# DHARMA Editorial Conventions draft 2024

current DHARMA display needs to be provided where missing

it may be a good idea to add physical and logical display columns

syntax highlighting in the XML and colouring in the display need to be provided for the public release

Specific references to EGD subsections could be added to each line, but if we want this, we’d better wait till we have the new EGD release.

## 1. Semantic and metrical structure (intrinsic structure)

EGD §2

| **Description** | **DHARMA markup** | **DHARMA full display** | **remark** |
| --- | --- | --- | --- |
| Stanza in named metre | <lg n="1" met="āryā">  <l n="a">\_</l>  </lg> |  |  |
| Stanza in metre without a known name | <lg n="2" met="+++++-++-+=">  <l n="a">\_</l>  </lg> | - - - - - ⏑ - - ⏑ - ⏓ |  |
| Verse line | <l n="a">kāritam idan nr̥patinā</l> | kāritam idan nr̥patinā |  |
| Word divided across verse lines (enjambement) | <lg n="3" met="āryā">  <l n="a" enjamb="yes">tat-sūnus sakalārāti</l>  <l n="b">-mada-ccheda-karāyudhaḥ</l>  </lg> | tat-sūnus sakalārāti-  -mada-ccheda-karāyudhaḥ | The hyphen at the beginning of the second line is editorial compound segmentation; that at the end of the first line is the display for enjambement. |

## 2. Layout (extrinsic structure)

EGD §3

daba note: I don’t think we need to add textpart divs here. we might want to add incipits (EGD §3.3.3) and forme work.

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **DHARMA markup** | **DHARMA display** | **remark** |
| Beginning of inscribed zone | <milestone type="pagelike" unit="zone" n="A"/>svasti śrī |  |  |
| Beginning of inscribed zone | <milestone type="pagelike" unit="item" n="A"/>svasti śrī |  |  |
| Beginning of surface or zone, with editorial label | <milestone type="pagelike" unit="item" n="N"/><label xml:lang="eng">Northern Doorjamb</label>svasti śrī |  |  |
| Beginning of page in copperplate inscription | <pb n="1r"/>svasti śrī |  |  |
| Beginning of line | <lb n="1"/>svasti śrī <lb n="2"/>kōpparakēcari | (1)svasti śrī (2)kōpparakēcari |  |
| Beginning of non-hierarchical division (fragment, masonry block, etc.) | <milestone unit="block" n="a"/>svasti śrī |  |  |
| Word divided across break | <lb n="1"/>…dhar  <lb n="2" break="no"/>ma… | (1)…dhar- (2)ma… | Same display and same attribute for breaks other than line beginnings. |

## 3. Palaeographic features

EGD §4.1-4.3

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **DHARMA markup** | **DHARMA display** | **remark** |
| Vowel | <seg type="component" subtype="prescript"/>  <seg type="component" subtype="postcript"/> |  |  |
| Consonant | <seg type="component" subtype="body"/> |  |  |
| Single symbol representing a number above 9 | <g type="numeral">200</g> |  |  |
| Symbol interpreted as punctuation | <g type="danda">.</g> |  | Display depends on @type. |
| Symbol interpreted as space filler | <g type="squiggleVertical">§</g> |  | Display depends on @type. |
| Uninterpreted symbol | <g type="floretComplex"/> |  | Display depends on @type. |
| Small space between words | <space/> |  | “Small” is defined as less than two typical character widths. |
| Large space between words or semantic units | <space quantity="3" unit="character"/> |  | Approximate width of space in typical-sized character. |
| Space left blank because it was not possible to write in a given spot | <space type="binding-hole"/> |  | Other values of @type for other reasons of inability to write. The size of these spaces is never encoded. |
| Space left blank presumably with the intent to fill later | <space type="vacat" quantity="5" unit="character"/> |  | Approximate width of space in typical-sized character. The size of these spaces must always be encoded. |
| Space of uninterpreted function | <space type="unclassified"/> |  | Size optionally encoded with @quantity and @unit. |

