# Health and Social Care Cloud Risk Framework

**This material is general guidance only. Recipients are responsible for exercising their own professional judgement in any use of the material.**

**Whilst efforts were taken to ensure that the information contained in this document is both clear and accurate at the time of publication, NHS Digital cannot guarantee that this information will be suitable for the recipient's own hosting and infrastructure requirements, or their procurement/commercial/legal context.**

**Accordingly, NHS Digital accepts no responsibility for any losses or damages arising from the use of this material.**

# 1 Introduction

The purpose of this document is to present a framework for assessing and managing risk around the use of public cloud technologies in the Health and Social Care sectors in England. This framework is intended to be treated as guidance and is recommended to be used by individual Data Controller organisations as they consider the use of public cloud facilities.

There are a wide variety of potential processing activities that may be successfully undertaken with the use of public cloud services, ranging from the use of public cloud to host reports that are intended for public distribution, to data analytics environments containing anonymised data across a region, to national-scale point-of-care clinical systems processing significant quantities of sensitive personal data to support direct care. The use of public cloud to support these scenarios – and indeed the use of any hosting facilities, public or private – can never be risk-free, and the degree of risk varies across these use cases.

Whilst any risk associated with the use of public cloud facilities remains with the Data Controller, this document provides a risk framework that enables a consistent assessment of those risks. This helps organisations to understand where their use of public cloud facilities aligns with their risk appetite.

This document is in five parts:

- The first part details the scope of this paper.
- The second part provides background and context around the need for this guidance.
- The third part provides an overview of risk classes that should be considered as part of each organisation's risk management process.
- The fourth part describes three separate dimensions of data use that need to be considered: the type of data being processed, the scale of the data, and its persistency. The overall risk depends on the degree to which each of these dimensions is applicable to any specific proposed use of public cloud facilities.
- The fifth part provides a model for assessing and managing risk.

# 2 Scope

This document is specifically intended to address:

- The processing of electronic assets that support information systems, including:
    - Data relating to individuals' contact with the health and social care system
    - Data processed across the health and social care sector (including, but not limited to, activities and processes carried out by NHS Digital itself)

Whilst not the primary target of this document, the risk-management principles are also relevant to:

- Board, Commercial, Financial, Contractual, Legal material generated or processed by NHS Digital or other health organisations
- Human Resources: personal data relating to members of staff

# 3 Background and Context

There is appetite across the Health and Social Care system in England to use public cloud computing. These facilities have emerged rapidly in recent years and now provide a cost-effective and agile means of provisioning infrastructure. However, uptake has been restricted, in part due to the lack of guidance on the use of such services, particularly in relation to security.

There is existing cross-government guidance[1] around the use of public cloud facilities. Whilst it provides an overall "permission statement" for the use of public cloud, that guidance is not intended to provide specific approval for the health and social care sector, nor give specific guidance on how to safeguard data.

Individual organisations within the Health and Social Care system hold Data Controller responsibilities and are therefore accountable for the systems they use and for the risk-based decisions that they must take. This document provides a framework that is specifically targeted to health and social care organisations to help them assess and manage the risk of using public cloud.

The framework provided in this document describes the kinds of risk that should be considered, the ways in which risk may be affected by different processing proposals and relates these to an individual organisation's risk appetite. This appetite may, reasonably, vary over time.

The use of this framework is intended:

- To provide more consistency in risk assessment.
- To help identify low risk scenarios which are suitable for initial adoption of public cloud. Over time, we would expect to see greater use of public cloud as we accrue demonstrable experience of safe and acceptable use.

