UKEOF
ENVIRONMENTAL OBSERVATION FRAMEWORK

The principles of good data and information management

Why is good data and information management important?

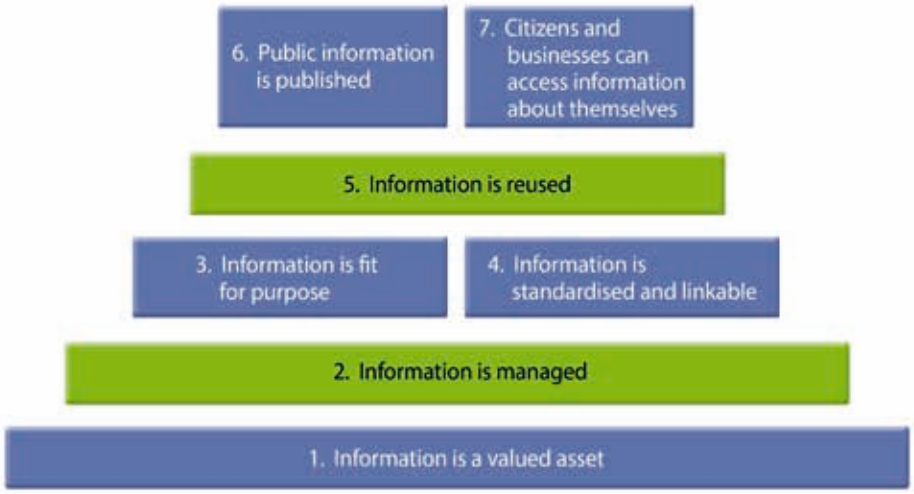
Data and information[1] are critical for making the right evidence-based decisions, whether for research, policy development or business investments, or to ensure accountability to parliament and the general public. Government departments and agencies collect, generate, store and use vast amounts of data, some of which have been obtained over a long period and at considerable cost.

Clear policies and procedures will help data and information managers look after and value data and information, whatever the media. This Advice Note provides guidance on how to do this well. It is based on the publication in 2012 of *Information Principles for the UK Public Sector* by the Cabinet Office.

ADVICE NOTE 1

[1] *Data and Information is sometimes collectively referred to in this document as 'information' in the sense of 'Information' in Information Technology (or the 'I' in IT).*

The principles of good data and information management



The Chief Information Officers Council has identified seven principles which build naturally into a hierarchy, as depicted in the diagram above. For example, it is unlikely that information can be reused (Principle 5) unless it is also valued, managed, fit for purpose and standardised (Principles 1-4).

Principle 1

Information is a valued asset

Information should be understood and valued as much as other organisational assets such as buildings, machinery, people or money.

This principle is the foundation for what follows and highlights the need for information to be valued in the same way as these other types of asset. It is important to note that the full value of information lies not just in its original purpose but in its potential to be reused for other purposes.

Principle 2

Information is managed

Information should be managed – stored, protected and exploited – according to its value.

Data and information managers need to consider the whole lifecycle of the information, from identification of need, creation, quality assurance, maintenance, reuse and ultimately to archiving or destruction once the information has ceased to be useful.

A range of best practices need to be in place, for example to ensure appropriate availability and integrity, avoid loss and ensure continuity across technology upgrades. It is particularly important that personal data are adequately protected.

Information also needs to be governed as it moves through its lifecycle, for example to make sure it's always clear who is responsible for it (ie an identifiable owner), and to comply with relevant legislation and regulation. The consistent assessment and ownership of these information risks is another important consideration when managing data and information.

The organisational culture must support best practice in data and information management, and make sure everyone responsible for processing these business assets is professionally qualified and appropriately skilled. This principle therefore also includes the processes, roles, responsibilities, training, and organisational structure and culture needed to ensure the effective and efficient use of information.

Principle 3

Information is fit for purpose

Information must be good quality and fit for both its primary purpose and potential secondary uses. It will not always be possible for the originator to foresee secondary uses, so it is important that the quality of the information is communicated consistently so future users can decide if it is suitable.

Quality includes factors such as accuracy, validity, reliability, timeliness, relevance

and completeness. The quality of data and information should also be regularly monitored to ensure that they at least meet the levels that have been assessed as necessary for their purposes.

A further aspect of this principle is to consider aligning the supporting technical platform and format with how information will be used. For example, if information is likely to be needed for online statistical analysis, it won't be appropriate to store it in a system or format that is only accessible to the originator, on back-up tapes or unstructured PDF format.

