CPNI

Centre for the Protection of National Infrastructure

DIGITALISATION INITIATIVES

ESTABLISHING HIGH-LEVEL INFORMATION NEED AND MANAGEMENT REQUIREMENTS

GUIDANCE DOCUMENT

This guidance document is intended for use by managers leading and teams delivering digitalisation initiatives. It sets out a process to determine high-level information need and any associated security and management requirements.

ESTABLISHING INFORMATION NEED

Digitalisation initiatives leverage the power of information in order to:

- drive better decision-making and service delivery across a whole range of existing public services and societal goals; and
- facilitate innovation and experimentation to create new products and services.

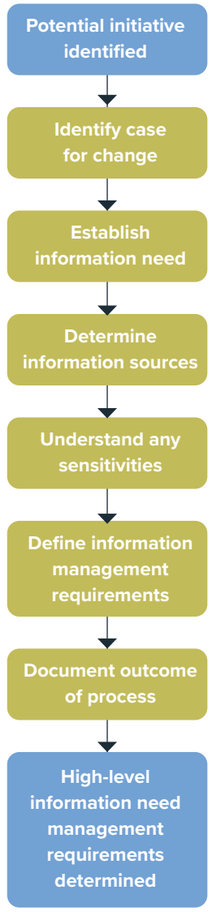
However, just increasing the volumes of data we collect and store will not necessarily deliver the benefits we seek. Storing, processing and consuming information also comes at a cost, both monetary and in terms of energy consumption, which is especially relevant in light of Net Zero Strategy targets.

It is therefore valuable to spend some time establishing the specific information that needs to be acquired, processed, stored and consumed in support of an activity or service. As part of this, any information which is likely to be sensitive and which may compromise security or privacy can be identified early. This allows appropriate decisions to be made about managing associated risks before any information is collected, aggregated or released.

The structured process set out in this guidance document, and outlined in the flow chart below, comprises 6 stages that will help an organisation to determine its high-level information need and the likely management requirements for that information. It will be of use to a team at the start of a digitalisation initiative of any size, whether in the public or private sector.

It does not replace the processes and outputs required for investment decision-making and project delivery, for example business cases, project plans and detailed information specifications.

PROCESS FOR ESTABLISHING HIGH-LEVEL INFORMATION NEED AND MANAGEMENT REQUIREMENTS



The output from this process will be useful evidence of the approach followed if undertaking a Security Considerations Assessment[1] in relation to the digitalisation initiative.

1. SET OUT THE CASE FOR CHANGE

The purpose of this stage is to define a high-level scope for the initiative and the planned outcome.

i. What is the challenge you are trying to address?

It is important to understand the precise nature of the problem that you are seeking to address and/or the new service or innovation that you are seeking to drive. This process should include an analysis of the current blockers.

> **Example** A local authority identifies that investors are reluctant to develop investment plans for brownfield sites due to the complexity and delays in obtaining relevant site information.

ii. What is the end-state that you are seeking to achieve?

This is in essence a statement of the policy or business driver for the project. For many projects, the intended end-state may need to be qualified by success criteria or performance metrics.

> **Example** The local authority, seeking to make it easier and quicker for potential developers or investors to identify sites, understand their suitability, including the capacity of existing services and surrounding environs, and to obtain the necessary planning consents, plans to develop a digital service that will provide information about its brownfield sites.

iii. What do you need to do to achieve this end-state?

Policy, process, organisational or system changes may be required to deliver the end-state you are seeking to reach. Early identification of these will help in establishing the information management and information security requirements later in the process.

> **Example** The local authority needs to build or acquire a new platform in order to provide a digital service that will make the relevant information available to potential developers.

2. ESTABLISH INFORMATION NEED

The purpose of this stage is to identify and prioritise the information required to achieve the proposed end-state.

i. What information do you need to support the end-state you are seeking to reach and what are the sources of this information?

You should begin by identifying the key information required to deliver the end-state you are trying to reach. You may already be the owner of this information or may be able to generate it, however, it is also possible that you may need to acquire and process information from other sources.

