CPNI

Centre for the Protection of National Infrastructure

TRIAGE PROCESS

FOR PUBLICATION OR DISCLOSURE OF INFORMATION

CONTENTS

01. INTRODUCTION

1.1 NEED FOR THE PROCESS

Good information management is expected within and between organisations, but if information is to be identified and agreed for wider dissemination, some additional consideration is needed. The process described in this document has been developed for the public sector and is relevant to those individuals and managers who are responsible for, or considering the potential of, the disclosure or publication of information, sometimes referred to as information or simply as data. It is important to distinguish between data, which can be a collection of items or records, and information which is data with a context.

The triage process provides a repeatable framework that may be employed to test whether information is sensitive in whole or in part, and if it is, what measures may be appropriate, proportionate and practical to reduce the sensitivity prior to its publication. The sensitivity may arise from the presence of personally identifiable information (sometimes referred to as personal data), or from information that has value from commercial (legal, contractual, economic or financial) or security perspectives.

It is not a one-off process as there is a continuing need to consider impact that access to the information has in light of developments in the wider information environment. For example, where the publication of new or amended information may undermine the measures taken to reduce the sensitivity of information that is currently available. It is important to recognise that an information owner has continuing responsibilities for information curation following its publication, including periodically re-visiting the triage process in light of developments in the wider information environment.

1.2 THE CONCEPT OF INFORMATION OWNERSHIP

The approach set out in this guidance acknowledges that enterprise information is ‘owned’ by the enterprise rather than individuals or silos within it. However, accountabilities must be assigned to roles in the organisation, and individuals (or teams) fill these roles. It is convenient to designate ‘information owners’ (typically from business groups rather than technical teams) who help coordinate those accountabilities. If the information in question is personal data then the information owner would be expected to fulfil the role of the Data Controller as set out in the UK Data Protection Act.

Information ownership involves having legal rights and complete control over a single piece or collection of information. It defines and provides information about the rightful owner of information assets and the acquisition, use and distribution policy implemented by the owner. Information ownership is primarily a governance process that details an organisation’s legal ownership of enterprise-wide information. A specific organisation or the owner has the ability to create, edit, modify, share and restrict access to the information.

Information ownership also defines the owner’s ability to assign, share or surrender all of these privileges to a third party. This concept is generally implemented in medium to large enterprises with huge repositories of centralised or distributed information. The Data Owner is responsible for claiming the possession and copyrights to such data to ensure their control and ability to take legal action if their ownership is illegitimately breached by an internal or external entity.

The main difference between an information owner and an information custodian is that the latter is responsible for the quality and availability of the information on a day-to-day basis. For example, it is likely that the custodian will draft the information quality rules by which their it is measured, and the owner will approve those rules. The custodian is likely to fulfil a role of Data Processor, as defined in the UK Data Protection Act, and be responsible for curation activities, such as handling subject access requests, managing the implementation of ‘right to be forgotten’ requests, and ensuring information quality is maintained, e.g. correcting errors and implementing updates to correct for changes or additions.

02.
OVERVIEW OF TRIAGE PROCESS

The process comprises four major steps:

1) **Establish information ownership (Section 3);**
2) **Establish scope of triage process (Section 4);**
3) **Undertake sensitivity analysis (Section 5);**
4) **Continuing governance and curation (Section 6).**

03.
ESTABLISH INFORMATION OWNERSHIP

Whilst an organisation may hold a collection of information, it may not own all of it, for example, where it incorporates licenced or copyright third party content. The process to establish information ownership is illustrated in Figure 1 and described in more detail below.