[1] https://www.gov.uk/guidance/public-sector-use-of-the-public-cloud; https://www.ncsc.gov.uk/guidance/implementing-cloud-security-principles

# 4 Risk Classes

This section provides a high-level overview of the risks that should be taken into account when the use of public cloud is being considered. Note that these risk classes are not exclusive to public cloud facilities; rather, they are relevant to all methods of data processing. The well-executed use of public cloud facilities may well *reduce* some classes of risk, compared to traditional on-premise computing environments.

| Risk class | Description |
|---|---|
| Confidentiality | Data may be subject to loss of confidentiality through breach, through unauthorised access, or through unintended or accidental leakage between environments. |
| Integrity | Data may be subject to loss of integrity through data loss or unintended manipulation. |
| Availability | Ensuring that access to your data is available when required. Network connectivity to cloud becomes a critical dependency and there is a risk of introducing a Single Point Of Failure (SPOF). Public cloud cannot be assumed to be permanently available; cloud availability and SLA must match the need. |
| Impact of breach | We cannot assume there can never be any breach, so we need to consider the *impact* of any unintended breach (unauthorised disclosure into an uncontrolled, or less-well-controlled than intended, environment) |
| Public perception | There is some degree of public concern over the use of public cloud given that these are widely available, shared, computing environments. |
| Lock-in | Flexibility may be impacted (resulting in increased levels of lock-in) by:<br>• the adoption of a specific public cloud provider's unique services.<br>• the difficulties involved in migrating large quantities of data may make it difficult, in time and/or cost, to migrate to an alternative in the event of future commercial or service changes.<br>• an architecture that is not sufficiently tailored to a public cloud model. |

Table 1

# 5 Dimensions that affect risk

The impact of risk is considered along three dimensions: data type, data scale and data persistence.

## 5.1 Data Type

The type of data being processed impacts risk. At the extremes: from managing reports intended for public distribution, to maintaining extremely sensitive PKI secrets. To support the range of potential types of health and social care data, we describe a health and social care classification scheme. To take advantage of cross-governmental principles around data classification, we also provide a mapping to the Government Security Classifications policy.

The Government Security Classifications Policy[2] came into force in April 2014, providing a policy that describes the classification of information assets into one of three high-level types, and provides a baseline set of security controls for each. It is intended to be used for all information assets across government departments, agencies, public sector delivery partners and the wider supply chain. The three high-level types are: OFFICIAL, SECRET and TOP-SECRET, with OFFICIAL having an additional handling caveat of OFFICIAL-SENSITIVE. Additional descriptors are possible to further classify assets. A few descriptors are provided as core (of which the main relevant to us are COMMERCIAL, PERSONAL), although it is permissible to introduce others, supported by local policies and business processes.

In addition, existing classification schemes are mostly concerned with securing various information assets, whereas we also need to consider the distinction along different axes: for example, how data can and should be shared and the legal bases for its processing. Public perception and potential concern is also heightened in the health sector, which needs to be taken into account when defining the approach to, and controls applied to the handling of health data assets. Statements as to how healthcare information is processed also exist, either as department policy (e.g. DH offshore processing policy), NHS Digital processes and practices (e.g. how Spine 2 is operated), or as part of existing commercial arrangements (both national, such as GPSoC, or local, such as Trusts' supplier contracts).

The new scheme described in this paper provides a health and social care sector-specific framework upon which an appropriate and proportionate set of security controls can be applied, dependent on the specific needs of different kinds of health and social care data. It is required because existing data classification schemes do

[2] https://www.gov.uk/government/publications/government-security-classifications

not achieve the level of granularity required to cover the variety of different data types that are processed across the health and care system, and there are specific needs and complexities in the processing of health data.

Note that additional controls can be added to any data-type in order to reduce any associated risks. For example, to address concerns regarding the Confidentiality, and integrity of data, such data may be separately encrypted before transfer to the cloud, using strong cryptography as defined by the current version of NIST SP800-57 and where the encryption keys are not stored with the cloud provider. In such circumstances the risk profile associated with the data being processed on public cloud is significantly mitigated.

