



---

dharma

## File Naming Conventions

Axelle Janiak & Adeline Levivier

Release Version 1.2, 2022-02-07

## Versioning History

Author(s)	Version	Changes	Date
Axelle Janiak & Adeline Levivier	0.1	Creation of the file and redaction of the first draft	2019-08
Emmanuel Francis, Arlo Griffiths & Annette Schmiedchen	0.2	Review	2019-09
Emmanuel Francis, Axelle Janiak & Adeline Levivier	0.3	New draft, split in two documents FNC and Version Control	2020-04
Annette Schmiedchen & Arlo Griffiths	0.4	Review	2020-07
Dániel Balogh	0.5	Review	2020-11
Emmanuel Francis & Annette Schmiedchen	0.6	Review	2020-11
Axelle Janiak & Adeline Levivier	0.7	Changing some identifier pattern and deleting the appendices B and C	2020-11
Emmanuel Francis, Axelle Janiak & Adeline Levivier	0.8	Formatting and final validation	2020-11
Axelle Janiak & Adeline Levivier	1.0	Official release	2020-11
Axelle Janiak	1.1	Clarifying the use of hyphen to provide a suffix for an item number, p. 4 and 6.	2021-02
Axelle Janiak	1.1	Adding some examples for manuscripts naming in 2.5	2021-04

Axelle Janiak	1.2.1	Adding the cross-reference links and adding a naming pattern for translation files	2021-06
Axelle Janiak	1.2.2	Adding the cross-reference links and adding a naming pattern for commentary files	2021-07
Axelle Janiak	1.2.3	Adding a naming pattern for external bibliography - changing the naming pattern for external files containing commentaries and translations.	2021-08
Axelle Janiak	1.2.4	Adding MON and CONART abbreviated for identifiers of conglomerate artefacts and monuments	2021-09
Axelle Janiak	1.2.5	Adding Somavarnśin corpus	2022-01
Axelle Janiak	1.2.6	Adding the identifier for EC	2022-02

# Contents

Summary	5
Introduction	7
1 Creating Identifiers	7
1.1 Identifiers for Inscriptions	7
1.2 Identifiers for Inscriptions from a Published Collection	8
1.2.1 South-Indian Inscriptions (SII) collection	8
1.2.2 Epigraphia Carnatica (EC and ECNS) collection	9
1.3 Identifiers for Manuscripts	9
1.4 Identifiers for Inscribed Artefacts	10
1.5 Identifiers for Inscribed Conglomerate Artefacts	10
1.6 Identifiers for Critical Editions	10
1.7 Identifiers for Diplomatic Editions of Manuscripts	10
1.8 Identifiers for Surrogates	11
1.9 Identifiers for Persons	11
1.10 Identifiers for Places	11
1.11 Identifiers for Monuments	11
1.12 Identifiers for Collections	12
2 Naming Files	12
2.1 Project name	12
2.2 Designation	12
2.2.1 File name for textual data	12
2.2.2 File name for visual data	13
2.2.3 Metadata files	14
2.2.4 External files for Manuscripts Critical and Diplomatic Editions	14
2.2.5 Files without identifier	14
2.2.6 Scholarly publications	15
2.3 Optional features	15
2.3.1 Version	15
2.3.2 Date	15
2.4 Extension	16
2.5 Some examples to help you name a file	16
2.5.1 Model for documentation files	16

2.5.2 Model for inscription files	16
2.5.3 Model for critical edition files	16
2.5.4 Model for diplomatic edition files	16
2.5.5 Model for visual documentation files	17
2.5.6 Metadata files	17
2.5.7 Archives files	18
Appendix A: Corpus designations and GitHub repositories	<b>19</b>
Task Force A	19
Task Force B	19
Task Force C	20

# Summary

The identifier system is organized according to the following patterns. **See [part 1 of the documentation for more details](#).**

