

ANALHITZAren erabilera

Mikel Iruskieta (Ixa Taldea)

<http://ixa2.si.ehu.es/clarink/analhitza.php?lang=eu>

Fitxategia igotzeko 2 aukera:

1A. TXT formatua igo

1B. Kopiatu eta itsatsi

ixa

Hasiera Gure produktuak Bideoak Ekimenak Zerbitzuak Kontakua

CLARIN CENTRE K

EU EN ES

ANALHITZA
Powered by CLARIN

ANALHITZAK euskarazko, gaztelaniako edo ingeleseko testu batetik ondorengo informazio linguistikoa erauzten lagunduko dizu:

- Izenak, adjektiboak, aditzak, adberbioak...
- pertsona-izenak, leku-izenak...
- binakako, hirunakako eta launakako hitz-sekuentziak
- ... eta gehiago!

Testua fitxategi batean duzuna, hemen bertan kopiatu duzuna edo web-orri batekoa izan daiteke, baina UTF8 kodeketan egon beharko du.

ANALHITZA erabiltzeko sartu aztertu nahi duzun testua azpiko 3 aukera hauetakoren bat erabiliz, eta jarraian, testuaren hizkuntza aukeratu (euskara, gaztelania edo ingelesa). Aipuri bat itxaron ondoren, emaitzak Excel fitxategi batean jasoko dituzu. Horrela, emaitzak modu errazean moldatu ahal izango dituzu zure beharretara.

Fitxategia igo (txt formatuan) Sartu testua Url-a sartu

ANALHITZAK testua automatikoki prozesatzen du IxaKat (euskararako) eta Ixa pipes (gaztelania eta ingeleserako) hizkuntza prozesatzailen kate modularrak erabiliz

ANALHITZA, hizkuntza-teknologiak ballatuz, humanitateetako eta giza zientzietako ikertzaileei datu linguistiko fidagarriak eta erraz manipula daitezkeenak eskaintzeko helburuarekin diseinatu da. Hauetako arloren batean zure ikergala definituta baduzu, testu-multzto bat eskura baduzu eta ANALHITZAren aztertzeko interesa baduzu, idatzi eta lagunduko dizugu: mikel.iruskieta@ehu.eus

ANALHITZA erabiltzen baduzu zure ikerketaren batean, ondorengo lan honen aipamena egin zure lan akademikoan mesedez:

Otegi, A. Imaz, O. Díaz de Ilarraza, A. Iruskieta, M. Uria, L. 2017 ANALHITZA: a tool to extract linguistic information from large corpora in Humanities research. *Procesamiento del Lenguaje Natural* 58: 77-84.

2. Aukeratu hizkuntza:

- Euskara

3. Sakatu “Bidali”

Testua fitxategi batean duzuna, hemen bertan kopiatuko duzuna edo web-orri batekoa izan daiteke, baina UTF8 kodeketan egon beharko du. ANALHITZA erabilzeko, sartu aztertu nahi duzun testua azpiko 3 aukera hauetakoren bat erabiliz, eta jarraian, testuaren hizkuntza aukeratu (euskara, gaztelania edo ingelesa). Apur bat itxaron ondoren, emaltzak Excel fitxategi batean jasoko dituzu. Horrela, emaltzak modu errazean moldatu ahal izango dituzure beharretara.

Fitxategia igo (txt formatuan) Sartu testua Uri-a sartu

Sartu testua:

Probarako testua erabili

Abereak batzarrean bildu ziren, lehoi zaharra hila zelako eta lehendakari berri bat izendatu behar zutelako. Izendatu zuten, eta honela esan zuen lehendakari berriak:
-Gure gizarteak hobetzeko, Igandeetan lan plaka bat egingo dugu.
Abere guztiek txalotu zituzten lehendakariaren hitzak. Lehendakari berria hontz zahar bat zen, jakintsua oso.
Baina guztiek ez; apoak, zapoak, ez zuen txalorik jo. Berekola eta alferra zen Izugarri:
-Bain! Baina zera ere...! Igandean, nik, lo eta Igeri!
Eta igandean denak abiatu ziren lan egitera, gizarte hobeago bat jasotzarren.
Baina apoak, zapoak, uretan, honela zioen:
-Ergel ez besteak! Ni ez na lanera! Ura freskoa dago, eta zeinen ederra den hementxe lo hartzea...
Horrela egin zuen zenbait igandetan, hontzak berriro batzarra deltu zuen arte.
-Lagunok, denok gaude pozik egin dako lanarekin, baina badago hemen gizartean bizi nahi ez duen abere bat, lagunei lagundu nahi ez dien abere bat, eta, horregatik, bota egin behar duguna, kanpora jaurti, gure gizartetik.
Marmarreen hasi ziren abere guztiak, baina apoak, zapoak, ez zuen hitzik atera.
Honela jarraitu zuen hontzak:
-Ez dut abere alfer horren izena alpatuko, baina jakin ezazu berdea dela, begi irtenak dituela eta uretan bizi dela.
Guztiak bat egin zuten eta abere guztiak, hontzak eta apoak, zapoak, ez zuten hitzik atera.


