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

SECURITYSERVICE
MI5

Policy for Bulk Data Acquisition, Sharing, Retention & Deletion

Policy Lead: [REDACTION]

Business sponsor: [REDACTION]

Policy Issue Date: October 2010

Review Date: October 2011

Policy Aim: To explain the Bulk Data Lifecycle from pre-Acquisition to Deletion and the responsibilities of users and managers

Summary

The Service obtains privileged access to large amounts of personal non-targeted data, known as 'Bulk Data', some of which may be regarded as sensitive e.g. financial data. To ensure that Service acquisition and exploitation of this data is necessary and proportionate, there are internal control and review mechanisms that users and managers must follow. These are explained in this policy document.

This policy is due for review in October 2011 but may be amended earlier to take into account any new legislation or oversight mechanisms affecting our work with Bulk Data.

Document date: 19 October 2010

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

What is Bulk Data?

The current working definition of non-targeted bulk data as agreed by Agency lawyers in early April 2010 specifies bulk data as:

> datasets acquired under Section 2(2)(a) of the Security Service Act 1989 or Sections 2(2)(a) or 4(2)(a) of the Intelligence Services Act 1994 which contain data about a wide range of individuals, including [non-adverse] data about individuals who are not of direct intelligence interest.'

In non-legal terms, Bulk Data may be described as a dataset or database containing data about a large range of individuals, including individuals who are of no intelligence interest to the Service, and which is too large to be susceptible to manual processing. Bulk data generally includes datasets that are not already covered by an oversight mechanism provided for by existing legislation. Such datasets might include ***travel data and population data (e.g. Passport records and records from government departments.)*** There may be a level of public assumption that the Service will not hold such data in bulk.

This means that datasets acquired under RIPA are not considered to be Bulk Data since there is a well-established oversight mechanism covering their acquisition. Similarly data acquired from our SIA partners' databases (e.g. GCHQ and SIS target records or datasets they have acquired through RIPA) is not regarded as Bulk Data. Finally, commercially and openly available datasets (e.g GB Info, Companies House etc) and data generated by corporate systems is not bulk data. The handling of these datasets is not covered by this policy.

Why does the Service need Bulk Data?

The Service faces increasingly sophisticated and diverse challenges in its work to counter terrorism and other threats to the UK's national security. We have many more potentially serious but sketchy leads and targets that are more mobile and that have access to new, capability-boosting technologies. To tackle this effectively, we need to be able quickly to ***identify and locate*** individual ***subjects of interest*** and target networks, including ones in which the majority of members may be unknown and on which there is little intelligence. To ensure that significant threats are identified quickly we need to be able to move cases quickly from initial lead to closure. **Part of this capability involves the acquisition of large sets of personal data.**

We need to hold very large sets of bulk data in order to make the right connections between disparate pieces of information to ***identify and locate subjects of interest***. Once a ***subject of interest*** has been identified, we can use bulk data to identify their linkages and associations, locate and track them [REDACTION]. A successful identification then becomes the jumping-off point for further enquiries and use of more intrusive resources, ensuring that such enquiries are as focussed as possible. By integrating Bulk Data [REDACTION] with information about individual ***subjects of interest*** from other sources of intelligence (liaison relationships, agent reporting, intercept, eavesdropping, surveillance), and from 'fusing' different data-sets in order to identify common links, we can better understand target networks, locations and behaviours, enabling a greater depth and breadth of target coverage. The fragmentary nature of many intelligence leads and the magnitude of the threat all mean that there is currently no effective method of resolving identities in a **timely** fashion without using Bulk Data.

Document date: 19 October 2010

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

Our requirement for bulk data is for us to hold the broadest and most accurate possible set of minimum identifying particulars for adult residents and visitors to the UK, so that we can resolve leads down to the minimum number of candidates promptly. This should allow us to resolve identities with high reliability, at pace, and with minimum intrusion. There is no single existing database out there that meets this requirement, so the Service fuses together a number of datasets (primarily from government and the other Agencies) to make sure we have the necessary coverage. Each of the sets we hold has to contribute to this central goal in order to justify its acquisition and retention.

There will always be a measure of duplicate data in our systems across a number of sets, but each set will provide a link to at least one unique identifying particular. Therefore, when drawing up the case for a new acquisition and subsequent retention, we need to have identified:

- a key gap in our ability to cover a certain type of scenario or set of circumstances or similar;
- a dataset which can fill that gap, possibly uniquely;
- the case for the acquisition being proportionate.

How is Bulk Data Acquired?

