

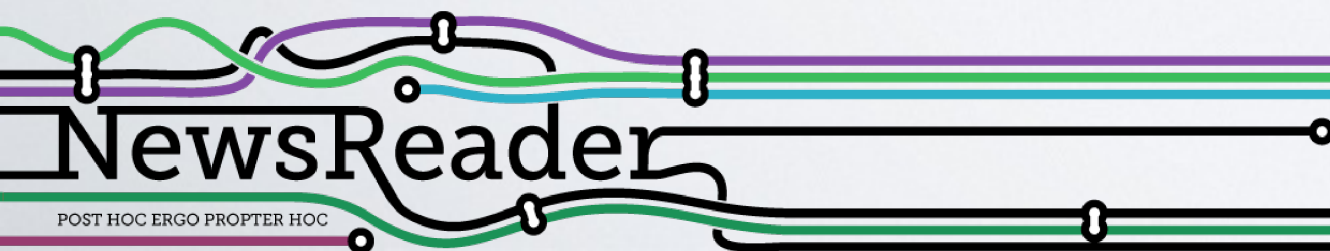


NewsReader Quick Start

Ian Hopkinson

Cars Hack Day

Amsterdam, 21st January. London 30th January 2015



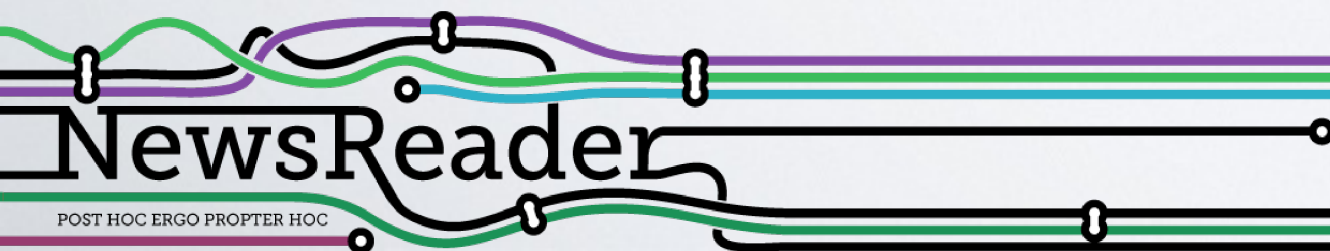
tl;dr

- API + docs:

<https://newsreader.scrapewiki.com/>

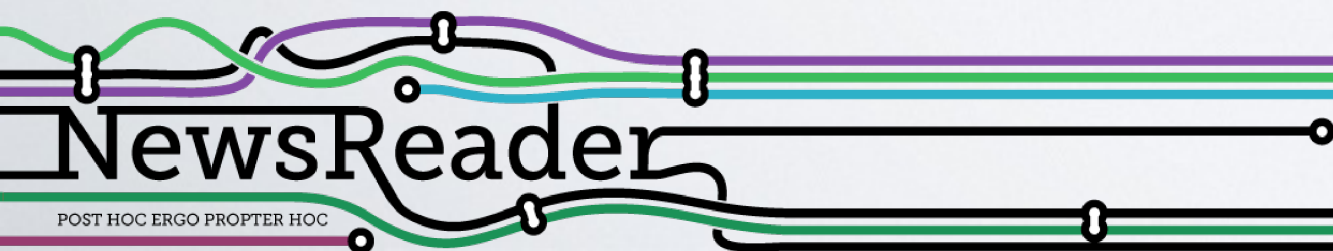
- URL Bundle (examples, BitBucket repo,...):

- <http://tab.bz/ydtco>



Overview

- A Simple NewsReader query API
 - Bitbucket repo – we want your feedback!
 - Warnings
- Demos
- Data on a disk
- Ideas for hacking



Simple API

- Easy access to the NewsReader data:

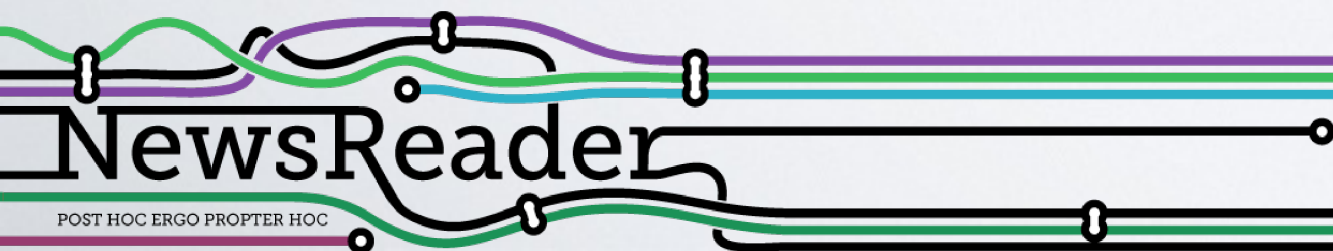
<https://newsreader.scrapewiki.com/>

- Visit the URL to see the documentation. In short:

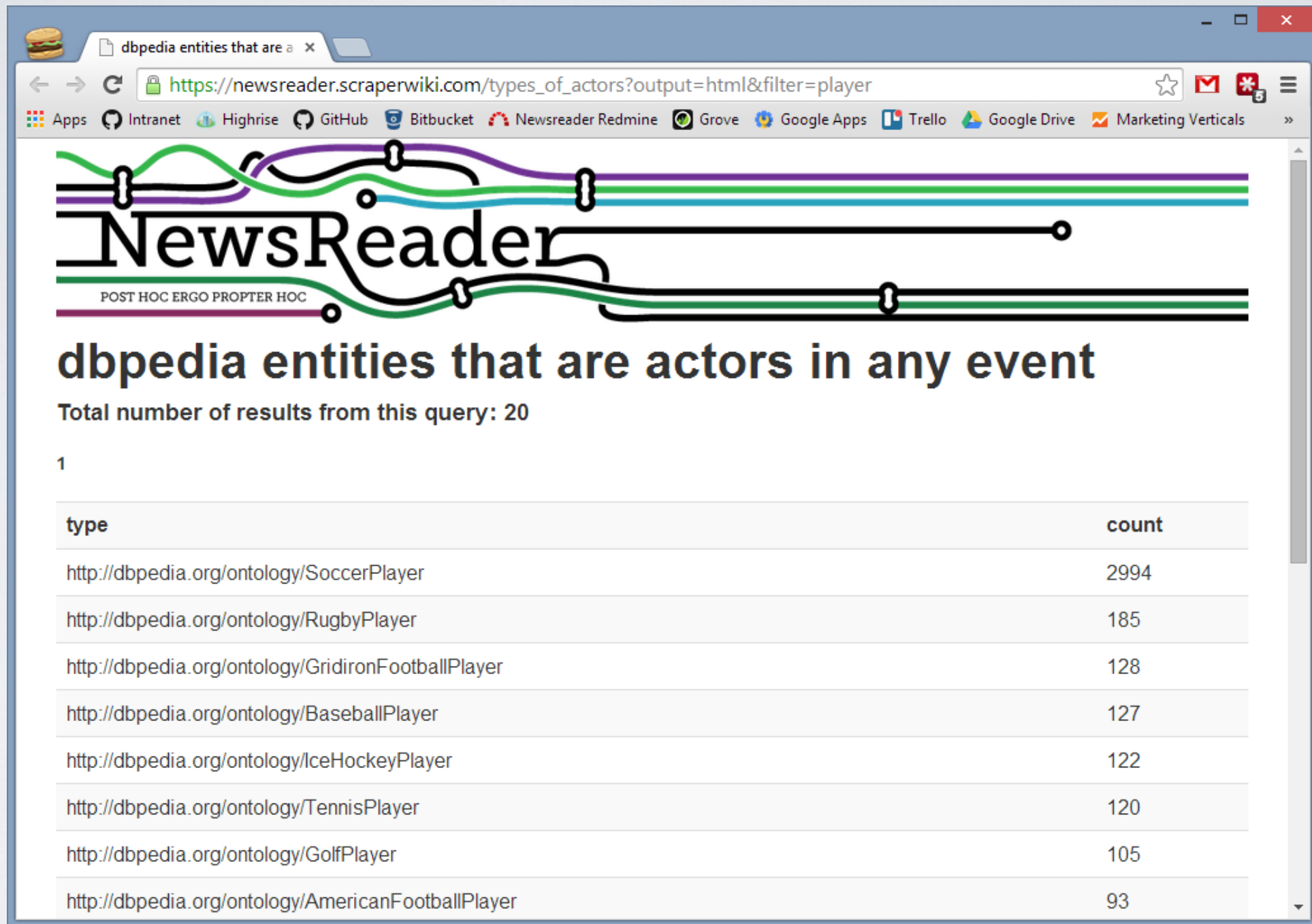
`{root}/{page/[n]/}query_name?param1=[string]¶m2=[string]`

