SECURITYSERVICE
MI5

Bulk Personal Data Guidance

Introduction

This guidance sets out the processes to be followed for the handling of Bulk Personal Data (BPD) throughout its lifecycle within MI5. It should be read in conjunction with the MI5 BPD Lifecycle policy.

At all stages of the lifecycle, the following is to be assessed:

- Business justification (necessity and proportionality)
- Intrusion into privacy (for guidance on assessing intrusion see Annex A)
- Corporate risk (for guidance on assessing corporate risk see Annex B)

Ethical considerations

Any person involved in this process, or using the data, may consult with the ***a senior MI5 official in the ethics team*** should they have any concerns regarding MI5's acquisition or use of data. Consultation may take place at any stage of the process and will be treated in strict confidence.

Bulk Personal Data Lifecycle

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

Definition of Data Categories

BPD Categories

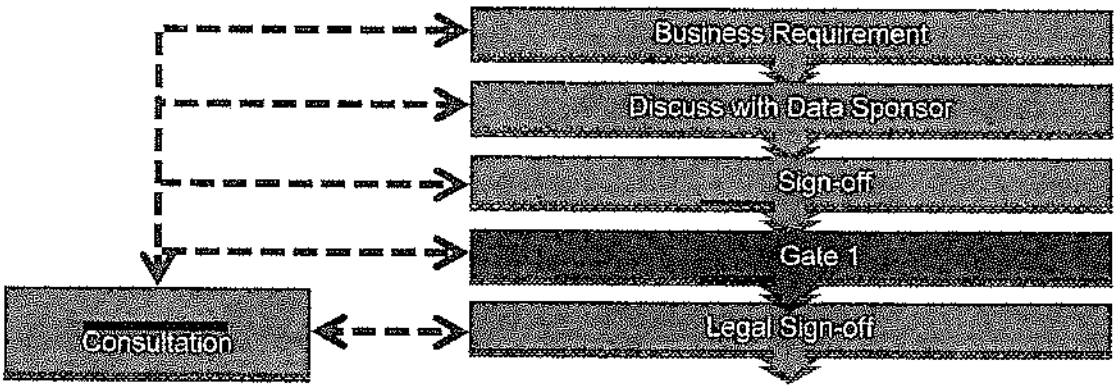
MI5 currently categorises its BPD holdings into the following:

Category	Description
LEA/Intelligence	***These datasets primarily contain operationally focussed information from law enforcement or other intelligence agencies.***
Travel	***These datasets contain information which enable the identification of individuals' travel activity.***
Communications	***These datasets allow the identification of individuals where the basis of information held is primarily related to communications data e.g. a telephone directory.***
Finance	***These datasets allow the identification of finance related activity of individuals.***
Population	***These datasets provide population data or other information which could be used to help identify individuals e.g. passport details.***
Commercial	***These datasets provide details of corporations/individuals involved in commercial activities.***

These BPD categories have been aligned with GCHQ. [REDACTION]

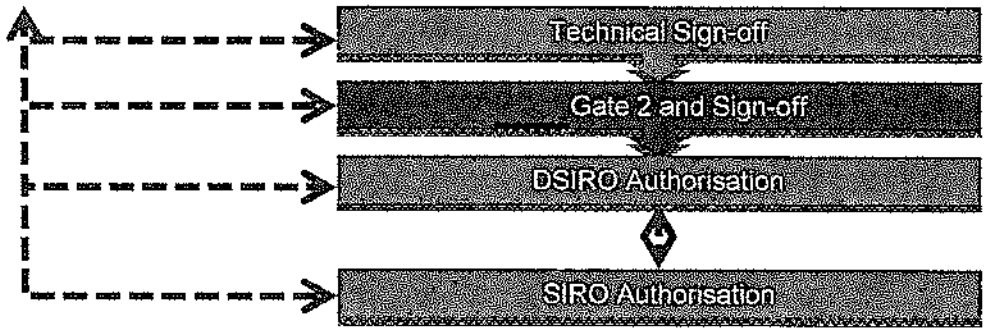
Authorisation

Summary of the Process



[REDACTION]

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



The authorisation to acquire BPD is managed via ***the relevant form***. All ***relevant forms*** must be supported by an ***senior MI5 official*** approved business case. Business cases are completed and endorsed initially by the relevant Data Sponsors prior to the ***senior MI5 official*** [REDACTION] listed below:

Business area	Data Sponsoring ***senior MI5 official***	Data Sponsors
[REDACTION]	[REDACTION]	[REDACTION]

How to complete a ***relevant form***

A ***relevant form*** must be used in any situation where it is the intention to acquire BPD. In essence, this is in any situation where our intention is to 'collect the haystack' rather than 'collecting needles'. ***Information Management*** and/or ***a legal adviser*** should be consulted in the event of uncertainty.