1. Inscriptions: INSPallava00236 (**INS + first word designation or acronym of the corpus + 5-digits incremental number**) and INSCIK01325-1 (**INS + first word designation or acronym of the corpus + 5-digits incremental number + hyphen + suffix number**)
2. Inscriptions in Published Collections: INSSII1400035 (**INS + acronym of the publication + volume number with 2-digits + item number with 5-digits**)
3. Manuscripts: LOr4530 (**Library's/repository's official abbreviation + collection if necessary + shelfmark**)
4. Artefacts: ART000001 (**ART + 6-digits incremental number**)
5. Conglomerate Artefacts: CONART000001 (**CONART + 6-digits incremental number**)
6. Critical Editions: CritEdSiksaGuru (**CritEd + workTitle**)
7. Diplomatic Editions: DiplEdSiksaGuruL633 (**DiplEd + workTitle + Manuscript ID**)
8. Surrogates: SUR000001 (**SUR + 6-digits incremental number**)
9. Persons: PSN00001 (**PSN + 5-digits incremental number**)
10. Places: PLC00001 (**PLC + 5-digits incremental number**)
11. Monuments: MON00001 (**MON + 5-digits incremental number**)
12. Collections: COL00001 (**COL + 5-digits incremental number**)

The generic file naming system works with this model. **See [part 2 of the documentation for more details](#).**

**DHARMA\_<designation>.<extension>**

The file naming pattern for metadata file is:

**DHARMA\_mdt\_<designation>.<extension>**

The file naming for documentation should have a declared version:

**DHARMA\_<designation>\_<version>.<extension>**

A special case: the manuscripts editions external files.

The file naming pattern for a translation:

**DHARMA\_<designation>\_trans<lgCode><translationNo>.<extension>**

The file naming pattern for a commentary:

**DHARMA\_<designation>\_com.<extension>**

The file naming pattern for a bibliography is:

**DHARMA\_<designation>\_biblio.<extension>**

# Introduction

In the collaborative environment of DHARMA, we will exchange many files in many formats using various tools. We all have had the experience of losing time looking for a file because it was not in the right folder, or of wondering what could be inside a file laconically named “document.pdf” or “note.docx”.

A set of File Naming Conventions (FNC) is a framework for naming files and folders, in a way that describes what they contain and how they relate to other files in a consistent, logical and predictable way. If we can all agree to follow such FNC for the remainder of the project, this will help us to save time and maintain peace of mind, while also helping to ensure successful data retrieval and facilitating the management of access rights at folder level.

This guide formulates rules on how to name your files for the best management of all your data that are meant to enter the “DHARMA-base”. But the Guide does not concern your work when you are writing your own articles or books, or are creating/editing any other file as long as it is not meant to be shared with other team members.

**Be aware of the difference between an “identifier” and a “filename”.** A filename is a name given to an electronically stored computer file, conforming to limitations imposed by the operating system, as in length or restricted choice of characters. While an identifier can be used to form a filename, it serves as a means to identify a unique object through a string with its own specifications.

## 1 Creating Identifiers

### 1.1 Identifiers for Inscriptions

Concerning a single inscription, the identifier will be composed of three elements: **INS + corpus designation + item No.**

It must always start with the three letters that form the abbreviation of “inscription”, written in uppercase letters: **INS**.

The **corpus designation** is a one-word designation or acronym of the corpus in question. We retain a number of acronyms from pre-existing projects, notably CIC, CIK, IDENK and EIAD, but changing its form to follow DHARMA format (e.g. the inscription



from Campa C0152 will become INSCIC00152). See [Appendix A](#) for the project's corpus designations.

Finally, the **item No** is the number given to identify the inscription in the corpus itself. Format the number with leading zeros so that it always forms a 5-digit number, e.g. 00001. For **item No** with a suffix, use a hyphen between the 5-digit number and the addition, e.g. INSCIK01325-1.

