

# CRMtex. A Semantic Model for the study of ancient texts. Non-alphabetical texts

Achille Felicetti & Francesca Murano  
Università degli Studi di Firenze



CRMtex is an extension of **CIDOC CRM** developed to support the study of **ancient documentation**. Based on a **semiotic approach** to the text, the model is intended to identify and define in a clear and unambiguous way the main entities involved in the study and edition of ancient handwritten texts and then to describe them by means of appropriate **ontological** instruments in a multidisciplinary perspective.

The extension provides tools for managing this kind of complexity by defining classes and properties for describing a hand written text in all its aspects, from its creation in the past, down to its present conservation, investigation and study by scholars, including its transcription, translation, interpretation and publication.

CRMtex allows to describe not only texts belonging to the major corpora (i.e. Latin and Greek) using **alphabetic** writing systems, but also corpora using different **not-alphabetic phonographic systems**, like the Persian cuneiform or the Mycenaean Linear B, the study of which requires multiple levels of interpretation and transcription by scholars.

The model consists of the following classes:

- TX1 Written Text.** Subclass of E25 Man-Made Feature
- TX2 Writing.** Subclass of F28 Expression Creation
- TX3 Writing System.** Subclass of E29 Design or Procedure
- TX4 Writing Field.** Subclass of E25 Man-Made Feature
- TX5 Reading.** Subclass of the CRMsci S4 Observation
- TX6 Transcription.** Subclass of E7 Activity
- TX7 Written Text Segment.** Subclass of TX1 Written Text
- TX8 Grapheme.** Subclass E90 Symbolic Object
- TX9 Glyph.** Subclass of E25 Man-Made Feature
- TX10 Style.** Subclass of E29 Design or Procedure

and properties:

- TXP1** used writing system (writing system used for)
- TXP2** includes (is included within)
- TXP3** rendered (is rendered by)
- TXP4** has segment (is segment of)
- TXP5** wrote (was written by)
- TXP6** encodes (is encoding of)
- TXP7** has item (is item of)
- TXP8** has component (is component of)
- TXP9** is encoded using (was used to encode) read (was read by)
- TXP10** transcribed (was transcribed by) has style (is style of)