If information needs to be obtained from third parties, any constraints around acquiring, processing and sharing this information should be explored at this stage.

> **Example** The local authority identifies that potential developers of brownfield sites require information on the location, nature and size of the different sites, any site constraints, infrastructure and facilities already in place and their current spare capacity and capability to support different possible developments. While the local authority already has information about the specific site, it needs to seek to obtain information from utility providers about their assets in and around the sites.

ii. Are there alternative ways of achieving the outcome without the information?

In some situations, the costs and/or risks associated with acquiring, processing and storing the information over its lifecycle may outweigh any financial or societal benefits - this is a counterfactual test. While it can be easy to assume that the outcome can only be achieved by increasing availability of information, it is worth checking whether there are viable alternative ways of achieving the same outcome.

1. The Security Considerations Assessment (SCA) process ensures security-related vulnerabilities are considered across a range of activities and processes within an organisation. This includes physical, personnel, cyber and cross-cutting security measures.

iii. Is it possible that you will acquire more information than is required to achieve the outcome sought?

It is often possible, especially with the introduction of multi-function devices, to acquire significantly more information than is required to deliver the proposed outcome.

Where the information relates to individuals, collection of irrelevant or excessive volumes of information is in contravention of the requirements relating to data minimisation set out in the Data Protection Act.

If any information has to be provided by a third party and is not already publicly available, there is likely to be cost associated with acquiring and processing this information. Avoiding speculative collection helps to reduce these costs and associated risks. If it is found that information is required at a later date, the information gathered will have the benefit of being the most up-to-date available. It is therefore best to limit the information you acquire and retain to the minimum required to achieve the desired outcome.

Example The local authority could create detailed maps of both the location and capacity of utility assets in the areas surrounding potential brownfield development sites. However, at this stage of the process, the precise locations of these assets are generally not relevant. Information on spare capacity and capability to support different potential developments and the costs and possible timescales for carrying out upgrade works is likely to be much more relevant.

iv. Of what quality does the information need to be to achieve the desired outcome?

The diagram shows the information quality considerations which determine its quality or its fitness for use. It differs slightly from one used in the Information Management Framework (IMF) as it is also intended to support remote data collection, for example, in connected places projects.

A description of each of these information properties is set out below:

1. **Relevance** – the information relates to the use to which the information is to be put

2. **Clarity** – the meaning of the information is unambiguous

3. **Accessibility** – the information is easily accessible to authorised users, but not to unauthorised users

4. **Consistency** – the information is uniform at a specific point in time and across transactions, both within and across computer programs

   The issue of consistency is critical where personal records from different services or systems are being processed. Incorrect association of individuals' records could lead to serious harm and result in enforcement action and/or large fines from the Information Commissioner. These inconsistencies can arise from variations in the way that names, addresses and relationships are captured, e.g., the shortening of first names, use of preferred (unregistered names), use of house names without available street numbers, etc.

5. **Provenance** – the origin of the information, in other words, where and when and from who did the information originate?

   Maintaining information about the source(s) of data and its processing enables users to determine whether it is, or continues to be, fit for purpose. For example, changes in frequency of update, sensor technology, as well as changes in processing algorithms can all affect the long-term utility of data.

6. **Timeliness** – the information is sufficiently up to date when it is needed

   Information that is out-of-date and incomplete undermines the effectiveness of decision-making, can reduce confidence and may result in breaches of the data protection legislation.

   The concept of "real-time" information availability needs to be carefully considered as there is a cost to achieving real-time communications and processing. In practice, the acceptable latency between capture of new or changed information and its use will depend on the rate of change and the decision-making use. For example, in monitoring natural environment trends changes may occur over years or decades. Information on storms and flooding may change by the minute or hour, while for a control system a delay of a few minutes in the information being received could have severe implications.

7. **Completeness** – the information does not have known gaps or omissions

   The existence of gaps or omissions in a data set can undermine information derived from it and any subsequent decision-making. There may be practical reasons why such gaps or omissions exist, for example, device failure or planned maintenance. Appropriate processes should be in place to ensure such issues are addressed in any analysis and service delivery.

