

AZ UTÓNÉVKERESŐ

Sass Bálint

sass.balint@nytud.mta.hu

MTA Nyelvtudományi Intézet
Budapest, 2015. január 29.

AZ UTÓNÉVKERESŐ

Sass **Bálint**

sass.balint@nytud.mta.hu

MTA Nyelvtudományi Intézet
Budapest, 2015. január 29.

AZ UTÓNÉVKERESŐ

Sass **Bálint**

sass.balint@nytud.mta.hu

<http://corpus.nytud.hu/utonevkereso>

MTA Nyelvtudományi Intézet
Budapest, 2015. január 29.

Az MTA Nyelvtudományi Intézete által
 anyakönyvi bejegyzésre alkalmasnak minősített utónevek jegyzéke: férfinevek – 2015. január 1.

A, Á	Áhim	Ámor	Arlen	Bajka	Bekes
Aba	Ahmed	Ámos	Árlen	Bajnok	Békés
Abád	Airton	Anakin	Armand	Bakács	Bekő
Abbás	Ajád	Anasztáz	Armandó	Baksa	Béla
Abdiás	Ajándok	Anatol	Ármin	Bakta	Belár
Abdon	Ajtony	Andon	Arnó	Balabán	Belián
Abdullah	Akács	Andor	Arnold	Baladéva	Belizár
Ábel	Ákos	Andorás	Arnót	Balambér	Ben
Abelárd	Aladár	Andos	Áron	Balár	Benája
Ábner	Aladdin	András	Árpád	Balassa	Bence
Abod	Aladin	André	Arszák	Balás	Bende
Abony	Alajos	Andrej	Árszen	Baldó	Bendegúz
Abos	Alán	Andzseló	Arszlán	Baldvin	Bendit
Abosa	Alap	Angelus	Artemon	Bálint	Bene
Ábrahám	Alárd	Angelusz	Artúr	Balló	Benedek
Ábrám	Alarik	Ángyán	Arvéd	Balmaz	Benedikt
Ábrán	Albert	Anicét	Arvid	Baltazár	Benediktusz
Ábris	Albin	Aníziusz	Arzén	Bán	Benett
Absa	Aldán	Antal	Aser	Bandó	Béni
Absolon	Áldás	Antigon	Áser	Bánk	Benignusz
Acél	Aldó	Anton	Ashraf	Bános	Benitó
Achilles	Áldor	Antónió	Asur	Barabás	Benjamin
Achillesz	Alek	Antos	Aszáf	Baracs	Benjámín
			Aszter		

KONCEPCIÓ

- 1 teljesség
= anyakönyvi bejegyzésre alkalmasnak minősített összes utónév
- 2 hiteles, egységes adatok az utónevekről
- 3 friss statisztikák, trendek
- 4 szabad hozzáférés bárki számára

KONCEPCIÓ – 2013. NOVEMBER

- 1 teljesség
= anyakönyvi bejegyzésre alkalmasnak minősített összes utónév
→ 2013. június: 3529 név
- 2 hiteles, egységes adatok az utónevekről
→ folyamatos javítás
→ *formalizált eredet* a nevek kb. 60%-ánál
- 3 friss statisztikák, trendek
→ 2012-es adatok alapján
- 4 szabad hozzáférés bárki számára
→ ahogy eddig is

KONCEPCIÓ – 2013. NOVEMBER – 2015. JANUÁR

- 1 teljesség
= anyakönyvi bejegyzésre alkalmasnak minősített összes utónév
→ 2013. június: 3529 név – nem változott, azóta +140 név
- 2 hiteles, egységes adatok az utónevekről
→ folyamatos javítás
→ *formalizált eredet* a nevek kb. 100%-ánál
- 3 friss statisztikák, trendek
→ 2012-es adatok alapján – jelenleg is
- 4 szabad hozzáférés bárki számára
→ ahogy eddig is

