

# Assignment 1: Evaluate name entity linking

**Remko Boschker s1282603**

This assignment compares the quality of named entity linking using [Open Calais](http://www.opencalais.com/opencalais-demo/) (<http://www.opencalais.com/opencalais-demo/>) and [DBpedia Spotlight](http://dbpedia-spotlight.github.io/demo/) (<http://dbpedia-spotlight.github.io/demo/>) for two fragments of English text (the Open Calais demonstrator does not support Dutch).

Sometimes recognizing amounts, dates, distances etc. is considered part of this task. It is not included here.

For Open Calais I have added the type of named entity that is encoded in the html in square brackets. Open Calais also does pronoun resolution. It works correctly and is not taken into account. Open Calais has a sophisticated recognition of positions and careers. However I have not included them in the named entity recognition task because they are not proper names or biological kinds or substances.

DBpedia Spotlight does not categorize the named entities found. If the entity found refers to the correct DBpedia page it is counted as correct. I have added the type in square brackets manually.

## First text fragment

[Unesco about the Silkroad](http://en.unesco.org/silkroad/about-silk-road) (<http://en.unesco.org/silkroad/about-silk-road>)

Indeed, the man who is often credited with founding the Silk Roads by opening up the first route from China to the West in the 2nd century BC, General Zhang Qian, was on a diplomatic mission rather than a trading

expedition. Sent to the West in 139 BC by the Han Emperor Wudi to ensure alliances against the Xiongnu, the hereditary enemies of the Chinese, Zhang Qian was captured and imprisoned by them. Thirteen years later he escaped and made his way back to China. Pleased with the wealth of detail and accuracy of his reports, the emperor sent Zhang Qian on another mission in 119 BC to visit several neighbouring peoples, establishing early routes from China to Central Asia.

## Open Calais

Indeed, the man who is often credited with founding the Silk Roads by opening up the first route from China [country] to the West in the 2nd century BC [province or state], General Zhang Qian [person], was on a diplomatic mission rather than a trading expedition. Sent to the West in 139 BC [province or state] by the Han Emperor Wudi [person] to ensure alliances against the Xiongnu, the hereditary enemies of the Chinese, Zhang Qian [person] was captured and imprisoned by them. Thirteen years later he [person] escaped and made his [person] way back to China [country]. Pleased with the wealth of detail and accuracy of his [person] reports, the emperor sent Zhang Qian [person] on another mission in 119 BC to visit several neighbouring peoples, establishing early routes from China [country] to Central Asia [region]

## DBpedia Spotlight

Indeed, the man who is often credited with founding the [Silk Roads](http://dbpedia.org/resource/Silk_Road) ([http://dbpedia.org/resource/Silk\\_Road](http://dbpedia.org/resource/Silk_Road)) [geographical feature] by opening up the first route from [China](http://dbpedia.org/resource/China) (<http://dbpedia.org/resource/China>) [country] to the West in the 2nd century BC, General [Zhang Qian](http://dbpedia.org/resource/Zhang_Qian)

