

The corpus header

Proposal for a generalised header for text corpora

CLARIN WP 2 Technical Report
NN

Latest revision: December 8, 2010

Abstract

This document describes how textual metadata of a DK-CLARIN corpus can be expressed by means of a TEI P5 header.

The structure of the header is oriented towards that one used by the BNC ? and PAROLE-DK ? but tries to avoid idiosyncracies not covered by TEI P5 as well as modifications of the TEI header schema.

1 Introduction

2 Main

3 Conclusion

Bibliography