The detailed process to be followed is:

1. In conjunction with the relevant Data Sponsor, you should draft Sections 1 and 2 (Data Description and Business Justification & Privacy Assessment) of ***the relevant form***. This should include an explanation of why you want the data, its intended use, how it will improve/add to an investigation etc.

2. The business justification also requires the requesting section(s) and the data sponsor 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, and on the other hand, any resultant interference with privacy, also referred to as 'intrusion'. In the context of BPD two aspects of intrusion must be considered;

   a. MI5 merely holding the data without any action being taken, particularly as the majority of individuals are not of direct intelligence or security interest – the **collateral intrusion**; and

   b. MI5 interrogating the data – the **actual intrusion** (Guidance on how to assess intrusion levels is available at Annex A.) [REDACTION]

If in doubt, you should consult ***a legal adviser*** for advice on these assessments.

3. Sections should consider whether the acquisition of unnecessary/extraneous data, such as a large proportion of minors (individuals under the age of 16), or sensitive personal data (as outlined in the MI5 BPD Lifecycle Policy) is proportionate with respect to the desired outcome. The threshold for acquisition of this type of data is necessarily higher and will require additional explicit justification when permission is sought to acquire it.

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

4. The business case must be approved by the relevant ***senior MI5 official*** (see table above) before being submitted to ***the data governance team*** who manage the authorisation process.

5. An initial assessment will be made by ***the information management team*** to determine whether further details are required before they pass the form to ***a legal adviser*** and the relevant technical team for sign-off.

6. ***Legal advisers*** will make their own assessment on the legality of acquiring the data. If they are not satisfied by the legality of the acquisition it will not be progressed further.

7. [REDACTION] the relevant technical team will consider the likely impact of the acquisition for the Service's IT.

8. In light of the responses from ***legal advisers and the relevant technical team, the information management team*** will then conduct a final assessment of the necessity and proportionality (which might result in recommending restricted access to part or all of the dataset). They will also make an assessment (high/medium/low) of the extent of political, corporate, or reputational risk and/or damage a compromise of the data would cause, including to the data supplier.

9. ***Legal advisers and the ethics team*** may be consulted by any party and at any stage of the [REDACTION] ***relevant form*** process, where the necessity and/or proportionality are unclear.

10. Where a relationship with the supplier is required to obtain the data, but the Service does not have one, the authorised ***relevant form*** allows contact to be established with a view to acquiring the data.

11. Written confirmation should be obtained from the data supplier that approval to provide the data has been granted at Board or senior management level (e.g. Senior Civil Service, ACPO rank, Chief Executive) from within the data providing organisation, department or agency. The providing organisation may seek further (higher) approval which in the case of government data may include the Permanent Under Secretary or a Minister. ***The information management team*** must receive a copy of this confirmation.

12. ***The information management team*** will escalate the completed ***relevant form*** to a ***senior MI5 official*** [REDACTION] on behalf of DSIRO[1]. ***The senior MI5 official*** will review the necessity and proportionality of acquiring the BPD and ensure it will assist MI5 in pursuing its statutory functions; and once satisfied they will authorise the acquisition. As part of this, ***the senior MI5 official*** must also 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.

Time Sensitive Acquisition

BPD should only be acquired once ***the relevant form*** has been authorised by ***the senior MI5 official*** on behalf of DSIRO. Where a time-sensitive business requirement is identified, ***the senior MI5 official*** can authorise acquisition verbally however the associated paperwork should be completed within 5 business days.

Unsolicited offers to provide a Bulk Personal Dataset

If staff are offered BPD by a contact, the relevant section's ***senior MI5 official*** must be informed and the relevant Data Sponsor consulted. The authorisation process should then be followed if the Service can identify a genuine requirement for the dataset.

[REDACTION]

---

[1] ***A senior MI5 official*** has the option to escalate to DSIRO or SIRO as necessary.

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

Acquiring BPD from SIA Partners

When a section becomes aware of BPD held by an SIA partner which may assist MI5 in progressing its work, the section concerned must discuss their requirements and potential access to the information with the relevant Data Sponsor. A formal request to acquire the data must be made on ***a relevant form***.

