Curious Code on on Wikipedia

David Branner
Hack && Tell, New York
20150826

English Wikipedia pages contain some strange code

English Wikipedia pages contain some strange code — a list that is always empty

English Wikipedia pages contain some strange code — a list that is always empty

```
<div id="p-variants"
     role="navigation"
     class="vectorMenu emptyPortlet"
     aria-labelledby="p-variants-label">
  <h3 id="p-variants-label">
    <span>Variants</span><a href="#"></a>
  </h3>
  <div class="menu">
    <ul></ul>
  </div>
</div>
```

English Wikipedia pages contain some strange code – a list that is always empty and a class to mark it as such:

```html
<div id="p-variants"
     role="navigation"
     class="vectorMenu emptyPortlet"
     aria-labelledby="p-variants-label">
  <h3 id="p-variants-label">
    <span>Variants</span><a href="#"></a>
  </h3>
  <div class="menu">
    <ul></ul>
  </div>
</div>
```

English Wikipedia pages contain some strange code – a list that is always empty and a class to mark it as such:

```
<div id="p-variants"
     role="navigation"
     class="vectorMenu emptyPortlet"
     aria-labelledby="p-variants-label">
  <h3 id="p-variants-label">
    <span>Variants</span><a href="#"></a>
  </h3>
  <div class="menu">
    <ul></ul>
  </div>
</div>
```

English Wikipedia pages contain some strange code – a list that is always empty and a class to mark it as such:

```
<div id="p-variants"
     role="navigation"
     class="vectorMenu emptyPortlet"
     aria-labelledby="p-variants-label">
  <h3 id="p-variants-label">
    <span>Variants</span><a href="#"></a>
  </h3>
  <div class="menu">
    <ul></ul>
  </div>
</div>
```

This p-variants **div generates a** "Variants" **list**

English Wikipedia pages contain some strange code – a list that is always empty and a class to mark it as such:

```
<div id="p-variants"
    role="navigation"
    class="vectorMenu emptyPortlet"
    aria-labelledby="p-variants-label">
  <h3 id="p-variants-label">
    <span>Variants</span><a href="#"></a>
  </h3>
  <div class="menu">
    <ul></ul>
  </div>
</div>
```

This `p-variants` div generates a "Variants" list, invisible on English pages because of class "emptyPortlet".

There is a similar `p-variants` div code block in the pages for many other languages:

There is a similar `p-variants` div code block in the pages for many other languages:

- French: …`<span>`Variantes`</span>`…

There is a similar `p-variants` div code block in the pages for many other languages:

- French: …`<span>`Variantes`</span>`…
- Nepali: …`<span>`बहुरुपहरु`</span>`…

There is a similar `p-variants` div code block in the pages for many other languages:

- French: `...<span>Variantes</span>...`
- Nepali: `...<span>बहुरुपहरु</span>...`
- Japanese `...<span>変種</span>...`

There is a similar `p-variants` div code block in the pages for many other languages:

- French: ...`<span>`Variantes`</span>`...
- Nepali: ...`<span>`बहुरुपहरु`</span>`...
- Japanese ...`<span>`変種`</span>`...
- Russian ...`<span>`Варианты`</span>`...

There is a similar `p-variants` div code block in the pages for many other languages:

- French: …`<span>`Variantes`</span>`…
- Nepali: …`<span>`बहुरुपहरु`</span>`…
- Japanese   …`<span>`変種`</span>`…
- Russian …`<span>`Варианты`</span>`…
- Georgian   …`<span>`ვარიანტები`</span>`…

There is a similar `p-variants` div code block in the pages for many other languages:

- French: …`<span>`Variantes`</span>`…
- Nepali: …`<span>`बहुरुपहरु`</span>`…
- Japanese   …`<span>`変種`</span>`…
- Russian …`<span>`Варианты`</span>`…
- Georgian   …`<span>`ვარიანტები`</span>`…
- and so on,

There is a similar `p-variants` div code block in the pages for many other languages:

- French: …`<span>`Variantes`</span>`…
- Nepali: …`<span>`बहुरुपहरु`</span>`…
- Japanese   …`<span>`変種`</span>`…
- Russian …`<span>`Варианты`</span>`…
- Georgian   …`<span>`ვარიანტები`</span>`…
- and so on,

none of which heads a list on the page in question.

