



# CESSDA ERIC Persistent Identifier Policy 2019

## Principles, Recommendations and Best Practices

Version 2.0

<https://doi.org/10.5281/zenodo.3611327>

Status: Public (CC-BY 4.0)  
Author: Hausstein, Brigitte (GESIS), Leader of the CESSDA PID Project; Horton, Laurence (Faculty of Information, University of Toronto)  
Date: 22 January 2020  
Document: CESSDA ERIC Persistent Identifier Policy 2019. Principles, Recommendations, and Best Practices  
Version: Final – Approved January 2020; v2.0  
Including input by Working Groups, CESSDA Main Office and Metadata Office  
Including feedback from consultation of CESSDA Service Provider Forum



## 1 Introduction

### 1.1 Purpose of the document

This document contains a revised version of the CESSDA PID Policy 2017. It includes the updated CESSDA PID Policy Principles and provides the latest recommendations and Best Practice examples. It addresses all CESSDA Service Providers (SPs).

### 1.2 Related Documents

CESSDA ERIC Persistent Identifier Policy. Version 1.0. 22 November 2017.  
<https://doi.org/10.5281/zenodo.3611317> (previous version of the PID Policy)  
CESSDA ERIC Persistent Identifier Policy. Best Practice Guidelines. Version 1.0. 22 November 2017. <https://doi.org/10.5281/zenodo.3611324>  
CESSDA ERIC Checklist for the Usage of Persistent Identifiers. Version 1.0, 2019.  
<https://doi.org/10.5281/zenodo.3611333>

### 1.3 Review

The ownership of this policy is with CESSDA. This document will be reviewed every two years by the CESSDA SPs.

## 2 Background

Data repositories must facilitate ways to identify and locate data. This poses complex technical and organisational challenges for data providers. PIDs provide a path to accessing data as well as a means for referencing and citing data sets. PIDs are a prerequisite for sustainable and reliable discovery and reuse. Providing and maintaining PIDs is a critical service a certified, and trustworthy data repository will offer.

PIDs are also an advertisement for data integrity, providing proof an object has not changed, or if it has, how. Additionally, PIDs help data repositories comply with FAIR data principles (findable, accessible, interoperable, reusable) set by FORCE 11 and provide future-proofing in case an archive relocates its holdings.<sup>1</sup>

The main task of CESSDA and its service providers is to provide documented, verifiable, and understandable data for research. One way of doing this is to assign a PID to a data set (and if desired to other parts of an object). The PID accompanies a specific version of the data set that allows tracking of which version is disseminated and gives users a simple way to cite the data creator and which version was used.

Therefore, CESSDA ERIC has adopted a clear policy with respect to persistent and unique identifiers that provide value to researchers, and are practical to implement across the CESSDA SPs. The first version together with Best Practice Guidelines, was published in 2017.

---

<sup>1</sup> FAIR Data Publishing Group: <https://www.force11.org/group/fairgroup>



### 3 Principles and Recommendations

#### 3.1 Principle 1: Identifying

**CESSDA SPs have to use a globally unique PID to identify their data holdings.**

The main PID systems<sup>2</sup> for data collections are **DOI<sup>®</sup>, Handle, URN and ARK**. DOI<sup>®</sup> and Handle are based on the same (Handle) system. DOI<sup>®</sup> and Handle are the most commonly used PIDs.

**CESSDA ERIC requires the use of PIDs.**

The following PID systems are accepted by CESSDA Tools and Services:

- 1) DOI
- 2) Handle (including *ePIC-handles*)
- 3) URN
- 4) ARK

This is the definitive list for fulfilling principle 10 of the CESSDA Data Access Policy.

#### Recommendations:

- CESSDA SPs commit to permanently maintaining the PID.
- An assigned PID always resolves to the same object.
- The PID should be embedded within the data file (e.g., as a variable).
- When choosing a PID system and a PID service provider, **credibility, long-term viability**, and use of **open standards** must be addressed. CESSDA SPs should use PID service providers with **clear and transparent** policy and business models.

#### Examples:

PID	Example PID name
ARK	<a href="http://n2t.net/ark:/13030/tf5p30086k">http://n2t.net/ark:/13030/tf5p30086k</a>
DOI <sup>®</sup>	<a href="https://doi.org/10.5878/002645">https://doi.org/10.5878/002645</a>
Handle	<a href="http://hdl.handle.net/11858/00-097C-0000-000E-011B-8">http://hdl.handle.net/11858/00-097C-0000-000E-011B-8</a>
ePIC Handle	<a href="http://hdl.handle.net/21.11998/0000-001A-3862-7">http://hdl.handle.net/21.11998/0000-001A-3862-7</a>
URN:NBN	<a href="http://urn.fi/urn:nbn:fi:fsd:T-FSD2117">http://urn.fi/urn:nbn:fi:fsd:T-FSD2117</a>

#### 3.2 Principle 2: Locating

**Data holdings of CESSDA SPs shall be findable online through a PID.**

#### Recommendations:

- A PID assigned by a CESSDA SPs shall resolve to a **landing page** at the website of the CESSDA SPs.
- This landing page **includes information about the data set**, such as its origin, version, availability and accessibility. It shall indicate if a newer version exists and preferably include information on differences between versions.

---

<sup>2</sup> For detailed information on available PID systems please see: <https://www.dpconline.org/handbook/technical-solutions-and-tools/persistent-identifiers>.



- The landing page also includes information about **related data collections** and (if known) their PID.

