

6 DATA ACCESS AGREEMENTS

6.1 Overview

Issues relating to data access are some of the most challenging, yet important, for an information network to manage. The challenge is rooted in the legitimate concerns of both custodians and users, as summarised in Box 4. These concerns are held by individuals, project groups, organisations and governments, and cover a wide range of issues. Some of these are genuinely important,² but others are largely spurious or secondary to other issues, for example mistrust of user intentions or lack of understanding of what the user requires.

Unless fully addressed, such concerns have the potential to hinder data access and, hence, reduce interest in cooperation. One way forward is to accompany transactions with formal agreements between parties, offering tangible assurances that the concerns of both sides will be met. A key principle is that data should be made accessible **through the custodian**, or their nominated distribution outlet, not from a secondary source.

Data access agreements need not be viewed as a defensive measure intended to limit access to data. Rather, they are a **positive means of increasing trust and fair dealing** amongst the network's or other partners. As confidence grows and data begin to be mobilised more easily, the need for formal agreements diminishes until, ultimately, they may no longer be required.

Data access agreements are perceived to be difficult to negotiate, but this need not be the case. A useful step is for the network hub to draft a **generic agreement** for distribution and adaptation by custodians. In the interests of simplicity, agreements should assume goodwill on the part of users, not bad faith. Similarly, administrative and cost impediments to data access should be kept to an absolute minimum.

2 Custodians need to ensure that data which are genuinely sensitive for reasons of privacy, confidentiality or security, are adequately protected. An example would be detailed descriptions of the locations of threatened species that are at risk from exploitation.

Custodians are responsible for advising users on the potential uses of a dataset. They may also recommend specific permitted uses or, conversely, excluded uses of a dataset. This may occur when, for example, the custodian believes the data are unsuitable for certain purposes due to uncertainties or ambiguities in the dataset. It also provides a means for custodians to safeguard their intellectual property. Varying conditions may be applied to different classes of user (e.g. government, NGO, research, commercial). Typical provisions to consider when drafting a data access agreement are presented in Box 5.

Box 4 Typical concerns over data access

Custodians

- Will the dataset be misused?
- Will intellectual property be respected?
- Will the cost of supply be recovered?
- Who will be liable in the event of a problem?
- Will professional credibility suffer from the release of the dataset?
- Could confidentiality be breached by the release of the dataset?
- Does the transaction comply with internal network guidelines?

Users

- Is the dataset fit for its intended use?
- Will it be available at the right time?
- Will onerous conditions be imposed on its use?
- Is it available in a form which can be easily handled?
- What, if anything, will it cost?

In the event that incorrect data are provided by a custodian, or they are used for an inappropriate purpose, liability could fall on the owner of the data, the custodian, a third party which has provided the data, or all of these. The situation is most serious when 'negligence' is detected, for instance when it is established that data were poorly maintained or falsely documented. Exposure to liability is an emerging issue. The most likely grounds for liability are where dataset documentation is misleading, deceptive or negligent, or where there has been a violation of the provisions of a data access agreement (see Onsrud 1989 for a discussion).

Box 5 Typical provisions in a data access agreement

- Whether data are available for single or multiple uses, and what those uses are.
- Whether data can be retained by the user following its designated use (risking independent updating or duplication), or whether they should be destroyed (incurring subsequent costs when the data are needed again).
- Whether the data are available for non-profit or profit-generating use and, if the latter, whether any royalties would flow back to the custodian.
- Whether users may give the data to third parties (the latter would normally be referred back to the custodian unless prior consent was arranged).
- Whether copies of any publications, products and other outputs derived (even in part) from the data should be sent to the custodian.
- What form of acknowledgement the user should employ (in cases where data have been interpreted several times before use, a long list of acknowledgements — known as an 'audit trail' — may be necessary).
- Some form of disclaimer that protects the custodian from legal liability in the event that their data prove to be unreliable or are used for an inappropriate purpose (legal advice should be obtained).