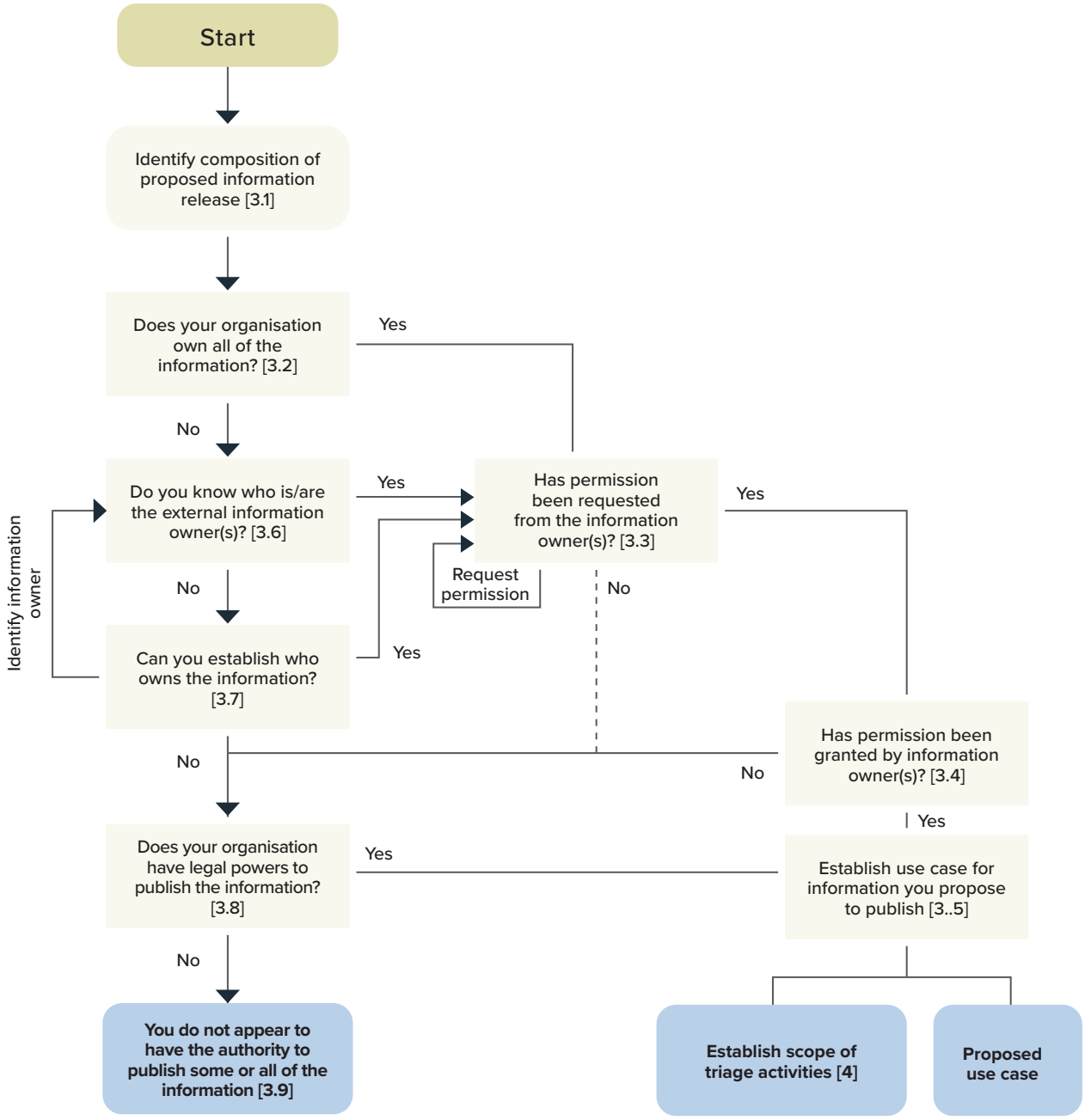

Figure 1 - Establish data ownership

Note: Figures in [] refer to the section number in this document

3.1 IDENTIFY THE COMPOSITION OF THE INFORMATION

Before considering the publication of information, its contents need to be identified and individual items defined in terms of their quality. As a guide the following properties are key to information quality:

a) accuracy – how close to the truth is the information? Is the format consistent with the accuracy of measurement or acquisition? For measured values is the precision of the result consistent with the accuracy of the sensors/scales used?

b) clarity – is the meaning of the information unambiguous and would users be able to access and interpret the underlying data model?

c) completeness – is the information complete or are there gaps in its coverage?

d) consistency – is the same thing identified and defined in the same way, i.e. is a consistent data model being employed? Where a data item represents options selected from a list, is there a risk that subjective judgments may have affected consistency?

e) provenance – who created the information, when and where it came from, and how has it been processed?

f) timeliness – is it clear when the information was collected and/or processed?

The output from this activity should be documented and available for release with the information. This quality information may be released as a set of metadata, i.e. information describing the content of the information as a whole and in parts, providing there is an agreed specification for doing so. The quality of any released information should be clear to the recipients, i.e. the release is accompanied by information regarding the key properties listed above.

3.2 DOES YOUR ORGANISATION OWN ALL OF THE INFORMATION?

Based on the description of the information developed in the previous step the ownership of individual elements or data items should be established. This provenance information should assist in determining the source and therefore potential ownership, however further due diligence may be required in respect of items where the master data has been automatically sourced, for example, address data derived from the Postal Address File (PAF).

3.3 HAS PERMISSION BEEN REQUESTED FROM THE INFORMATION OWNER(S)?

The explicit consent of the information owner(s) should be sought prior to taking any steps towards publication. Often within organisations, ownership may be separate from custody, the latter being the person or team responsible for the day-to-day information management, whilst the owner may be in a different part of the organisation.

Before starting to develop one or more use cases for publication or disclosure, the information owner's consent should be sought in principle, and the owner should be involved in decision-making about its sensitivity as well as the governance and curation arrangements following any subsequent decision to publish it.

In some circumstances it may not be appropriate to seek the information owner's permission, for example, in the case of Freedom of Information or Environmental Information Regulations requests, the decision to publish or not lies with the organisation holding the data.

3.4 HAS PERMISSION BEEN GRANTED BY INFORMATION OWNER(S)?

At this stage, permission from the information owner should be considered to be provisional subject to completing the triage process and where applicable agreeing the governance and curation arrangements for any published information.

It should not be assumed that because the information owner has consented to publication for one use case then this consent applies to all potential use cases. The characteristics of the information may vary between use cases as a result of protective measures and the audiences and access arrangements may differ significantly between use cases.