#### Examples:

PID	Landing Page
<a href="http://urn.fi/urn:nbn:fi:fsd:T-FSD2117">http://urn.fi/urn:nbn:fi:fsd:T-FSD2117</a>	<a href="https://services.fsd.uta.fi/catalogue/FSD2117?lang=en&amp;study_language=en">https://services.fsd.uta.fi/catalogue/FSD2117?lang=en&amp;study_language=en</a>
<a href="https://doi.org/10.5878/002350">https://doi.org/10.5878/002350</a>	<a href="https://snd.gu.se/sv/catalogue/study/snd0739/001">https://snd.gu.se/sv/catalogue/study/snd0739/001</a>
<a href="https://doi.org/10.17026/dans-x7r-395c">https://doi.org/10.17026/dans-x7r-395c</a>	<a href="https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:56502">https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:56502</a>

### 3.3 Principle 3: Resolving

**CESSDA SPs shall use global PID services that ensure the resolvability of PIDs.**

#### Recommendations:

- Use a trusted and well-known PID system running an uninterrupted, reliable global resolver.
- Guarantee **resolvability** of PID by maintaining **up-to-date landing pages** even if the data collection is no longer accessible.

#### Examples

DOI<sup>®</sup>, Handle, URN, and ARK are trusted global PID systems that comply with the CESSDA PID Principle 3 (Resolving).

### 3.4 Principle 4: Referencing and Citation

**PIDs should be used to ensure the referencing and citation of data.**

#### Recommendations

- Include the PID in information on **how to cite** the data.
- **Include the PID** in a recommended citation standard.

#### Examples

- Recommended citation in the GESIS Data Catalogue: de Winter, Lieven; Bryder, Tom; Linch, Philip; Norris, Pippa; Chavel, Cecile; Weßels, Bernhard; Kielhorn, R.; Marsh, Michael; Sola, Giorgio; Gardella, Louisa; Thomassen, Jacques; et al. (1999): European Candidates Study 1994. GESIS Data Archive, Cologne. ZA3077 Data file Version 1.0.0, <https://doi.org/10.4232/1.3077>
- Recommended citation in the UK Data Archive catalogue: University of Essex, Institute for Social and Economic Research. (2018). Understanding Society: Waves 1-8, 2009-2017, and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 11th Edition. UK Data Service. SN: 6614, <http://doi.org/10.5255/UKDA-SN-6614-12>



### 3.5 Principle 5: Visibility

**PIDs must be included in resource-discovery metadata provided by CESSDA SPs for the CESSDA Data Catalogue (CDC).**

#### Recommendations

- Provide PIDs in **resource-discovery** and **harvestable** metadata.
- Implement **version control** and assign a new PID to major revisions of a data collection.
- If a data collection is no longer accessible, a **tombstone landing page** must be maintained by the CESSDA SPs, including citation metadata and information about the data provider.

#### Examples

- **Versioning**  
GESIS (three-digit version number: major.minor.revision)  
Data collection: German General Social Survey - ALLBUS 2010  
Version 1.0.1. <https://doi.org/10.4232/1.11782>  
Version 1.0.0 <https://doi.org/10.4232/1.11692>  
<https://dbk.gesis.org/dbksearch/sdesc2.asp?no=4612&db=e&doi=10.4232/1.11782>
- **Tombstone page**  
Scientific Use File German Ageing Survey (SUF DEAS) 1996, Version 2.0  
<https://doi.org/10.5156/DEAS.1996.M.002>  
Migrants to Stockholm 1880-1925  
[https://snd.gu.se/en/catalogue/study/SND0236.\(https://doi.org/10.5878/000143\)](https://snd.gu.se/en/catalogue/study/SND0236.(https://doi.org/10.5878/000143))



## 4 Definitions

**Unique identifiers** are individual references resolvable on a global scale allowing large numbers of unique identifiers.

**Persistent identifiers** resolve as HTTP URIs with support for content negotiation, and these HTTP URIs should be persistent.

**Data holdings:** Data collections for which the CESSDA SPs are responsible.

**Resolution:** A PID must be actionable, meaning that the identifier leads the user to information about the object. This is also called the resolution of an identifier.

## 5 References

CESSDA ERIC Persistent Identifier Policy. Version 1.0. 22 November 2017.

<https://doi.org/10.5281/zenodo.3611317>

CESSDA Persistent Identifier Policy. Best Practice Guidelines. Version 1.0. 22 November 2017.

<https://doi.org/10.5281/zenodo.3611324>

Consortium of European Social Science Data Archives (CESSDA). STATUTES of CESSDA ERIC. Updated version 13 May 2019.

<https://www.cessda.eu/content/download/1466/20924/file/STATUTES%20ERIC%20CESSDA%20UPD%2013.05.19.pdf>

Consortium of European Social Science Data Archives. CESSDA Data Access Policy. Bergen, June 2016.

<https://www.cessda.eu/content/download/963/8608/file/CESSDA%20Data%20Access%20Policy.pdf>

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014. <https://www.force11.org/group/joint-declaration-data-citation-principles-final>.

## 6 Acknowledgments

This document was created by Brigitte Hausstein and Laurence Horton. It is based on the CESSDA ERIC Persistent Identifier Policy (version 1.0; <https://doi.org/10.18448/16.0040>) which was compiled by the CESSDA PID Task Force: Brigitte Hausstein (lead) and Kerrin Borschewski (both from GESIS), Birger Jerlehag (SND), René van Horik and Lilian van der Vaart (both from DANS) in 2017.

The work has been supported by CESSDA since 2015.