8. **Accuracy** – the closeness of the information to the truth

   The presence of biases or errors means the information is not representative of the phenomena or entities it represents. The accuracy of the information can be influenced by the method of collection, the choices available for information items selected from lists, or differences in interpretation of the meaning of terms. When presenting daily statistics, it is important to differentiate between the number of events, incidents, cases occurring on a specific day in comparison to the number occurring over a period. The latter is particularly important where there are measurement and/or reporting delays of more than a few hours.

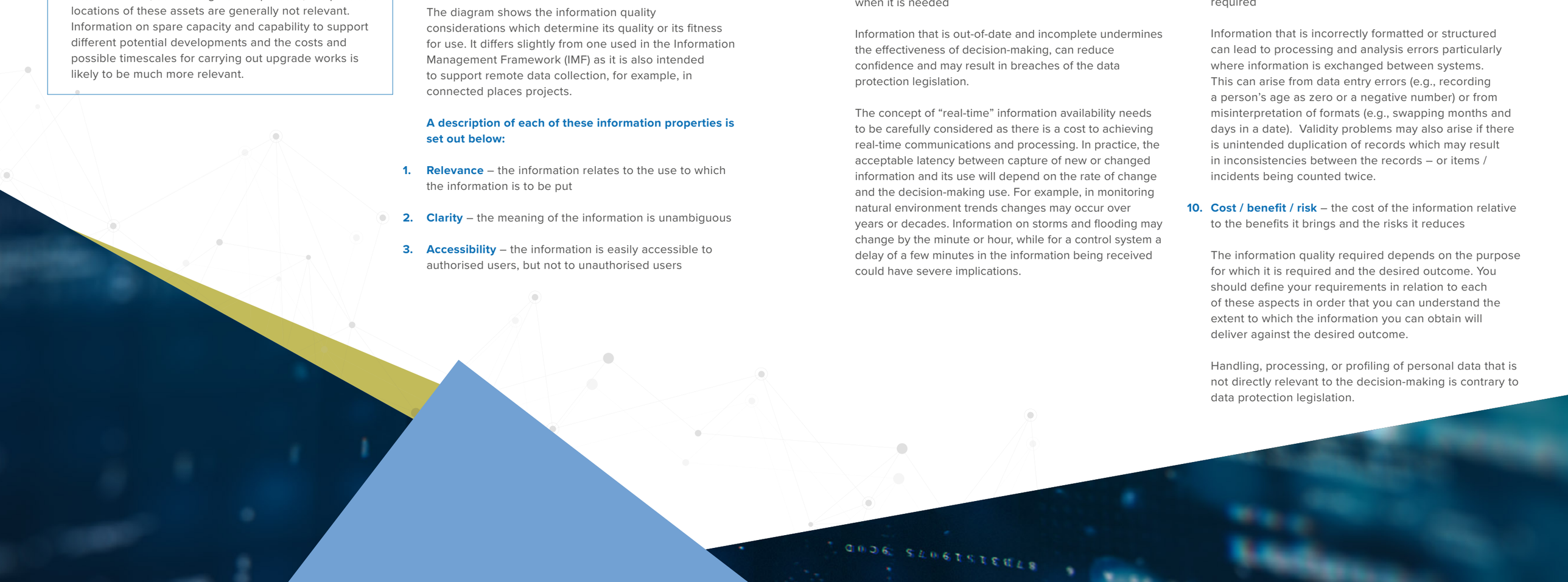
9. **Validity** – the information is within the range and format required

   Information that is incorrectly formatted or structured can lead to processing and analysis errors particularly where information is exchanged between systems. This can arise from data entry errors (e.g., recording a person's age as zero or a negative number) or from misinterpretation of formats (e.g., swapping months and days in a date). Validity problems may also arise if there is unintended duplication of records which may result in inconsistencies between the records – or items / incidents being counted twice.

