CaribData Decision Guide

Data Sharing for Data Producers in Small Islands

Selvi Jeyaseelan, Ian R Hambleton

2026-01-26

Table of contents

1 How to use this decision guide **2**

2 Start with the data you actually have **2**

3 Clarify the purpose before choosing how to share **2**

4 Think about risk in real-world terms **3**

5 Consider the available sharing models **3**

6 Aim for minimum viable sharing **4**

7 Recognise when not sharing is the right decision **4**

8 Document decisions and plan to revisit them **4**

9 Regional considerations for small states **5**

10 Worked decision scenarios **5**
10.1 Scenario 1: Sharing national health survey microdata 5
10.2 Scenario 2: Sharing administrative education data 6
10.3 Scenario 3: Sharing disaster impact data during an emergency . 6

11 Appendix A: Data Sharing Decision Worksheet **7**
11.1 Section 1: Dataset overview 7
11.2 Section 2: Purpose of sharing 7
11.3 Section 3: Risk and harm 8
11.4 Section 4: Sharing options considered 8
11.5 Section 5: Chosen approach 8
11.6 Section 6: Decision record 9

1 How to use this decision guide

This guide is designed to support *real-world judgement*. It does not offer rules, thresholds, or automatic answers. Instead, it provides a structured way to think through data sharing decisions in small island developing states (SIDS), where scale, risk, and capacity shape what is realistic and responsible.

You might use this guide when:

- responding to a request for data,
- considering proactive data sharing,
- designing a new data system, or
- reviewing an existing sharing arrangement.

The guide can be read from start to finish, but it is equally useful as a reference when a specific decision arises. A reusable *Decision Worksheet* is provided in Appendix A and can be used to document decisions and support consistency over time.

2 Start with the data you actually have

Data sharing decisions should always begin with the dataset itself, not with abstract commitments to openness or reuse.

At this stage, it helps to describe the data in simple terms:

- What type of data is this (survey, administrative, operational, research)?
- Who or what does it describe?
- How detailed is it (individual-level, aggregated, geographic)?
- How current is it, and how often is it updated?
- How confident are you in its quality and completeness?

In SIDS, datasets are often small but information-rich. This combination increases both their potential value and their potential risk. Being explicit about what the data are—and what they are not—prevents later confusion and supports defensible decisions.

If you cannot clearly explain the dataset to a non-specialist colleague, it is usually too early to consider sharing it.

3 Clarify the purpose before choosing how to share

The most important data sharing question is *why*, not *how*.

Sharing should always be linked to a clear purpose. Common purposes include:

- informing national or regional policy,
- supporting research,

- improving service delivery,
- enabling coordination during emergencies, or
- increasing transparency and accountability.

At this stage, consider:

- Who is requesting the data, if anyone?
- What question are they trying to answer?
- What decision, action, or insight would the data support?
- Is this use time-sensitive or ongoing?

Different purposes justify different sharing models. A request framed broadly or vaguely should prompt further clarification. A clear purpose can shape the sharing decision.

4 Think about risk in real-world terms

Risk in data sharing is best understood as the *potential for harm*, not as a technical property of the data alone.

Harms may affect:

- individuals (loss of privacy, stigma, legal consequences),
- communities or groups (labelling, discrimination),
- institutions (loss of trust, reputational damage), or
- the data system itself (misuse, loss of confidence).

In small island contexts, risk is often amplified by:

- small populations,
- strong local knowledge,
- close social networks, and
- high visibility of institutions and decisions.

Risk assessment does not require precision. It requires *reasonable judgement* about likelihood, severity, and reversibility. Importantly, risk also exists in *not* sharing data: poor decisions, duplication of effort, and external misrepresentation are real harms that must be weighed alongside disclosure risk.

5 Consider the available sharing models

Once the data, purpose, and risks are understood, the next step is to consider the full range of sharing options.

Common models include:

- open release,
- aggregated or summary outputs,
- registered access,

- controlled or restricted access,
- time-limited or purpose-specific access, and
- not sharing (for now).

There is no default or preferred option. The aim is not maximum openness, but *appropriate openness*. Listing the options explicitly helps avoid the false choice between "share everything" and "share nothing".

6 Aim for minimum viable sharing

A particularly useful principle in SIDS is **minimum viable sharing**.

This means sharing only what is needed to meet the purpose, and no more. It recognises that partial sharing is often good practice, not a compromise.

Minimum viable sharing might involve:

- sharing indicators rather than raw data,
- reducing geographic or demographic detail,
- applying time limits or embargoes,
- requiring analysis plans in advance,
- involving local review of outputs.

This approach allows data to be used while keeping risk proportionate and manageable.

7 Recognise when not sharing is the right decision

Not all data should be shared, and not all requests should be approved.