3.5 ESTABLISH USE CASE FOR PROPOSED INFORMATION THAT YOU PROPOSE TO PUBLISH

This is an important step as aspects of the use case will need to be considered as part of the triage process, for example, in determining what treatments may be applied to reduce the sensitivity whilst maintaining the overall integrity of the information. Specific aspects to consider in developing the use case are:

a) Who are the potential users of the information?

b) How are they likely to use the information?

c) What are the consequences for users if the information is of poor quality or incorrect?

d) What decisions are they trying to make using the information?

e) Does the granularity of the information affect its potential uses?

f) How will the information be accessed? Will it be:

   i) open information, i.e. published on the Internet and downloadable without registration or logging into an account?

   ii) public information, i.e. readily accessible over the Internet, but users have to register or log into an account to download it? This is appropriate where licence conditions may be associated with the information, for example, the Open Government Licence;

   iii) shared information, i.e. accessible only to a group or approved users and/or organisations?

One the use case has been established the triage process can move on to the next stage, see Section 4.

3.6 DO YOU KNOW WHO IS/ARE THE EXTERNAL INFORMATION OWNER(S)?

If the contents of the information as a whole are not owned by the organisation that is planning to publish or disclose it, then further steps need to be taken to establish the ownership and any conditions of use of it as a whole or subsets of the information. Once any third-party ownership has been identified it is necessary to obtain permission from the information owner(s) for publication to be considered (see 3.3).

3.7 CAN YOU ESTABLISH WHO OWNS THE INFORMATION?

This may require some research to establish beyond reasonable doubt who the owner is for any information which the organisation does not consider it owns.

3.8 DOES YOUR ORGANISATION HAVE LEGAL POWERS TO PUBLISH THE INFORMATION?

It may be necessary to consider whether your organisation has the necessary legal powers to publish the information, where:

a) the information owner no longer exists;

b) it is not possible or practical to trace the information owner;

c) exceptionally, it is not in the public interest to seek permission from the information owner(s).

It is likely that this only needs to be covered where the organisation has statutory or regulatory powers or obligations that will allow it to publish the information.

3.9 NO AUTHORITY TO PUBLISH

In the event that the organisation does not have the information owner's consent to consider publication, or that the owner has not given consent and the organisation lacks the legal capacity to publish, then the organisation should not publish the information.

04.
ESTABLISH SCOPE OF TRIAGE PROCESS

Having established the owner's consent to consider publication of the information, in whole or in part, the next step is to assess the required scope of the triage process. This step is illustrated in Figure 2 and described below.