The proposed data classification scheme is illustrated in Table 2 below:

| Type | | Description | Example |
|---|---|---|---|
| Publicly available information | | Statistical material that is intended for public distribution. Identification from these materials, with or without any other materials, is not feasible. | The number of diabetics in Sheffield, or location information for health-care providers. |
| Synthetic (test) data | | Synthetic (test) data is fictional data, engineered to be representative of real data, that is created in order to avoid the need to use real data when developing and testing IT systems. Synthetic data must pose zero risk of contributing to the revealing of any personal data. | Fabricated dummy HES data set, used for testing purposes, risk assessed to ensure that there is no risk of the data contributing to the access to any personal data. |
| Aggregate data | | Summarised and anonymised data, but which is not suitable for public distribution, for example due to the risk that it may be used with other material to contribute to the re-identification of individuals. The risk of such re-identification is not necessarily significant, but does exist (especially in the presence of a sustained and skilled attack). | Summarised records of activity of a particular hospital. |
| Already encrypted materials | | Materials that are already encrypted before they touch the cloud, using strong cryptography as defined by the current version of NIST SP800-57 and where the encryption keys are not stored with the cloud provider. | Scanned hospital patient notes which are encrypted by an application before being uploaded to the cloud for archive purposes |
| Personal Data (PID) | | Information about an identified individual | |
| includes | Demographic data | Information about the individual rather than their clinical details | A person's address details and NHS Number |
| includes | High Risk demographic data | Demographic data where, in the event of a breach, there is a high risk of significant harm | The address details of a person under the care of the UK Protected Persons |

| | | | | |
|---|---|---|---|---|
| | | | | Service[3], likely to be reflected in an S-flag applied to their PDS details. |
| | Personal confidential data (PCD) | | PCD is based on the ICO definition of sensitive personal data, extended within health and social care to include:-<br>• deceased persons<br>• information that is given in confidence and is owed a duty of care, such as:<br>○ Social care records / child protection / housing assessments<br>○ DNA / finger prints<br>○ Bank / financial / credit card details<br>○ National Insurance number / Tax, benefit or pension records<br>○ Travel details (for example at immigration control, or Oyster records)<br>○ Passport number / information on immigration status / travel records<br>○ Work record or place of work / School attendance / records | A person's medication history |
| | includes | Legally-restricted PCD | Sensitive personal data that are subject to additional regulations or statute, under the<br>• Gender Recognition Act 2004[4],<br>• Human Fertilisation & Embryology Act 2008[5] | Details of a person's previous gender |
| | includes | Extra-delicate PCD | Sensitive personal data that are sometimes seen to be additionally delicate, but for which there are no legal restrictions. This determination is often not consistent, but is commonly-held, and is often related to conditions that attract, or are considered to attract, stigma. For example, HIV status, mental health conditions, other conditions contained within the SCR "sensitive code" list. Whilst many patients see information on these kinds of condition to be particularly private | Details that a person has asked not to be shared |

[3] http://www.nationalcrimeagency.gov.uk/about-us/what-we-do/specialist-capabilities/uk-protected-persons-service
[4] http://www.legislation.gov.uk/ukpga/2004/7/pdfs/ukpga_20040007_en.pdf
[5] http://www.legislation.gov.uk/ukpga/2008/22/pdfs/ukpga_20080022_en.pdf

| | | | | |
|---|---|---|---|---|
| | | | and not to be shared under any circumstances, others see them as important to share, and for any stigmas to be removed. Note that there is no legal distinction between PCD and Extra-delicate PCD. | |
| Anonymised data | | | Sensitive personal data that has been subject to de-identification and/or other privacy-enhancing techniques, in line with the ICO Anonymisation Code of Practice. Risk of re-identification is remote (and would be based on activities that are illegal and/or break contractual arrangements). No way of authorised linking with other data-sets. | Extract from a research database where all pseudonyms have been removed |
| Pseudonymised data | | | Sensitive personal data that has been subject to de-identification and/or other privacy-enhancing techniques, in line with the ICO Anonymisation Code of Practice, containing a pseudonym that allows for linking with other data-sets where that is permitted through business justification and legal basis. Otherwise, risk of unauthorised re-identification is remote (and would be based on activities that are illegal and/or break contractual arrangements). | HES data set |
| includes | Reversibly pseudonymised data | | Pseudonymised data where the pseudonym is also intended to be used to facilitate re-identification where that is supported by business purpose and legal basis. | Data dissemination to support risk stratification (where individuals may subsequently be usefully re-identified to support their direct care) |
| includes | Irreversibly pseudonymised data | | Pseudonymised data where re-identification is not intended. | Data dissemination to support a research project that never requires re-identification |
| Patient account data | | | Account credentials (including any recovery materials) for citizen accounts for patient-facing online health tools | A person's account details for NHS Choices |
| Patient choices | | | Statements / preferences made by patients regarding the use of their data | A person's expressions of their wishes recorded in their GP's clinical system or on the Spine |
| Patient meta-data (identifiable) | | | Information about how identified patients have used patient-facing online health tools | History of an identified person's use of NHS Choices' symptom information |
| Patient meta-data (linkable) | | | Information about how patients have used patient-facing online health tools (not identified, but linkable across sessions) | History of an unknown (but linkable) person's use of NHS Choices' symptom information |