There is a similar `p-variants` div code block in the pages for many other languages:

- French: …`<span>`Variantes`</span>`…
- Nepali: …`<span>`बहुरूपहरू`</span>`…
- Japanese …`<span>`変種`</span>`…
- Russian …`<span>`Варианты`</span>`…
- Georgian …`<span>`ვარიანტები`</span>`…
- and so on,

none of which heads a list on the page in question.

That is the situation on pages for most languages.

Where does this `p-variants` div actually have an effect?

Where does this `p-variants` div actually have an effect?

For one, on Serbian pages:

https://sr.wikipedia.org/wiki/Пјатра_Њамц

Where does this `p-variants` div actually have an effect?

For one, on Serbian pages:

https://sr.wikipedia.org/wiki/Пјатра_Њамц

```
… <span>sr</span><a href="#"></a> …
<div class="menu">
  <ul>
    <li id="ca-varlang-0" class="selected"><a
href="/sr/Пјатра_Њамц" lang="sr" hreflang="sr" >sr</
a></li>
    <li id="ca-varlang-1"><a href="/sr-ec/
Пјатра_Њамц" lang="sr-EC" hreflang="sr-EC" >Ћирилица</
a></li>
    <li id="ca-varlang-2"><a href="/sr-el/
Пјатра_Њамц" lang="sr-EL" hreflang="sr-EL" >Latinica</
a></li>
  </ul>
</div>
```

Where does this `p-variants` div actually have an effect?
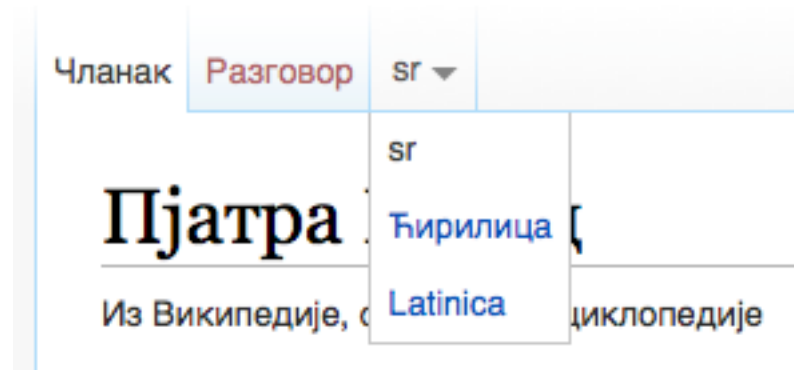
For one, on Serbian pages:

https://sr.wikipedia.org/wiki/Пјатра_Њамц

```
… <span>sr</span><a href="#"></a> …
<div class="menu">
   <ul>
      <li id="ca-varlang-0" class="selected"><a
href="/sr/Пјатра_Њамц" lang="sr" hreflang="sr" >sr</
a></li>
      <li id="ca-varlang-1"><a href="/sr-ec/
Пјатра_Њамц" lang="sr-EC" hreflang="sr-EC" >Ћирилица</
a></li>
      <li id="ca-varlang-2"><a href="/sr-el/
Пјатра_Њамц" lang="sr-EL" hreflang="sr-EL" >Latinica</
a></li>
   </ul>
</div>
```
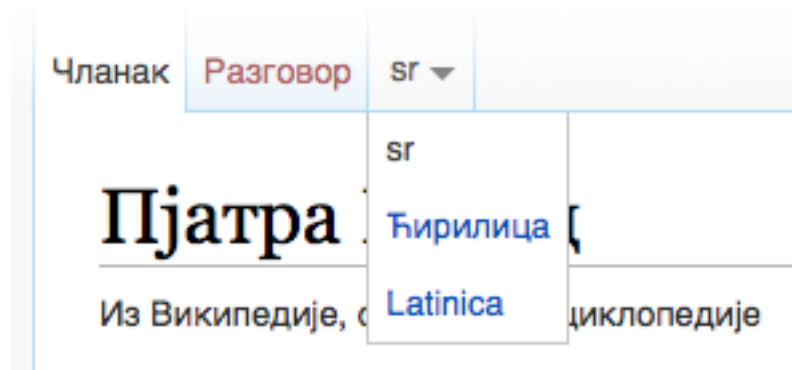
This code creates a menu