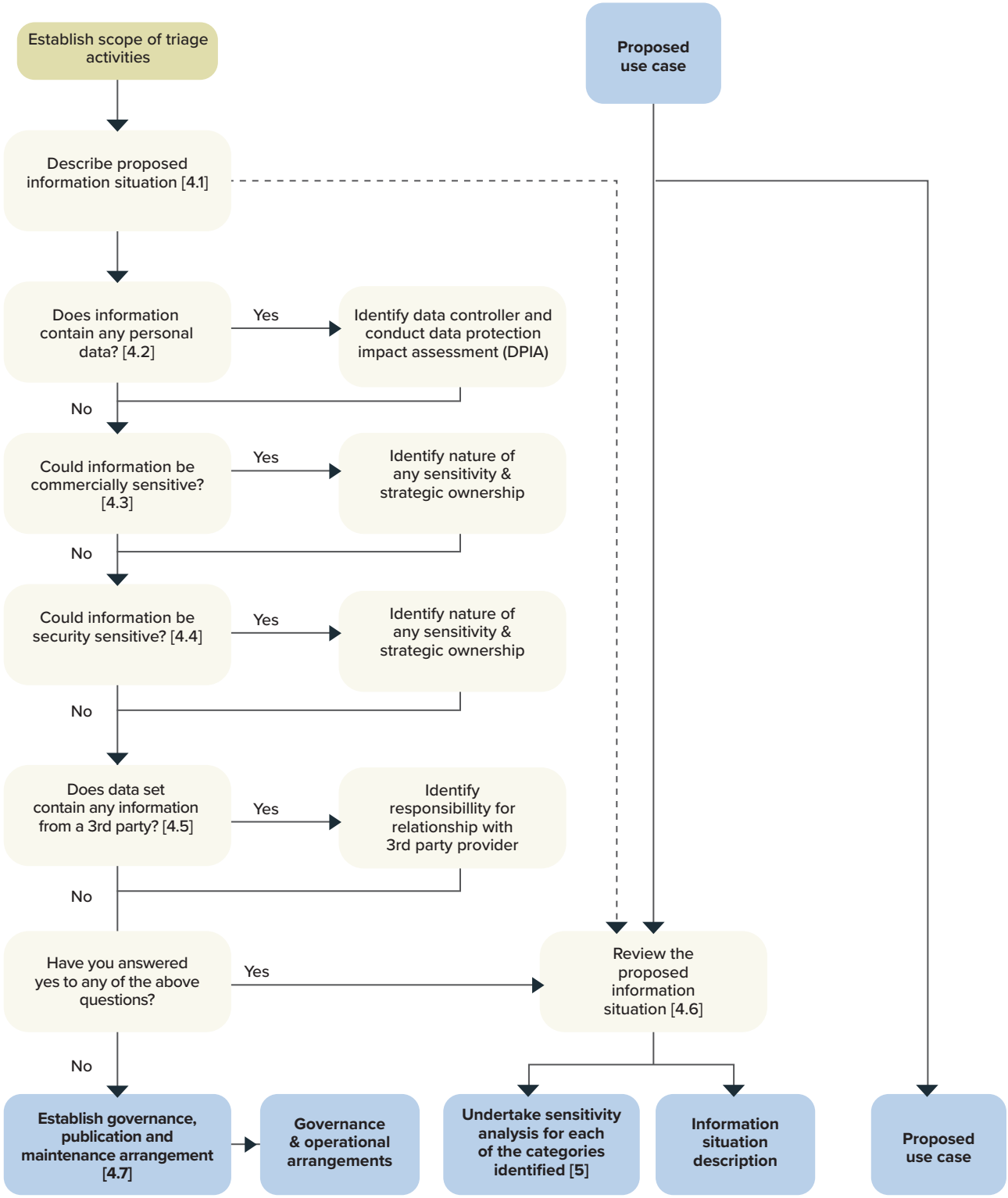

Figure 2 - Scoping the triage process

Note: Figures in [] refer to the section number in this document

4.1 DESCRIBE PROPOSED INFORMATION SITUATION

Before any assessment can be made as to whether the information may be sensitive it is necessary to establish the information situation and the environment into which it is proposed to release it. The information environment should be considered as comprising four components:

1) **Information:** What (other) information exists in the information environment in which the intended release will be made? How does the existing published information overlap with or connect to the information that you propose to publish? This information is needed to identify what elements (key variables) are potentially risky. The risks may arise from aggregation or in statistically matching one dataset with another thereby improving the conditions for statistical disclosure.

2) **Agency:** An agent may be a human or an automated process (e.g. an application crawling and indexing open or publicly available information), that is capable of acting on and in the information environment.

3) **Governance processes:** This relates to how relationships between users and/or agents and the information is managed. It may include formal governance (e.g. laws, regulations, data access controls or confidentiality agreements, licensing arrangements and policies which prescribe and/or proscribe user behaviour). It also encompasses behavioural or cultural norms and/or practices, and users' pre-dispositions (e.g. individuals who analyse information with a view to identifying sensitive information by aggregating what was apparently innocuous information). From a governance perspective the accountable party is the information owner who agreed to its release. This is an ongoing responsibility as the information owner should be prepared to answer questions regarding their approval to release it and should view it as a risk-based decision.

4) **Infrastructure:** This encompasses how infrastructure and wider societal and economic structures shape the information environment. Infrastructure includes the set of interconnecting structures (physical, technical) and processes (organisational, managerial, contractual, legal) that comprise the information environment. Where applicable, it also includes the information processing systems.

Depending on the information access arrangements in the proposed use case established in section 3.5, the level of control the data owner and curators has over the information environment may vary considerably.

Environments can exist inside other environments, for example as part of its overall organisational information environment an organisation may operate a secure environment that it employs to publish shared information to trusted 3rd parties, with a subset of this information released into a publicly accessible repository. The organisation relies upon its governance and security infrastructure to reduce the risk of information leaking between these environments or becoming openly available data.

An information situation addresses the concept of the relationship between information and its environment. Information situations can be:

a) static, i.e. the information environment is fixed, and few if any changes will be made to the other or published information; or

b) dynamic, i.e. the information and the environment are subject to change, which may include streaming of information, periodic updates of the information and the publication of new/additional information.

From information protection and security perspectives, static information environments potentially present less risk, but in practice few situations are likely to be static. The process of sharing or releasing information in a dynamic situation represents a greater risk as the triage process represents a snapshot, which the governance arrangement may require to be periodically repeated to address emergent aggregation risks.

4.2. DOES INFORMATION CONTAIN ANY PERSONAL DATA?

Personal data only includes information relating to natural persons who:

a) can be identified or who are identifiable, directly from the information in question; or

b) who can be indirectly identified from that information in combination with other information.

To assess whether information contains personal data it is necessary to know what other information is or could be co-present. It is not sufficient to attempt to judge whether information is personal or not using absolute criteria (e.g. the non-relative properties of the data themselves). Thus, in assessing whether the information contains personal data, the assessment should consider in what context, and more specifically in what information environment it would be published.

Article 4 of GDPR and Section 3 of the UK Data Protection Act 2018 provide the following definition of personal data:

> "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

This is a broad definition, particularly as through aggregation, information which in isolation does not identify of make identifiable a person, can through accumulation or association result in identification or make the person identifiable.

A decision that information contains personal data does not mean that the information may not be published, instead it requires further consideration of what steps should be taken to maintain confidentiality of the personal aspects of the data. These measures are considered further in Section 5.

4.3 COULD INFORMATION BE COMMERCIALLY SENSITIVE?

As with the identification of personal data, the review should consider the information in the environment in which it would be published. Whilst there is no settled definition for the concept of commercially sensitive information, it is generally recognised as information that, if disclosed, could prejudice a supplier's commercial interests (for example, trade secrets, profit margins, innovations or new ideas) or a buyer's purchasing interests (for example, budgets, competitors' tender price or contents, and financial terms of a contract or deal). These interests are sometimes generalised as intellectual property and legally binding restrictions.

From a supplier's perspective, a trade secret may comprise of a formula, pattern, device, or compilation of information which is used in the supplier's business, and which gives the supplier an opportunity to obtain competitive advantage over others who do not know or use it. The presence of commercially sensitive information does not necessarily preclude publication of some or all of the information, but it may require additional processing, as described in Section 5, to reduce its sensitivity.