10. **Cost / benefit / risk** – the cost of the information relative to the benefits it brings and the risks it reduces

    The information quality required depends on the purpose for which it is required and the desired outcome. You should define your requirements in relation to each of these aspects in order that you can understand the extent to which the information you can obtain will deliver against the desired outcome.

    Handling, processing, or profiling of personal data that is not directly relevant to the decision-making is contrary to data protection legislation.



3. DETERMINE INFORMATION SOURCES

The purpose of this stage is to establish where the information of the required quality could, or will, be obtained from and whether any new information will need to be generated.

What are the information sources likely to be?

It will be necessary to establish where the information identified as being required to deliver the outcome will be sourced from so that any constraints can be identified early and taken into consideration.

If you will need information to be provided by a third party, and this information is not already in the public domain, it is advisable to start discussions with any relevant third parties as early as possible. This will allow you to understand and factor into your decision-making any specific requirements they have relating to the processing, storage and consumption of their information. You should also consider the impact to the outcome you are trying to achieve if consent to use the information was withdrawn by the third party. Implementation of formal information sharing agreements can be used to manage risks associated with third party supply and/or use of information. These agreements establish the legal relationships and governance arrangements that can address issues such as cost, liabilities, licencing, etc.

Information may also come from sensors and applications which you or others deploy. Where sensors and/or applications are provided as a service, consideration should be given to how information access is maintained on termination or expiry of the service agreement.

More information on the security-minded deployment of IoT will be available on the CPNI website shortly at: www.cpni.gov.uk/security-minded-approach-digital-engineering.

4. UNDERSTAND ANY SENSITIVITIES

The purpose of this stage is to establish whether any of the information that is needed to achieve the desired outcome is sensitive, for example from a commercial, personal or security perspective, and where it is, to determine the appropriate and proportionate measures that need to be put in place to protect it.

i. Is any of the information sensitive?

Information should be regarded as sensitive if its loss, misuse or modification or unauthorised access can:

- **adversely affect the privacy, welfare or safety of an individual or individuals;**
- **compromise intellectual property or trade secrets of an organisation;**
- **cause commercial or economic harm to an organisation or country; and/or**
- **jeopardise the security, internal and foreign affairs of a nation.**

You should also take into consideration that when information is aggregated, new sensitivities can arise or existing sensitivities increase. For example, compromise of certain types of information on their own may have little or no impact, but when associated with other information, there could be significant implications. Further, the volume of information that is stored together can increase the level of impact that would occur if the information was compromised.

CPNI has developed a triage process that provides a repeatable framework which may be employed to test whether information is sensitive in whole or in part (see www.cpni.gov.uk/security-minded-approach-open-and-shared-data).

If information is being provided by a third party, that party may highlight aspects which are sensitive and require certain measures to be in place to protect that information before it shared. You should make sure you fully understand what information is sensitive and the protection measures required prior to obtaining the information. Typically, the governance and information handling arrangements should be specified in a licence or information sharing agreement.

ii. What risks may result from unauthorised access to, and use of, sensitive information and what security risk mitigation measures are required?

Where any information is identified as being sensitive, it is important to understand the security risks that arise in order for you to determine appropriate and proportionate measures for mitigating them.

Detailed information on understanding the security risks and implementing a security-minded approach is available in the following standards:

- **ISO 19650-5:2020 – Security-minded approach to information management**

  This standard is intended for use by any organisation involved in the use of information management and technologies in the creation, design, construction, manufacture, operation, management, modification, improvement, demolition and/or recycling of assets or products, as well as the provision of services, within the built environment.

- **PAS 185:2017 - Smart Cities – Specification for establishing and implementing a security-minded approach**

  This publicly available specification (PAS) is intended for use by decision-makers in smart cities from the public, private and third sectors as well as smart city data officers. It might also be of relevance to those who are interested in utilising information to deliver smart city objectives.