Branner, Curious Code on Wikipedia (20150826) p. 21

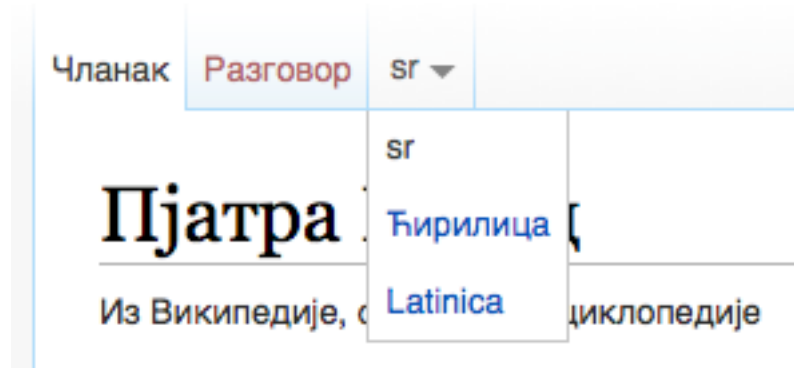# This code creates a menu

This code creates a menu
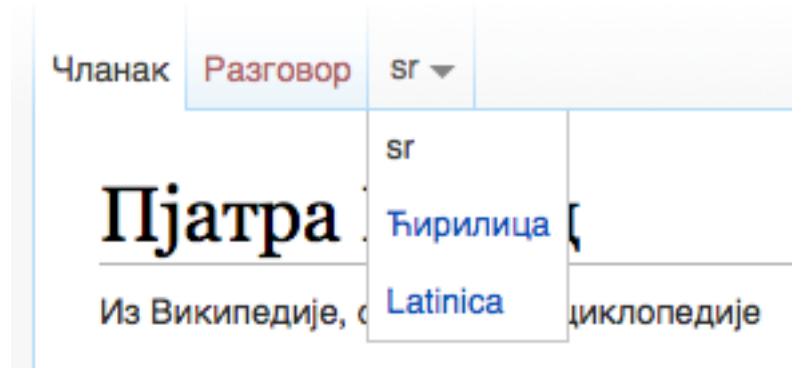
allowing the user to select which of two scripts

This code creates a menu

allowing the user to select which of two scripts — "Ћирилица" (Cyrillic) or "Latinica" (Roman) the content appears in.

Branner, Curious Code on Wikipedia (20150826) p.24

This code creates a menu

allowing the user to select which of two scripts — "Ћирилица" (Cyrillic) or "Latinica" (Roman) the content appears in.

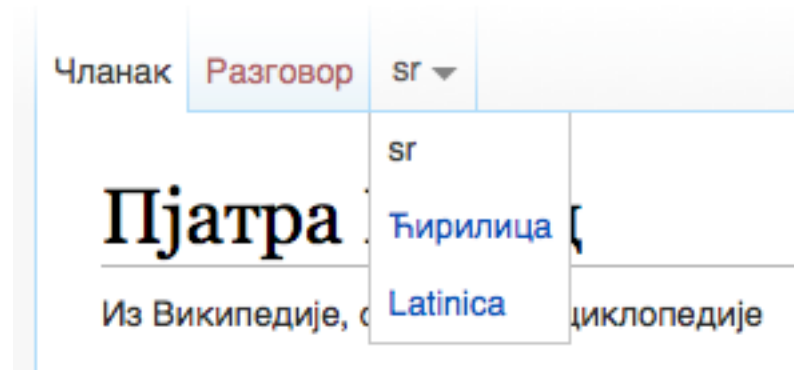The two forms of the page have identical content

This code creates a menu

allowing the user to select which of two scripts — "Ћирилица" (Cyrillic) or "Latinica" (Roman) the content appears in.