4.4 COULD INFORMATION BE SECURITY SENSITIVE?

Sensitivity from a security perspective relates to information whose loss, misuse, modification, or unauthorised access to, could:

a) adversely affect the:

   i) welfare or safety of an individual or individuals;

   ii) safety, security, integrity or operation of critical assets at national and regional levels;

b) cause commercial or economic harm to the country; and/or

c) jeopardise the security, internal and foreign affairs of a nation.

The degree of sensitivity will depend on nature and extent of the information.

Of particular concern and relevance is information that may:

a) assist in hostile reconnaissance, for example, the identification of vulnerable individuals or the identification of sensitive assets;

b) enable the planning and execution of hostile or malicious acts.

As with the identification of both personal data and commercially sensitive information, the review should consider the information in the environment in which it would be published.

4.5. DOES DATA SET CONTAIN ANY INFORMATION FROM A 3RD PARTY?

The presence of 3rd party information in the information raises licensing, copyright and consent issues. The presence of such information is likely to be identified as provenance of the information is documented and analysed. This is particularly important where the information has been aggregated from multiple sources, which in turn may be collating or aggregating information from further sources.

The issue of consent may arise where information has been collected from individuals for one purpose and the data processor is now proposing to use it for another purpose. Whilst the information may not be personal data as outlined in section 4.2, the data protection principle of informed consent should be considered.

4.6 HAVE YOU ANSWERED YES TO ANY OF QUESTIONS 4.2 TO 4.5?

If negative responses apply to all of the questions posed in the titles of the above four sections, then the matters outlined in section 4.7 should be addressed prior to publishing the information. However, If the answer to one or more of the questions was yes, then there is a need to consider the potential sensitivity of the proposed disclosure and the steps in section 5 should be addressed in parallel with the work outlined in section 4.7.

4.7 ESTABLISH GOVERNANCE, PUBLICATION AND MAINTENANCE ARRANGEMENTS

4.7.1 GOVERNANCE

It is reasonable to expect and anticipate that the information environment into which the information is published will change over time, this may be as a result of technological advances, an increasing amount of available information, development of new analysis tools and techniques, or changes in the user population. The responsibilities of the information owner and custodians do not stop when the data is released, appropriate governance arrangements should be put in place. Where any of the information is considered to be personal data, the provisions of the Data Protection Act will need to be included in the governance arrangements.

Whilst the information you have published may be considered not to present a security or privacy risk at the time of its disclosure this may not remain the case in the medium term. With regards to personal data the ICO takes the view:

> "Means of identifying individuals that are feasible and cost-effective, and are therefore likely to be used, will change over time. If you decide that the data you hold does not allow the identification of individuals, you should review that decision regularly in light of new technology or security developments or changes to the public availability of certain records."

The same is true of other types of sensitive information, new means and/or changes in the information environment may enable security or commercially sensitive information to be inferred or identified.

From a governance perspective it is prudent to adopt at least the following measures:

a) identify and continue to monitor and assess the quality of the data being released (see 3.1), and make this quality information readily available to all users;

b) maintain a register of all the information you have shared or released;

c) use the process outlined in this document to compare proposed publishing activities with past disclosures to take account of the possibility of linkage between

releases, or variations in the releases leading to an unintended disclosure of sensitive information;

d) monitor and be aware of changes in the information environment and how these may impact on published information, by:

   i) maintaining awareness of developments in new technologies and security that may affect your information situation

   ii) monitoring changes in the law, regulation and guidance on the types of information that you are publishing, sharing and disseminating

   iii) keeping track of current and new public information sources, both available on the internet and as appropriate available from more traditional sources such as public registers, local community records, estate agents' lists, professional registers, the library, etc.

e) wherever possible you should also keep track of how your information is used.

Whilst it may be suggested that published data should be open, this presents two challenges for information owners:

a) you have no concept of who is downloading and using the information or ability to measure whether it is of practical or beneficial use;

b) it is not possible to contact an information user to inform them of corrections to the information in the event that errors or omissions are discovered post publication.

A prudent approach is to treat the information as public information rather than as an open data, by adopting a simple process whereby users register, i.e. provide their name, email address and their intended use before downloading. This registration information may be crucial later when you are seeking to demonstrate the value or impact of publishing the information, considering the next release, refining its use case or developing new use cases, and considering the risk and utility trade-offs prior to publication of new or updated information.

As part of the governance arrangements the security policy regarding the information should be considered. Whilst this can be handled on a case-by-case basis a prudent approach is for the organisation to establish an information security strategy that provides a co-ordinated and consistent approach to the:

a) sensitivity and security assessment of information that the organisation:

   i) is considering for publication, disclosure or sharing; and

   ii) has already published, disclosed or shared;

b) decisions regarding appropriate and proportionate treatment of information following the assessment to:

   i) reduce the sensitivity or security concerns to an acceptable level; or

   ii) withhold publication;

c) review and approval of information-sharing agreements, where access to and use of the information is subject to constraints set by the information owner;

d) handling changes in circumstances affecting the information environment.

The security strategy should also address management responsibility and accountability for information security decisions and the publication processes.

4.7.2 MAINTENANCE

This encompasses activities that serve to deliver the information ready for publication and usage in an appropriate form and manner for the purposes for which it is being disclosed. Maintenance activities are likely to include combinations of the following measures:

a) validation and verification;

b) cleansing; reformatting;

c) enrichment; movement;

d) integration from multiple systems;

e) updating of published information; and where required

f) the withdrawal of the published information.