Legitimate reasons to pause or decline include:

- high risk with no feasible safeguards,
- unclear or weak public benefit,
- poor data quality or incomplete documentation,
- legal or ethical constraints, or
- lack of capacity to manage sharing responsibly.

A decision not to share is not a failure. It should, however, be documented and revisited. Circumstances change, and today's "no" may become tomorrow's "yes".

8 Document decisions and plan to revisit them

Data sharing decisions should be *visible, consistent, and defensible.*

Documentation protects both individuals and institutions by showing that decisions were considered, proportionate, and grounded in context. In small teams, documentation also preserves institutional memory.

At a minimum, record:

- the purpose of sharing,
- risks considered,
- options assessed,
- the chosen approach and safeguards, and
- a review date.

9 Regional considerations for small states

Regional data sharing offers both opportunity and risk for SIDS.

Pooling data can:

- increase analytical power,
- strengthen regional visibility, and
- reduce duplication of effort.

But it can also:

- increase disclosure risk,
- dilute local context, and
- shift interpretation away from national priorities.

Regional sharing works best when it:

- respects national control,
- includes local expertise,
- is transparent about use and interpretation, and
- delivers visible benefit back to participating countries.

10 Worked decision scenarios

10.1 Scenario 1: Sharing national health survey microdata

A Ministry of Health has conducted a nationally representative health and lifestyle survey. The dataset includes age, sex, district, education, self-reported conditions, and health behaviours. The survey was expensive and will not be repeated for several years.

A regional research group requests access to the individual-level data to study noncommunicable disease risk across several Caribbean countries.

Decision process

The data are individual-level and sensitive, with indirect identifiers and small district counts. The purpose—regional public health research—is legitimate and aligned with national priorities. Risks include re-identification and misinterpretation without local context.

Several options are considered. Open release is ruled out as high risk. Aggregated tables alone would not support the proposed analysis.

The ministry chooses minimum viable sharing:

- national reporting is completed first,
- microdata are de-identified and geographic detail reduced,
- access is granted under a data use agreement,
- analysis plans are reviewed in advance, and
- outputs are shared with the ministry prior to publication.

The decision is documented and set for review after two years.

What this shows
Controlled, time-limited sharing can deliver public benefit while managing risk.

10.2 Scenario 2: Sharing administrative education data

A Ministry of Education maintains an administrative dataset covering enrolment, attendance, examination results, and special educational needs. The data include school identifiers and small subgroups.

A national NGO requests access to analyse educational inequality by district.

Decision process

The data concern children and include strong indirect identifiers. Risks include identification of students or schools and potential stigmatisation.

The ministry determines that the purpose can be met without sharing microdata. Instead, it provides:

- aggregated district-level indicators,
- suppression of small cell counts,
- agreed interpretation notes, and
- shared responsibility for public communication.

What this shows
Aggregation is a powerful sharing tool, especially for sensitive administrative data.

10.3 Scenario 3: Sharing disaster impact data during an emergency

Following a major hurricane, a national emergency agency compiles rapid data on housing damage, displacement, and service disruption. The data are incomplete

and evolving.

Regional partners request access to support coordination and funding appeals.

Decision process

Speed is critical, but data quality is variable. Premature public release could cause confusion or undermine trust.

The agency decides to:

- share provisional summaries with regional partners,
- clearly label data as preliminary,
- update figures regularly, and
- delay public release of granular data.

What this shows
Preliminary data are a valid state, and controlled sharing can precede public release.

11 Appendix A: Data Sharing Decision Worksheet

This worksheet is intended to support and record judgement. It can be adapted to local context.

11.1 Section 1: Dataset overview

Table 1: Dataset Overview

Item	Notes / Completion
Dataset name	
Data type	
Population covered	
Key variables	
Level of detail	
Data quality concerns	

11.2 Section 2: Purpose of sharing

Table 2: Purpose of Sharing

Item	Notes / Completion
Requestor	
Intended use	

Item	Notes / Completion
Public interest or benefit	
Time sensitivity	

11.3 Section 3: Risk and harm

Table 3: Risk and Harm

Item	Notes / Completion
Potential individual harms	
Potential group or community harms	
Institutional or reputational risks	
Re-identification risk (low / medium / high)	

11.4 Section 4: Sharing options considered

For each option, note suitability and concerns.

Table 4: Sharing Options

Option	Suitability, concerns, or reasons
Open release	
Aggregated outputs	
Registered access	
Controlled access	
Time-limited access	
Not shared	

11.5 Section 5: Chosen approach

Table 5: Chosen Approach

Item	Notes / Completion
Selected sharing model	
Safeguards applied	
Conditions of use	

Item	Notes / Completion
Duration or review date	

11.6 Section 6: Decision record

Table 6: Decision record

Item	Notes / Completion
Decision-maker(s)	
Date	
Rationale summary	
Review schedule	
Notes	