

대한민국 인맥지도

2018-10-07

-[]- Box and Whisker

강규영

alankang@boxnwhis.kr

발표자 및 발표내용

발표자: 데이터 분석회사 박스앤위스커에서 개발자로 일하는 중

발표 주제: “대한민국 인맥지도” 홈페이지 제작기. 구체적으로?

- 위키데이터 프로젝트 및 SPARQL 쿼리 API 소개
- 파이썬 NetworkX 라이브러리를 이용한 네트워크 분석
- 자바스크립트(d3js)를 이용한 네트워크 시각화
- 파이썬 + 자바스크립트 + 깃헙 페이지스 활용하여 무료로 그럴듯한 홈페이지 만들기

전 대통령 이명박은 전 대통령 전두환의 아들 전재만의 동서 효성그룹 사장 조현준의
사촌 전 한국타이어 사장 조현범의 장인이다.

응??



분리된 혼맥

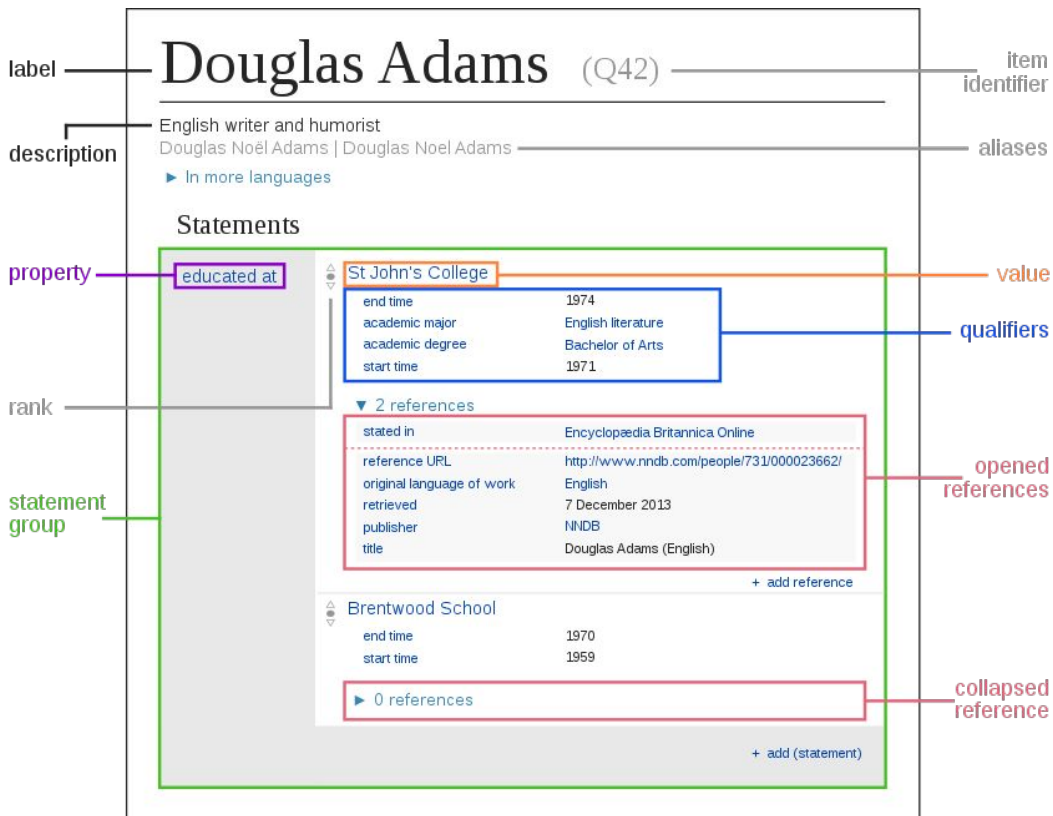
서로 연결되지 못하고 단절된 혼맥 덩어리들의 상위 10개 목록입니다. 데이터를 추가하여 단절된 혼맥들을 서로 연결시켜주세요.

1. 구인회 (LG그룹의 창업주) 외 482명
2. 조선 인조 (조선의 제16대 임금) 외 276명
3. 김정일 (조선민주주의인민공화국의 국가 원수) 외 26명

궁금할만한 지점들

- 데이터는 어디서 났을까?
- 각 혼맥 덩어리의 주요 인물들 어떻게 찾았을까?
- 그래프 탐색기는 어떻게 만들었을까?
- 웹사이트는 어떻게 만들었을까?

데이터 획득: 위키데이터 + SPARQL



데이터 획득: 위키데이터 + SPARQL

```
SELECT ?child ?childLabel
WHERE
{
  # ?child father Bach
  ?child wdt:P22 wd:Q1339.
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }
```


주요 인물 찾기: 파이썬 + NetworkX

```
g = nx.Graph()
g.add_edges_from(
    (source, target)
    for rel, source, target in links
    if rel in KINSHIP
)
```

```
sorted(
    nx.betweenness_centrality(g).items(),
    key=lambda x: x[1],
    reverse=True
)
```

그래프 탐색기: 자바스크립트 + d3js



웹사이트 제작: 깃헙 페이지스 + 지킬



Free hosting with GitHub Pages

Sick of dealing with hosting companies? **GitHub Pages** are *powered by Jekyll*, so you can easily deploy your site using GitHub for free—**custom domain name** and all.

[Learn more about GitHub Pages →](#)

마무리

- 위키데이터 프로젝트? 매우 유망/유용해요. 함께 기여해 보아요 :)
- 파이썬? 사랑입니다.
- 자바스크립트? 어딘지 모르게 이상하지만 유용해요.
- d3js? 초기 진입 장벽이 높지만 일단 한 번 잡썰보세요.
- 깃헙? 우리 마이크로소프트가 달라졌어요.

참고자료

- 대한민국 인맥지도 <https://akngs.github.io/smallworld/>
- 위키데이터 프로젝트 <https://wikidata.org>
- 파이썬 NetworkX 라이브러리 <https://networkx.github.io>
- 자바스크립트 d3js 라이브러리 <https://d3js.org>
- 깃헙 페이지스 <https://pages.github.com/>
- 지킬 사이트 생성기 <https://jekyllrb.com/>