Once ***the relevant form*** has been authorised by ***a senior MI5 official***, Data Sponsors will complete ***a relevant form*** outlining the business case for acquiring the data, details of the data fields required, update frequency and intended use of the data. This is sent to the relevant SIA partner and once they are satisfied the business case is justifiable and sharing the data will not breach any sensitivity considerations they may have, arrangements will be made to share the data. Timescales are dependent on agreeing a number of necessary procedures, such as:

- The frequency and timings of supply,
- How access to the data will be controlled within MI5,
- Filtering out any unnecessary data (where possible),
- Safe and secure transportation of the BPD,
- Automation of the extract, delivery and ingest process.

Collection and Storage

Transfer of Data from Suppliers

To ensure the security and integrity of BPD which MI5 relies heavily upon, and to reassure data providers their data will be handled securely, it is essential the appropriate physical controls are in place. These safeguard against unauthorised access to, or loss of, BPD during transportation to and subsequent storage in MI5 premises.

[REDACTION]

Permitted Users and Usage

Users and Systems

BPD is currently accessed primarily via MI5's corporate analytical systems [REDACTION]

The size of the user community for analytical systems has a direct impact on intrusion, which will increase as the number of users grows. Owing to the inherent sensitivity associated with BPD, it must be carefully matched to the analytical system it will be loaded into. [REDACTION]

Before access is granted to corporate analytical systems, all users must read and sign a Code of Practice on [REDACTION]. Once this is signed [REDACTION] users must also complete a mandatory training course before being granted access to these systems. There is no formal course for the ***specialist*** user community. Users of these systems are instead mentored by experienced colleagues with expertise in these systems and the datasets held within them.

In addition Privileged Users of these analytical systems must also sign the Privileged User Security Operating Procedures (SyOPs) on [REDACTION] and there is line manager responsibility for their conduct and training. [REDACTION]

Usage

Permitted queries are typically focused on fully identifying an individual that is subject of a lead or an ***subject of interest*** for whom we hold limited information. By extension it is often also necessary to identify the associates of an ***subject of interest*** to determine if they also pose a threat to national security.

[REDACTION]

Sharing Bulk Personal Data

The sharing of BPD is carefully managed to ensure that disclosure only takes place when it is justified on the basis of the relevant statutory disclosure gateway. The decision to share a BPD outside the Security Service rests with ***a senior MI5 official*** on behalf of DSIRO.

Sharing within the SIA

To the extent the SIA all have a common interest in acquiring information for national security purposes, it may be lawful for MI5 to share BPD with SIS or GCHQ. Within the SIA, the relevant gateways for these purposes are (i) section 2(2)(a) so far as disclosure by the Security Service is concerned, and (ii) sections 2(2)(a) and 4(2)(a) respectively of Intelligence Services Act so far as acquisition by SIS and GCHQ are concerned.

In relation to each dataset, there are two sides to the information transaction, whereby both the disclosing and receiving agency have to be satisfied as to the necessity and proportionality of sharing a particular dataset. MI5 need to establish in each case that both (i) disclosure by the Security Service under section 2(2)(a) is necessary for the proper discharge of the Security Service's statutory function of protecting national security, and also (ii) that acquisition by SIS and GCHQ is necessary for their respective statutory functions in respect of national security under sections 2(2)(a) and 4(2)(a) respectively of ISA.

In circumstances where GCHQ or SIS identifies a requirement, they should discuss their requirements with the relevant MI5 data sponsor. If the requesting agency and the MI5 data sponsor believe there is a business case to share the data a formal request must be made to MI5 via ***a relevant form.*** [REDACTION] The relevant data sponsor is then responsible for submitting ***the relevant form***.

The relevant form.

The relevant form outlines the business case submitted by the requesting Agency, detailing the data requested, the necessity and proportionality case for disclosure of that data and the proposed data handling arrangements. This must be approved by the relevant data sponsoring ***senior MI5 official*** before being submitted to ***the relevant team*** who will consult ***a legal adviser*** on the legality of disclosure and the relevant technical feasibility.

A senior MI5 official will confirm the strength of the business case for sharing data is sufficient, and any security, ethical and reputational risks have been adequately considered. [REDACTION]

Once ***the relevant form*** is approved by ***a senior MI5 official***, arrangements will be made for the data to be shared with the relevant Agency using suitably accredited network for electronic transfer. Where electronic transfer is not possible, physical transfer must be conducted in accordance with policy.