Hizkuntza aukeratu : Euskara

Bidali Garbitu

ANALHITZAK testua automatikoki prozesatzen du IxaKat (euskararako) eta Ixa pipes (gaztelania eta ingeleserako) hizkuntza prozesatzaleen kate modularrak erabiliz.
ANALHITZA, hizkuntza-teknologiak ballatuz, humanitateetako eta giza zientzietako ikertzaileei datu linguistiko fidagarriak eta erraz manipula daitezkeenak eskaintzeko helburuarekin diseinatu da. Hauetako arloren batean zure ikerria definituta baduzu, testu-multzo bat eskura baduzu eta ANALHITZArekin

4. Gorde excel fitxategia:

EU EN ES



ANALHITZAK euskarazko, gaztelaniako edo ingeleseko testu batetik ondorengo informazio linguistikoa erauzten lagunduko dizu:

- Izenak, adjektiboak, aditzak, adberbioak...
- pertsona-izenak, leku-izenak...
- binakako, hirunakako eta launakako hitz-sekuentziak
- ... eta gehiago!


Testua fitxategi batean duzuna, hemen bertan kopiatuko duzuna edo web-orri batekoa izan daiteke, baina UTF8 kodeketan egon beharko du.

ANALHITZA erabiltzeko, sartu batere nahi duzun testua azpiko 3 aukera hauetakoren bat erabiliz, eta jarraian, testuaren hizkuntza aukeratu (euskara, gaztelania edo ingelesa). Apur bat itxaron ondoren, emaitzak Excel fitxategi batean jasoko dituzu. Horrela, emaitzak modu errazean moldatu ahal izango dituzue zure beharretara.

Fitxategia igo (txt formatuan) Sartu testua Uri-a sartu

Emaitza

Analizatutako fitxategi kopurua :

Sarrera Fitxategia	Sortua	Irteera Fitxategia
2019-07-09_14:43:36_BDcmGW.txt	July 09 2019 14:43:56	emaitza_2019-07-09_14:43:36_BDcmGW.xlsx 

ANALHITZAK testua automatikoki prozesatzen du ixaKat (euskararako) eta Ixa pipes (gaztelania eta ingelesezako) hizkuntza prozesatzaileen kate modularrak erabiliz.

ANALHITZA, hizkuntza-teknologiak baliatuz, humanitateetako eta giza zientzietako ikertzaileei datu linguistiko fidagarriak eta erraz manipula daitezkeenak eskaintzeko helburuarekin diseinatu da. Hauetako arloren batean zure ikergala definituta baduzu, testu-multzo bat eskura baduzu eta ANALHITZArekin aztertzeke interesa baduzu, idatzi eta lagunduko dizugu: mikel.iruskietala@ehu.es

ANALHITZA erabiltzen baduzu zure ikerketaren batean, ondorengo lan honen aipamena egin zure lan akademikoan mesedez:

Otegi, A. Imaz, O. Díaz de Ilarraza, A. Irukieta, M. Uria, L. 2017 ANALHITZA: a tool to extract linguistic information from large corpora in Humanities research. *Procesamiento del Lenguaje Natural* 58: 77-84.

5. Joan 11.
kalkulu-orrira:
"11_lemak.csv"

Dokumentua soilik irakurtzeko moduan ireki da.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	14	lukan															
2		9	izan														
3		8	eta														
4		8	ez														
5		7	egin														
6		6	abere														
7		5	bat														
8		5	baina														
9		4	gizarte														
10		4	igande														
11		4	lan														
12		4	apo														
13		4	lehendakari														
14		4	honela														
15		4	hori														
16		4	zapo														
17		3	berri														
18		3	ur														
19		3	guzti														
20		3	egon														
21		3	hitz														
22		2	alfer														
23		2	lo														
24		2	bai														
25		2	hontza														
26		2	lagun														
27		2	ni														
28		2	batzar														
29		2	orduan														
30		2	kanpo														
31		2	nahi														
32		2	*edun														
33		2	behar														
34		2	gu														
35		2	izendatu														
36		2	esan														
37		2	zahar														
38		1	zenbait														
39		1	krokodilo														

01_aurkibidea.csv 02_izenak.csv 03_adjetivoak.csv 04_aditzak.csv 05_adberbioak.csv 06_determinatzaileak.csv 07_konjuntzioak.csv iteak.csv 11_lemak.csv 12

Bilatu Bilatu denak Formatudun bistaratzea Bereizi maiuskula/minuskulak

10 / 16 orria PageStyle_11_lemak.csv

EAGLES (eta FREELING) etiketak

ANALHITZAko kategoriak ulertzeko

Position	Attribute	Values
0	category	N : <i>noun</i>
1	type	C : <i>common</i> ; P : <i>proper</i>
2	case	N : <i>nominative</i> ; G : <i>genitive</i> ; D : <i>dative</i> ; F : <i>accusative</i> ;
3	gen	F : <i>f</i> ; M : <i>m</i> ; C : <i>c</i>
4	num	S : <i>s</i> ; P : <i>p</i> ; N : <i>n</i>

Position		Values
0	A	adjective
0	D	determiner
0	F	Punctuation
0	N	noun
0	P	pronoun
0	R	adverb
0	S	adposition
0	V	verb
0	W	date
0	Z	number

[Freeling](#) etiketak

Eskerrik asko ANALHITZA
erabili izanagatik