11.7.16

[REDACTION]

NOTE: REDACTIONS ARE INDICATED [REDACTION] AND GISTS ARE IN BOLD, DOUBLE-UNDERLINED AND ITALICS

LOOSE MINUTE

To: **Bulk Data Review Panel Members**

From: ***MI5 Official***

Ext. No: [REDACTION]

Date: 19 September 2014

File Ref.: [REDACTION]

Copied To: ***Legal adviser and MI5 officials***

SUBJECT: Bulk Data Policy: Implementing a Flexible Review Period for BPD

1. This LM seeks Bulk Data Review Panel (BDRP) endorsement for a change in policy to move to flexible periodicity for reviewing Bulk Personal Data (BPD) (BDRP actions 33 & 34).

Current MI5 approach

2. The current policy and process for reviewing BPD is:

- All BPD datasets are reviewed on paper every 6 months, ahead of IS Commissioner visits.
- BDRP is required to review every dataset at least once every 2 years (on a rolling basis)
- BDRP also reviews any datasets referred to the panel by the business, ***the relevant team*** or Panel members.
- A thematic review has been implemented since April 2014 (with the theme for subsequent meetings to be agreed at each BDRP)

Case for change

3. The paper review places a heavy burden on the business and ***the relevant team***, currently requiring in the order of [REDACTION] datasets to be reviewed every six months. The business section has to prepare the case for retention, including gathering relevant evidence of use of changes to the dataset; ***the relevant team*** carries out a ***warrantry team***-style quality control of the review, forms a judgement on overall levels of intrusion and corporate risk, and makes a recommendation to ***the Senior MI5 Official***; ***the Senior MI5 Official*** signs-off the completed ***Form for Retention*** (or not, as the case may be).

4. Currently, the review periodicity take no account of the levels of intrusion or sensitivity of the dataset, and is arguably heavy handed in relation to datasets judged to be of low intrusion or low sensitivity (it is right that we review high intrusion and high sensitivity datasets frequently, but do we really need to review low sensitivity datasets (eg telephone directories) every 6 months?).

5. GCHQ already uses a flexible review period. The GCHQ review panel has discretion to assign a 6 month or 12 month review period for each dataset, determined primarily by levels of intrusion and sensitivity. It can also require datasets to be reviewed more frequently on the basis of lack of use. SIS is moving towards a flexible review period based on 12/24/36 months (or 18/32/46 months after first review); this has been approved by the SIS ***relevant team***, but not yet implemented. The IS Commissioner indicated to SIS during

NOTE: REDACTIONS ARE INDICATED [REDACTION] AND GISTS ARE IN BOLD, DOUBLE-UNDERLINED AND ITALICS

discussions around this change in policy that he is completely comfortable with flexible review periods provided there are clear criteria for how review periods are determined.

A new approach

6. It is proposed that review periods for MI5 BPD should determined by:

   i. ***Intrusion*** - the level of intrusion associated with the database

   ii. ***Corporate Risk*** – the level of corporate risk associated with the dataset

   iii. ***Usage*** – low levels of usage means that D/SIRO and BDRP can require datasets to be reviewed more frequently.

   iv. ***Theme*** – as determined in advance by the BDRP

7. The assessments of **intrusion** and **corporate risk** should be the primary determinants of the review period applied to a dataset. The periodicity proposed is:

High Intrusion and/or **High Corporate Risk**	-	**6 months**
Medium Intrusion and/or **Medium Corporate Risk**	-	**12 Months**
Low Intrusion and/or **Low Corporate Risk**	-	**2 years**

8. MI5's criteria for assessing intrusion in relation to BPD are at Annex A, and the criteria for assessing corporate risk are at Annex B. Where assessments of intrusion and corporate risk differ, the higher level of assessment will determine the review period (eg 'medium' intrusion and 'low' corporate risk would result in a review period of 12 months, not 2 years). If MI5 adopts this approach, the number of datasets falling into each review period based on current figures would be as follows:

	Risk	Intrusion	Total
6 months	[REDACTION] High	[REDACTION] High	[REDACTION] High
12 months	[REDACTION] Medium	[REDACTION] Medium	[REDACTION] Medium
2 years	[REDACTION] Low	-	[REDACTION] Low
			[REDACTION]