The two forms of the page have identical content — the only difference is which script that content is displayed in.
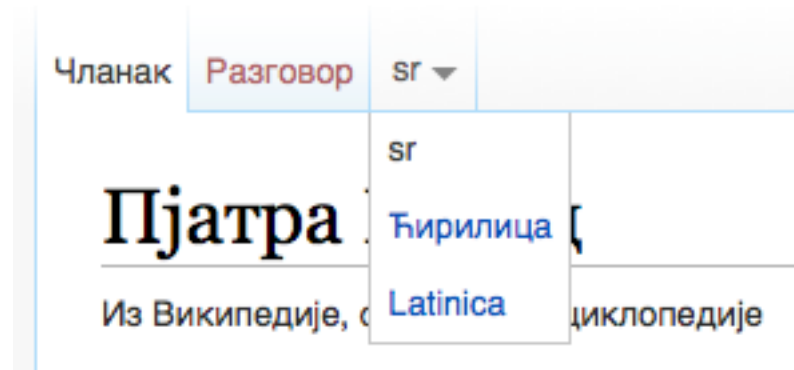
This code creates a menu

allowing the user to select which of two scripts — "Ћирилица" (Cyrillic) or "Latinica" (Roman) the content appears in.

The two forms of the page have identical content — the only difference is which script that content is displayed in. Different Serbian-speakers may read in either script.

Cyrillic:

Чланак   Разговор   Ћирилица ▾

Пјатра Њамц

Из Википедије, слободне енциклопедије

Пјатра Њамц (рум. *Piatra Neamț*) је град у Румунији, у источном делу земље, у ист управно средиште округа Њамц.

Пјатра Њамц има површину од 77,4 km, на којој је по последњем попису из 2002. г

Cyrillic:

Пјатра Њамц

Из Википедије, слободне енциклопедије

Пјатра Њамц (рум. *Piatra Neamț*) је град у Румунији, у источном делу земље, у ист
управно средиште округа Њамц.

Пјатра Њамц има површину од 77,4 km, на којој је по последњем попису из 2002. г

Roman

Pjatra Njamc

Из Википедије, слободне енциклопедије

Pjatra Njamc (rum. *Piatra Neamț*) je grad u Rumuniji, u istočnom delu zemlje, u ist
središte okruga Njamc.

Pjatra Njamc ima površinu od 77,4 km, na kojoj je po poslednjem popisu iz 2002. g

Branner, Curious Code on Wikipedia (20150826) p.29

A change to the content of the page is reflected immediately on both versions:

A change to the content of the page is reflected immediately on both versions:

I doubled a word in the Cyrillic version:

A change to the content of the page is reflected immediately on both versions:

I doubled a word in the Cyrillic version:

на којој је по последњем попису из 2002.

на којој је по последњем последњем попису из 2002.

A change to the content of the page is reflected immediately on both versions:

I doubled a word in the Cyrillic version:

на којој је по последњем попису из 2002.

на којој је по последњем последњем попису из 2002.

and it immediately doubled itself in the Roman version:

A change to the content of the page is reflected immediately on both versions:

I doubled a word in the Cyrillic version:

на којој је по последњем попису из 2002.

на којој је по последњем последњем попису из 2002.

and it immediately doubled itself in the Roman version:

na kojoj je po poslednjem popisu iz 2002.

na kojoj je po poslednjem poslednjem popisu iz 2002.

A change to the content of the page is reflected immediately on both versions:

I doubled a word in the Cyrillic version:

на којој је по последњем попису из 2002.

на којој је по последњем последњем попису из 2002.

and it immediately doubled itself in the Roman version:

na kojoj je po poslednjem popisu iz 2002.

na kojoj je po poslednjem poslednjem popisu iz 2002.

(Yes, I deleted it afterwards. And yes, it disappeared immediately from both versions.)

