

Euskal Herriko Ahotsak

“Euskal Herriko hizkerak eta ahozko ondarea”

- 356 herri.
- 6003 hizlari eta 8500 ordu grabatuta.
- %50 landuta (71.594 pasarte) eta %5 transkribatuta (400 ordu)
- 3.803 audio-pasarte online (135 ordu)
- 55.813 bideo-pasarte online (2.000 ordu)

Digitalizazioan eta prozesamenduan aldaketak

Gure oinarriak 1999 urte hartan.

- Dena bideoz.
- Bideoa ez da bidea, helburua baizik.
- 2.0 filosofia (irekia, lankidetzan oinarritua).
- Dena online.
- Dena libre, doan eta lizentzia librearekin. Berrerabilgarria. CC-BY-SA.

Digitalizazioan eta prozesamenduan aldaketak

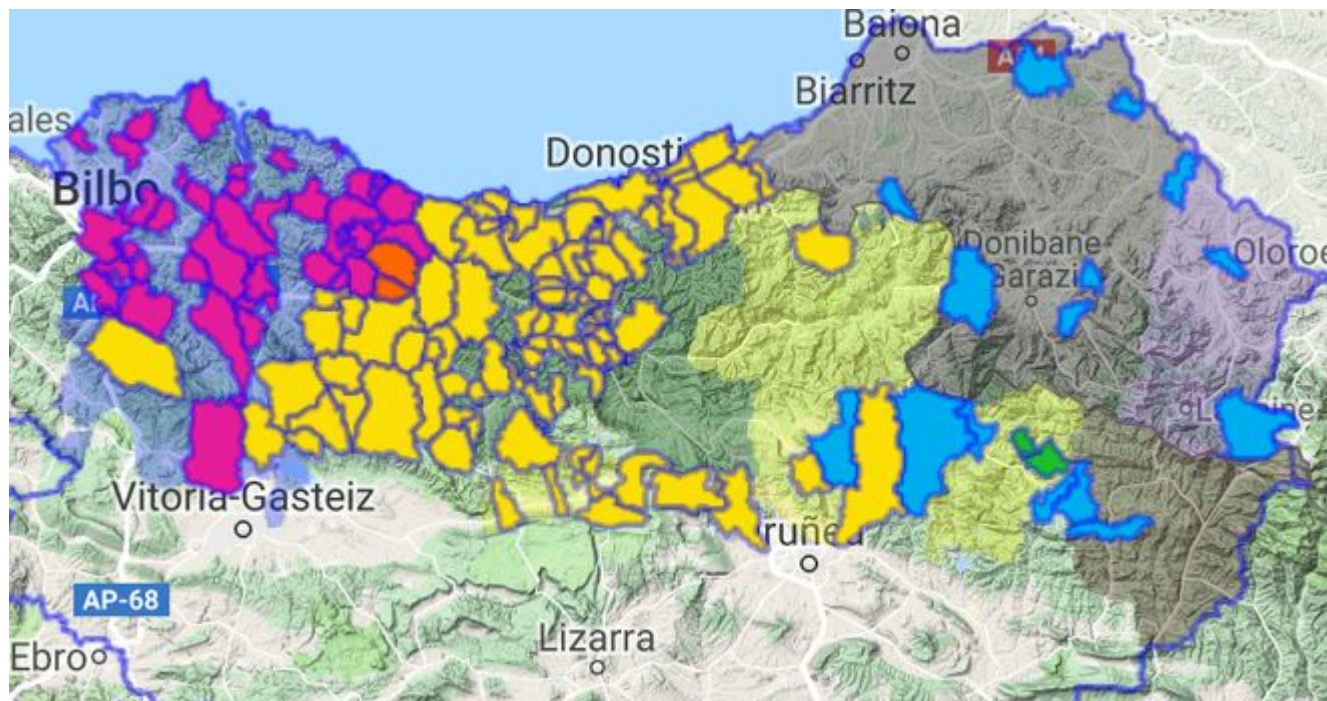
Nola aldatu diren gauzak 20 urte hauetan.

- Baliabide teknikoak izugarri hobetu eta merketu.
- Sarearen aukerak ugaritu.
- CC lizentziak zabaldu.
- Formatu eta euskarriekin zoramena.
- Terak eta Terak pilatzen dira.

Ahozko Tradiziozko Corpora

- Azken belaunaldi elebakarraren (?) hizkera daukagu jasota.
- Helburua: ahozko materialean oinarritutako corpus linguistiko bat osatzea.
 1. lema bakoitzaren aldaerak;
 2. aldaeren eta lemen banaketa geografikoa;
 3. datu estatistikoak.

Informazioa hor zegoen, baina nola erauzi? >> Sistema semi-automatikoa.



Ahozko Tradiziozko Corpusa

Eraitza politak.

- 10.000 pasarte, 300 ordutik gora.
- 3M hitzeko corpusa (%5).

Hanka motz:

- Transkripzio gehiago behar ditugu.
- Lematizazioa arindu eta hobetu behar dugu.

Lankidetzaren indarra

- Lankidetzaren gure proiektuaren funtsa.
- Garapen teknikoaren beharra.
- Erakunde ofizial nazionalik ez. “Ematea ere zaila”.