Sharing data and applications in-situ

[REDACTION] Sharing data in this way requires both the requesting and disclosing agencies to assess the necessity and proportionality of the access and use being sought [REDACTION

The senior MI5 official should be consulted in relation to any proposals to access data on other SIA systems, or to allow SIA access into MI5 systems.

Sharing outside the SIA

MI5 neither confirms, nor denies the existence of specific BPD holdings to organisations outside the SIA or a limited number of individuals within OSCT. Therefore any request to share BPD with an organisation other than GCHQ or SIS should reiterate this position as the requestor should approach the provider themselves. Attempts to ascertain MI5 BPD holdings by non-SIA organisations should be reported to ***the relevant team***.

In the event that a formal request is made to MI5 for BPD to be shared, the same legal disclosure tests would need to be applied as when sharing with SIA partners. The requestor would also require a legal gateway to acquire the data, which the Security Service would need to be satisfied met the test of necessity and proportionality. All enquiries should be directed to ***the senior MI5 official.***

[REDACTION]

Retention and Review

The Review Process

The Bulk Personal Data Review Panel (BPDR Panel) meets every 6 months to review BPD based on its review category. The aim of the Panel is to ensure BPD has been properly acquired and its retention remains necessary and proportionate to enable MI5 to carry out its statutory function to protect national security. Panel members must satisfy themselves the level of intrusion generated by a dataset is justifiable under Article 8(2) of the ECHR and is in line with the requirements of the Data Protection Act 1998.

The BPDR Panel operates under the authority of the Executive Board. The BPDR Panel Terms of Reference are available [REDACTION].

The BPD review categories dictate when each dataset will be reviewed (See Bulk Data Policy for details). The review of BPD retention must be captured on ***a relevant form***. [REDACTION]

At the review the Panel decides whether to retain the dataset for a further review period or to delete it. 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 SIRO, the Executive Board or DG as necessary for a decision.

The BPDR Panel will also review sharing of data, applying similar tests to those for retention. It will also commission and review thematic work in relation to BPD to inform policy development and effective risk management as it judges appropriate.

High Sensitivity Datasets

Specific arrangements are in place for particularly sensitive datasets.

[REDACTION]

Deletion of Data

Deletion process

If data is no longer required, the relevant Data Sponsor should request its deletion via ***the senior MI5 official***, and not wait for the next review. If agreed, ***the information management team*** will authorise the deletion of the relevant data and ***the senior MI5 official*** will pass the requirement for deletion to the relevant technical section. Further detail is included in the MI5 Bulk Data Policy.

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

Annex A – Frequently Asked Questions

What is intrusion?

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 the dataset. 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...'.

In relation to BPD, MI5 recognises a key distinction in levels of intrusion between (i) the simple holding of data (Inherent intrusion) and (ii) the use of that data (Actual and Collateral intrusion). The level of intrusion arising from the holding of data is generally assessed to be very limited. The level of intrusion rises significantly when data is used. Analytical processes are aimed at minimising the collateral intrusion, and distilling out the ***subject of interest*** relevant information as quickly as possible.

How do I assess intrusion?

The overall level of intrusion associated with a bulk personal dataset represents a combination of the following factors, each of which must be assessed on acquisition and review;

Holding Data:

Inherent Intrusion – Level of intrusion inherent in the data, i.e. that which arises from the simple holding of the data, including;

- The extent of 'metadata'[2] v 'content'[3]
- The extent to which the data is publically available
- The extent to which sensitive and personal data are present

Using Data:

Actual Intrusion – the level of intrusion resulting from the analysis and exploitation of data in relation to ***subjects of interest***.

Collateral Intrusion – the level of intrusion into the privacy of individuals who are not the subject of national security interest (***people of no intelligence interest***), once safeguards to minimise collateral intrusion have been implemented.

The following table illustrates the relationship between Inherent, Actual and Collateral Intrusion, and the characteristics of intrusion at each stage.

	Subjects of interest	***People of no intelligence interest***
Analysis of Data Intrusion arising from	**Actual Intrusion** • Intrusion levels vary depending on types of analysis • Intrusion levels likely to be highest but deemed necessary and proportionate	**Collateral Intrusion** • Intrusion levels may be high initially but greatly reduced when analyst identifies this data relates to ***people of no intelligence interest*** • Intrusion should always be minimised

[2] Meaning the combination of 'Communications Data' and 'Content Meta-data' [REDACTION]

[3] Meaning Narrative Data [REDACTION]

analysis and exploitation of data	• Intrusion should always be minimised when conducting analysis	• when conducting analysis
Holding Data Intrusion arising from holding data	**Inherent Intrusion** • Level of intrusion determined by; metadata v content; availability (public/private); presence of sensitive data • Level of intrusion the same for ***subjects of interest*** and ***people of no intelligence interest.*** • Levels of intrusion limited, until data is accessed and used.	