Other languages offering you this sort of choice:

Other languages offering you this sort of choice:

	Cyrillic	Roman
Tajik:	кирилликӣ	lotinī

Other languages offering you this sort of choice:

	Cyrillic	*Roman*
Tajik:	кирилликӣ	lotinī
Uzbek:	кирилл	lotin

Other languages offering you this sort of choice:

	Cyrillic	*Roman*	*Arabic*
Tajik:	кирилликӣ	lotinī	
Uzbek:	кирилл	lotin	
Kurdish:		Tîpên latînî	Tîpên erebî

Other languages offering you this sort of choice:

	Cyrillic	Roman	Arabic
Tajik:	кирилликӣ	lotinī	
Uzbek:	кирилл	lotin	
Kurdish:		Tîpên latînî	Tîpên erebî
Kazakh:	Кирил	Latın	توته

Other languages offering you this sort of choice:

	Cyrillic	Roman	Arabic
Tajik:	кирилликӣ	lotinī	
Uzbek:	кирилл	lotin	
Kurdish:		Tîpên latînî	Tîpên erebî
Kazakh:	Кирил	Latın	توته

ᑎᑎ� ᓄᑦᕿ

Other languages offering you this sort of choice:

	Cyrillic	*Roman*	*Arabic*	**?**
Tajik:	кирилликӣ	lotinī		
Uzbek:	кирилл	lotin		
Kurdish:		Tîpên latînî	Tîpên erebî	
Kazakh:	Кирил	Latın	توته	
?				ⴖⴖ�⁴ⴖ ⴑⵀⴕ⁴ⴖ

Other languages offering you this sort of choice:

	Cyrillic	Roman	Arabic	Canadian Aboriginal Syllabics
Tajik:	к и р и л л и к ӣ	lotinī		
Uzbek:	к и р и л л	lotin		
Kurdish:		Tîpên latînî	Tîpên erebî	
Kazakh:	К и р и л	Latın	توته	
Inuktitut:		ilisautik		ᑎᑎᕋᐅᓯᖅ ᓄᑖᖅ

Other languages offering you this sort of choice:

	Cyrillic	*Roman*	*Arabic*	*Canadian Aboriginal Syllabics*
Tajik:	кирилликӣ	lotinī		
Uzbek:	кирилл	lotin		
Kurdish:		Tîpên latînî	Tîpên erebî	
Kazakh:	Кирил	Latın	توته	
Inuktitut:		ilisautik		ᐅᓇᓕᐊᕐᔪᖅ ᓄᑖᖅ

and Chinese…

Chinese has up to eight options in the `p-variants` menu-div:

Chinese has up to eight options in the `p-variants` **menu-div:**

不转换　　　 [no transformation — may be mixed forms]

简体　　　　 [generic simplified forms]

繁體　　　　 [generic traditional forms]

大陆简体　　 [Mainland China simplified forms]

香港繁體　　 [Hong Kong traditional forms]

澳門繁體　　 [Macao traditional forms]

马新简体　　 [Malaysia-Singapore simplified forms]

台灣正體　　 [Taiwan "official" forms (avoids "繁體")]

Chinese has up to eight options in the `p-variants` menu-div:

不转换 　　　[no transformation — may be mixed forms]
简体 　　　　[generic simplified forms]
繁體 　　　　[generic traditional forms]
大陆简体 　　[Mainland China simplified forms]
香港繁體 　　[Hong Kong traditional forms]
澳門繁體 　　[Macao traditional forms]
马新简体 　　[Malaysia-Singapore simplified forms]
台灣正體 　　[Taiwan "official" forms (avoids "繁體")]

Why does Chinese need the last five regional forms? Aren't generic "simplified" and "traditional" enough?

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation]

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation] — not only the scripts, but also actual words are being exchanged.

Branner, Curious Code on Wikipedia (20150826) p.49

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation] — not only the scripts, but also actual words are being exchanged. But there is still only a single page.

Branner, Curious Code on Wikipedia (20150826) p.50

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation] — not only the scripts, but also actual words are being exchanged. But there is still only a single page.

As an example, here are the three names registered for the title of the movie "The Matrix":

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation] — not only the scripts, but also actual words are being exchanged. But there is still only a single page.

As an example, here are the three names registered for the title of the movie "The Matrix":

generic simplified:
　　黑客帝国 [Hacker Empire]

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation] — not only the scripts, but also actual words are being exchanged. But there is still only a single page.