## 4. Scribal alterations in text

EGD §4.4

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **DHARMA markup** | **DHARMA display** | **remark** |
| Characters added by premodern scribe | dha<add>rma</add> | dha⟨⟨rma⟩⟩ | Locus of insertion and eventual addition mark encoded by attributes. |
| Characters deleted by premodern scribe but still legible | dharma<del>ma</del> | dharma⟦ma⟧ | Manner of deletion encoded by attributes. If the deleted text is unclear or only tentatively legible, this is encoded and displayed normally (section 5). |
| Characters deleted by premodern scribe and no longer legible | <del><gap reason="illegible" quantity="1" unit="character"/></del> |  | Manner of deletion encoded by attributes. Number of illegible characters encoded as in other lacunae (section 5). |
| Scribal correction | bottun<subst><del rend="corrected">u</del><add place="overstrike">a</add></subst> | bottun⟦u⟧⟨⟨a⟩⟩ | Manner of deletion and locus of addition may be encoded with different attributes. |

## 5. Imperfectly legible text

EGD §5

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **DHARMA markup** | **DHARMA display** | **remark** |
| Unclear reading (characters would be ambiguous outside of their context) | dha<unclear>rma</unclear> | dha(rma) | This is without regard to the degree of damage and used in all cases where a character or vestige that would be ambiguous or illegible without its context is confidently read in context. |
| Tentative reading | dha<unclear cert="low">rma</unclear> | dha(rma?) | This indicates reduced confidence in the reading, not reduced legibility of the original. |
| Ambiguous reading with two or more possible alternatives | <choice>  <unclear>a</unclear>  <unclear>o</unclear>  </choice> | (a/o) |  |
| Lacuna: unknown number of characters irretrievably lost | <gap reason="lost" extent="unknown" unit="character"/> | […] |  |
| Lacuna: confidently estimated number of characters irretrievably lost | <gap reason="lost" unit="character" quantity="3"/> | [3+] |  |
| Lacuna: tentatively estimated number of characters irretrievably lost | <gap reason="lost" unit="character" quantity="3"precision="low"> | [ca. 3+] |  |
| Lacuna: unknown number of characters illegible, but vestiges are present | <gap reason="illegible" extent="unknown" unit="character"/> | […] |  |
| Lacuna: confidently estimated number of characters illegible, but vestiges are present | <gap reason="illegible" unit="character" quantity="3"/> | [3×] |  |
| Lacuna: tentatively estimated number of characters illegible, but vestiges are present | <gap reason="illegible" unit="character" quantity="3"precision="low"/> | [ca. 3×] |  |
| Lacuna: unknown number of characters (encoded from previous edition that makes no distinction between loss and illegibility) | <gap reason="undefined" extent="unknown" unit="character"/> | […] |  |
| Lacuna: confidently estimated number of characters (encoded from previous edition that makes no distinction between loss and illegibility) | <gap reason=”undefined" unit="character" quantity="3"/> | [3\*] |  |
| Lacuna: tentatively estimated number of characters (encoded from previous edition that makes no distinction between loss and illegibility) | <gap reason="undefined" unit="character" quantity="3" precision="low"/> | [ca. 3\*] |  |
| Lacuna: confidently estimated number of lines | <gap reason="lost" quantity="3" unit="line"/> | [3 lines lost] | Similarly with reason “illegible” and “undefined” |
| Lacuna: tentatively estimated number of lines | <gap reason="lost" quantity="3" unit="line" precision="low"/> | [ca. 3 lines lost] | Similarly with reason “illegible” and “undefined” |
| Lacuna: unknown number of lines | <gap reason="lost" extent="unknown" unit="line"/> | [unknown number of lines lost] | Similarly with reason “illegible” and “undefined” |
| Lacuna in verse with known prosody | <seg met="+++-++"><gap reason="lost" quantity="6" unit="character"> | [– – – ⏑ – –] | Similarly with reason “illegible” and “undefined” |
| Lacuna confidently restored by editor | dha<supplied reason="lost">r</supplied>ma | dha[r]ma | Similarly with reason “undefined”, but NOT with reason “illegible”. If vestiges are present, a restoration is to be shown as an unclear reading. Basis of restoration may be encoded with @evidence. |
| Lacuna tentatively restored by editor | dha<supplied reason="lost" cert="low">r</supplied>ma | dha[r?]ma dha[r?]ma | Similarly with reason “undefined”, but NOT with reason “illegible”. If vestiges are present, a tentative restoration is to be shown as a tentative reading. Basis of restoration may be encoded with @evidence. |
| Lacuna with differently treated parts | <gap reason="illegible" quantity="10" unit="character" precision="low"/><seg met="+++-++"><gap reason="illegible" quantity="6" unit="character"/></seg><supplied reason="lost">abc</supplied> | [ca.10x – – – ⏑ – – abc] |  |