The following publicly available specification may also be of use:

- **PAS 7040:2019 – Digital manufacturing. Trustworthiness and precision of networked sensors. Guide**

  This PAS is intended for use by organisations that design, build, sell and maintain networked sensors for digital manufacturing and that acquire, integrate and maintain them in operational deployments. The principles set out in the PAS are generally applicable to a number of fields, including smart cities, the built and natural environments.

Further information and guidance is available at www.cpni.gov.uk/security-minded-approach-digital-engineering and

www.ncsc.gov.uk/blog-post/connected-places-new-ncsc-security-principles-for-smart-cities

iii. Are there any privacy implications relating to the information?

It is essential to assess whether any of the information will comprise personal data as defined under the Data Protection Act 2018 as this will impact on what can be acquired and how, as well requirements around information processing, storage, sharing, retention and destruction.

Further information and guidance is available at www.ico.org.uk

In addition to the requirements of the Act, consideration should be given as to whether the data may enable the pattern-of-life of an individual or identifiable group to be determined and thus require protection.

Example A local authority is collecting information about electric vehicle charging to identify areas requiring further provision of chargers. The collected information uniquely identifies individual vehicles and the locations at which they are charged. This builds up a pattern-of-use of some vehicles and the pattern-of-life of the vehicles' drivers or passengers. Measures would be required to suitably anonymise the data and prevent it being de-anonymised using other data sets.

iv. Who will need access to what information?

Where information is sensitive, only those with a genuine need-to-know should be able access it. It is therefore necessary to determine which organisations and types of role will be likely to need access to this sensitive information and for what purpose. For example, in relation to specific sensitive information:

- **who needs to be able to create, update, configure or delete it;**
- **who can see it once it has been processed;**
- **how long does someone need access to the information; and**
- **are tailored views required specific purposes?**

Access to sensitive information should be limited using appropriate role-based access controls, both for information users and those with system administration permissions. The need for access to sensitive information should be regularly reviewed and access removed when no longer necessary. In addition to any security policies or processes governing access, consideration should be given in the technical and information architecture to control measures that separate the more sensitive information, for example, separate storage or other technical means of limiting access.

Example Rather than providing detailed information about the current capacity of utilities by time of day around its brownfield sites, which could provide sensitive information about the usage of neighbouring built assets, the local authority provides a geospatial 'heat map' indicating whether there is a low, medium or high-level capacity available. This provides potential developers with the information that they need but not the sensitive information which, in this case, is irrelevant to the task they are performing.

v. Who should not need access to information?

It is equally important to understand who does not need access to information. You should therefore consider whether there is a legitimate need to make the data more widely available and if so, whether this would be done through:

- **Publication – issued for sale or distribution to the public in printed or digital form**
- **Disclosure – a deliberate, authorised action taken as part of a statutory process, for example, planning application submissions and responses to Freedom of Information requests**
- **Sharing – provision of data from one organisation to another and can involve the reciprocal exchange of data between them.**

Publication allows you to control what information is released and when, but once published you have no control over how people use it.

Disclosure differs from information sharing in that it is not accompanied by, and contingent on establishing, a legal agreement between the parties involved.

Legal agreements that accompany information sharing may impose specific confidentiality requirements and restrictions on information use.

vi. For what unauthorised uses might the information be employed, and what would be the implications?

It is essential that the value of the data is considered from the perspectives of both permitted and unauthorised use. In some cases, its value could be significantly greater to an unauthorised user, or third party, than is perceived by the data owner or legitimate users.

For example, depending on the nature of the data and/or information it may:

- **enable hostile reconnaissance to determine how best to disrupt, destroy or gain unauthorised access to sites or built assets, including neighbouring sites or built assets;**
- **allow determination of how best to sabotage or otherwise interfere with the operation of plant, machinery and systems;**
- **support targeting of individuals or groups to harass, harm or mislead them;**
- **jeopardise commercially or economically valuable intellectual property or provide economic advantage by reducing the time and effort involved in collecting data.**

5. DEFINE INFORMATION MANAGEMENT REQUIREMENTS

The purpose of this stage is to establish the information management processes required to assure information quality and to manage it over its lifecycle.