As an example, here are the three names registered for the title of the movie "The Matrix":

generic simplified:

黑客帝国 [Hacker Empire]

Hong Kong:

廿二世紀殺人網絡 [22nd Century Web of Murder]

The ideal being followed on Chinese Wikipedia is to have "字词转换" [character or word transformation] — not only the scripts, but also actual words are being exchanged. But there is still only a single page.

As an example, here are the three names registered for the title of the movie "The Matrix":

generic simplified:

黑客帝国 [Hacker Empire]

Hong Kong:

廿二世紀殺人網絡 [22nd Century Web of Murder]

Taiwan:

駭客任務 [Hacker Duty] (different word for "hacker")

Here are transcriptions of Hillary Clinton's name:

Mainland China: 希拉里 [Xīlālǐ]
Hong Kong:　　　希拉莉 [Xīlālì]
Taiwan:　　　　希拉蕊 [Xīlāruǐ]

Here are transcriptions of Hillary Clinton's name:

Mainland China: 希拉里 [Xīlālǐ]
Hong Kong:　　　希拉莉 [Xīlālì]
Taiwan:　　　　希拉蕊 [Xīlāruǐ]

Here is Oprah Winfrey:

Mainland China: 奥普拉·温弗里 [Àopǔlā Wēnfúlǐ]
Hong Kong:　　　奧花·雲費　　[Àohuā Yúnfèi]
Taiwan:　　　　歐普拉·溫芙蕾 [Ōupǔlā Wēnfúléi]

"highway junction / interchange":

Mainland China: 立交桥 [elevated intersection bridge]
Hong Kong: 交匯處 [convergence place]
Taiwan: 交流道 [exchange road]

"highway junction / interchange":

Mainland China: 立交桥 [elevated intersection bridge]
Hong Kong:　　　交匯處 [convergence place]
Taiwan:　　　　交流道 [exchange road]

"motor":

Mainland China: 电动机 [electric machine]
Hong Kong:　　　摩打　 [módǎ (phonetic transcription)]
Taiwan:　　　　馬達　 [mǎdá (phonetic transcription)]

Observations:

Observations:

- Most examples involve proper nouns and other words of foreign origin.

Observations:

- Most examples involve proper nouns and other words of foreign origin.
- Hong Kong often but not always reflects Cantonese pronunciation, but it can be read off in Mandarin.

Observations:

- Most examples involve proper nouns and other words of foreign origin.
- Hong Kong often but not always reflects Cantonese pronunciation, but it can be read off in Mandarin.
- Most often it is Taiwan and Mainland China that have different words; Hong Kong appears less often, and Macao and Malaysia-Singapore are very rare.

Observations:

- Most examples involve proper nouns and other words of foreign origin.
- Hong Kong often but not always reflects Cantonese pronunciation, but it can be read off in Mandarin.
- Most often it is Taiwan and Mainland China that have different words; Hong Kong appears less often, and Macao and Malaysia-Singapore are very rare.

I attribute this to the Chinese tradition of viewing Chinese language as being a single entity, however varied regionally.

I used to teach a course in varieties (accents, syntax, and diction) of Mandarin, and these regional synonyms were a useful resource.

I used to teach a course in varieties (accents, syntax, and diction) of Mandarin, and these regional synonyms were a useful resource. The material is much fuller now than when I was teaching Mandarin regularly, and it seems someone is paying attention to making it available.

I used to teach a course in varieties (accents, syntax, and diction) of Mandarin, and these regional synonyms were a useful resource. The material is much fuller now than when I was teaching Mandarin regularly, and it seems someone is paying attention to making it available.

Not only page titles but sometimes many other words are assigned regional synonyms.

I used to teach a course in varieties (accents, syntax, and diction) of Mandarin, and these regional synonyms were a useful resource. The material is much fuller now than when I was teaching Mandarin regularly, and it seems someone is paying attention to making it available.