These activities are necessary to ensure that publication is not simply a release and forget issue, errors may need to be corrected and users may have questions regarding the information quality.

Where information is being withdrawn consideration should be given as to whether there is a need for its archival, i.e.

replication or placement of data into an archive where it is stored but where no maintenance, usage or publication occurs. Once the information has been withdrawn and where applicable archived, the information should be purged. This involves removal of every known copy of the information to prevent inadvertent re-publishing of the information.

4.7.3 PUBLICATION

The publication process that makes the information available outside the organisation should include appropriate independent reviews to ensure any modifications to the data required as part of the release process have been correctly undertaken. These checks should apply on initial release and to any subsequent updates.

05.
UNDERTAKE SENSITIVITY ANALYSIS AND INFORMATION PREPARATION FOR PUBLICATION

5.1 KNOW YOUR INFORMATION AND ITS RELEVANCE TO USE CASE

In the preceding sections 3 and 4, information ownership has been established and the scope of the triage process accessed. The nature of the information has been established and its relevance to the use case assessed. The review in section 4 may have identified the need for further sensitivity assessment as part of the pre-publication risk analysis and control and one or more of the detailed assessments in section 5.2 may be required. The process to be followed is illustrated in Figure 3.

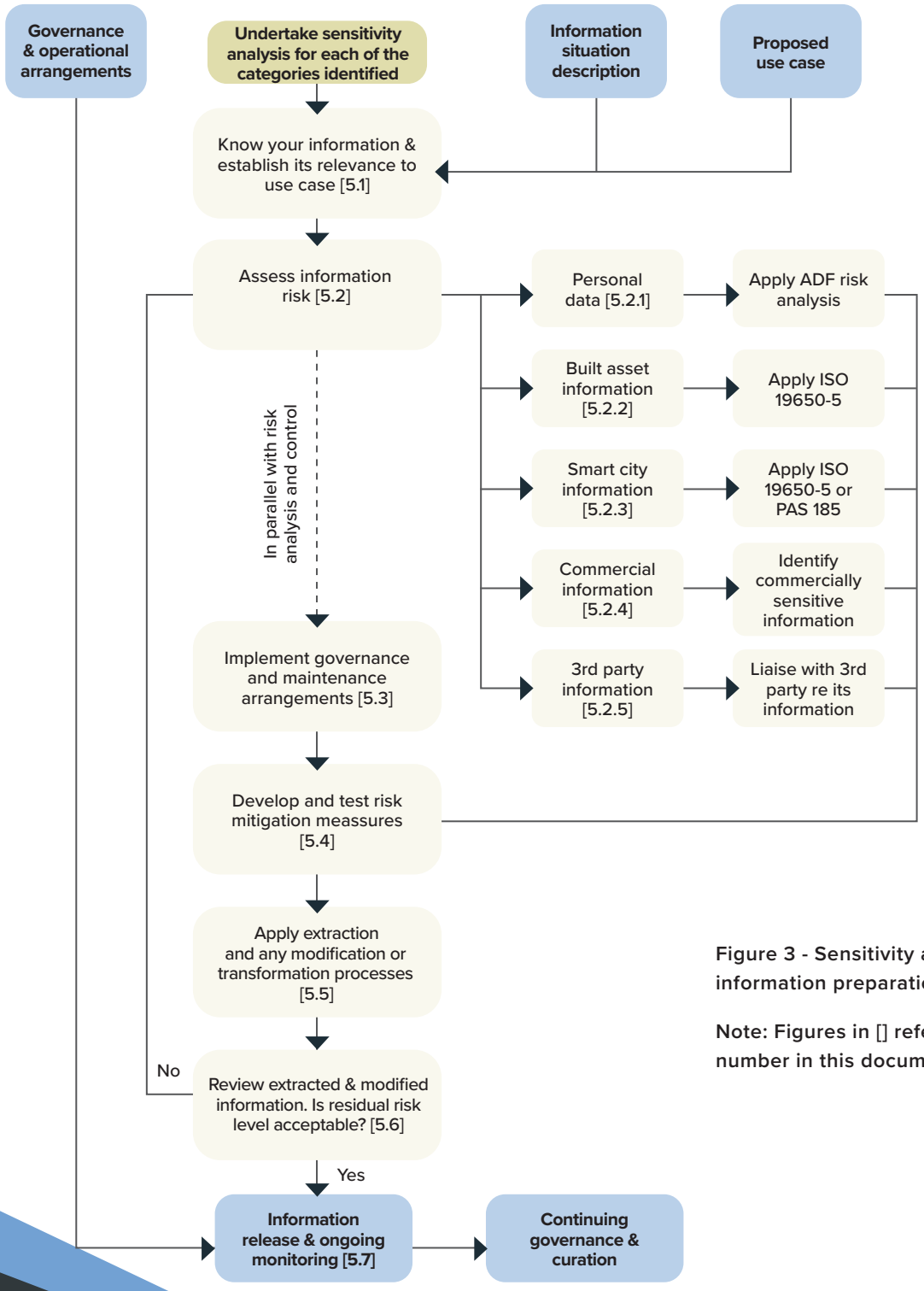

Figure 3 - Sensitivity analysis and information preparation for publication

Note: Figures in [] refer to the section number in this document

In undertaking the sensitivity analysis, one needs to adopt the mindset of an investigator or hostile actor, who may draw upon other open or publicly available information to interpret or attack the information that you plan to release. The analysis should consider plausible motivations and assess whether the objective of the investigator or hostile actor is achievable more easily by other means. Specific points to consider are:

a) motivation: what is the investigator or hostile actor trying to achieve?

b) means: what resources (including other information) and skills do they or may they possess?

c) opportunity: how will they access the information you plan to release?

d) target information: what information would they be seeking to accomplish their objective?

e) is their objective more easily achievable by other means?

f) effect of data divergence — the investigator or hostile actor may only have access to old or a subset of the information. Would the proposed release improve the accuracy or reliability of the information they are using?