Some inscriptions are not clearly attributable to any corpus, yet clearly attributable to an area covered by DHARMA. In such rare cases, use Varia as corpus designation, e.g. INSVaria00001. Such a designation can also be used if the provenance/area of origin of the inscription is not clearly known. The EpiDoc files of such inscriptions will be stored in a GitHub repository common to the whole project, that is not related to a specific task-force. For these inscriptions only, the identifier can be changed during the project, in case new knowledge makes it possible to connect them to a more specific corpus.

A single inscription can thus be recorded as belonging to several different corpora, e.g. INSPallava0001, stored in the Github repository tfa-pallava-epigraphy, will also be identified as INSEIAD00139 with an element <altIdentifier>.<sup>1</sup> However, there should be only one single EpiDoc file for this specific inscription in the entire DHARMA framework. Such instances should thus be recorded in order to keep track of them and avoid duplicate EpiDoc editions. You can use the readme file of the repository on Github to point out that the inscription belongs to a secondary corpus. You can also add a link toward the main edition in the repository to which the “alternative identifier” belongs.

## 1.2 Identifiers for Inscriptions from a Published Collection

### 1.2.1 *South-Indian Inscriptions* (SII) collection

This section is an exception to the identifier pattern given in §1.1. It concerns inscriptions published in the *South-Indian Inscriptions* (SII) collection as well as, if this is found more convenient for other collections.

For corpora of inscriptions defined by such publications --- that is, inscriptions, not associated with any of the DHARMA regional/dynastic corpora and named according to an existing compendium --- it is possible, but not compulsory, to attribute, with the approval of your task-force leader, an ID recording the position in the collection rather than giving an ID related to a corpus defined on other grounds.<sup>2</sup>

---

<sup>1</sup> It will be dealt with in the metadata spreadsheets.

<sup>2</sup> This option is particularly meant for TF-A, in which a large portion of the SII will be just digitally encoded (without revision by want of access to visual documentation). The IDs following the pattern “INSSII#####” will enable easy location of the files named after it.

In such cases the pattern must be: **INS + abbreviated title + volume No + item No.**

We keep the abbreviation of “inscription” in uppercase letters: **INS**, already used to identify inscriptions in §1.1.

The **abbreviated title** is an acronym for the collection as a whole, e.g. SII.

The **volume No** must be given with a 2-digit number, e.g. 01. The number is used to identify the volume of the collection. For SII, it must be contained between 01 and 41.

Finally, the **item No** is the number given to identify the inscription in the volume itself. Format the number with leading zeros so that it always forms a 5-digit number, e.g. 00001.

### 1.2.2 *Epigraphia Carnatica* (EC and ECNS) collection

For ***Epigraphia Carnatica*** (EC and ECNS)<sup>3</sup>, the model is slightly different. Indeed, we will add a reference to the Taluq between the volume number and the item number. We recommend you use the usual pattern made available by EC, e.g. Bn for Bangalore. So all in all, the pattern will be: **INS + abbreviated title + volume No + Taluq reference + item No.**

We keep the abbreviation of “inscription” in uppercase letters: **INS**.

The **abbreviated title** is an acronym for the collection as a whole, e.g. EC, or ECNS.

The **volume No** for *Epigraphia Carnatica* old series will be numbered from 01 to 12 and 14 to 18, and the new series from 01 to 13.

The **Taluq reference** will always be two letters. Be aware that it will be the only case in which a lower case letter is tolerated in a DHARMA identifier.

And finally, the **item No** is the number given to identify the inscription in the volume itself. Format the number with leading zeros so that it always forms a 3-digit number, e.g. 001. Note that in other case, it would have been 6-digits number, however to keep the length consistent with the addition of a Taluq abbreviated form, we must reduced it.

## 1.3 Identifiers for Manuscripts

For a manuscript, keep the official pattern from the library/repository: **Library's/repository's official abbreviation + collection if necessary + shelfmark.**

---

<sup>3</sup> The volumes of the old series of *Epigraphia Carnatica* will be identified as EC, while the ones published in the new series should be abbreviated with ECNS.