Parameters are strings, date filters (i.e. 1999-08), or URIs (i.e.

http://dbpedia.org/resource/David_Beckham, dbpedia)



Simple API



dbpedia entities that are actors in any event

Total number of results from this query: 20

1

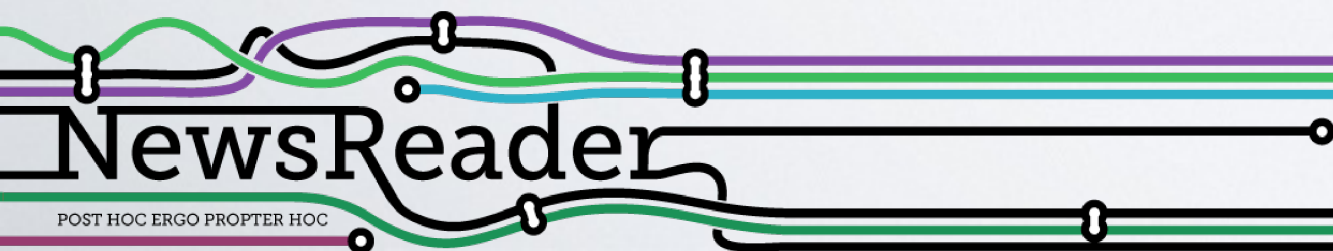
type	count
http://dbpedia.org/ontology/SoccerPlayer	2994
http://dbpedia.org/ontology/RugbyPlayer	185
http://dbpedia.org/ontology/GridironFootballPlayer	128
http://dbpedia.org/ontology/BaseballPlayer	127
http://dbpedia.org/ontology/IceHockeyPlayer	122
http://dbpedia.org/ontology/TennisPlayer	120
http://dbpedia.org/ontology/GolfPlayer	105
http://dbpedia.org/ontology/AmericanFootballPlayer	93

We Need Your Feedback!

- Bitbucket repo:

https://bitbucket.org/scrapewikids/newsreader_api_flask_app

- Use the issue tracker
- All and any feedback welcome!



Warnings

- Result counts inaccurate for some queries
 - Can't page beyond 500, use filter or datefilter to reduce
- Triplestore is large, queries may be slow
 - Let us know about timeouts
- “Fragments” don't get passed
 - Make {some_url}#thingEvent
 - Into {some_url}%23thingEvent

Accessing LexisNexis articles

- <http://www.newsreader-project.eu/LNdata/...>
- Get username, then set password:
 - <https://www.lexisnexis.com/trialid>
 - Usernames on piece of paper
- Links should redirect...
 - If they fail then change `csi=138620` to `csi=8399` in the URL
 - Search using `Ini({document_id})`

Data on a disk

- In case of boredom / emergency
 - Pre-computed data for event labels, types of actor and FrameNet labels, ESO labels
 - Raw RDF input (derived data)