A KERESŐFELÜLET

Funkció: tájékoztatás, segítség a névválasztáshoz

Keresési szempontok:

① „belső” adatok:

- név eleje/vége, szótagszám, hangrend, magánhangzók

② „külső” adatok:

- nem
- névnap
- becézés
- gyakorisági adatok

③ „külső” szabad szöveges leírások:

- eredet
- jelentés

A keresések eredményében *minden* kattintgatható! :)

A KERESŐFELÜLET

Funkció: tájékoztatás, segítség a névválasztáshoz

Keresési szempontok:

① „belső” adatok:

- név eleje/vége, szótagszám, hangrend, magánhangzók

② „külső” adatok:

- nem
- névnap
- becézés
- **gyakorisági adatok**

③ „külső” szabad szöveges leírások:

- **eredet** – *kereshetőség*: szabadszavas ↔ formalizált
- jelentés

A keresések eredményében *minden* kattintgatható! :)

Cél: „szemantikus keresés” az utónevek adataiban.

Nehézség: az eredetleírás szabad szöveg.

MEGOLDÁS:

a leírást szigorúan meghatározott formális alakra hozzuk, mely minden információt explicit tartalmaz, és géppel könnyen értelmezhető.

Formalizált alak:

nyelvi *elemek* +

az egyik elemet a másikba alakító *műveletek*.

- a *Bálint* formalizálása:

eredetleírás

A latin Valentinus családnévből származik

formalizált alak

latin:Valentinus"családnév" [névalkotás] [>] ~

Eredmény: egyértelmű, kereshető leírások.

Formalizált alak:

nyelvi *elemek* +

az egyik elemet a másikba alakító *műveletek*.

- a *Bálint* formalizálása:

eredetleírás

A latin Valentinus családnévből származik

formalizált alak

latin:Valentinus"családnév" [névalkotás] [>] ~

Eredmény: egyértelmű, kereshető leírások.

MEGTALÁLJUK...

- az összes *családnévből* származó nevet,
- az összes *névalkotással* létrejött nevet.

Formalizált alak:

nyelvi *elemek* +

az egyik elemet a másikba alakító *műveletek*.

- a *Bálint* formalizálása:

eredetleírás

A latin Valentinus családnévből származik

formalizált alak

latin:Valentinus"családnév" [névalkotás] [>] ~

Eredmény: egyértelmű, kereshető leírások.

MEGTALÁLJUK...

- az összes *családnévből* származó nevet,
- az összes *névalkotással* létrejött nevet.

A formalizálás jelentős részben *Horváth Ágnes* munkája.

eredetleírás	<i>A latin Valentinus családnévből származik</i>
formalizált alak	latin:Valentinus"családnév" [névalkotás] [>] ~
	elem [művelet] elem [művelet] elem

elem = nyelv(attrib):névalak"definíció"

műveletek

- [>] = „származik”
- [alakvált] = közelebből nem meghatározott típusú alakváltozata
- [formája] = toldalékos forma (kicsinyítőképzős, latinosított, továbbképzett)
- [női=párja] = adott férfinév női párja
- [rövidülése]
- [becézője]
- [névalkotás] = nem utónévből utónév (virágnév, közszó, családnév)
- [megfelelője] = más nyelvű megfelelő

[>] → [alakvált] → [formája], [női=párja], [rövidülése] → [becézője]

GYAKORISÁGOK ÉRTELMEZÉSE

Alap: össznépeség ↔ újszülöttek (2012)

Trend: ezek összevetése alapján

kategória	szöveges	rang
1	első tízben	-10
2	rendkívül gyakori	-20
3	nagyon gyakori	-50
4	közepesen gyakori	-100
5	elég ritka	-200
6	nagyon ritka	-500
7	rendkívül ritka	-1000
8	néhány előfordulás	a többi
<hr/>		
1-4	„mindennapos”	
1-5	„közismert”	
1-6	„különlegesebbek is”	