The acquisition of Bulk Data is tightly controlled and is subject to internal scrutiny by ***senior MI5 officials***, operating under the authority of ***senior MI5 officials*** and DDG, and by the Service's legal and ethics advisors. There are standard processes intelligence sections **must** follow in order to acquire Bulk Data. All acquisition requests must be supported by a business case approved by a ***senior MI5 official*** from the intelligence section and must be supported by a data analysis and exploitation expert, known as 'data sponsors'. ***A senior MI5 official*** will authorise the acquisition once he is satisfied that it is both necessary and proportionate for the Service to hold and use the dataset in pursuit of the Service's statutory function to protect national security. As part of this, ***the senior MI5 official*** must be satisfied that any resulting interference with individuals' right to privacy, as enshrined in Article 8(1) European Convention on Human Rights (ECHR), is justifiable under Article 8(2) for the purpose of protecting national security.

Bulk Data is acquired by ***the relevant team*** on behalf of the Service. ***The relevant team*** will only acquire data once a business case, supported by data sponsors, has been approved by an intelligence section SMG and has been signed off by ***a senior MI5 official***. All acquisition requests must be submitted on ***the relevant form***. Within the data providing organisation, department or agency, the Service will seek the agreement at Board or senior management level (e.g. Senior Civil Service, ACPO rank). Some departments may seek the approval of their Minister to pass data to the Service.

When an investigative section identifies a bulk dataset it requires in order to progress work or where it is offered access to a bulk dataset by contacts or by a third party, the investigative section must discuss their requirements and potential access to information with data sponsors. ***The relevant team*** must also be consulted since they will eventually acquire the information from the donor organisation. Investigative sections and data sponsors must be able to justify the acquisition and subsequent retention and/or updates of a dataset as necessary and proportionate by weighing up, on the one hand, the business gains of having the information against, on the other hand, any resultant interference with privacy, also referred to as 'intrusion'. If in doubt, ***Legal advisers*** advice should be sought when evaluating whether the acquisition of a dataset is necessary and proportionate.

Document date: 19 October 2010

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

Guidance on how to assess intrusion levels is available at Annex A [REDACTION]

In some cases it may be necessary for ***the relevant team*** to approach the data provider to examine whether any unnecessary/extraneous parts of the dataset can be removed prior to acquisition. Such extraneous data might include large numbers of minors, details of earnings or medical information. Once these initial discussions have taken place, the investigative section must complete and submit ***the relevant form.***

The relevant team will also make a further assessment on the necessity and proportionality of acquiring the data, and will take into account the extent of political, corporate, or reputational embarrassment and/or damage that compromise of the data would cause, including to the bulk data supplier.

Timescales for acquisition and use vary.

[REDACTION]

Bulk Datasets may not be used without appropriate authorisation from senior MI5 officials. **[REDACTION] Failure to follow the processes described in this policy may result in disciplinary action being taken.**

Physical Collection & Storage of Bulk Data

In order to ensure the security and integrity of the datasets that the Service relies upon for its enhanced analytical capabilities and to reassure data providers that their data will be handled securely, it is essential that the necessary physical controls are in place to mitigate against unauthorised access to, or loss of, this information during transportation to and subsequent storage in Thames House.

[REDACTION]

Permitted Use

Access to Bulk Data is limited to those with a business need, e.g. investigators, operational staff, data analysts and system administrators. Before access is granted all users must read and sign the relevant Code of Practice. They must also attend a compulsory training course that lasts two days (full time or integrated in other courses).

[REDACTION]

All users must ensure that Bulk Data searches are necessary and proportionate to enable the Service to carry out its work and searches must be structured and targeted in a way that is most likely to select information relevant to the enquiry.

[REDACTION]

The Review Process

Document date: 19 October 2010

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

The Bulk Data Review Panel (BDR Panel) meets every 6 months to review all bulk datasets. The aim of the Panel is to ensure that Bulk Data has been properly acquired and its retention remains necessary and proportionate to enable the Service to carry out its statutory duty to protect national security for the purposes of s.2(2)(a) Security Service Act 1989. Panel members must satisfy themselves that the level of intrusion is justifiable under Article 8(2) of the ECHR and is in line with the requirements of the Data Protection Act 1998.

The Bulk Data Review Panel is chaired by a senior MI5 official and operates under the authority of DDG (senior official responsible for the Service's operational capacity).

The Panel weighs up each dataset's usage over the 6 month period against necessity, proportionality, level of intrusion and the potential corporate, legal, reputational and political risk. The Panel also considers the frequency of acquisition and updates and whether such information could be acquired elsewhere by, for example, commercial means. The Panel decides whether to retain the dataset for a further 6 months or to delete it. In particularly sensitive cases, the Panel may recommend an earlier review. ***When the panel cannot agree on retention or deletion, the case will be referred to one of the Directors and ultimately the Director General for a decision.***

Sharing Bulk Data with SIA Partners

Where GCHQ or SIS identify that a dataset owned by the Service would enable them to discharge their statutory functions,, they should discuss their requirements with data sponsors [REDACTION. The data sponsor will complete ***the relevant form***, which outlines the business case submitted by the requesting Agency, detailing the actual data requested, the necessity and proportionality of holding that data and data handling proposals. This must be approved by ***senior MI5 officials.***

