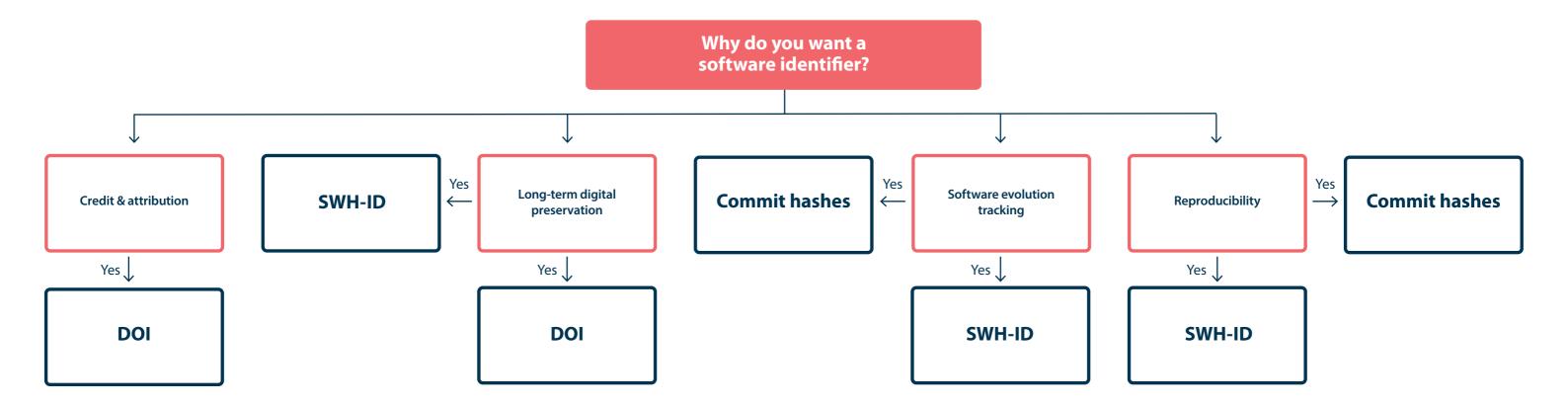# Software Identifiers

A decision tree to distinguish between the different types of persistent identifiers typically used for software to use in a research data repository, software development platform or other type of administrative system. The tree covers various considerations which were identified by FREYA partners as effecting their choice of identifier. Some identifiers are suitable for more than one purpose and some identifiers can be implemented in a very specific way and have similar coverage. Therefore, some of these options repeat.

Why do you want a software identifier?

Credit & attribution

SWH-ID ← Yes ← **Long-term digital preservation**

Commit hashes ← Yes ← **Software evolution tracking**

Reproducibility — Yes → **Commit hashes**

Yes ↓ **DOI**

Yes ↓ **DOI**

Yes ↓ **SWH-ID**

Yes ↓ **SWH-ID**

Software Identifiers

A decision tree to distinguish between the different types of persistent identifiers typically used for software to use in a research data repository, software development platform or other type of administrative system. The tree covers various considerations which were identified by FREYA partners as effecting their choice of identifier. Some identifiers are suitable for more than one purpose and some identifiers can be implemented in a very specific way and have similar coverage. Therefore, some of these options repeat.

Commit hashes

Commit hashes are commonly used internal identifiers for software commits using Git, but their persistence greatly depends on the preservation of services like GitHub. An example for commit hashes are SHA-1 hashes.

Format: 35e32b6a00dec02ae7d7c45c6b7106779a124685

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

SWH-ID

The SoftWare Heritage persistent IDentifier is used by the Software Heritage Archive to identify software artifacts. SWH-IDs are cryptographic intrinsic hashes computed on the entire set of artifact properties to form a Merkle structure. They can refer to the software content but also to specific revisions, releases, snapshots, or directories.

Format: swh:1:cnt:94a9ed024d3859793618152ea559a168bbcbb5e2

For more information on Software 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