[http://dbpedia.org/resource/Zhang\\_Qian](http://dbpedia.org/resource/Zhang_Qian) [**person**], was on a [diplomatic mission](http://dbpedia.org/resource/Diplomatic_mission) ([http://dbpedia.org/resource/Diplomatic\\_mission](http://dbpedia.org/resource/Diplomatic_mission)) [**not a named entity**] rather than a trading expedition. Sent to the West in 139 BC by the Han [Emperor Wudi](http://dbpedia.org/resource/Emperor_Wu_of_Han) ([http://dbpedia.org/resource/Emperor\\_Wu\\_of\\_Han](http://dbpedia.org/resource/Emperor_Wu_of_Han)) [**person**] to ensure alliances against the [Xiongnu](http://dbpedia.org/resource/Xiongnu) (<http://dbpedia.org/resource/Xiongnu>) [**people**], the hereditary enemies of the Chinese, [Zhang Qian](http://dbpedia.org/resource/Zhang_Qian) ([http://dbpedia.org/resource/Zhang\\_Qian](http://dbpedia.org/resource/Zhang_Qian)) [**person**] was captured and imprisoned by them. Thirteen years later he escaped and made his way back to [China](http://dbpedia.org/resource/China) (<http://dbpedia.org/resource/China>) [**country**]. Pleased with the wealth of detail and accuracy of his reports, the emperor sent [Zhang Qian](http://dbpedia.org/resource/Zhang_Qian) ([http://dbpedia.org/resource/Zhang\\_Qian](http://dbpedia.org/resource/Zhang_Qian)) [**person**] on another mission in 119 BC ([http://dbpedia.org/resource/119\\_BC](http://dbpedia.org/resource/119_BC)) [**not a named entity**] to visit several neighbouring peoples, establishing early routes from [China](http://dbpedia.org/resource/China) (<http://dbpedia.org/resource/China>) [**country**] to [Central Asia](http://dbpedia.org/resource/Tarim_Basin) ([http://dbpedia.org/resource/Tarim\\_Basin](http://dbpedia.org/resource/Tarim_Basin)) [**region incorrect**].

## Manual Annotation

Indeed, the man who is often credited with founding the Silk Roads [**geographical feature**] by opening up the first route from China [**country**] to the West [**region**] in the 2nd century BC, General Zhang Qian [**person**], was on a diplomatic mission rather than a trading expedition. Sent to the West [**region**] in 139 BC by the Han [**people**] Emperor Wudi [**person**] to ensure alliances against the Xiongnu [**people**], the hereditary enemies of the Chinese [**people**], Zhang Qian [**person**] was captured and imprisoned by them. Thirteen years later he escaped and made his way back to China [**country**]. Pleased with the wealth of detail

and accuracy of his reports, the emperor sent Zhang Qian [person] on another mission in 119 BC to visit several neighbouring peoples, establishing early routes from China [country] to Central Asia [region].

## Second text fragment

[the Guardian about a new series of the Cold Feet tv series](https://www.theguardian.com/media/2016/sep/06/cold-feet-gets-warm-welcome-as-6m-tune-in-for-new-series)

[\(https://www.theguardian.com/media/2016/sep/06/cold-feet-gets-warm-welcome-as-6m-tune-in-for-new-series\)](https://www.theguardian.com/media/2016/sep/06/cold-feet-gets-warm-welcome-as-6m-tune-in-for-new-series)

It was 13 years since Cold Feet last aired new episodes, but the sitcom's fans proved as loyal as ever with more than 6 million tuning in to its return on Monday night.

The return of the comedy, which stars James Nesbitt, Hermione Morris, John Thomson and Robert Bathurst, drew an average audience of 6.1 million and a 29% share of all TV viewing between 9pm and 10pm.

The show, which tells the tale of three thirtysomething couples in Manchester, originally ran for five series between 1998 and 2003.

It was one of ITV's biggest shows at the time, described by some as the UK's answer to Friends or Thirtysomething, with almost 11 million viewers tuning into its tear-stained finale.

ITV will be pleased with the popularity of the first episode of the eight-part series even if Kevin Lygo, the broadcaster's new director of television, told an audience at the Edinburgh TV festival that he was not completely sure he would have commissioned its revival.

The show easily won the 9pm slot on Monday night, albeit not against the strongest of competition with BBC1 airing Crimewatch, BBC2 Ripper Street and Channel 4 999: What's Your Emergency?

## Open Calais

It was 13 years since Cold Feet last aired new episodes, but the sitcom's fans proved as loyal as ever with more than 6 million tuning in to its return on Monday night.

The return of the comedy, which stars James Nesbitt [person], Hermione Morris [person], John Thomson [person] and Robert Bathurst [person], drew an average audience of 6.1 million and a 29% share of all TV viewing between 9pm and 10pm.

The show, which tells the tale of three thirtysomething couples in Manchester [city], originally ran for five series between 1998 and 2003.

