

Hizkuntzen sailkapena, Wikipediako funtsezko 10.000 artikuluen arabera

Definizioa

Wikipediak osatutako hizkuntza sailkapena neurtzen da. Sailkapen hori egiteko, hizkuntza bakoitzari puntuazio bat egozten zaio, hizkuntza horretako Wikipediak osatutakoa artikuluko kopuruaren arabera. Neurtu egiten da ea Wikipediako funtsezko 10.000 artikuluen artean zenbat dituen osatuta Wikipedia bakoitzak, eta zein tamainakoak diren osatuta dauzkan artikuluko horiek. Zenbat eta artikuluko gehiago izan sortuta eta zenbat eta luzera handiagoak, orduan eta puntuazio altuagoa izango du Wikipedia bakoitzak.

Hedadura

Mundu osoa

Aldagaiak eta kategoriak

- **Hizkuntza:**
Wikipediaren sailkapeneko aurreneko 50 hizkuntzak.
- **Tamaina:**
Falta direnak: osatzeke dauden artikuluko kopurua.
< 8k: 8.000 karaktere baino gutxiagoko artikulukoak. Artikulu motzak dira.
8k – 16k: 8.000 eta 16.000 karaktere bitarteko artikulukoak. Artikulu ertainak dira.
>16K: 16.000 karaktere baino gehiagoko artikulukoak. Artikulu luzeak dira.
- **Puntuazioa:**
Wikipedia bakoitzak osatuta dauzkan artikuluen araberako puntuazioa, artikuluen tamaina ere kontutan hartzen duena.
Puntuazio hori kalkulatzeko, ondorengo formula erabiltzen da:

$$\text{Emitza gordina} = \text{artikulu motzak} * 2 + \text{artikulu ertainak} * 3 + \text{artikulu luzeak} * 4$$
$$\text{Puntuazioa} = \text{emaitza gordina} / (10.000 * 0,04)$$

Iturria	
Wikipedia	
Iturriaren ezaugarriak	https://meta.wikimedia.org/wiki/List_of_Wikipedias_by_expanded_sample_of_articles
Lehenengo datua eta maiztasuna	Lehenengo datua 2020ko da eta urtero berrituko da.
Lagina eta akats-tartea	Unibertso osoa neurtzen da. Wikipedia bakoitzak (hizkuntzaren arabera banatuta) 10.000 artikulua garrantzitsuenetatik osatuta dituen guztiak jasotzen dira. Neurketa urte amaieran egiten da.
Informazio osagarria	
Argitalpenak eta txostenak	