

Building and Querying Semantic Layers for Web Archives

Pavlos Fafalios, Helge Holzmann, Vaibhav Kasturia and Wolfgang Nejdl
L3S Research Center, Hannover, Germany
{fafalios, holzmann, kasturia, nejdl}@l3s.de

Analysis of Evaluation Results (False Positives)

No.	Query	No. of False Positives	Article URL with Reason
2	Los Angeles Lakers players	7	<ul style="list-style-type: none">• http://query.nytimes.com/gst/fullpage.html?res=950DE5DD123FF930A3575BC0A96F948260 Cannot find mention of any L.A. Lakers Player• http://query.nytimes.com/gst/fullpage.html?res=950DEFDD1731F932A0575BC0A96F948260 Cannot find mention of any L.A. Lakers Player• http://query.nytimes.com/gst/fullpage.html?res=950DEEDC173EF930A15754C0A96F948260 Cannot find mention of any L.A. Lakers Player• http://query.nytimes.com/gst/fullpage.html?res=950DE6DB133FF933A1575BC0A96F948260 Cannot find mention of any L.A. Lakers Player• http://query.nytimes.com/gst/fullpage.html?res=950DE5D81739F935A2575BC0A96F948260 Cannot find mention of any L.A. Lakers Player• http://query.nytimes.com/gst/fullpage.html?res=950DEEDF1E30F93BA3575BC0A96F948260 Cannot find mention of any L.A. Lakers Player• http://query.nytimes.com/gst/fullpage.html?res=950DE5DB1F38F93AA35754C0A96F948260 Cannot find mention of any L.A. Lakers Player
3	African-american film producer	4	<ul style="list-style-type: none">• http://query.nytimes.com/gst/fullpage.html?res=9E0CEFDA153CF93AA3575BC0A964958260 No mention of African-american film producer• http://query.nytimes.com/gst/fullpage.html?res=9E0CE4DA103FF932A0575BC0A964958260 Incorrect mapping of Michael Jackson to some Michael J. Jackson, a Mercedes-Benz Executive• http://query.nytimes.com/gst/fullpage.html?res=9E0CE2DF123CF935A3575BC0A964958260 Can't find mention of any African-American film producer• http://query.nytimes.com/gst/fullpage.html?res=9E0CE7DB103EF934A3575BC0A964958260 Can't find mention of any African-American film producer

5	Ferrari Formula One drivers	2	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9E0CEED91F3FF932A1575BC0A964958260 Unable to find any Ferrari Formula One driver • http://query.nytimes.com/gst/fullpage.html?res=9E0CE7DD1631F935A3575AC0A964958260 Article mentions Tony Brooks who is NFC American footballer but has been incorrectly mapped to Formula One driver of the same name.
---	-----------------------------	---	--

7	American Crime Thriller films	16	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9C0CE7DC1231F933A25755C0A966958260 Unable to find any American crime thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE1DA1E38F93BA25755C0A966958260 Unable to find any American crime thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE4D71E31F934A25755C0A966958260 Unable to find any American crime thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE6D6163FF934A25755C0A966958260 Unable to find any American crime thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE4D81130F932A35755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE0DD103EF937A35755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE0DE153AF933A25755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE2DD153BF933A25755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE5DD1530F936A35755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE5DD173DF933A25755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE4DB163FF930A35755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE3DB1E31F933A25755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0DE1DF153AF936A25755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE5DB1E3CF934A25755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE4D91238F930A35755C0A966958260 Unable to find any American Crime Thriller film • http://query.nytimes.com/gst/fullpage.html?res=9C0CE2D71531F933A25755C0A966958260 Unable to find any American Crime Thriller film
---	-------------------------------	----	--

9	Australian cricketers one day internationals	10	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9504EFDC173FF93BA35754C0A962958260 Article has incorrect mapping to Kevin Wright, an Australian Cricketer but in the article referring to some other person. • http://query.nytimes.com/gst/fullpage.html?res=9A03E1D71F3CF935A35754C0A962958260 Article has incorrect mapping to Kevin Wright, an Australian Cricketer but in the article referring to Deputy Mayor of New York John Dyson by the same name. • http://query.nytimes.com/gst/fullpage.html?res=9C05E7D7143FF933A25754C0A962958260 Article has incorrect mapping to Kevin Wright, an Australian Cricketer but in the article referring to Deputy Mayor of New York John Dyson by the same name. • http://query.nytimes.com/gst/fullpage.html?res=9A0DE5DF113FF936A25754C0A962958260 Article has incorrect mapping to Matthew Elliott, an Australian Cricketer but in the article referring to some other person. • http://query.nytimes.com/gst/fullpage.html?res=9C01E7D8163EF93AA25754C0A962958260 Article has incorrect mapping to Kevin Wright, an Australian Cricketer but in the article referring to Deputy Mayor of New York John Dyson by the same name. • http://query.nytimes.com/gst/fullpage.html?res=9507E2D91530F931A2575BC0A962958260 The article mentions transactions of American Baseball. Might contain baseball player incorrectly matched to Australian Cricket Player • http://query.nytimes.com/gst/fullpage.html?res=9C0DE5DD1339F936A1575BC0A962958260 Article is about some takeover for Canadian Company • http://query.nytimes.com/gst/fullpage.html?res=9D01EFDD1339F936A1575BC0A962958260 Article is about video game console makers Nintendo • http://query.nytimes.com/gst/fullpage.html?res=9800E6DB1738F932A3575AC0A962958260 Article has incorrect mapping to Alan Connolly, an Australian Cricketer but in the article referring to some other person. • http://query.nytimes.com/gst/fullpage.html?res=9B05E0DC133BF935A2575AC0A962958260 Article is about Microprocessor Chips
10	Companies listed on NYSE	1	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=990CEFDA1F31F937A35754C0A963958260 No mention of any American Company

