service provided by means of a telecommunications system, or service provided by means of a "relevant system" and are thus systems data, regardless of how revealing they may be of the meaning of the communication. For example, a swipe left or right on an online dating site is how users control and use the service (ie enables the functioning of the service), and yet is highly revealing of sexual preference. The same applies to any content entered into a website that is processed as part of the service.** **The Defendants seek to avoid this conclusion by the suggestion that "*unstructured data*" is "*considered to be and treated as content*". The legal basis of this**

			suggestion is unclear and unexplained. It is incorrect as a result of the very wide definition of "systems data" in s 263(4)–(5) and the statutory deeming provision in s. 261(6)(b). It is also inconsistent with the 21 June Note. Such data is "systems data" and is therefore never content, regardless of how revealing it may be of meaning. **The Defendants' comment in relation to passwords is not understood and not explained by reference to the statutory definitions. If the Defendants mean to suggest it is not "systems data", that is incorrect: whether or not communicated, a password will necessarily enable or facilitate the operation of a "relevant system" (any system containing communications or information, such as a mobile phone or computer) or a service provided by means of a "relevant system".** **It is unclear what the Defendants' mean by their references to "structured data". However, this does not matter, as that term forms no part of the statutory definitions.**
8	Images of speakers on a video call **(Identifying data)** — this might be separated and used for facial recognition of the participants	The Defendants do not agree with this example. Images of speakers on a video call are not capable of being logically separated from the remainder of the communication: they would need to be extracted from the video call by way of the call being examined. The examination of such images, where acquired by bulk interception or bulk equipment interference, would therefore be subject to the safeguards applicable to the examination of content, including the British Islands safeguard. It is possible that some image data, for example in a BPD, could be identifying data but not in the facts given in this example.	**The Defendants appear to accept that images of the speakers themselves are identifying data, but then say that they would not be "secondary data" (etc) because they cannot logically be separated from the rest of the communication. They therefore accept this example in principle, but seem to say it is not (yet) technically possible.** **If automated facial recognition can logically separate out a unique identifier for each face present**

			on a call, the result is identifying data and would not reveal meaning, so they would not be subject to the British Islands safeguard. For example, if automated facial recognition processes can work out that Mr A spoke to Ms B on a Skype call, that fact is secondary data, not "content". There is no evidence that this is not technically possible. **Indeed, the same would apply to any photograph obtained by the agencies through bulk EI.**
9	Voices of callers on telephone calls **(Identifying data)** — these might be separated and used for voiceprint speaker identifications	The Defendants do not agree with this example. Voices of callers on telephone calls are not capable of being logically separated from the remainder of the communication: they would need to be extracted from the telephone call by way of the call being examined. The examination of such information, where acquired by bulk interception or bulk equipment interference, would therefore be subject to the safeguards applicable to the examination of content, including the British Islands safeguard.	**See example 8 above.** **The Defendants again appear to accept that a recording of a telephone call is identifying data, but again appear to suggest it is not technically possible logically to separate the identity of the caller.** **There is no evidence of this, and it seems very unlikely. If automated voiceprint speaker recognition can identify each speaker from the recording, the result is identifying data, which can be logically separated without revealing meaning, and is therefore not subject to the British Islands safeguard. For example, if automated voice recognition works out that Mr A spoke to Ms B on a mobile telephone call, that fact is secondary data, not content.**
10	"*a person's name, address, occupation, dietary preferences and country of birth*": BPD CoP ¶4.39 **[AB/3/23]**. Therefore: anything tending to identify race or ethnic background, citizenship, nationality, sex, gender, age, preferences for	Agreed that the specific items of personal data listed at §4.39 of the BPD CoP could, subject to the introductory comments above, amount to identifying data and – if logically separable etc. – would not be protected data within the BPD context. However, the suggestion that "***anything*** *tending to identify race or ethnic background [etc.]*" would constitute non-'protected data' is not accepted. Any 'identifying data' would have	**It is agreed that any identifying data would have to be capable of being logically separated without revealing meaning to constitute non-"protected data", etc. However, this is true also of the examples given in the CoP, which do tend to reveal ethnic origin, citizenship, etc.**

	literature/source of news/music, relatives, friends or associates **(Identifying data)**.	to be capable of being logically separated from the remainder of the communication without revealing meaning. If it is not so capable, it would only be protected data under s.203 if it is not 'private information' within the meaning of s. 203(1)(c) and 203(4) (and the examples given by the Claimant almost certainly would be private information).	**The Claimant believes that a "not" may be missing from the Defendants' final sentence.**
11	"*information that is not private information which may be attached to the email, such as a publicly disseminated electronic magazine, would <u>not be protected material</u>*": EI CoP ¶2.7 (box) (emphasis added). Therefore: electronic copies of the *Holy Quran*, *Playboy*, *Socialist Worker* and *The Daily Mail*, along with copies of anything else a person has read which is in the public domain, even if the fact of having read it may be highly revealing.	Agreed, subject to the introductory comments above (including as to the need for factual enquiry as to whether particular information is 'private', as the extract from the EI CoP makes clear).	**Agreed.**