5.2 ASSESS INFORMATION RISK

5.2.1 PERSONAL DATA

If the information contains personal data guidance provided by the Information Commissioner's Office should be followed. There is extensive guidance available on the anonymisation of personal data, for example, The Anonymisation Decision-Making Framework published by the UK Anonymisation Network.

In addition to the protection of personal data about individuals there is a need to consider the pattern-of-life issues that may enable the identification and location of groups of individuals to prevent them being targeted by those seeking to perform, hostile or malicious acts. A pattern-of-life relates to the identification of habits, routines and preferences of an individual or group of individuals that enable prediction of future actions and/or behaviour.

Variables that are particularly relevant in establishing patterns-of-life are generally spatial and temporal ones, i.e. placing the individual or group at a location on a predictable basis. From a spatial perspective the increased tagging of assets, e.g. photographs, tweets, etc with geospatial information and the collection and processing of network related identifiers, e.g. computer IP addresses, Wi-Fi SSIDs, Bluetooth devices and beacons, etc., increases the risk that they contain variables relating to a person's location or their personal devices' locations over time. Unless such information is essential in respect of the proposed use case it is prudent to remove these variables from the information to reduce the risk of inadvertently leaking pattern-of-life data.

5.2.2 BUILT ASSET INFORMATION

Within the UK, a built asset should be considered sensitive, in whole or in part, if it:

a) is a designated site under sections 128 or 129 of the Serious Organised Crime and Police Act 2005;

b) forms part of the critical national infrastructure (only the asset owner, the lead government department and CPNI will be aware of its status);

c) fulfils a defence, law enforcement, national security or diplomatic function;

d) is a commercial site involving the creation, processing, trading or storage of valuable materials, currency, pharmaceuticals, chemicals, petrochemicals, or gases or the provision or production of enablers for production of these materials;

e) constitutes a landmark, nationally significant site or crowded place (as determined by NaCTSO); and/or

f) is used or is planned to be used to host events of security significance.

In addition, an asset or product should be considered sensitive if there is sufficient risk that it has been judged that it could be used to significantly compromise the integrity of the built asset as a whole, or its ability to function. The specific assets or asset attributes which shall be considered include, as a minimum:

a) location, routes, cabling, configuration, identification and use of control systems;

b) location and identification of permanent plant and machinery;

c) structural design details;

d) location and identification of security or other control rooms;

e) location and identification of regulated spaces, or areas housing regulated substances (e.g. nuclear isotopes and biohazards) or information; and

f) technical specification of security products and features.

The nature and type of information that is sensitive in respect of these assets will vary between built assets, but typically will relate to the information required by an attacker undertaking hostile reconnaissance. Examples of such information include but are not limited to: security measures (physical, technical and handling of incidents): security guarding (numbers, patrol patterns, response times, etc); site/building/floor plans showing information about the use or function of spaces; the location of specific operations, organisational units or users.

For built assets the timeliness and granularity of information availability may significantly increase the risk. This is particularly an issue where such information contributes to the ability to conduct hostile reconnaissance remotely or allows more detailed analysis of the underlying operation of a built asset and its relationships to other assets.

5.2.3 SMART CITY INFORMATION

The increased use of digital geospatial information regarding built assets and smart city operations requires careful consideration before such information is published. For smart cities services the timeliness, granularity and level of geospatial information available may significantly increase the risk. This is a sensitive issue where such information contributes to the ability to conduct hostile reconnaissance enabling operation of city-related services to be analysed and/or geospatial profiling of services and service users. In the case of service users, care should be taken to ensure that the published information does not facilitate the identification and targeting of users or groups of users, except where such analysis is being undertaken by appropriately authorised service personnel as part of their smart city duties.

For legacy assets where the geospatial data has been derived from digitising analogue records, a false level of accuracy may be assumed, particularly if the original data was based on hand-drawn records on hardcopy maps. This is both a quality and security issue as it can lead to unintended damage to underground assets due to inaccurate information. From a security perspective there are issues regarding the potential remote use of geospatial information in hostile reconnaissance and the potential for knowledge of specific asset adjacencies being used to increase the impact of hostile or malicious acts. Where it is proposed to publish geospatial information about assets that are not readily observable, the asset owner(s) and operator(s) should be consulted about what security measures may be required to protect the assets.

5.2.4 COMMERCIAL INFORMATION

As noted in section 4.3, commercially sensitive information generally relates to trade secrets, whose nature and sensitivity will vary between organisations. Examples of trade secrets include:

a) customer lists, if individuals this information may also be personal data;

b) marketing and product development strategies and analysis;

c) manufacturing information, including formulas for specific products;

d) supplier lists and the nature of the materials or services supplied;

e) cost, price or other quantitative data relating to the organisation's operations;

f) data or information covered by commercial confidentiality or non-disclosure agreements.