The generic information lifecycle comprises:

- **Capture – the activity associated with the creation and initial storage of a piece of information, including its metadata.**
- **Maintenance – the activities that deliver the information ready for use in an appropriate form and manner. These activities include: validation and verification; cleansing; reformatting; enrichment; movement; integration; and updating.**
- **Synthesis – the creation of additional information by combining information sources.**
- **Usage – the application of information to activities, functions or tasks.**
- **Archival – the placement of information in an archive where it is stored but where no maintenance, usage or publication occurs.**
- **Publication – the release of information for sale or distribution to the public in printed or digital form.**
- **Purging – the removal of every known copy of a piece of information from an organisation, and its contractors/supply chain, partners, etc. Purging does not necessarily mean destruction of all copies of the information, for example, retention of a copy may be required under the UK Public Records legislation, but all other copies of sensitive information may be purged.**

i. What information management is needed to, at a minimum, meet the defined information quality and information security requirements?

It is important that the information management measures developed can be applied across the lifecycle of the information. The presence of sensitive information in a system or application has implications regarding its lifecycle management. The system or application should therefore be applied in a way that supports the security risk mitigation measures decided upon in Stage 4.

For example, where an application or platform is procured as an outsourced service, the supplier's personnel and their supply chain may have access to the information. In these circumstances appropriate supply chain security[1] measures should be addressed in the contract(s), which should also explicitly state what happens to the information on contract expiry or termination. Where a service is developed, operated, and supported by a public authority, appropriate personnel security measures should be in place for those individuals in high-risk roles, e.g., system administration and development teams with access to the sensitive information.

ii. Once the information has been used for its intended purpose, how long will it be retained and for what purpose(s)?

The retention period and the form the information is in both have implications for how the information is managed, including the security required. It is good information management practice to implement a documented process for the review of retained information, with a view to securely disposing of, or purging, it when no longer required. The periodicity of these reviews should be determined by legal and contractual requirements regarding data retention and audit. For example, personal data related to an identifiable living individual needs to be afforded Data Protection Act 2018 complaint protection over its lifecycle, however statistical data that is anonymised does not require that level of protection. Some information related to built assets may be required for the life of the asset, which can be 60 – 100 years, whilst other information may only be required for a warranty period.

Decisions about the retention period and potential uses of the information therefore need to be taken during the planning period so that it can be retained in a format that maintains its long-term utility and with appropriate and proportionate security in place.

1. Supply chain security - www.cpni.gov.uk/supply-chain-security

iii. How will the information lifecycle be managed, and by whom?

Information ownership should be clearly established prior to the award of any contract or approval of a data sharing agreement. The information owner should specify the accountability and responsibility arrangement for the lifecycle management in any contract or data sharing agreement. The liability of parties handling the information should be clear, including, where appropriate, the flow down through supply chains.

iv. Where is the data and/or information going to be stored and processed?

Service or solution providers will often utilise 'cloud-based' processing and storage platforms which could mean that the information is being processed or stored by that provider at a location of their choice which, subject to contract, may be changed without notification. In many cases the supplier or service provider will not be providing the hardware or software delivering the cloud platform, but will be procuring these from a cloud service provider, who may in turn be further sub-contracting aspects of service delivery. Due diligence[1] should be exercised to understand how and where the information is going to be processed and stored.

6. DOCUMENT OUTCOME OF PROCESS

This final stage involves creating a document summarising the findings and referencing any detailed documentation supporting the analysis. It is recommended that as part of the documentation a process is agreed, and responsibility assigned, for a periodic review of information needs and management. This document should provide a sound basis for project initialization as well as acting a useful reminder of the assumptions that were made at this early stage of the initiative.