## 1.4 Identifiers for Inscribed Artefacts

The data model of the DHARMA project requires us to treat artefacts, i.e. text supports, as different entities from inscriptions, i.e. inscribed texts, and to assign to each artefact its own identifying number: **ART + item No.**

Use the abbreviation **ART** (in uppercase) to determine the identifier as related to an artefact.

For artefacts the **item No** and is to be formatted with leading zeros so that it always forms a **6-digit number**, e.g. 000001. These item numbers will be pregenerated for you.

## 1.5 Identifiers for Inscribed Conglomerate Artefacts

**ADD a definition for conglomerate Artefacts.**

Use the abbreviation **CONART** (in uppercase) to determine the identifier as related to an artefact.

For artefacts the **item No** and is to be formatted with leading zeros so that it always forms a **6-digit number**, e.g. 000001. These item numbers will be pregenerated for you.

## 1.6 Identifiers for Critical Editions

To identify a critical edition, the pattern will be composed of the following two elements: **CritEd + WorkTitle.**

Start with the abbreviation of “critical edition” with an uppercase letter at the beginning of each part: **CritEd**.

The **WorkTitle** starts with an uppercase letter. Don’t use any diacritics. Please consider shortening long titles, to **obtain a string of no more than 15 characters**.

Note that this identifier will be used to name the main edition file as well as its translation.

## 1.7 Identifiers for Diplomatic Editions of Manuscripts

To identify a diplomatic edition of a manuscript, the pattern will be composed of the following three elements: **DiplEd + WorkTitle + manuscriptID.**

Start with the abbreviation of “diplomatic edition”: **DiplEd**.

The **WorkTitle** starts with an uppercase letter and is without diacritics. Please consider shortening long titles, in order to **obtain a string of no more than 15 characters**.

The **manuscriptID** should be established following the pattern given in § 1.3: **Library's/repository's official abbreviation + collection if necessary + shelfmark**.

## 1.8 Identifiers for Surrogates

To identify a surrogate (i.e. estampage, plaster cast, drawing, analog photograph), use **SUR + item No.**

Use the three uppercase letters that form the abbreviation of “surrogate” to start the ID: **SUR**.

The **item No** is to be formatted with leading zeros so that it always forms a **6-digit** number, e.g. 000001.

## 1.9 Identifiers for Persons

Persons have to be declared as a canonical reference inside an authority file.<sup>4</sup> An identifier must be provided for each entry. To build the identifier, use **PSN + item No.**

Start with the three uppercase letters that form the abbreviation of “person”: **PSN**.

The **item No** is to be formatted with leading zeros so that it always forms a 5-digit number, e.g. 00001. These numbers will be pregenerated for you.

## 1.10 Identifiers for Places

Places have to be declared as a canonical reference inside an authority file.<sup>5</sup> An identifier must be provided for each entry. To build the identifier, **PLC + item No.**

Start with the three uppercase letters that form the abbreviation of “place”: **PLC**.

The **item No** is to be formatted with leading zeros so that it always forms a 5-digit number, e.g. 00001. These numbers will be pregenerated for you.

## 1.11 Identifiers for Monuments

Monuments can be declared inside an authority file<sup>6</sup> An identifier must be provided for each entry. To build the identifier, **MON + item No.**

---

<sup>4</sup> The workflow regarding authority file for persons will be described in a dedicated guide (workflow under construction).

<sup>5</sup> The workflow regarding authority files for places is managed first at the project level and will be imported in the [Pleiades](#) gazetteer with a batch upload. It will help to avoid manually typing each entry. The workflow will be described in a dedicated guide (workflow under construction).

<sup>6</sup> The workflow regarding authority files for places is managed first at the project level and will be imported in the [Pleiades](#) gazetteer with a batch upload. It will help to avoid manually typing each entry. The workflow will be described in a dedicated guide (workflow under construction).

