Plan rada za Vrančića u MHR

Polazište: tekst iz PDF-a. Cilj: TEI XML za svaki pojedinačni dokument (pisma - po godinama).

0. Provjeriti jesu li svi brojevi stranica ispravno pročitani.

1. XQuery: pdf2tei.xq

2. XSL: 14pdfxml.xsl

3. Unaccent

4. Korektura - ispravljati prema izvorniku, uz spell check

5. (paralelno) TEI označavanje: div, p, note; elemente lb/@left brisati

5.1 prvo div

5.2 p - između milestone/lb (i note)

5.3 spell check - ako je nešto nepravilno u izvorniku, staviti sic (očite tiskarske greške ispraviti i označiti s corr). OPREZ: brojke ne ulaze u spell check, ako je nešto OCR pročitao kao brojke, a da nisu, treba ručno ispraviti! Eventualne žive glave i kustode (na dnu stranica) ne treba korigirati - staviti u <note type="typographia">.

## Format

0. korekture se rade prema pdf (v. Dropbox)

00. zadržati <lb/>

000. provjeriti je li pismo već u CroALa - ako jest, ne treba ga korigirati, samo označiti <div type="croala">

0000. pisma Vrančića članovima obitelji također ne treba korigirati (njih je priredila D. Sorić), ali pisma članova obitelji Vrančiću DA!!

1. svako pismo je <div>

2. Početak je <opener> (uključujući i mađarski naslov)

3. Unijeti nove odlomke (<p>); pomoć je <lb rend>

4. Označiti <note>

5. Dodati crticu na kraju retka, ako je riječ rastavljena, a crtica nije pročitana