14 HÓNAP KERESÉSI ADATAIBÓL

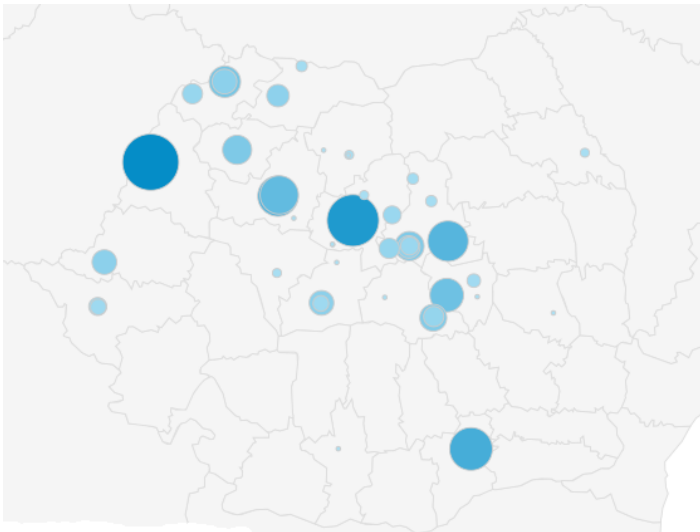
- 253 lájk
- 40000 felhasználó
- közel 200000 lekérdezés

14 HÓNAP KERESÉSI ADATAIBÓL

Területi megoszlás

14 HÓNAP KERESÉSI ADATAIBÓL

Területi megoszlás: *Románia*



14 HÓNAP KERESÉSI ADATAIBÓL

Esemény

14 HÓNAP KERESÉSI ADATAIBÓL

Esemény

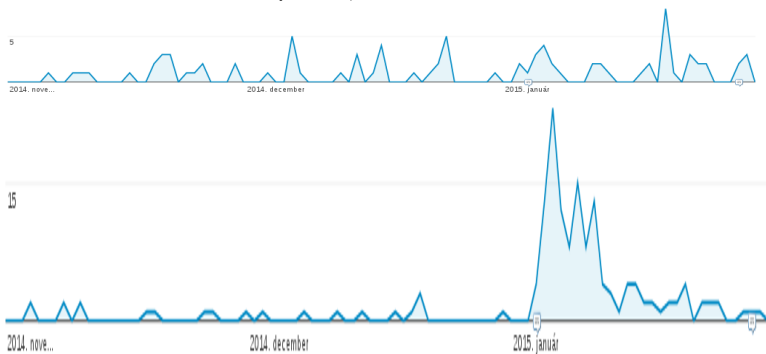
Bence az elmúlt 3 hónapban ↓



14 HÓNAP KERESÉSI ADATAIBÓL

Esemény

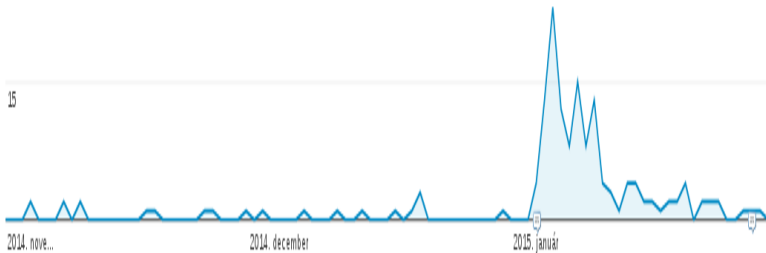
Bence az elmúlt 3 hónapban ↓



14 HÓNAP KERESÉSI ADATAIBÓL

Esemény

Bence az elmúlt 3 hónapban ↓



Rikárdó az elmúlt 3 hónapban ↑

AZ UTÓNÉVKERESŐ

Sass Bálint

sass.balint@nytud.mta.hu

<http://corpus.nytud.hu/utonevkereso>

MTA Nyelvtudományi Intézet
Budapest, 2015. január 29.