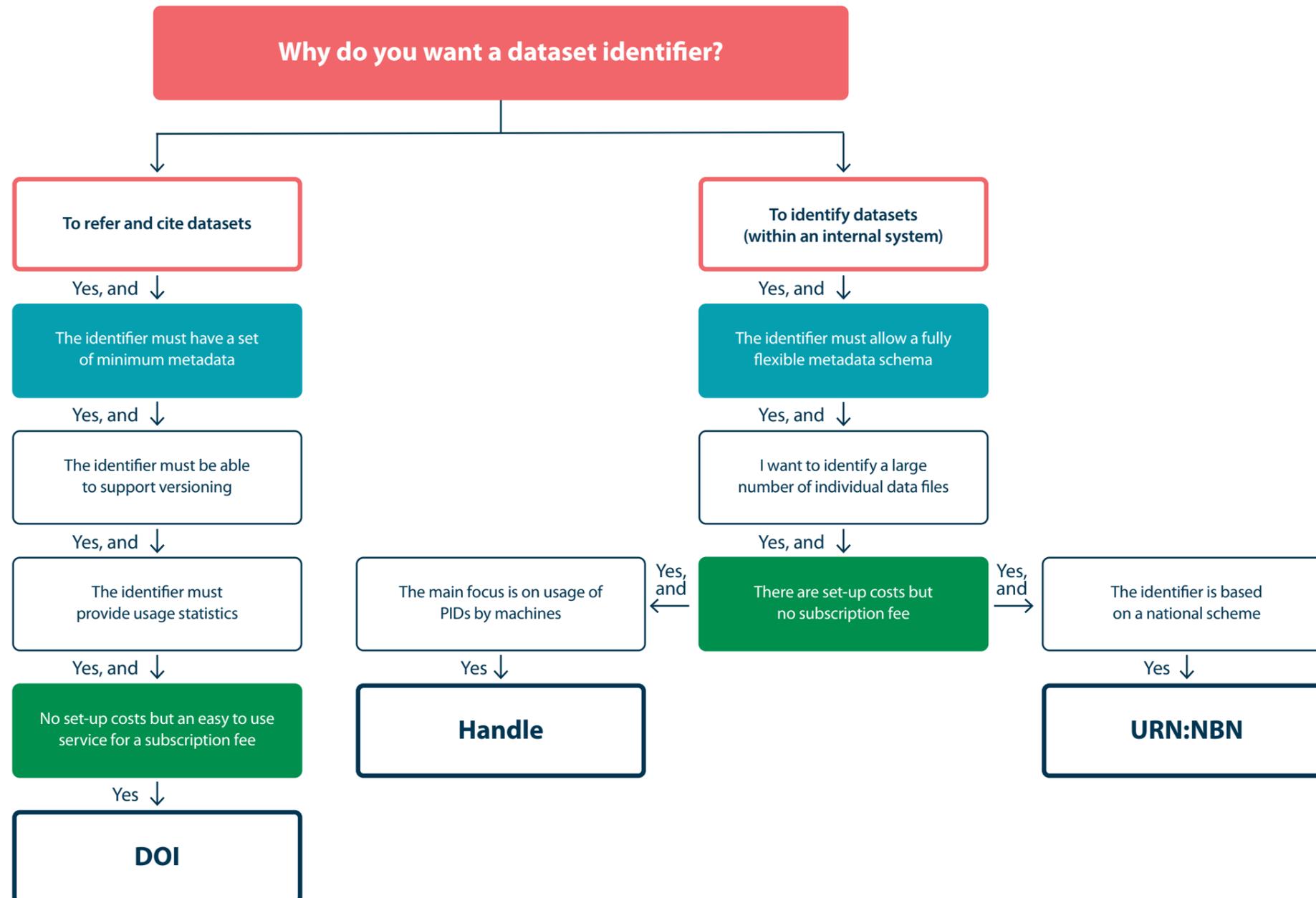# Dataset Identifiers

A dataset is a collection of data. The data can be one or more tables, documents or files. This diagram provides some initial guidance on selecting a dataset identifier.

This decision tree addresses inbuilt/'out of the box' functionality. Identifiers such as Handle and URN:NBN can be configured to adopt much of the functionality of DOIs but this requires technical resources.

**Why do you want a dataset identifier?**

**To refer and cite datasets**

Yes, and ↓

The identifier must have a set of minimum metadata

Yes, and ↓

The identifier must be able to support versioning

Yes, and ↓

The identifier must provide usage statistics

Yes, and ↓

No set-up costs but an easy to use service for a subscription fee

Yes ↓

**DOI**

**To identify datasets (within an internal system)**

Yes, and ↓

The identifier must allow a fully flexible metadata schema

Yes, and ↓

I want to identify a large number of individual data files

Yes, and ↓

The main focus is on usage of PIDs by machines

Yes, and ←

There are set-up costs but no subscription fee

Yes, and →

The identifier is based on a national scheme

Yes ↓

**Handle**

Yes ↓

**URN:NBN**

# Dataset Identifiers

A dataset is a collection of data. The data can be one or more tables, documents or files. This diagram provides some initial guidance on selecting a dataset identifier.

This decision tree addresses inbuilt/'out of the box' functionality. Identifiers such as Handle and URN:NBN can be configured to adopt much of the functionality of DOIs but this requires technical resources.

## URN:NBN

National persistent identifier infrastructure based on the Uniform Resource Name (URN), that is a Uniform Resource Identifier (URI) that uses the URN scheme as a namespace and in which the implementation is based on National Bibliographic Number for which the National Library is the Registration Agency.

*Format: urn:nbn:nl:ui:13-y0h-pdg*

## DOI

A Digital Object Identifier is a persistent identifier or Handle used to identify objects, based on an ISO standard. The developer and administrator of the DOIs is the International DOI Foundation (IDF). The DOI system is implemented through a federation of registration agencies, such as DataCite, coordinated by the IDF. DataCite provides services to create, find, cite, connect, and use datasets and other research objects.

*Format: 10.17026/dans-x5j-6ruy*

## Handle

The Handle system (www.handle.net) was developed by Corporation for National Research Initiatives (CNRI). It is a framework for managing digital information and provides a naming scheme for unique identifiers, called 'Handles'. A resolution system translates the handles into location-related data. A centrally administered registry service manages the resolving naming authorities. A Handle consists of two parts: a naming authority and a unique string that identifies an object. The ePIC consortium provides PID services for the European research community for the allocation and resolution of persistent identifiers.

*Format: 11304/69544d65-3ef6-45ca-84a6-04152313872e*

## For more information on dataset identifiers

- FREYA Knowledge Hub page on Identifiers for Publications and Data
  *https://www.pidforum.org/t/pids-forpublications-and-data/297*
- D3.1 Survey of Current PID Services Landscape
  *https://doi.org/10.5281/zenodo.3554254*