

TEIP5-WP2 skema

På <http://dkclarin.dk/schemas/WP2> findes WP2's TEIP5-skema. Skemaet kan bruges ved validering af WP2-data ved udvikling og aflevering af tekster fra WP2's arbejds pakker. Skemaet er udviklet på baggrund af den aktuelle version af <http://korpus.dsl.dk/clarin/corpus-doc/text-header.pdf>.

Skemaet findes i følgende versioner:

- W3C schema
- RELAX NG Schema (XML syntax)
- RELAX NG Schema (Compact syntax)

Skemaet er dokumenteret vha. en ODD-fil som specificerer ændringer i forhold til den fulde TEIP5-specifikation. Hvis man ønsker at læse direkte i en skema-specifikation, så anbefales RELAX NG Schema i "compact syntax", som den mest "læselige" version.

Testfiler

Testfiler ligger i http://dkclarin.dk/schemas/WP2/test_files. De kan benyttes sammen med et rng- eller et xsd-skema.

Testfilen http://dkclarin.dk/schemas/WP2/test_files/formatsample_xsd.xml validerer med xsd-skema, de øvrige validerer vha. rng-skema. Testfilerne er valideret vha. Oxygen.

2010-05-19: Version 1.0

Version 1.0 af skemaet implementerer den struktur af elementer som er beskrevet i kapitel 2 af text-header.pdf, samt den delmængde af begrænsninger i værdier fra kapitel 3 som følger den generelle TEI-specifikation.

- W3C schema: http://dkclarin.dk/schemas/WP2/TEIDKCLARINv1_0/xml.xsd
- RELAX NG Schema (XML syntax): http://dkclarin.dk/schemas/WP2/TEIDKCLARINv1_0.rng
- RELAX NG Schema (Compact syntax):
http://dkclarin.dk/schemas/WP2/TEIDKCLARINv1_0.rnc
- ODD-dokument, der beskriver ændringer i forhold til det generelle TEIP5-format:
<http://dkclarin.dk/schemas/WP2/teip5-dkclarin-20100519.odd>

http://dkclarin.dk/schemas/WP2/README_TEIP5DKCLARIN_validation_v1_0.pdf

Kort om skemaet

Skemaet medtager alle elementer fra modulerne: *core* og *textstructure*. Fra *namesdates* inkluderes kun elementet "birth", fra *dictionaries* inkluderes kun elementet "lang". Fra modulerne: *analysis*, *corpus*, *header* og *linking* er elementerne delvist inkluderet. Se ODD-dokumentet for yderligere detaljer.

Tidligere versioner

2010-03-03: Version 0.9

Version 0.9 af skemaet implementerer den struktur af elementer som er beskrevet i kapitel 2 af *text-header.pdf* pr 2010-03-03, samt den delmængde af begrænsninger i værdier fra kapitel 3 som følger den generelle TEI-specifikation.

- W3C schema: <http://dkclarin.dk/schemas/WP2/version0.9/TEIDKCLARINv0.9/xml.xsd>
- RELAX NG Schema (XML syntax):
<http://dkclarin.dk/schemas/WP2/version0.9/TEIDKCLARINv0.9.rng>
- RELAX NG Schema (Compact syntax):
<http://dkclarin.dk/schemas/WP2/version0.9/TEIDKCLARINv0.9.rnc>