Acquiring Bulk Data from SIA Partners

When an investigative section becomes aware of a bulk dataset held by SIS or GCHQ that might assist the Service in progressing our work, the investigative section must discuss their requirements and potential access to information with data sponsors. ***Formal applications for acquisition of bulk data from SIA partners must be submitted on the appropriate form and authorised before being sent to the SIA partner. Once the relevant SIA partner is satisfied that the business case is justified and that sharing the data will not breach any security considerations that they may have, arrangements will be made to share the data.***

Informing Data Providers about Sharing Bulk Data with SIA Partners

Beyond assuring data providers that their information will be handled securely and used to protect UK national security, the Service will NOT routinely volunteer any special conditions/limitations regarding sharing. The Service is effectively acquiring data in the UK on behalf of UK Intelligence.

[REDACTION]

Deletion of Data

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

Where the BDR Panel decides that a dataset should be deleted, the appropriate team will delete it within a reasonable timescale as specified by the Panel. The dataset will be deleted in its entirety. Deleted Bulk Data will not be archived.

[REDACTION]

The Role of the Data Sponsor

Data sponsors are business representatives with a thorough knowledge of all Service data holdings. They provide advice to investigative sections on current data holdings that may meet a business requirement, assist with the formulation of a business case to acquire new bulk datasets and they represent business sections at the 6-monthly bulk data review meetings.

It is important that the data sponsor is involved at the earliest opportunity when considering acquisition, retention or deletion of bulk data as they are required to endorse business decisions, and ensure the correct procedures are followed.

[REDACTION]

External Oversight

All the Agencies have accepted that some form of external oversight is desirable. What shape, and how this should be accomplished, is not yet agreed. This is complicated by the wider movement towards oversight reform which may result in the replacement of the Intelligence Services Commissioner with some form of Inspector General and is likely to require some form of legislation. Given the uncertainty on timings there is broad consensus across the Agencies, Home Office and Cabinet Office that an interim solution involving the existing arrangements is needed.

Currently the Intelligence Services Commissioner is responsible for oversight of Agency use of bulk data. The purpose of this oversight is to review and test our judgements on the necessity and proportionality of acquiring and using bulk datasets and to ensure that our policies and procedures for the control of, and access to, bulk datasets in our keeping are both sound and strictly observed. Although we brief the Home Secretary on the Service's use of these techniques; independent oversight will provide a third party view of the arrangements that have been agreed and an endorsement of our judgements that will be valuable both to the Service and to ministers.

Our aim is for the Commissioner to be able to report positively to the Prime Minister and Home Secretary on our arrangements for working with and handling of Bulk Data. This will bolster Ministerial confidence in the integrity of our operational work. All papers requested by the Commissioner must be made available to him.

Annex A

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

Guidance for Investigators & Managers

[REDACTION]

What is intrusion?

In this context, 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 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...' .

Intrusion levels must be considered as part of the assessment of necessity and proportionality of acquiring and/or retaining a dataset.

How do I assess intrusion?

For datasets falling within the definition of Bulk Data, you will need to assess the interference with privacy resulting from:

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

b. the Service interrogating that data – the actual interference

Collateral interference with privacy refers to the intrusion resulting from the Service holding information on an individual who is of no intelligence interest, prior to that information being interrogated or looked at in any way. This sort of interference with privacy is the 'price' we pay for being able to use bulk datasets to find those who are of intelligence interest. Due to the measures which the Service takes to only acquire those parts of a dataset which are really necessary, hold bulk datasets securely, and restrict access to bulk datasets, the collateral interference with privacy will almost always be very low and, in any event, lower than the actual interference (as to which see below).

Actual interference with privacy refers to the intrusion which takes place when analysts or investigators perform a search on [REDACTION], 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.

Both collateral and actual interference with privacy are assessed at 3 levels: LOW, MEDIUM and HIGH and each type of interference should be assessed separately since they will not usually be the same.

Document date: 19 October 2010

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

As a general guideline, you should bear in mind that collateral interference will almost always be LOW and that actual interference will almost always be higher than collateral interference.

When making your assessment, be it of collateral or actual interference, you should assess the expectation of privacy that the average person would have in the data within the dataset. As a general guideline, the higher the expectation of privacy, the higher the level of interference with privacy. When assessing expectation of privacy, there are a number of factors you will need to take into account which will require you to understand the nature of the data you are acquiring, as follows:

- 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)?
- 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 (e.g. relating to finances or medical conditions), albeit 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 to/from the UK?
- is the information contained in the dataset anonymous?
- does the dataset include a disproportionate number of minors?
- What amount of data about individuals in 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. Check that you have considered all the relevant factors and that you have given appropriate weight to those factors. Ask yourself whether the intrusion level you have arrived at sounds reasonable.