6.2 Roles of the Hub, Custodians and Users

The primary role of the hub is to enable the network's partners to realise the benefits of cooperation. In terms of data access, these benefits comprise value to the user through access to essential data; value to the custodian for being of service (paving the way for future exchanges and access to value-added products); and value to the network for having enabled the transaction.

The hub can help to realise these benefits by facilitating the negotiation and management of data access agreements. This could be achieved through the preparation of a generic agreement embodying the principles of cooperation it wishes to invoke, i.e. that **environmental data are available to all stakeholders**, for non-commercial purposes, at minimum cost and with minimum administrative or other impediments to access and use, whilst protecting the legitimate interests of custodians.

In cases where access to an important dataset is not being provided satisfactorily, the hub may wish to scrutinise the restrictions imposed by the custodian and recommend new operating procedures. Experience has shown that many restrictions imposed by custodians are unnecessary, even arbitrary, and can impose severe constraints on data access.

The hub's advice will need to be consistent with government policies relating to information access and exchange. Where current policies are inconsistent with the principle of wide information availability, with environmental decision-making likely to be less effective as a consequence, network hubs have a role in advising the government accordingly.

Custodians are responsible for developing policies on data access which are consistent with the broader policy frameworks in which they operate. They need to ensure that the provisions in their data access agreements are **necessary to protect their legitimate interests**, and do not have the unintended consequence of inhibiting reasonable use of the data by other users.

Users are obliged to **comply with conditions** prescribed by custodians. This is crucial if the spirit of trust, so necessary to the successful operation of a network, is to be built and maintained. Users may also provide feedback to custodians in the form of advice of any errors or deficiencies encountered in the data, and an indication of their future requirements, enabling custodians to continually improve their service.

Finally, there are two responsibilities which users would ordinarily uphold even if not specified by agreement. Where practicable, copies of any information products derived fully or in part from a dataset are provided to the custodian; and the source(s) of the data used should be **fulsomely acknowledged**.

6.3 Cost Recovery

Very few organisations have a corporate objective to collect data merely to sell data. Data-rich organisations are generally in the business of selling the **value they add** to data. In other words, the collection and management of data is a means to a larger corporate end. As such, the cost of building and managing data is a cost of doing business or, more simply, a business overhead. Certainly these costs need to be recovered by selling products and services to clients, but not necessarily by charging total costs to external data users.

When raw data are provided to a user, without any significant added value, then the question arises as to what proportion of the data management overhead to charge to that user. Normally, there is no argument that the marginal costs, to the custodian, of providing those data can be charged to the user; the issue is what is a **reasonable price** for data that already exist (particularly where collected at public expense), but which were not created for the user or in the expectation that the user would require them.

Unfortunately, the development of information networks is sometimes inhibited by market-driven ideologies which authorise the cost recovery in data transactions. The consequence of this is that organisations which could contribute to the resolution of important environmental concerns are unable to do so because they lack the resources to purchase the relevant data. The resources available to environmental organisations, particularly NGOs, seldom reflect the significance to the community of the issues at stake. This is largely due to a failure of the market to internalise environmental costs and benefits.

Within a network, **cost barriers should be kept as low as practicable** to facilitate data exchange. However, depending on the financial positions of the organisation concerned, and the types of user who require access to the data, different strategies for cost recovery may be applied. Not all of these are monetary-based, as illustrated in Box 6.

One way of facilitating data access in hostile policy environments (e.g. over-bureaucratic or market-driven) is to develop datasets in partnership with other organisations, or to barter data and other services (see Volume 4). There is no reason why such arrangements should not include commercial partners.

Box 6 Strategies for cost recovery

- Supply data totally free.
- Supply data free to reciprocating users (i.e. barter).
- Recover immediate cost of supply (i.e. time, consumables, energy etc. to process the request).
- Recover incremental costs (i.e. of managing the data and combating depreciation).
- Recover development costs (i.e. the initial investment in building the dataset).
- Recover 'market value' (i.e. the maximum cost which the user will pay).