It was one of ITV [organisation]'s biggest shows at the time, described by some as the UK [country]'s answer to Friends or Thirtysomething, with almost 11 million viewers tuning into its tear-stained finale.

ITV [organization] will be pleased with the popularity of the first episode of the eight-part series even if Kevin Lygo [person], the broadcaster's new director of television, told an audience at the Edinburgh TV festival [event] that he [person] was not completely sure he [person] would have commissioned its revival.

The show easily won the 9pm slot on Monday night, albeit not against the strongest of competition with BBC1 airing Crimewatch, BBC2 Ripper Street and Channel 4 999: What's Your Emergency?

## DBpedia Spotlight

It was 13 years since [Cold Feet](#) ([http://dbpedia.org/resource/Pilot\\_\(Cold\\_Feet\)](http://dbpedia.org/resource/Pilot_(Cold_Feet))) [product] last aired new episodes, but the [sitcom \(http://dbpedia.org/resource/Sitcom\)](http://dbpedia.org/resource/Sitcom) [not a

**named entity**]'s fans proved as loyal as ever with more than 6 million [tuning \(http://dbpedia.org/resource/Tuner\\_\(radio\)\)](http://dbpedia.org/resource/Tuner_(radio)) **[not a named entity]** in to its return on Monday night.

The return of the comedy, which stars [James Nesbitt \(http://dbpedia.org/resource/James\\_Nesbitt\)](http://dbpedia.org/resource/James_Nesbitt) **[person]**, [Hermione \(http://dbpedia.org/resource/Hermione\\_Granger\)](http://dbpedia.org/resource/Hermione_Granger) Morris **[person incorrect]**, [John Thomson \(http://dbpedia.org/resource/John\\_Thomson\\_\(comedian\)\)](http://dbpedia.org/resource/John_Thomson_(comedian)) **[person]** and [Robert Bathurst \(http://dbpedia.org/resource/Robert\\_Bathurst\)](http://dbpedia.org/resource/Robert_Bathurst) **[person]**, drew an average audience of 6.1 million and a 29% share of all TV viewing between 9pm and 10pm.

The show, which tells the tale of three [thirtysomething \(http://dbpedia.org/resource/Thirtysomething\\_\(TV\\_series\)\)](http://dbpedia.org/resource/Thirtysomething_(TV_series)) **[product incorrect]** couples in [Manchester \(http://dbpedia.org/resource/Manchester\)](http://dbpedia.org/resource/Manchester) **[city]**, originally ran for five series between 1998 and 2003.

It was one of [ITV \(http://dbpedia.org/resource/ITV\\_\(TV\\_network\)\)](http://dbpedia.org/resource/ITV_(TV_network)) **[organisation]** 's biggest shows at the time, described by some as the [UK \(http://dbpedia.org/resource/United\\_Kingdom\)](http://dbpedia.org/resource/United_Kingdom) **[country]**'s answer to Friends or [Thirtysomething \(http://dbpedia.org/resource/Thirtysomething\\_\(TV\\_series\)\)](http://dbpedia.org/resource/Thirtysomething_(TV_series)) **[product]**, with almost 11 million viewers [tuning \(http://dbpedia.org/resource/Tuner\\_\(radio\)\)](http://dbpedia.org/resource/Tuner_(radio)) **[not a named entity]** into its tear-stained finale.

[ITV \(http://dbpedia.org/resource/ITV\\_\(TV\\_network\)\)](http://dbpedia.org/resource/ITV_(TV_network)) **[organisation]** the first episode of the eight-part series even if [Kevin Lygo \(http://dbpedia.org/resource/Kevin\\_Lygo\)](http://dbpedia.org/resource/Kevin_Lygo) **[person]**, the broadcaster's new director of television, told an audience at the [Edinburgh TV festival](#)

[http://dbpedia.org/resource/Edinburgh International Television Festival](http://dbpedia.org/resource/Edinburgh_International_Television_Festival))  
[event] that he was not completely sure he would have [commissioned](http://dbpedia.org/resource/Green-light)  
(<http://dbpedia.org/resource/Green-light>) [not a named entity] its revival.