Start with the three uppercase letters that form the abbreviation of “monument”:  
**MON.**

The **item No** is to be formatted with leading zeros so that it always forms a 5-digit number, e.g. 00001.

## 1.12 Identifiers for Collections

To identify a collection or an institution, use **COL + item No**.

Start with the three uppercase letters that form the abbreviation of “collection”:  
**COL.**

The **item No** is to be formatted with leading zeros so that it always forms a 5-digit number, e.g. 00001. These numbers will be pregenerated for you.

# 2 Naming Files

Most files related to DHARMA should be named following this model: **DHARMA\_<designation>.<extension>**.<sup>7</sup> However, in some special cases, e.g. documentation files, you will need a specific pattern containing a version statement, so **DHARMA\_<designation>\_<version>.<extension>**.

Additional date or version can be added, wherever necessary. See §[2.3](#).

Metadata will require the addition of the prefix “**mdt**” between underscores after the project name and before the <designation>:  
**DHARMA\_mdt\_<designation>.<extension>**.

## 2.1 Project name

The acronym of the project should always be included in the name of a file created in the framework of the project. Always start with it and spell it with capital letters: **DHARMA**. It should always be followed by an underscore.

## 2.2 Designation

### 2.2.1 File name for textual data

Always use the text identifier in the <designation> segment.

- For a file containing the edition of an inscription:  
**DHARMA\_<InscriptionID>.<extension>**

---

<sup>7</sup> In this part, every element in angle brackets should be replaced depending on the context and on your usage.

- For a file containing the edition of an inscription reproduced from a published collection or compendium (see §1.2):  
**DHARMA\_<InscriptionIDwithinCorpus>.<extension>**
- For a file containing a critical edition: **DHARMA\_<CritEdID>.<extension>**
- For a file containing a diplomatic edition: **DHARMA\_<DiplEdID>.<extension>**

### 2.2.2 File name for visual data

For visual data (i.e. photos and 3D models), we need the possibility to quickly identify the content of the image, which can be an inscription, an artefact, a manuscript, a surrogate (e.g. an estampage), or the view of a place. The model for the visual data is the same as for any file: it contains successively the project name, the designation (designation of inscription, artefact, manuscript, place), the <view No> with leading zero, when necessary, so that it forms a 3-digit number and, finally the extension defining the format of the file (jpg, tif, raw, nef, ply, obj): **DHARMA\_<designation>\_<view No>.<extension>**

- For a photo file concerning an inscription:  
**DHARMA\_<InscriptionID>\_<view No>.<extension>**
- For a photo file concerning an artefact: **DHARMA\_<ArtefactID>\_<view No>.<extension>**
- For a photo file concerning a manuscript: **DHARMA\_<ManuscriptID>\_<view No>.<extension>**
- For a photo file concerning a surrogate: **DHARMA\_<SurrogateID>\_<view No>.<extension>**
- For a photo file concerning a view of a place (general views, archaeological sites): **DHARMA\_<PlaceID>\_<view No>.<extension>**

The 3D and the RTI models are managed in a different way given that 3D and RTI processing produce various sets of files named by the software itself, that can not be changed. However, the photo files used to create those models are to be named following the DHARMA convention. Each set is to be stored in a main folder named according to this model:

**DHARMA\_rti\_<InscriptionORArtefactID>**

OR

**DHARMA\_3d\_<InscriptionORArtefactID>**

It has then to be compressed (the extension for a compressed file is .zip or .rar). The compressed folder includes all the photos needed to process RTI and 3D models.

### 2.2.3 Metadata files

Metadata require the addition of the prefix “**mdt**” between underscores after the project name and before the <designation>:  
**DHARMA\_mdt\_<designation>.<extension>.**

The <designation> should always be the identifier regarding the inscription, artefact, surrogate, person, place or collection described in the file.