9. In relation to **usage**, datasets meeting the following criteria will also be referred to the BDRP for discussion:

- Any datasets with **no demonstrable usage during a review period**, or where there are issues or concerns around usage; a lack of usage may require that dataset to be **placed on 6 monthly review thereafter;**
- Any datasets **held by MI5 but not ingested** within 6 months to be submitted to the next Panel;
- Any datasets referred to the panel by the **business, *legal advisers*, *the relevant team*, or Panel members** during the process of authorisation, review, sharing or transfer;

NOTE: REDACTIONS ARE INDICATED [REDACTION] AND GISTS ARE IN BOLD, DOUBLE-UNDERLINED AND ITALICS

- Any dataset **approved for deletion** by the BDRP but **not deleted within 6 months**;

10. The BDRP will also review datasets on the basis of **themes** [REDACTION] where datasets falling under a chosen theme are reviewed together. This will enhance consistency and enable strategic issues to be explored by the Panel. Each meeting of the Panel will decide what theme will be addressed at the next panel so that business and compliance teams can prepare the appropriate paperwork.

11. The D/SIRO and BDRP may choose to **vary the review period** by exception (eg to require a dataset to be reviewed in six months rather than two years, if there is a lack of usage). The review period may be increased (eg from 1 year to 2), or reduced (eg 2 years to 6 months). Whenever a review period is varied, the reason must be recorded.

12. BDRPs will continue to be held every six months ahead of IS Commissioner visits. All datasets submitted for review on paper will be submitted to the BDRP, and to the IS Commissioner.

Conclusion

13. Adopting such an approach will significantly reduce the compliance overhead for the business, and the workload for compliance teams. This is important because some aspects of data management and oversight remain weak and require enhanced oversight. Some of the resource saved can be invested in improved governance for other aspects to data management, and some freed to deliver front line business activity. Adopting a flexible approach using the criteria defined above would represent a logical and proportionate risk-based approach to the management and oversight of bulk personal data. It would mean we review the most intrusive and most sensitive datasets more frequently than less intrusive and less sensitive datasets. It would enable DSIRO and BDRP to assign a different review period in the event of any concerns around usage arising. It would also bring MI5 in line with the flexible approaches adopted (or about to be adopted) by GCHQ and SIS.

14. The risks associated with adopting this approach are primarily related to oversight. Ministers and the IS Commissioner may feel that 'liberalising' MI5's oversight regime for bulk personal data is not appropriate in the current, post-SNOWDEN climate. However, IS Commissioner comments to SIS, and the fact that the other agencies have (or are about) to use a flexible review period, suggests that this risk is limited. We will need to agree this change in policy with the IS Commissioner before implementation.

15. If this change is endorsed, ***the relevant team*** would wish to implement the new policy as soon as possible after the November 2014 BDRP. However, there is a strong possibility of the three agencies being able to align their respective review periods, given there is little difference in philosophical approaches and all three agencies have indicated a willingness to flex their current arrangements in order to achieve aligned review periods and criteria (this is to be discussed at an SIA meeting on 30 September). If necessary, we propose delaying implementation for a short time if it would assist the three agencies to align.

Recommendation: BDRP members are asked to endorse this change in policy with immediate effect.

[REDACTION]

NOTE: REDACTIONS ARE INDICATED [REDACTION] AND GISTS ARE IN BOLD, DOUBLE-UNDERLINED AND ITALICS
Annex A

Assessing intrusion into privacy in relation to Bulk Personal Data

1. In the context of BPD, intrusion relates to the level of interference with the privacy of individuals (and, in particular, those individuals of no national security interest) caused by the acquisition, retention and use of bulk personal data. The legal framework is set out in ECHR 8(2) which states that 'there shall be no interference by a public authority with the exercise of this right [to privacy] except such as is in accordance with the law and is necessary in a democratic society in the interests of national security...'.

2. The interference with privacy arising from the holding and use of datasets falling within the definition of bulk personal data is assessed on the basis of:

a. the Security Service **merely holding** that data without any action being taken – the **collateral interference**; and

b. the Security Service **interrogating** that data – the **actual interference**

Collateral interference refers to the intrusion resulting from the Security Service holding the personal data prior to that information being interrogated or looked at in any way. This is particularly important in the case of BPD as most will datasets contain significant quantities of information about individuals who are of no intelligence interest. Due to the measures which the Service takes to only acquire those parts of a dataset which are really necessary, to hold bulk personal datasets securely, and to restrict access to bulk personal datasets, the collateral interference with privacy will usually be lower than the actual interference.

