PressMint

Egunkari Historikoen Corpus Elkarreragingarriak

Tomaž Erjavec, Matyáš Kopp, Maciej Ogrodniczuk,
Petya Osenova, German Rigau

Ainara Estarrona, Aritz Farwell, Xabier Goenaga

PressMint

Erronka

Prentsa historikoa altxor ikaragarria da Giza eta Gizarte Zientzietarako, baina sarritan baliabide digitalak ez dira elkarreragingarriak eta metodo digitalen erabilera oraindik mugatua da.

Gure Helburua

Europako XX. mendearen hasierako prentsa historikoaren corpusa osatzea ikerketarako eta analisi aurreratuetarako:

- ▶ ParlaMint proiektuaren oinarri teknikoa erabilita
- ▶ Eleaniztuna eta elkarreragingarria
- ▶ Hainbat formatutan deskargatzeko eskuragarria

Ezaugarri Nagusiak

- ▶ FAIR printzipiotan oinarritua (*Findable*, *Accessible*, *Interoperable*, *Reusable*)
- ▶ Informazio linguistikoarekin etiketatua (PoS etiketatzea, lematizazioa eta gaien sailkapena)
- ▶ Faksimileak, hizkuntza normalizazioa eta entitateen arteko loturak gehitzeko aukera

Eragina eta Onurak

- ▶ Ikerketa transnazionala, eleaniztuna, eta diziplinartekoa bultzatu
- ▶ Giza eta Gizarte Zientzietan metodo digitalen erabilera sustatu
- ▶ Prestakuntza eta elkarlanaren bitartez komunitatea indartu
- ▶ CLARIN azpiegitura aberastu kalitate handiko baliabide berriekin

Europa Mailako Ekimena

Jarraitu gure aurrerapenak eta sartu corpusetara hemen:

https://www.clarin.eu/pressmint

CLARIN

HiTZ
Hizkuntza Teknologiako Zentroa
Basque Center for Language Technology

Universidad del País Vasco — Euskal Herriko Unibertsitatea

CLARIAH-EUS