### 2.2.4 External files for Manuscripts Critical and Diplomatic Editions

Translations, commentaries and bibliographies of critical editions are to be kept in separate files from the main edition. To name those files, add an underscore after the identifier and follow the set of instructions below.

Translations are the most complicated pattern regarding the critical editions. We need to be able to distinguish between languages and anticipate the possibility of several translations being provided with the edition. After the identifier of the edition followed by the underscore, give the abbreviated code “trans”, then add the language code ISO 639-3. The first letter of the 3-letters code should be given in upper-case. Remember it should be formed by 3 letters, i.e “Eng” for English and “Fra” for French. To finish, give the translation a 2-digit number:  
**DHARMA\_<designation>\_trans<lgCode><translationNo>.<extension>**

Commentaries of critical editions will have a suffixed filename with the abbreviated expression “com” after the identifier of the edition:  
**DHARMA\_<designation>\_com.<extension>**

Bibliographies filenames are working as commentaries with the suffix “biblio”:  
**DHARMA\_<designation>\_biblio.<extension>**

### 2.2.5 Files without identifier

For entities without an identifier, you are free to choose the <designation> segment of the filename. We just ask you to follow some basic rules:

- Avoid blank spaces and underscores
  - use hyphens and “camelCase” in separating parts of file names
- Avoid the following special characters / \ : \* ? “ ” < > | [ ] { } \$ + %
  - note that only the special character & is allowed
- Create meaningful but brief names
  - your entire file path, including folder names, should never contain more than 256 characters
- Avoid diacritics

## 2.2.6 Scholarly publications

The scholarly publications collected and stored in ShareDocs to create the DHARMA Library have their own naming system:

**<authorName>\_<year>\_<abbreviatedTitle>.<extension>.**

If there are two authors, the pattern is:

**<authorName1&authorName2>\_<year>\_<abbreviatedTitle>.<extension>.**

If there are more than two authors, the pattern is:

**<authorName1&al>\_<year>\_<abbreviatedTitle>.<extension>.**

Finally, in the case of multi-volume works referred to by their abbreviated title (e.g. ARIE), follow the following pattern: **<acronym>\_<volume>\_<year>.<extension>.**

Presentations from the project's workshops and conferences can be referenced like this:

**<authorName>\_<year><eventName>\_<abbreviatedTitle>.<extension>.**

e.g. Morlock\_2019KickOff\_IntroductionXmlTeiEpidoc.pptx

Consult the *ShareDocs Memo* for more information or contact the Visual Data Manager.

8

## 2.3 Optional features

The optional features can be used after an underscore for any file format except txt and xml since for those two the versioning is taken care of by git.

### 2.3.1 Version

The version is to be indicated with the lowercase letter “v” followed by a 2-digit number, e.g. **v01**.

### 2.3.2 Date

The date in this block has to follow the pattern **YYYYMMDD**. It is mandatory only for archiving purposes but, if it seems useful, it is possible to record a date associated with a version. Note that you must type the version first followed by an underscore and then the date according to this pattern: **v00\_YYYYMMDD**.

---

<sup>8</sup> Guidelines for using ShareDocs are available in the presentation made during the project's Kick-off meeting (Berlin, 2019). See <https://bit.ly/3nOb9n0>.



## 2.4 Extension

The extension defining the format of the file is mandatory. This extension starts with a full stop and is followed by a code of 3 or 4 letters (e.g. .xml, .jpg, ...).