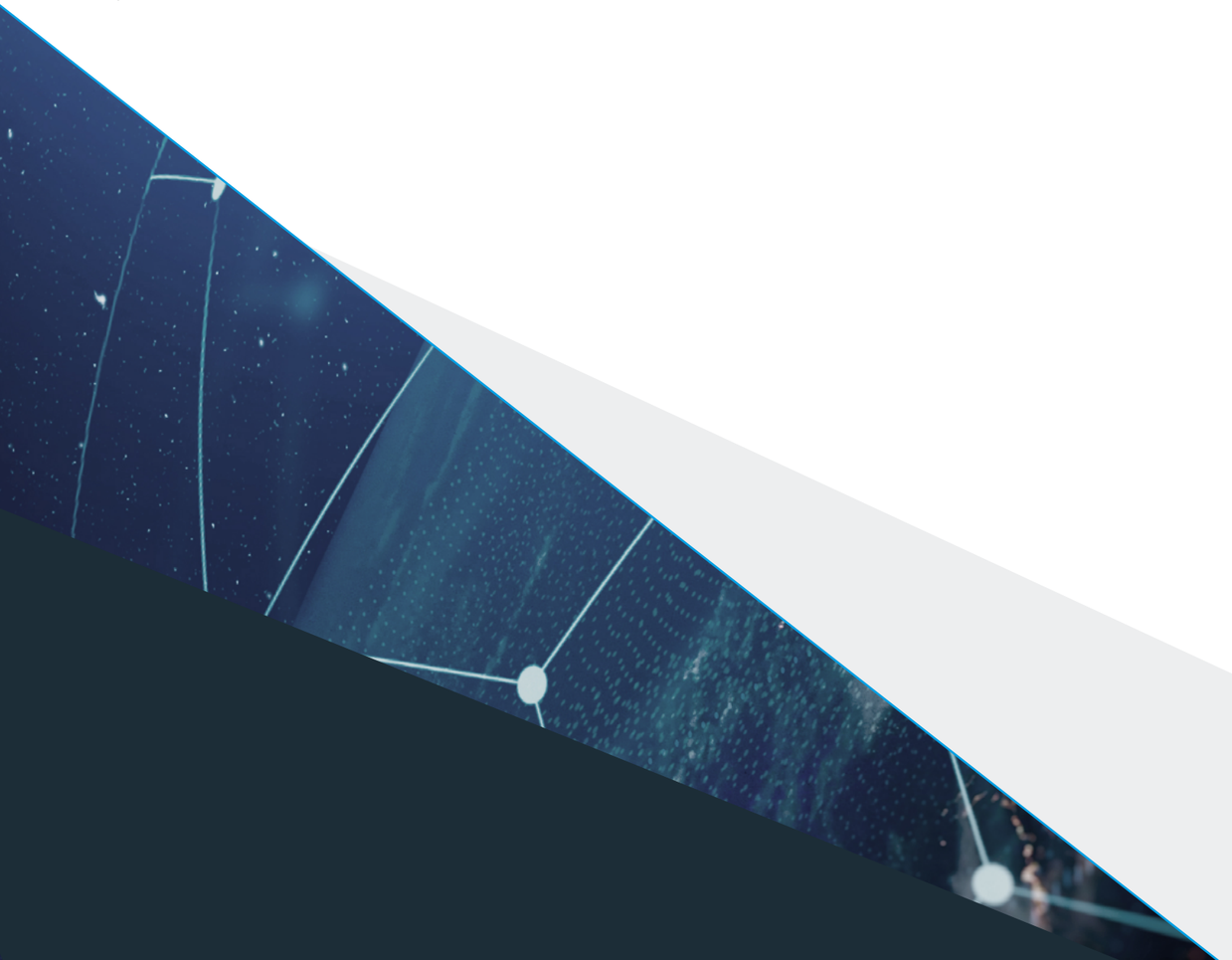
Further information on security-minded handling of data and information is available on the CPNI website at:

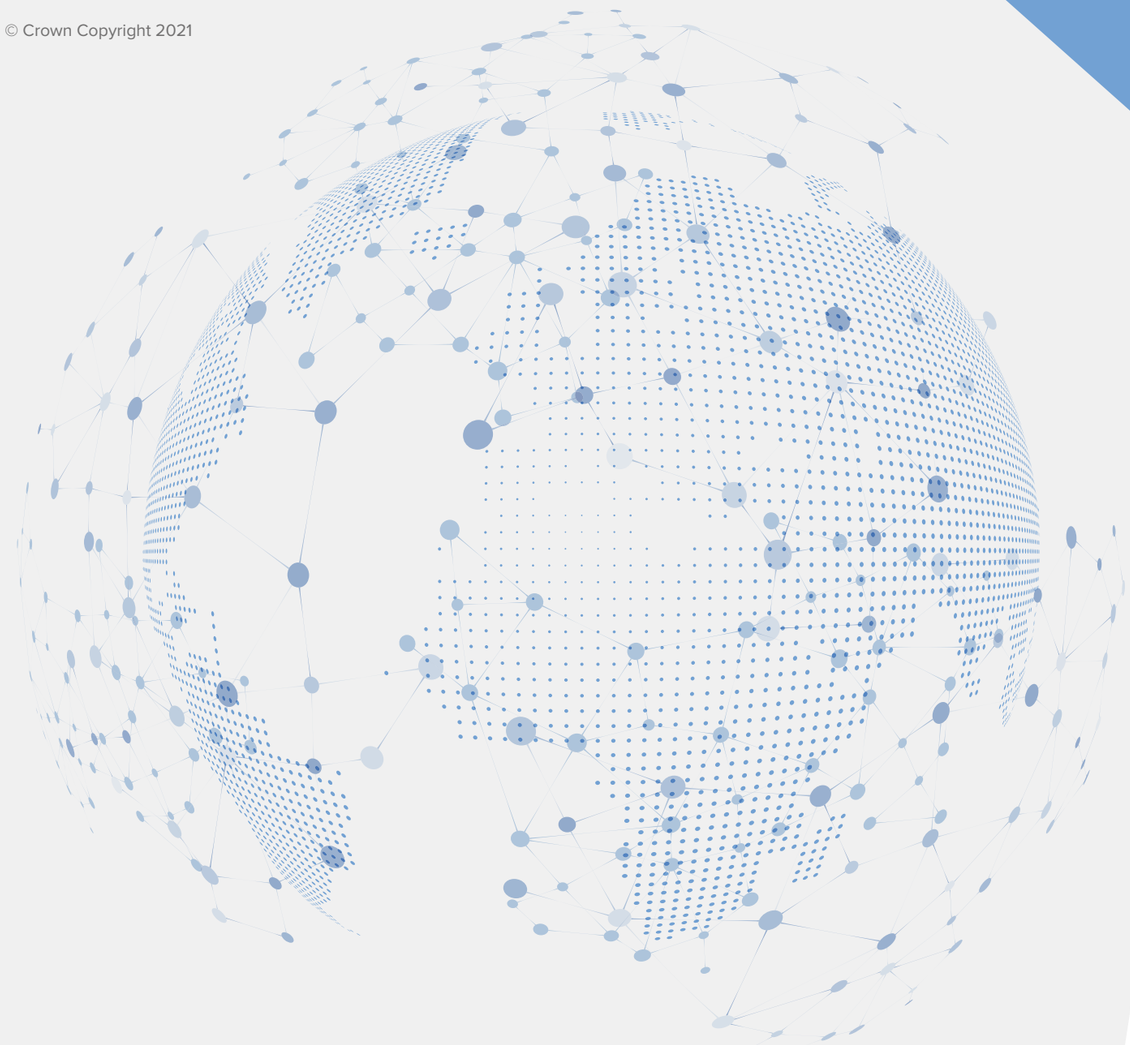
www.cpni.gov.uk/security-minded-approach-open-and-shared-data

www.cpni.gov.uk/security-minded-approach-information-management

www.cpni.gov.uk/security-minded-approach-digital-engineering

www.cpni.gov.uk/security-minded-approach-developing-smart-cities





CPNI
Centre for the Protection of National Infrastructure

DIGITALISATION INITIATIVES
ESTABLISHING HIGH-LEVEL INFORMATION NEED AND MANAGEMENT REQUIREMENTS