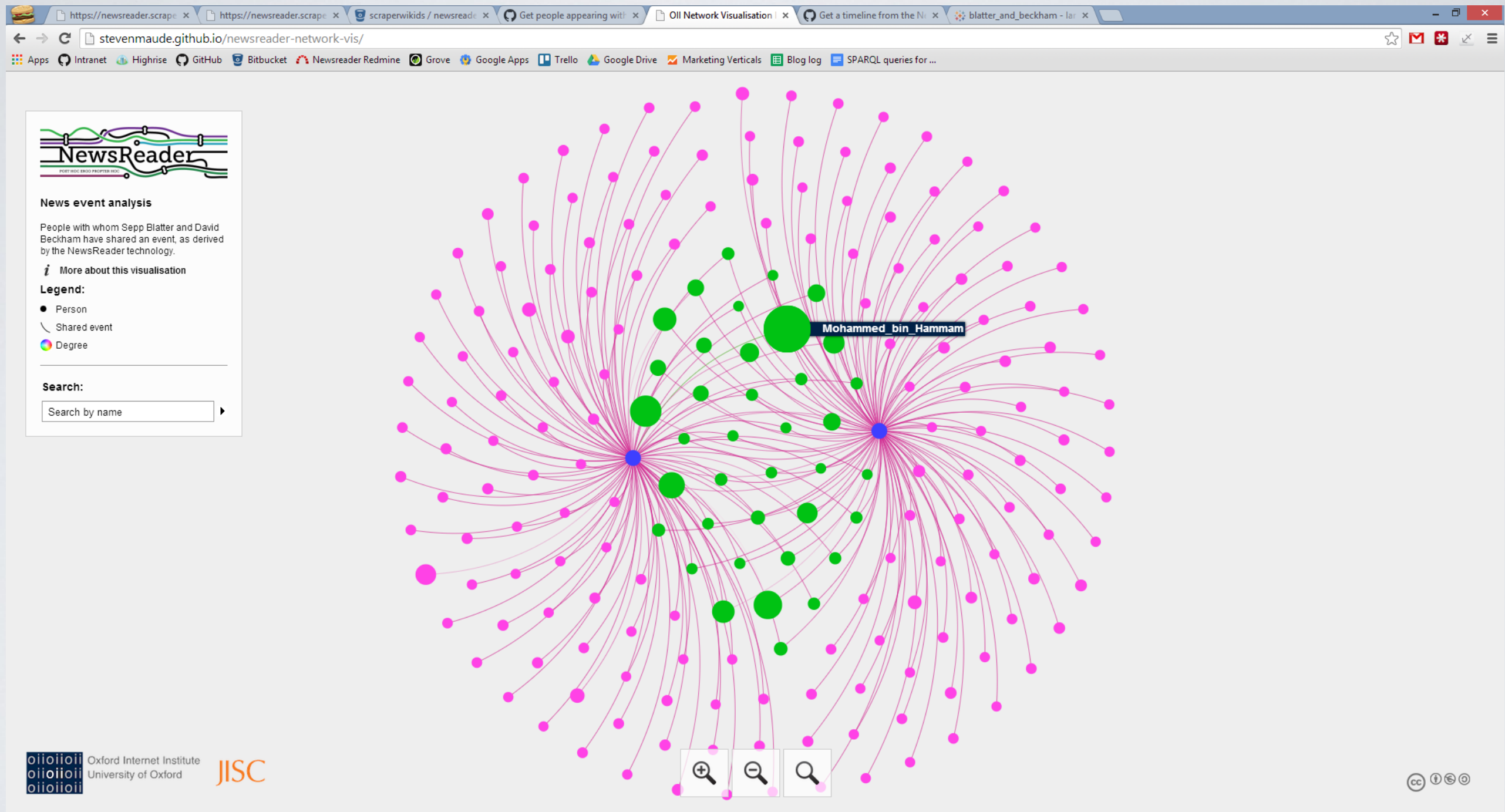
Patterns

- Find events
 - Labels, types of actor, FrameNet, ESO queries
 - Probe events in detail using precis, situation graph and describe queries
- FrameNet, ESO and situation graphs are where the exciting new stuff is at
 - Semantic role labelling and reasoning

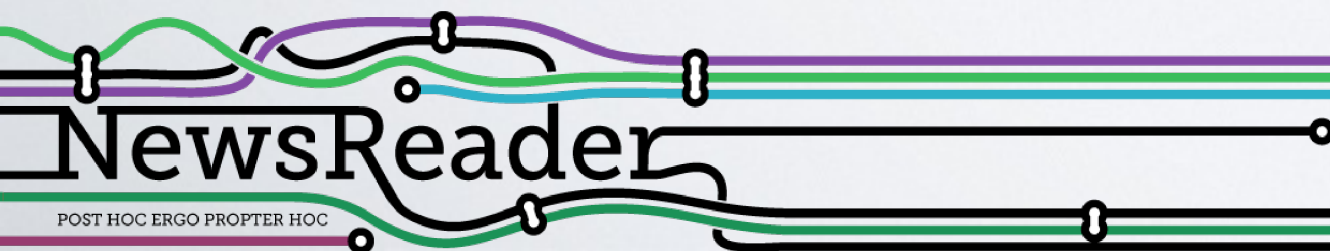
Suggestions

- Exploring stories
- Visualising events and timelines
- Making an interface to the Simple API
- Improving the Simple API
- Roll your own SPARQL query

Example 1



Example 2



Thank You!
#newsreader

<https://newsreader.scrapewiki.com/>

<http://tab.bz/ydtco>