## 2.5 Some examples to help you name a file

### 2.5.1 Model for documentation files

**DHARMA\_<designation>\_<version>.<extension>**

DHARMA\_ZoteroGuide\_v01.pdf

DHARMA\_TransliterationGuide\_v01.md

DHARMA\_TransliterationGuide\_v01.pdf

### 2.5.2 Model for inscription files

**DHARMA\_<InscriptionID>.<extension>**

DHARMA\_INSGangaEast00001.xml

DHARMA\_INSUttiramerur00007.txt

DHARMA\_INSSII1400035.xml

### 2.5.3 Model for critical edition files

**DHARMA\_CritEd<WorkTitle>.<extension>**

DHARMA\_CritEdSvayambhu.xml

DHARMA\_CritEdRsisasana.xml

**DHARMA\_CritEd<WorkTitle>\_trans<lgCode><translationNo>.<extension>**

DHARMA\_trans\_CritEdSvayambhu\_transEng01.xml

DHARMA\_trans\_CritEdRsisasana\_transFra02.xml

**DHARMA\_CritEd<WorkTitle>\_com.<extension>**

DHARMA\_CritEdSvayambhu\_com.xml

DHARMA\_CritEdRsisasana\_com.xml

**DHARMA\_CritEd<WorkTitle>\_biblio.<extension>**

DHARMA\_CritEdSvayambhu\_biblio.xml

DHARMA\_CritEdRsisasana\_biblio.xml

### 2.5.4 Model for diplomatic edition files

**DHARMA\_DiplEd<WorkTitle>< manuscriptID>.<extension>**

DHARMA\_DiplEdPurvadhigamaL881\_transEng01.xml

DHARMA\_DiplEdKutaraManavaLeidenOr2215\_transFra01.xml

**DHARMA\_DiplEd<WorkTitle>\_trans<lgCode><translationNo>.<extension>**

DHARMA\_DiplEdPurvadhigamaL881.xml

DHARMA\_DiplEdKutaraManavaLeidenOr2215.xml

**DHARMA\_DiplEd<WorkTitle>\_com.<extension>**

DHARMA\_DiplEdPurvadhigamaL881\_com.xml

DHARMA\_DiplEdKutaraManavaLeidenOr2215\_com.xml

**DHARMA\_DiplEd<WorkTitle>\_biblio.<extension>**

DHARMA\_DiplEdPurvadhigamaL881\_biblio.xml

DHARMA\_DiplEdKutaraManavaLeidenOr2215\_biblio.xml

### 2.5.5 Model for visual documentation files

**DHARMA\_<InscriptionID>\_<view No>.<extension>**

DHARMA\_INSUttiramerur00007\_002.jpg

DHARMA\_INSBengalCharters00032\_001.ply

**DHARMA\_<ArtefactID>\_<view No>.<extension>**

DHARMA\_ART000243\_001.tif

DHARMA\_ART000167\_001.nef

**DHARMA\_<ManuscriptID>\_<view No>.<extension>**

DHARMA\_LOr4530\_002.jpg

**DHARMA\_<SurrogateID>\_<view No>.<extension>**

DHARMA\_SUR000007\_002.jpg

DHARMA\_SUR000032\_001.raw

**DHARMA\_<PlaceID>\_<view No>.<extension>**

DHARMA\_PLC00065\_023.jpg

DHARMA\_PLC00178\_003.tif

**DHARMA\_rti\_<InscriptionID>.<extension>**

DHARMA\_rti\_INSUttiramerur00007.zip

DHARMA\_rti\_INSBengalCharters00032.rar

### 2.5.6 Metadata files

**DHARMA\_mdt\_<InscriptionID>.<extension>**

DHARMA\_mdt\_INSGangaEast00001.xml  
DHARMA\_mdt\_INSSII1400035.xml

**DHARMA\_mdt\_<artefactID>.<extension>**

DHARMA\_mdt\_ART000243.xml  
DHARMA\_mdt\_ART000167.xml