| | | | |
|---|---|---|---|
| Professional user account data | | Account credentials (including any recovery materials) for professional user (e.g. clinician, health professional, etc) accounts that control access to any personal data (including PCD) | A Clinical Application logon. |
| Professional account data (less-sensitive) | | Account credentials (including any recovery materials) for professional user (e.g. clinician, health professional) accounts that control access to anonymised information | Authentication details to portal providing access to anonymised data. |
| Audit data | | Data that records the use of a system and the provenance of the data that system manages | Clinical system audit trail |
| includes | Professional user meta-data | Information about how users have used clinical or administrative tools that process personal data | History of a GP's use of their clinical system, or of SCR |
| includes | Audit data (personal) | Data describing the use of a clinical or administrative system that processes personal data, where that audit data itself includes or references PCD | The audit trail of a GP system showing all users' interactions and use of the system |
| includes | Audit data (non-personal) | Data describing the use of a clinical or administrative system, where that audit data itself does not include or reference PCD | History of logins to a clinical system |
| Key Materials | | Material that provide long-lived linkage between reversibly-pseudonymised data and personal data, or provides a similarly significant security function | Look-up tables or decryption keys |
| includes | Very short-lived | One-time decryption keys | A decryption key generated to support (and only usable within) a specific re-identification activity within an individual user session |
| includes | Rotatable | Material that provide linkage between reversibly-pseudonymised data and personal data, that persists over time and over user sessions but is generally rotatable | An encryption key used by a DSCRO to re-identify pseudonyms included in many data disseminations |
| includes | Long-lived, persistent | Material that provide long-lived and persistent linkage between reversibly-pseudonymised data and personal data, or provides a significant security function | A root certificate private key for a widespread PKI |

Table 2

Table 3 below provides an agreed[6] mapping between the health data types and the Government Security Classification Policy. This enables us to take advantage of cross-government policy statements and published principles (such as the 14 NCSC

[6] Agreed by the Healthcare Cloud Working Group, including NHS Digital, NHS England, DH, GDS.

Cloud security principles[7]) around that classification, whilst treating those statements as necessary but not necessarily sufficient in a health and social care context.