In assessing the sensitivity of commercial information account should be taken of what information is readily made available in a public report or marketing material published by the organisation. If for example press releases are issued for acquisition of major new customers, the customer list in general may not be sensitive, but there may be individual customers that have required non-disclosure agreements to be signed preventing any unauthorised publicity regarding the relationship. In these circumstances unless a waiver is given by the customer, then the customer's details should be treated as sensitive and not for publication.

5.2.5 THIRD PARTY INFORMATION

The provenance of the proposed information release should be investigated to ascertain whether it contains Information sourced from a 3rd party, which:

a) is subject to licence conditions limiting its use or further dissemination; or

b) was collected from individuals or organisations and where consent has not been given for the publication or disclosure of the raw data.

If either of the above conditions are met for any of the information, then further steps should be taken to ascertain whether the data may legitimately be published and if so in what form it would be acceptable to so do.

5.3 IMPLEMENT GOVERNANCE AND MAINTENANCE ARRANGEMENTS

If not already in place the governance and maintenance arrangements should be established.

5.4 DEVELOP AND TEST RISK MITIGATION MEASURES

Based on the required assessments as determined in section 4.6, and undertaken as outlined in section 5.2, the data owner should:

a) assess the risks associated with the publication;

b) based on the sensitivity assessments consider what mitigation measures may be appropriate and proportionate to reduce the risk to a more acceptable level;

c) document the proposed use case, information situation, risk analysis and any decisions regarding measures to be applied to the information before it can be published.

If in light of the risks a decision is taken not to publish the information the reasons should be recorded for future reference and to guide the organisation's information management policy.

The nature of the measures applied may include:

a) modification of the data to reduce the risks of identification, or re-identification from anonymised information, specific individuals, groups of individuals, built assets or organisations. This can include redaction or omission of specific sensitive information;

b) reducing the level of granularity through aggregation of spatial, temporal or variables/categories;

c) delaying the publication of the information to permit some historical trend analysis without exposing current/real-time data about individuals, built assets or their operation;

d) procedural measures such as limiting access, use of information sharing agreements to specify acceptable uses and the security measures required to protect the information if downloaded onto the users' computers.

Specialist advice should be sought regarding the use of the above types of measures to ensure that the required effect is delivered and that the modifications do not create artefacts which highlight the information that the data owner is seeking to protect.

5.5 APPLY EXTRACTION AND ANY MODIFICATION OR TRANSFORMATION PROCESSES

Based on the measures decided on in section 5.6, a process should be defined, configured and tested to extract the information that is to be published, modify or transform it as required and format it for publication. The process and testing arrangements should be documented to enable repeatable extracts and provide a clear audit trail of how the information has been manipulated.

5.6 REVIEW EXTRACTED AND MODIFIED INFORMATION TO ASSESS RESIDUAL RISK

On completion of the activities outlined in section 5.7 the information prepared for publication should be reviewed to ensure that the measures taken achieve the objectives regarding the protection of sensitive information. If there are still concerns from the information owner regarding the level of residual risk, a further iteration of the process in section 5 may be required.

As part of this review consideration should be given to what documentation will be made available regarding the information and the extent to which, if at all, any explanation

is provided regarding any modification or transformation processes that have been applied.

The value and utility of the published information will be affected by its quality (see 3.1). The information owner and the publisher would ensure that information quality is documented and made available for release with the information.

Where the quality information is released as metadata it should be published to an open specification and in a recognised format. Examples of the set of information which may be provided as metadata include:

a) source information, i.e. how and when the information is collected or obtained and any subsequent processing;

b) the nature of the content of the release, e.g. variables it contains;

c) information about the accuracy, completeness and coverage of the release;

d) information format (i.e. structure and coding); and user rights.

The security and sensitivity considerations outlined in this document apply to both the information release and any associated metadata. Where information is shared with rather than publicly released, access to the metadata should be restricted to those who have access to the information and not made publicly available.

5.7 INFORMATION RELEASE AND ONGOING MONITORING

Where the information is to be shared with a defined audience or user group, prior to publication the access control mechanisms should be verified first, to ensure that:

a) an information sharing agreement is in place addressing the information governance and security requirements for the published data;

b) appropriate processes are in place to manage access to the information, i.e. the granting, monitoring and revoking of access according to users' needs;

c) the technical controls limiting access to the information have been subjected to and successfully passed formal security testing.

Once all of the preparations for publication have been completed the content can be released into the approved information environment. As there is a cost to publishing data, it is prudent to monitor its use. Depending on how the information was published this may involve collecting information on the number of downloads and or logins by individual users. As noted in section 4.7.1, the usage information is important evidence for decisions regarding future use cases and the impact of publishing the information. It may also be useful in instances where the publisher needs to notify users of errors or omissions discovered after the data has been published.

06.
CONTINUING GOVERNANCE AND INFORMATION CURATION

The information owner, or the information custodian acting on the owner's behalf, is responsible for managing the continuing information governance and ensuring that the release is appropriately curated to address errors or issues reported by users.

Disclaimer

CPNI
Centre for the Protection of National Infrastructure