This principle doesn't require information to be perfect, only that it is the right quality for its intended use and that its quality characteristics are clear to future users.

Principle 4

Information is standardised and linkable

There will be many more opportunities for exploiting information if it is available in standardised and linkable forms.

Standardisation is important for structured information such as dataset definitions, and unstructured information such as metadata tags applied to documents. Standardisation within an organisation is important for staff to fully exploit the information; if an organisation uses widely accepted open standards it will unlock even more value for other users.

Standardisation is important both for the way information is recorded and in the way concepts are defined:

- Format, eg date always being entered as yyyy-mm-dd
- Content, eg forename, surname, address, etc.
- Concepts, eg defining roles such as patient, offender, learner, claimant, driver

Even further value can be unlocked if information can be linked. A good example is document references and citations that allow the reader to draw on a wealth of associated information (this is the basis of the 'world wide web'). A similar concept can be applied to structured data, based on an understanding of the relationships between items and the use of consistent identifiers to reference authoritative sources (the basis of the 'semantic web'). For example, tagging spending information with an authoritative code for the organisation involved would allow it to be unambiguously linked with details of the organisation itself and third-party information about that organisation (eg service satisfaction measures).

Principle 5

Information is reused

Information is even more valuable if it can be used more than once or for more than one purpose. A good data manager will proactively look for opportunities for reuse.

These could include:

- Internal reuse – making the most of information for its primary purpose and identifying secondary uses. For example, operational data can sometimes be reused to support performance improvement or research.
- External reuse – sharing information with other organisations, either within the public sector or with private businesses and citizens.
- Holding master data – ensuring an organisation's data is the only authoritative source for business information (eg an authoritative list of organisation codes), which is nominated, maintained and promoted as such.

Reuse involves considering what information an organisation can make available to others, and looking at how an organisation might reuse information held by others.

Whilst this principle strongly encourages reuse, it is important to appreciate that reuse does require a careful risk-based judgement to be made with regard to exploiting versus protecting information, as well as consideration to the costs and benefits involved, and any rights or other commercial considerations.

Information which initially appears unsuitable may be reusable if it can be reformatted. For example, operational information that identifies individuals can be 'anonymised' or aggregated and then be of wider value. Also, in cases where the partner organisation is known beforehand, concerns over security or privacy can sometimes be mitigated by means of negotiation, joint-working and data-sharing agreements.

Principle 6

Public information is published

Public information includes the objective, factual information on which public services run and are assessed, on which policy decisions are based, or which is collected or generated in the course of public service delivery. Public information should be published wherever practicable, unless there are overriding reasons not to.

This principle goes beyond adhering to minimum legal requirements and advocates a proactive approach to presenting, formatting and promoting information in useful formats for wider consumption, without it needing to be specifically requested or mandated in legislation.

Consider the different channels available to publish information to the public. This includes internal publication processes, the use of publication hubs such as data.gov.uk and relationships with third-party 'information intermediaries' such as commercial publishers.

The benefits of publishing information should be balanced against possible risks and sensitivities, such as information which could compromise individuals' privacy, commercial and legally privileged information, and information that is required to maintain security.

Principle 7

Citizens and businesses can access information about themselves

Citizens and businesses should be able to access information about themselves, along with an explanation of how that information is used by others. This may be either on request or, preferably, by making it available by default. In effect, such information should be considered as belonging to the citizen, although entrusted to the care of a public body.

This principle goes beyond minimum legal requirements. It advocates a proactive approach which makes it easy for citizens to access information about themselves, without having to make a request and even when access is not mandated in legislation. This might be achieved, for example, by making it securely available online. Information managers need to consider how this can work in practice, to enable users to view information and perform transactions, for example correcting inaccuracies.

Clearly the desire to publish information does need to be balanced against constraints which may prevent this. Exclusions would include, for example, legally privileged information and information that is required to maintain security.

Summary

Data and information cannot be shared effectively and used openly unless they have been managed throughout their life cycle as valuable assets. The above principles provide a firm foundation for doing this.

UKEOF
ENVIRONMENTAL OBSERVATION FRAMEWORK

www.ukeof.org.uk

This is a series of advice notes prepared by UKEOF's Data Advisory Group.

UKEOF works to improve coordination of the observational evidence needed to understand and manage the changing natural environment. It is a partnership of public sector organisations with an interest in using and providing evidence from environmental observations. Contact us at office@ukeof.org.uk.