## 6. Editorial intervention

EGD §6

| **Description** | **DHARMA markup** | **DHARMA display** | **remark** |
| --- | --- | --- | --- |
| Non-standard lexicon or spelling | <orig>dhamma</orig> | ¡dharma! |  |
| Non-standard lexicon or spelling normalised by the editor | <choice><orig>dhamma</orig><reg>dharmma</reg></choice> | ¡dharma!⟨dharmma⟩ |  |
| Erroneous or unintelligible original | <sic>dhoma</sic> | ¿dhoma? |  |
| Erroneous or unintelligible original corrected by the editor | dha<choice><sic>t</sic><corr>r</corr></choice>ma | dha¿t?⟨r⟩ma |  |
| Character(s) omitted by scribe and restored by the editor (editorial addition) | dha<supplied reason="omitted">rma</supplied> | dha⟨rma⟩ |  |
| Character(s) omitted by the scribe and not restored by the editor | catvāriṁśat samā<seg met="-="><gap reason="omitted" quantity="2" unit="character"/></seg> |  | Unrestored omissions tend to occur in metrical text, since without a known prosody it is hard to be certain that something was omitted. But the same encoding may also be used without <seg> and @met, and also with @extent="unknown" instead of @quantity. |
| Superfluous characters engraved by scribe (e.g. dittography) and suppressed by the editor | dharma<surplus>ma</surplus> | dharma{ma} |  |

## 7. Editorial markup in translations and commentary

EGD §9.2.7-14 and a few other sections

The markup and display applicable to lacunae and their restoration may also appear in translations.

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **DHARMA markup** | **DHARMA display** | **remark** |
| Text in language other than the surrounding text, modern language written in Latin script | <foreign xml:lang=”fra”>vis-à-vis</foreign> |  |  |
| Text in language other than the surrounding text and the primary language of the edition, modern or premodern language transliterated from non-Latin script | <foreign xml:lang="tel-Latn">cuvikaṇṭhi-</foreign> |  |  |
| Explanatory insertion in translation | Kantacēṉaṉ <supplied reason="explanation">Skt. Skandasena</supplied> | Kantacēṉaṉ (Skt. Skandasena) |  |
| Tentative explanatory insertion | the Raṭṭa lord <supplied reason="explanation" cert="low">Amoghavarṣa I</supplied> |  |  |
| Corresponding original word inserted in translation | Temple <supplied reason="explanation"><foreign>tēvakulam</foreign></supplied> | Temple (tēvakulam) |  |
| Words inserted into translation for clarity or syntactical correctness | Fiftieth <supplied reason="subaudible">year</supplied> of Nantippōttaracar | Fiftieth [year] of Nantippōttaracar |  |
| Words tentatively inserted for clarity or syntactical correctness | <supplied reason="subaudible" cert="low">The donee</supplied> has been allocated a residence |  |  |
| Unintelligible segment in the original | <gap reason="ellipsis"/> |  | Usually followed by the original segment in <supplied reason="explanation"><foreign></foreign></supplied> and optionally also by an explanatory note. |
| Untranslated segment | <gap reason="ellipsis"/> |  | Followed by a note explaining why this text has not been translated. |
| Alternative translation of bitextual phrase (śleṣa) | who is attended by mercy as Droṇa <seg rend="pun">is followed by Kr̥pā</seg> |  |  |

## 