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

What is Corporate Risk?

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

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. ***The relevant team has the 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.

Using these criteria, the fact that the Service holds bulk financial, albeit anonymised, data is assessed to be a HIGH corporate risk since there is no public expectation that the Service will hold or have access to this data in bulk. ***Were it to become widely known that the Service held this data the media response would most likely be unfavourable and probably inaccurate.*** On the other hand, we have assessed the holding of [REDACTION] passport data to be a LOW Corporate Risk since 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.

Is there a review process?

Yes. The Bulk Data Review Panel (BDR Panel) meets every 6 months to review all acquisitions. New acquisitions will be reviewed at the first BDR meeting following acquisition. Sections will be notified of the date of the BDR meeting around 8 weeks in advance. Data Sponsors must then complete ***the relevant form*** which covers 1) necessity, 2) proportionality, 3) level of intrusion in relation to the Article 8 right to privacy and 4) corporate/ legal/ reputational/ political risk. You may be asked to provide the following information:

- Commentary on how often a dataset is searched, its usefulness and impact on specific operations

- Commentary on whether the dataset provides a unique picture or aggregates existing information in the Service's records

- The number of rows in a dataset, i.e. how many individuals can be identified, and the proportion that are adverse/non-adverse

- The type and depth of the data about individuals which the dataset contains

Document date: 19 October 2010

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

- Commentary, if applicable, on how difficult it would be to acquire again [REDACTION] or whether it would be possible to get the same data in a different way.

- An indicator of the intrusiveness of the data, against a (reasonable) expectation of privacy. We use a high, medium, low classification. For guidance on assessing intrusiveness see 'How do I assess intrusion?' above.

- If applicable, the steps that have been taken in order to minimise intrusion in relation to that dataset.

The relevant team will also make an assessment of the potential corporate risk in the Service holding the particular dataset.

What is the Bulk Data Review Panel?

The aim of the Panel is to ensure that the bulk data has been properly acquired and its retention remains necessary and proportionate to enable the Service to carry out its statutory duty to protect national security for the purposes of s.2(2)(a) Security Service Act 1989. Panel members must satisfy themselves that the level of intrusion is justifiable under Article 8(2) of the ECHR and is in line with the requirements of the Data Protection Act 1998.

A ***senior MI5 official*** chairs the Bulk Data Review Panel and is advised by ***other senior MI5 officials.*** The meeting is attended by bulk data users/sponsoring sections and staff who acquire the data and who ingest it into our central database. ***A senior MI5 official*** attends as an observer. The Bulk Data Review Panel operates under the authority of DDG (senior officer responsible for the Service's operational capability) and ***another senior MI5 official.*** [REDACTION]

The Panel weighs up each dataset's usage over the 6 month period against necessity, proportionality, level of intrusion and potential corporate, legal, reputational and political risk. The Panel also considers the frequency of acquisition/update and whether such information could be acquired elsewhere by, for example, commercial means. The Panel then decides whether to retain the dataset for a further 6 months. In particularly sensitive cases, the Panel may recommend an earlier review. Where the Panel cannot agree on retention or deletion, the case will be referred to ***a senior MI5 official*** and DDG and, ultimately DG for a decision.

What If I want to delete the dataset outside the BDR Meeting Schedule?

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

You should submit a Loose Minute to the appropriate MI5 official. The Loose Minute should include a description of the data, its value/use since the last BDR meeting and the reason for deletion. Legal Advisers will be consulted to ensure that the deletion does not breach any disclosure obligations in the context of legal proceedings. Investigative sections will also be consulted. Assuming there is no requirement to retain then data, it will be authorised by a senior MI5 official and the data will then be deleted. The BDR Panel will be informed of the deletion outside of committee.

How do I acquire Bulk Data from SIS or GCHQ?

When you become aware of a bulk dataset held by SIS or GCHQ that might assist us in furthering investigations, you should first discuss the requirement and potential business case with your data sponsor. ***Formal applications for acquisition of bulk data from SIA partners must be submitted on the appropriate form and authorised before being sent to the SIA partner. Once the relevant SIA partner is satisfied that the business case is justified and that sharing the data will not breach any security considerations that they may have, arrangements will be made to share the data.***

[REDACTION]

What if SIS or GCHQ want to acquire our Bulk Datasets?

Where GCHQ or SIS identify that a dataset owned by the Service would enable them to discharge their statutory functions, they should discuss their requirements with data sponsors. The data sponsor will complete an internal ***relevant form***, which outlines the business case submitted by the requesting Agency, detailing the actual data requested, the necessity and proportionality of holding that data and data handling proposals. This must be approved by ***senior MI5 officials.***

Document date: 19 October 2010