11	Video game consoles	14	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9402EFD8163FF93BA35754C0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9C02E6DF113FF936A25754C0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9E01E3DB133FF930A25754C0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=980CE7DD1639F937A15754C0A962958260 Article has wrong mapping of New England Electric System (NES) to Nintendo Game console NES. • http://query.nytimes.com/gst/fullpage.html?res=9B02E6DD133EF930A15754C0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9C03E7D91526E231BC4F53DFB26F958A Article has wrong mapping of New England Electric System (NES) to Nintendo Game console NES. • http://query.nytimes.com/gst/fullpage.html?res=9F0CE4D81531F934A3575BC0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9401E6D71730F932A2575BC0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9C03E0DC1E31F934A3575BC0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9C03E4DB1030F936A2575BC0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9D04E7D91131F934A3575BC0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9B0DEFD6103CF930A35754C0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=9C04E2DF1E3EF934A15754C0A962958260 Article doesn't mention any video game consoles • http://query.nytimes.com/gst/fullpage.html?res=980CEFD81639F937A3575BC0A962958260 Article has wrong mapping of Weatherford International (WII) to Nintendo Game Console Wii.
----	---------------------	----	---

12	Indian Padma Shri recipients	4	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9E0CEFDDB1E31F936A35754C0A964958260 Article mentions Ben Kingsley, who was awarded the Padma Shri but is an English actor and has Indian descent • http://query.nytimes.com/gst/fullpage.html?res=9E0CE5DF113DF932A3575AC0A964958260 Article mentions Ben Kingsley, who was awarded the Padma Shri but is an English actor and has Indian descent • http://query.nytimes.com/gst/fullpage.html?res=9E0CE6D6103EF93AA3575AC0A964958260 Article mentions Ben Kingsley, who was awarded the Padma Shri but is an English actor and has Indian descent • http://query.nytimes.com/gst/fullpage.html?res=9E0CE7DA1F3CF936A2575AC0A964958260 Article mentions Ben Kingsley, who was awarded the Padma Shri but is an English actor and has Indian descent
15	Grammy Award winners	6	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9C00EEDB1038F932A3575BC0A96E958260 No mention of any Grammy Award Winner • http://query.nytimes.com/gst/fullpage.html?res=9D00E0DA1038F932A3575BC0A96E958260 No mention of any Grammy Award Winner • http://query.nytimes.com/gst/fullpage.html?res=9C00E6D81038F932A3575BC0A96E958260 No mention of any Grammy Award Winner • http://query.nytimes.com/gst/fullpage.html?res=9506EED71038F932A3575BC0A96E958260 No mention of any Grammy Award Winner • http://query.nytimes.com/gst/fullpage.html?res=9806E5D61038F932A3575BC0A96E958260 No mention of any Grammy Award Winner • http://query.nytimes.com/gst/fullpage.html?res=9500EFDA1038F932A3575BC0A96E958260 No mention of any Grammy Award Winner
16	Indian meat dishes	2	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=950DE1D61739F935A35752C0A96F948260 Article contains no Indian dish, Incorrect mapping of butter, chicken.... with butter chicken • http://query.nytimes.com/gst/fullpage.html?res=950DEFDF1F3CF93BA35755C0A96F948260 Article contains no Indian dish, Incorrect mapping of butter, chicken.... with butter chicken
17	Mammals of India	2	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9A02E4DA143AF930A1575AC0A962958260 Article has no mention of any Indian Mammalian Animal • http://query.nytimes.com/gst/fullpage.html?res=9A03E2D8113FF93BA2575AC0A962958260 Article has no mention of any Indian Mammalian Animal

18	US Fast food chains	2	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=950DE1D71F3EF930A35754C0A96F948260 No mention of any US fast food chain • http://query.nytimes.com/gst/fullpage.html?res=950DE3D9163EF934A35754C0A96F948260 No mention of any US fast food chain
19	NASA civilian astronauts	1	<ul style="list-style-type: none"> • http://query.nytimes.com/gst/fullpage.html?res=9502E3DC1F3AF932A05754C0A961958260 Incorrect mapping of Gregory Johnson, a NASA civilian astronaut to Greg Johnson, president of Franklin Templeton Investments