| Type | | | Map to Govt. Security Classification | Notes |
|---|---|---|---|---|
| Publicly available information | | | *No applicable mapping* | The most obvious mapping is to something like UNCLASSIFIED but this is no longer part of the model |
| Synthetic (test) data | | | OFFICIAL | |
| Aggregate data | | | OFFICIAL | |
| Already encrypted materials | | | OFFICIAL | |
| Personal Data (PID) | | | OFFICIAL-SENSITIVE | |
| includes | Demographic data | | OFFICIAL-SENSITIVE | |
| includes | High Risk demographic data | | OFFICIAL-SENSITIVE | |
| includes | Personal Confidential Data (PCD) | | OFFICIAL-SENSITIVE | |
| includes | include | Legally-restricted PCD | OFFICIAL-SENSITIVE | |
| includes | include | Extra-delicate PCD | OFFICIAL-SENSITIVE | |
| Anonymised data | | | OFFICIAL-SENSITIVE | |
| Pseudonymised data | | | *Maximum of variants* | |
| Includes | Reversibly pseudonymised data | | OFFICIAL-SENSITIVE | |
| Includes | Irreversibly pseudonymised data | | OFFICIAL-SENSITIVE | |
| Patient account data | | | OFFICIAL-SENSITIVE | |
| Patient choices | | | OFFICIAL-SENSITIVE | |
| Patient meta-data (identifiable) | | | OFFICIAL-SENSITIVE | |
| Patient meta-data (linkable) | | | OFFICIAL-SENSITIVE | |
| Professional user account data | | | OFFICIAL-SENSITIVE | |
| Professional user account data (less-sensitive) | | | OFFICIAL-SENSITIVE | |
| Audit data | | | *Maximum of variants* | |
| Includes | Professional user meta-data | | OFFICIAL-SENSITIVE | |
| Includes | Audit data (personal) | | OFFICIAL-SENSITIVE | |
| Includes | Audit data (non-personal) | | OFFICIAL-SENSITIVE | |
| Key Materials | | | *Maximum of variants* | |
| Includes | Very short-lived | | OFFICIAL-SENSITIVE | |
| Includes | Rotatable | | OFFICIAL-SENSITIVE | Whilst we need such data to be treated to the highest standards, they do not fit into the government policy criteria for SECRET or TOP-SECRET. |
| Includes | Long-lived, persistent | | OFFICIAL-SENSITIVE | |

Table 3

[7] https://www.ncsc.gov.uk/guidance/implementing-cloud-security-principles

Whilst we can (mostly) demonstrate an appropriate mapping from health data type to the Government Security Classification Policy, there are some limitations that emerge:

- Many data types map to OFFICIAL-SENSITIVE, but there are many kinds of data in this category that we will control, and disseminate, in different ways
- We cannot, through the Government Security Classification Policy, indicate the very highly sensitive NHS materials such as PKI secrets as needing any greater control than many other kinds of information

## 5.2 Data Scale

There are two dimensions when considering scale: taking account of the depth (e.g. scope of data for any one individual) and the breadth (e.g. how many individuals are included).

**For depth**, data should be treated the same whether there is a single data item that causes a particular classification to apply, or whether there are many.

**For breadth**, the scale[8] is:

| Scale | Description | Example scale |
|---|---|---|
| Extra-Small (XS) | Very low volume | less than 10,000 records or events |
| Small (S) | Local scale, such as an individual Trust | between 10,000 and 1m records or events |
| Medium (M) | Regional scale, such as county or ACO | between 1m and 5m records or events |
| Large (L) | National scale | over 5m records or events |

Table 4

This approach recognises the difference in potential harm given the scale of breach; it provides a wider recognition of very large datasets that are commonly processed across the health system (both inside and outside of NHS Digital). However, it is recognised that this banding is still somewhat artificial, requiring a degree of judgement.

[8] Note that the scale in question here is the number of patients / events, and how that is analogous to geographic indicators of scale, not the specific geographic spread of a particular data set: for example, a data-set covering 1000 individuals spread across the country represents less risk than a data-set covering 1 million individuals in a specific city.

## 5.3 Data Persistence

A public cloud facility can be used to process data in many ways, ranging from, at one extreme, processing that requires long-term persistence of data, to the opposite extreme where data may be purely transient (is never persisted). The range of levels that is used in Part 5 is as follows:

| Persistency | Description | Example |
|---|---|---|
| Persistent | Data is deliberately placed into persistent physical storage (for example using databases or file stores) for long term / indefinite use. | Clinical System holding long-lived patient clinical information |
| Temporary | Data is deliberately placed into persistent physical storage (for example using databases or file stores) for a short-defined period, typically for a specific project. | Dissemination environment providing access to national pseudonymised data to support a specific research project |
| Cached | Data may be persisted into persistent physical storage as part of the required processing but it is kept only to support time-bound transactions, rather than long-term | Message queue |
| Transient | Data transits the facility but is never intentionally persisted out-of-memory | Web interface capturing data that is immediately transferred outside of public cloud |

Table 5

Note that transient data is not risk-free: rather, different risks exist depending on the level of data persistence. In general, the level of overall risk reduces between persistent and transient.