Not only page titles but sometimes many other words are assigned regional synonyms. The words are placed in a `div` with a special attribute, `data-noteta-code`:

```
<div class="noteTA-title"
    data-noteta-code="zh-cn:贝拉克·奥巴马；zh-tw:巴拉克·歐巴馬；zh-hk:巴拉克·奧巴馬；">
</div> ...
<div class="noteTA-local">
<div data-noteta-code="zh-cn:肯尼亚；zh-tw:肯亞；zh-hk:肯雅；"></div>
<div data-noteta-code="zh-cn:可卡因；zh-tw:古柯鹼；zh-hk:可卡因；"></div>
<div data-noteta-code="zh-cn:伊利诺伊州；zh-tw:伊利諾州；zh-hk:伊利諾伊州；"></div>
<div data-noteta-code="zh-cn:火奴鲁鲁；zh-tw:檀香山；zh-hk:檀香山；"></div>
<div data-noteta-code="zh-cn:新罕布什尔州；zh-tw:新罕布夏州；zh-hk:新罕布什爾州；"></div>
<div data-noteta-code="zh-cn:切尔西·克林顿；zh-hk:切爾西·克林頓；zh-tw:雀兒喜·柯林頓"></div>
<div data-noteta-code="zh-cn:希拉里；zh-hk:希拉莉；zh-tw:希拉蕊"></div>
<div data-noteta-code="zh-cn:米歇尔；zh-tw:蜜雪兒；zh-hk:米歇爾"></div>
<div data-noteta-code="zh-cn:奥普拉·温弗里；zh-hk:奧花·雲費；zh-tw:歐普拉·溫芙蕾"></div>
<div data-noteta-code="zh-cn:乔治·克鲁尼；zh-tw:喬治·克隆尼；zh-hk:佐治·古尼；"></div>
<div data-noteta-code="zh-cn:格莱美奖；zh-tw:葛萊美獎；zh-hk:格林美獎"></div> ...
```

It seems possible that this content is being supplied automatically from an external reference (rather than being assembled descriptively from actual pages), although if so I haven't located such a source yet.

It seems possible that this content is being supplied automatically from an external reference (rather than being assembled descriptively from actual pages), although if so I haven't located such a source yet.

In the mean time, it's a simple matter to scrape this content using the `lxml` module:

Notice how many more words than just the title are supplied:

Notice how many more words than just the title are supplied:

```
In [1]: import scrape as S
ln [2]: S.get_words('黑客帝国') # English: 'The Matrix'
Out[2]:
[{'zh-hk': '廿二世紀殺人網絡', 'zh-hans': '黑客帝国', 'zh-tw': '駭客任
務'},
 {'zh-hant': '尼歐', 'zh-hans': '尼奥'},
 {'zh-hant': '母體', 'zh-hans': '矩阵'},
 {'zh-cn': '史密斯特工', 'zh-hk': '史特工', 'zh-tw': '史密斯探員'},
 {'zh-cn': '托马斯·安德森', 'zh-hk': '安湯武', 'zh-tw': '湯瑪斯·安德
森'},
 {'zh-cn': '艾波克', 'zh-hk': '天啟', 'zh-tw': '艾巴'},
 {'zh-cn': '开关', 'zh-hk': '變體', 'zh-tw': '蘇薇琪'},
 {'zh-cn': '塞弗', 'zh-hk': '暗碼', 'zh-tw': '塞佛'},
 {'zh-cn': '坦克', 'zh-hk': '戰車', 'zh-tw': '坦克'},
 {'zh-cn': '多泽', 'zh-hk': '推土機', 'zh-tw': '道瑟'},
 {'zh-cn': '耗子', 'zh-hk': '滑鼠', 'zh-tw': '茂史'},
 {'zh-cn': '先知', 'zh-hk': '祭司', 'zh-tw': '先知'},
 {'zh-cn': '重裝上陣', 'zh-hk': '決戰未來', 'zh-tw': '重裝上陣'},
 {'zh-cn': '矩阵革命', 'zh-hk': '驚變世紀', 'zh-tw': '最後戰役'},
 {'zh-hant': '程式', 'zh-hans': '程序'}]
```

```
In [3]: s.get_words('馬達') # English: 'motor'
Out[3]: [{'zh-hk': '摩打', 'zh-hans': '电动机', 'zh-tw': '馬達'}]
```

```
In [3]: s.get_words('馬達') # English: 'motor'
Out[3]: [{'zh-hk': '摩打', 'zh-hans': '电动机', 'zh-tw': '馬達'}]

In [4]: S.get_words('程式') # English: 'computer program'
Out[4]:
[{'zh-hant': '電腦程式', 'zh-hans': '计算机程序'},
 {'zh-hant': '執行', 'zh-hans': '运行'},
 {'zh-hant': '程式設計', 'zh-hans': '编程'},
 {'zh-hans': '计算机', 'zh-mo': '電腦', 'zh-sg': '计算机', 'zh-cn':
'计算机', 'zh-hant': '電腦', 'zh-tw': '電腦', 'zh-hk': '電腦'},
 {'zh-hant': '程式', 'zh-hans': '程序'},
 {'zh-hk': '軟件', 'zh-cn': '软件', 'zh-hans': '软件', 'zh-tw': '軟
體'}]
```

```
In [3]: s.get_words('馬達') # English: 'motor'
Out[3]: [{'zh-hk': '摩打', 'zh-hans': '电动机', 'zh-tw': '馬達'}]

In [4]: S.get_words('程式') # English: 'computer program'
Out[4]:
[{'zh-hant': '電腦程式', 'zh-hans': '计算机程序'},
 {'zh-hant': '執行', 'zh-hans': '运行'},
 {'zh-hant': '程式設計', 'zh-hans': '编程'},
 {'zh-hans': '计算机', 'zh-mo': '電腦', 'zh-sg': '计算机', 'zh-cn':
'计算机', 'zh-hant': '電腦', 'zh-tw': '電腦', 'zh-hk': '電腦'},
 {'zh-hant': '程式', 'zh-hans': '程序'},
 {'zh-hk': '軟件', 'zh-cn': '软件', 'zh-hans': '软件', 'zh-tw': '軟
體'}]

In [5]: S.get_words('滑鼠') # English: 'computer mouse'
Out[5]:
[{'zh-hant': '滑鼠', 'zh-hans': '鼠标', 'zh-sg': '滑鼠'},
 {'zh-hant': '滑鼠', 'zh-hans': '鼠标', 'zh-sg': '滑鼠'},
 {'zh-hk': '鐳射', 'zh-hans': '激光', 'zh-tw': '雷射', 'zh-sg': '鐳
射'},
 {'zh-hant': '簡報', 'zh-hans': '演示'},
 {'zh-hant': '光學', 'zh-hans': '光电', 'zh-sg': '光学'}]
```

Branner, Curious Code on Wikipedia (20150826) p.75

```
In [3]: s.get_words('馬達')  # English: 'motor'
Out[3]: [{'zh-hk': '摩打', 'zh-hans': '电动机', 'zh-tw': '馬達'}]

In [4]: S.get_words('程式')  # English: 'computer program'
Out[4]:
[{'zh-hant': '電腦程式', 'zh-hans': '计算机程序'},
 {'zh-hant': '執行', 'zh-hans': '运行'},
 {'zh-hant': '程式設計', 'zh-hans': '编程'},
 {'zh-hans': '计算机', 'zh-mo': '電腦', 'zh-sg': '计算机', 'zh-cn':
'计算机', 'zh-hant': '電腦', 'zh-tw': '電腦', 'zh-hk': '電腦'},
 {'zh-hant': '程式', 'zh-hans': '程序'},
 {'zh-hk': '軟件', 'zh-cn': '软件', 'zh-hans': '软件', 'zh-tw': '軟
體'}]

In [5]: S.get_words('滑鼠')  # English: 'computer mouse'
Out[5]:
[{'zh-hant': '滑鼠', 'zh-hans': '鼠标', 'zh-sg': '滑鼠'},
 {'zh-hant': '滑鼠', 'zh-hans': '鼠标', 'zh-sg': '滑鼠'},
 {'zh-hk': '鐳射', 'zh-hans': '激光', 'zh-tw': '雷射', 'zh-sg': '镭
射'},
 {'zh-hant': '簡報', 'zh-hans': '演示'},
 {'zh-hant': '光學', 'zh-hans': '光电', 'zh-sg': '光学'}]
```

Lesson? Be surprised at what is hidden on the Web...

End