The overall assessment of the level of intrusion (HIGH, MEDIUM and LOW) associated with a dataset is based on consideration of the following criteria:

Intrusion	
High	Dataset: • Contains highly intrusive data • Contains significant amounts of sensitive and personal data • Contains significant amounts of content as well as metadata
Medium	Dataset: • Contains limited amounts of highly intrusive data • Contains limited amounts of sensitive personal data • Metadata and moderate amounts of content • Majority of records are non-adverse
Low	Dataset: • Does not contain highly intrusive data • Contains little or no sensitive personal data • Contains mostly metadata and little or no content • Mostly adverse records (dataset contains a high proportion of adverse records)

When making an assessment of intrusion, the assessment should be based on the expectation of privacy an average member of the public would have about the data within the dataset. In general, the higher the expectation of privacy, the higher will be the level of interference with privacy. When assessing expectation of privacy, a number of factors need to be taken into account, and the nature of the data needs to be understood:

- 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 the data had been collected by the data provider?
- would the individual be aware the data provider might share their data with other bodies?

- does the dataset contain sensitive personal information (please see the MI5 BPD Lifecycle Policy), 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?
- is the information contained in the dataset anonymous?
- 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 which takes into account all the characteristics of the dataset in the round. Understanding the above will enable you to make an assessment of whether the intrusion is LOW, MEDIUM or HIGH.

Examples of Intrusion Assessments

Actual Intrusion Level	**LOW**	**MEDIUM**	**HIGH**
Dataset	OLYMPIC ACCREDITATION	Travel Data [REDACTION]	[REDACTION]
Commentary	The dataset has been knowingly provided to UK HMG for security reasons. There will be an expectation this data would be shared with MI5, and 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 of 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]

What is Corporate Risk?

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

Corporate Risk refers to the potential for political embarrassment and/or damage to the reputation of MI5 and its SIA partners, data providers and HMG were it to become public knowledge MI5 holds certain datasets in bulk. It is ***the information management team's*** responsibility to assess the level of risk, be it LOW, MEDIUM or HIGH by taking the following factors into account:

- the general expectation of privacy in any given dataset, and the assessed levels of collateral and actual intrusion (see 'Intrusion' above),
- the public foreseeability of MI5 holding the data, and the possible media and public response were it to become known that MI5 held certain datasets in bulk;
- The impact on MI5 capabilities, including the potential compromise of sensitive sources and techniques, the impact on investigations and operations, or the identification of MI5 staff;

  [REDACTION]
- the impact on the reputation of the data providers and our relationship with them [REDACTION];
- the impact on liaison partners and our relationship with them [REDACTION];
- the resulting reputational and operational damage to MI5 and HMG more widely.

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

Corporate Risk	
HIGH	Dataset: • Is not publically available and/or not avowed in public and/or viewed as highly protected by the owner • It is not publically foreseeable that MI5 would hold the data or have access to it • [REDACTION]
MEDIUM	Dataset: • is not publically available; and viewed as moderately sensitive by the owner. • it is partially foreseeable to the public that MI5 would be interested in (and may hold or have access to) such data in bulk. • [REDACTION]
LOW	Dataset: • Is generally available (publically or nearly publically available) • It is publically foreseeable MI5 would have access to the data (or possibly hold it) to support their statutory functions. • [REDACTION]

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

Examples of Corporate Risk Assessments

	LOW	MEDIUM	HIGH
Dataset	[REDACTION] passport data [REDACTION]	[REDACTION]	[REDACTION]
Corporate Risk Explanation	The corporate risk is LOW as the public has a reasonable expectation MI5 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]

How long will it take before I can access the data?

Following approval, timeliness will in part depend on the priority of the acquisition. The acquisition and ingest phases of data require necessary procedures to be followed before it is exploitable, such as:

- Defining the business requirements (scope, frequency and priority of the dataset)
- Cover arrangements for the MI5 relationship for this provision
- Prior agreement for any payment relating to data provision
- Ensuring the data owner can supply the data as securely as possible,
- Agreeing the frequency and timings of supply with the provider,
- Organising the data so it can be ingested into MI5 systems as efficiently as possible,
- Filtering out any unnecessary data (where possible)