The show easily won the 9pm slot on Monday night, albeit not against the strongest of competition with [BBC1](http://dbpedia.org/resource/BBC_One)  
([http://dbpedia.org/resource/BBC One](http://dbpedia.org/resource/BBC_One)) [product] airing [Crimewatch](http://dbpedia.org/resource/Crimewatch)  
(<http://dbpedia.org/resource/Crimewatch>) [product], [BBC2](http://dbpedia.org/resource/BBC_Two)  
([http://dbpedia.org/resource/BBC Two](http://dbpedia.org/resource/BBC_Two)) [product] [Ripper Street](http://dbpedia.org/resource/Ripper_Street)  
([http://dbpedia.org/resource/Ripper Street](http://dbpedia.org/resource/Ripper_Street)) [product] and [Channel 4](http://dbpedia.org/resource/Channel_4)  
([http://dbpedia.org/resource/Channel 4](http://dbpedia.org/resource/Channel_4)) [product] 999: What's Your  
Emergency?

## Manual Annotation

It was 13 years since [Cold Feet](#) [product] last aired new episodes, but the sitcom's fans proved as loyal as ever with more than 6 million tuning in to its return on Monday night.

The return of the comedy, which stars [James Nesbitt](#) [person], [Hermione Morris](#) [person], [John Thomson](#) [person] and [Robert Bathurst](#) [person], drew an average audience of 6.1 million and a 29% share of all TV viewing between 9pm and 10pm.

The show, which tells the tale of three thirtysomething couples in [Manchester](#) [city], originally ran for five series between 1998 and 2003.

It was one of [ITV's](#) [organisation] biggest shows at the time, described by some as the UK's [country] answer to [Friends](#) [product] or [Thirtysomething](#) [product], with almost 11 million viewers tuning into its tear-stained finale.

ITV **[organisation]** will be pleased with the popularity of the first episode of the eight-part series even if Kevin Lygo **[person]**, the broadcaster's new director of television \_\_\_\_\_, told an audience at the Edinburgh TV festival **[event]** that he was not completely sure he would have commissioned its revival.

The show easily won the 9pm slot on Monday night, albeit not against the strongest of competition with BBC1 **[product]** airing Crimewatch **[product]**, BBC2 **[product]** Ripper Street **[product]** and Channel 4 **[product]** 999: What's Your Emergency **[product]**?

## Counts and discussion

In the table below the annotations are counted. Values are presented in a table cell as [ correct/wrong/missed ].

type	text1 manual	text1 Open Calais	text1 DBpedia Spotlight	text2 manual	text2 Open Calais	text2 DBpedia Spotlight
person	4	4/0/0	4/0/0	5	5/0/0	4/1/0
people	3	0/0/3	1/0/2	-	-	-
city	-	-	-	1	1/0/0	1/0/0
region	3	1/0/2	0/1/2	-	-	-
province or state	-	0/2/0	-	-	-	-
country	3	3/0/0	3/0/0	1	1/0/0	1/0/0
geographical feature	1	0/0/1	1/0/0	-	-	-



event	-	-	-	1	1/0/0	1/0/0
organization	-	-	-	2	2/0/0	2/0/0
product	-	-	-	9	0/0/9	7/1/1
not a named entity	-	0/0/0	0/2/0	-	0/0/0	0/3/0
<b>total</b>	<b>14</b>	<b>8/2/6</b>	<b>9/3/4</b>	<b>19</b>	<b>10/0/9</b>	<b>17/5/1</b>

Recognizing people, cities, countries and organizations appears to be working well with almost no errors. However the more abstract *the Chinese* or *the West* are missed by both systems.

Open Calais mistakes BC (before Christ) in a date for British Columbia and does not recognize the names of television series or broadcast channels at all.

DPpedia Spotlight appears to sometimes create a reference to a page for a word that is not a named entity, but it does not create a reference for any noun that presumably has a DPpedia entry.