**DHARMA\_mdt\_<SurrogateID>.<extension>**

DHARMA\_mdt\_SUR000007.xml  
DHARMA\_mdt\_SUR000032.rdf

**DHARMA\_mdt\_<PlaceID>.<extension>**

DHARMA\_mdt\_PLC00065.xml  
DHARMA\_mdt\_PLC00178.xml

**DHARMA\_mdt\_<PersonID>.<extension>**

DHARMA\_mdt\_PSN00065.xml  
DHARMA\_mdt\_PSN00178.xml

**DHARMA\_mdt\_<CollectionID>.<extension>**

DHARMA\_mdt\_COL00065.xml  
DHARMA\_mdt\_COL00178.xml

2.5.7 Archives files

**DHARMA\_<designation>\_<version>\_<date>.<extension>**

DHARMA\_FNC\_v01\_20201106.pdf

## Appendix A: Corpus designations and GitHub repositories

### Task Force A

Name of the corpus	Corpus Designation	GitHub repository
Accālpuram	<b>Accalpuram</b>	tfa-accalpuram-epigraphy
Cempiyaṇ Mahādevī	<b>CempiyanMahadevi</b>	tfa-cempiyan-mahadevi-epigraphy
Cīrkāḷi	<b>Cirkali</b>	tfa-cirkali-epigraphy
Koṭumpālūr	<b>Kotumpalur</b>	tfa-kotumpalur-epigraphy
Mēlappaḷuvūr & Kīlappaḷuvūr	<b>MelKil</b>	tfa-melappaluvur-kilappaluvur-epigraphy
Pallava	<b>Pallava</b>	tfa-pallava-epigraphy
Pāṇḍya	<b>Pandya</b>	tfa-pandya-epigraphy
South Indian Inscriptions	<b>SII</b>	tfa-sii-epigraphy
Tamil Nadu (varia)	<b>TamilNadu</b>	tfa-tamilnadu-epigraphy
Uttiramērūr	<b>Uttiramerur</b>	tfa-uttiramerur-epigraphy

### Task Force B

Name of the corpus	Corpus Designation	GitHub repository
Arakan	<b>Arakan</b>	tfb-arakan-epigraphy
Bādāmi Caḷukya	<b>BadamiCalukya</b>	tfb-badamicelukya-epigraphy
Bengal Charters on Copper Plates, Stones and Other Supports	<b>BengalCharters</b>	tfb-bengalcharters-epigraphy
Bengal Dedication Inscriptions	<b>BengalDED</b>	tfb-bengalded-epigraphy

Bhaumakara	<b>Bhaumakara</b>	tfb-bhaumakara-epigraphy
Dakṣiṇa Kosala	<b>DaksinaKosala</b>	tfb-daksinakosala-epigraphy
Early Andhra	<b>EIAD</b>	tfb-eiad-epigraphy
Eastern Gaṅga	<b>GangaEast</b>	tfb-gangaeast-epigraphy
Epigraphia Carnatica	<b>EC</b>	tfb-ec-epigraphy
Kalyāṇa Cālukya	<b>KalyanaCalukya</b>	tfb-kalyanacalukya-epigraphy
Licchavi	<b>Licchavi</b>	tfb-licchavi-epigraphy
Maitraka	<b>Maitraka</b>	tfb-maitraka-epigraphy
Raṣṭrakuṭa	<b>Rastrakuta</b>	tfb-rastrakuta-epigraphy
Śailodbhava	<b>Sailodbhava</b>	tfb-sailodbhava-epigraphy
Sātavāhana	<b>Satavahana</b>	tfb-satavahana-epigraphy
Somavaṃśin	<b>Somavamsin</b>	tfb-somavamsin-epigraphy
Telugu Inscriptions	<b>Telugu</b>	tfb-telugu-epigraphy
Veṅgi Cālukya	<b>VengiCalukya</b>	tfb-vengicalukya-epigraphy
Viṣṇukunḍin	<b>Visnukundin</b>	tfb-visnukundin-epigraphy

## Task Force C

<b>Name of the corpus</b>	<b>Corpus Designation</b>	<b>GitHub repository</b>
Campā	<b>CIC</b>	tfc-campa-epigraphy
Khmer	<b>CIK</b>	tfc-khmer-epigraphy
Nusantara	<b>IDENK</b>	tfc-nusantara-epigraphy