

Additional Wikipedia Based Features

- This feature fires when both strings redirect or lead directly to the same Wikipedia page.

United States

From Wikipedia, the free encyclopedia

(Redirected from [Untied States](#))

- Examples:
 - U.S. and Untied States redirect to United States
 - President Clinton redirects to Bill Clinton
 - Irish Potato Famine redirects to Great Irish Famine

Aliases

- This feature takes advantages of the links in Wikipedia's articles:

“Perot was not a fan of [President](#) George H. W. Bush and vigorously opposed ...”

“The device was hurriedly put together in an attempt to find the bullet in the body of [U.S. President](#) James Garfield.”

- Both “U.S. President” and “President” link to the article “President of the United States”.

Aliases

- This feature also introduces some noise: “British” links to “United States” somewhere in Wikipedia.
- To solve this problem, matches are weighted by how frequently the string links to a page; “British” links to “Britain” far more often than “United States”.

Lists

- This feature takes advantage of Wikipedia's list pages. If X belongs to a list of Ys, X is assumed to be a type of Y.

List of cities in Bosnia and Herzegovina

From Wikipedia, the free encyclopedia

This is a list of [cities](#) and [towns](#) in [Bosnia and Herzegovina](#):

FBH marks that the city is in the [Federation of Bosnia and Herzegovina](#), while RS marks that

The only places in Bosnia-Herzegovina with the legal status of a City are the capital [Sarajevo](#)

Only [Sarajevo](#) and [Banja Luka](#) have more than 150,000 inhabitants (according to the last census) that due to massive migrations during the 1992-1995 war and lack of a state-wide official census

- [Banja Luka](#) (de facto capital of [Republika Srpska](#))
- [Bihać](#) (FBH)
- [Berkovići](#) (RS)
- [Bijeljina](#) (RS)
- [Kozarska Dubica](#) (RS)
- [Bosanska Gradiška](#) (RS) - known 1992-2004 as [Gradiška](#)*
- [Bosanska Krupa](#) (FBH)
- [Bosanski Petrovac](#) (FBH)
- [Brčko](#) (Brčko is the seat of Brčko District, which is part of neither entity)

Gender

- When identifying the gender of a proper name, we compare the counts of male and female pronouns in the name's Wikipedia article to determine the gender.

Picabo Street

From Wikipedia, the free encyclopedia

Picabo Street (born [April 3, 1971](#), in [Triumph, Idaho](#)) is an [American skier](#), now retired and living in [Park City, Utah](#). She was raised on a small [farm](#) in Triumph, several miles southeast of [Sun Valley, Idaho](#), where she learned to ski and race.

She first joined the [United States Ski Team](#) in [1989](#), at the age of 17. She has written an [autobiography](#) entitled *Nothing to Hide* ([ISBN 0-07-140693-X](#)). She appeared on the TV show [Nickelodeon GUTS](#) in [1994](#).

Additional Examples

- Alias:
 - “Body” - “Corpse”
 - “Clinton” - “President Clinton”
- Redirect:
 - “Lewinsky” - “Monica Lewinsky”
- Gender:
 - “Bono” - Male
 - “Khun Sa” - Male

Results (ACE-02 BNews)

	Recall	Precision	F-Score
Baseline	51.2%	69.9%	59.1%
WikiCats	56.8%	68.0%	61.9%
WikiOther	56.5%	67.9%	61.7%
All	60.7%	65.4%	63.0%

Results (ACE-02 NWire)

	Recall	Precision	F-Score
Baseline	49.9%	68.5%	57.8%
WikiCats	57.6%	63.5%	60.4%
WikiOther	56.8%	63.6%	60.0%
All	N/A	N/A	N/A

Results (ACE-02 NPaper)

	Recall	Precision	F-Score
Baseline	54.9%	73.3%	62.7%
WikiCats	59.7%	68.4%	63.7%
WikiOther	59.8%	68.5%	63.9%
All	64.1%	67.7%	65.8%