Actual interference refers to the intrusion which takes place when analysts or investigators perform a search on against the dataset, resulting in a 'hit' which then prompts them to look at the information on a specific individual and take action. The level of interference with privacy will rise at this point; the extent to which it will rise will depend upon the factors set out below.

Assessment Levels - Both collateral and actual interference with privacy are assessed at 3 levels: LOW, MEDIUM and HIGH, and each type of interference is assessed separately since they will not usually be the same. Collateral interference will almost always be lower than actual interference.

Criteria - When an assessment is made, be it of collateral or actual interference, the test should be the expectation of privacy that the average person would have in relation to the data contained within the dataset. In general, the higher the expectation of privacy, the higher will be the level of interference. Factors that need to be taken into account include the following:

- Has the data been provided willingly by the individual to another government department or agency?
- Has the data been provided by the individual to a non-governmental body (e.g. within the commercial sector)?
- Has the data been made publically available by the individual (e.g. published on-line)?
- Would the individual be aware that the data had been collected by the data provider?
- Would the individual be aware that the data provider might share their data with other bodies?
- Does the dataset contain sensitive personal information (as per the Data Protection Act criteria, eg. relating to finances or medical conditions), even in a non-detailed format ?
- Does the dataset consist of more than basic personal details (e.g. more than name, date of birth, address etc)?
- Does the dataset include details of travel movements?
- Is the information contained in the dataset anonymous?

[REDACTION]

NOTE: REDACTIONS ARE INDICATED [REDACTION] AND GISTS ARE IN BOLD, DOUBLE-UNDERLINED AND ITALICS

- Does the dataset include a disproportionate number of minors?
- What amount of data about individuals is contained within the dataset?

As well as consideration of the expectation of privacy, the assessment of intrusion process should always include a "common sense" test that takes into account all the characteristics of the dataset in the round.

Understanding the nature of the data acquired, coupled with the common-sense test outlined above will enable an assessment of whether the intrusion (or interference with privacy) is LOW, MEDIUM or HIGH.

Examples of Intrusion Assessments

Actual Intrusion Level	**LOW**	**MEDIUM**	**HIGH**
Dataset	OLYMPIC ACCREDITATION	[REDACTION] – Travel Data	[REDACTION]
Commentary	The dataset has been knowingly provided to the UK government for security reasons. There will be an expectation that this data would be shared with the Service, and that tracing would be conducted against it in the interest of national security. The intrusion is therefore low. However any intrusion is still minimised through limiting access and ensuring that all searches are specific and subject to audit.	Results of a query would identify the movements the individuals subject to the query. Due to limited intelligence it is common for queries to be conducted and return data on people of no intelligence interest. Intrusion is minimised through limiting access and ensuring that all searches are specific and subject to audit. Handling caveats are also imposed to limit risk.	[REDACTION]

Collateral Intrusion Level	**LOW**	**MEDIUM**	**HIGH**
Dataset	[REDACTION]	[REDACTION]	[REDACTION]
Commentary	[REDACTION]	[REDACTION]	[REDACTION]

[REDACTION]

NOTE: REDACTIONS ARE INDICATED [REDACTION] AND GISTS ARE IN BOLD, DOUBLE-UNDERLINED AND ITALICS
Annex B

Assessing Corporate Risk in relation to Bulk Personal Data

Corporate Risk refers to the potential for political embarrassment and/or damage to the Service's reputation and that of its partners and data providers were it to become public knowledge that the Service holds certain datasets in bulk.

It is ***the relevant team's*** responsibility to assess the level of that risk, be it LOW, MEDIUM or HIGH. In order to assess Corporate Risk, ***the relevant team*** will take into account:

- the general expectation of privacy in any given dataset,
- the assessed levels of collateral and actual intrusion,
- the possible media and public response were it to become known that we held certain datasets in bulk,
- the impact (adverse) publicity would have on the reputation of the data providers and our relationship with them,
- the impact (adverse) publicity would have on our partners and our relationship with them; and
- the resulting reputational and operational damage to the Service.

Were it to become widely known that the Service held this data the media response would most likely be unfavourable and probably inaccurate.

Example of Corporate Risk Assessments

	LOW	MEDIUM	HIGH
Dataset	[REDACTION] passport data	[REDACTION]	[REDACTION]
Corporate Risk Explanation	The corporate risk is LOW as the public has a reasonable expectation that the Service holds travel-related data and may hold it in bulk. Moreover, passport forms state that details may be passed to other departments and agencies when it is in the 'public interest' to do so.	[REDACTION]	[REDACTION]