# 6 Risk Framework

Two important aspects to risk management are risk analysis and risk appetite. With those aspects measured, a consistent approach to managing risk can be taken.

## 6.1 Risk Appetite

Risk appetite may be influenced by a number of factors; for example:

- The degree to which an organisation believes it may be subject to challenge, perhaps as a result of public fears over the ways in which personal data is processed (see Part 3 above)
- The degree to which an organisation wishes to "play safe" in its use of public cloud facilities, or alternatively is comfortable in operating at the "cutting edge"
- The available budget: a constrained budget will, other factors being equal, drive additional use of elastic public cloud facilities
- The degree of risk associated with an organisation's existing infrastructure services

## 6.2 Risk analysis and management

This section describes a risk analysis tool that is based upon the three dimensions described in Part 4 above, to aid in a consistent approach to the assessment of risk in the use of public cloud facilities.

Note:

- This is not intended to be an overly prescriptive model; rather it is to inform an organisation's assessment and promote consistency within an individual organisation and across organisations
- It is assumed that the controls in place by any selected public cloud facility satisfy the NCSC Cloud Security principles[9] and that therefore such use is "well-executed" as described in recent guidance from GDS[10]
- When assessing processing scenarios, consider the most sensitive aspect where there is more than one involved

The Risk Framework tool is available separately. An initial impact score is assigned to each data type. That score is then scaled separately by the scaling factors assigned to each measure of scale and persistence, resulting in a "Risk Impact Score" value.

[9] https://www.ncsc.gov.uk/guidance/implementing-cloud-security-principles
[10] https://www.gov.uk/guidance/public-sector-use-of-the-public-cloud

The tool maps the generated Risk Impact Score to one of five “Risk Profile Levels”. This provides an overall view that reflects the “degree of risk or contentiousness” of the described use of public cloud.

In general, all potential uses – and risks – should be weighed against the benefits of public cloud facilities (for example in terms of cost, time-to-launch, flexibility, etc). The Risk Appetite Level may also be subsequently affected by a privacy-enhancing technique, or additional controls, that are added to a processing environment (at either an infrastructure or application layer, or both).

Following those steps, the table below provides an overview of the impact of the resulting level in terms of governance. This takes into account the degree to which, at present, there is relatively little use of high-profile public cloud take-up across health and care, but with the expectation that, over time, we would expect to modify these expectations given greater experience.

| Risk Profile Level | Governance Expectation |
|---|---|
| Class I | All organisations are expected to be comfortable operating services at this level. |
| Class II | Whilst there may be some concerns over public perception and lock-in, most organisations are expected to be comfortable operating services at this level. |
| Class III | At this level, risks associated with impact of breach become more significant, and the use of services at this level therefore requires specific risk management across all risk classes described in Section 4, requiring approval by CIO / Caldicott Guardian level. |
| Class IV | At this level, it is likely to become more difficult to justify that the benefits of the use of public cloud outweigh the risks. However, this case may still be made, requiring approval by CIO / Caldicott Guardian, and would be required to be made visible to the organisation’s Board. Specific advice and guidance may be provided by NHS Digital on request. |
| Class V | Operating services at this level would require board-level organisational commitment, following specific advice and guidance from NHS Digital. |

Table 6

In addition, the Risk Profile Level drives the level of controls that are required to be implemented by the public cloud provider. The description of these controls is provided separately.

# 7 Document Control

## 7.1 Copyright



## 7.2 Related References, Links and Documents

These documents will provide additional information:-

| |
|---|
| NHS and social care use of public cloud services |
| Health and Social Care Cloud Security Good Practice Guide |
| Health and Social Care Data Risk Model |