

Social Network

Dan Fanelli

November 22, 2016

Social Network Investigation



Figure 1:

Investigate a social network that is of interest to you: *GENI*

Carefully define what the vertices represent and what the edges represent.

EDGES: *In this model, the edges represent a parent-child relationships*

NODES: *In this model, each node represents a person*

Are there any new modeling techniques that you had to employ?

No new modelling techniques, just the scraping, the graph, and the Neo4j storage of the graph.

Start Node: Kevin Bacon:

```
## [ 1 ] Kevin Bacon
## [ 2 ] Edmund Norwood Bacon
## [ 3 ] Ellis Williams Bacon
## [ 4 ] Thomas Pryor BACON
## [ 4 ] Anne Elizabeth BACON
## [ 3 ] Helen Atkinson Bacon
## [ 4 ] Robert Comly
## [ 4 ] Lydia Townsend Comly
## [ 2 ] Ruth Hilda Holmes
## [ 3 ] Artemas Holmes
## [ 4 ] Artemus Henry Holmes
## [ 4 ] Lillian Holmes
## [ 3 ] Dorothy Frances Smith
## [ 4 ] George Campbell Smith
## [ 4 ] Annie Schwartz
## [1] "Finished Kevin Bacon"
```



Kevin Norwood Bacon MP

Gender: Male

Current Location: New York, NY, USA

Birth: July 8, 1958
North Fayette, PA, USA

Immediate Family: Son of [Edmund Norwood Bacon](#) and [Ruth Hilda Holmes](#)
Husband of [Kyra Sedgwick](#)
Father of [<private> Bacon](#) and [<private> Bacon](#)
Brother of [<private> Bacon](#); [<private> \(Bacon\)](#); [<private> \(Bacon\)](#); [<private> \(Bacon\)](#); [<private> \(Bacon\)](#) and 2 others

Added by: [M](#) on June 24, 2009

Managed by: [Michael Reid Delahunt, PhD](#) and 4 others

Curated by: [Jenna](#), Volunteer Curator

[Overview](#)
[Media](#)
[Timeline](#)
[Discussions \(1\)](#)
[Sources](#)
[Revisions](#)
[DNA](#)

Figure 2:

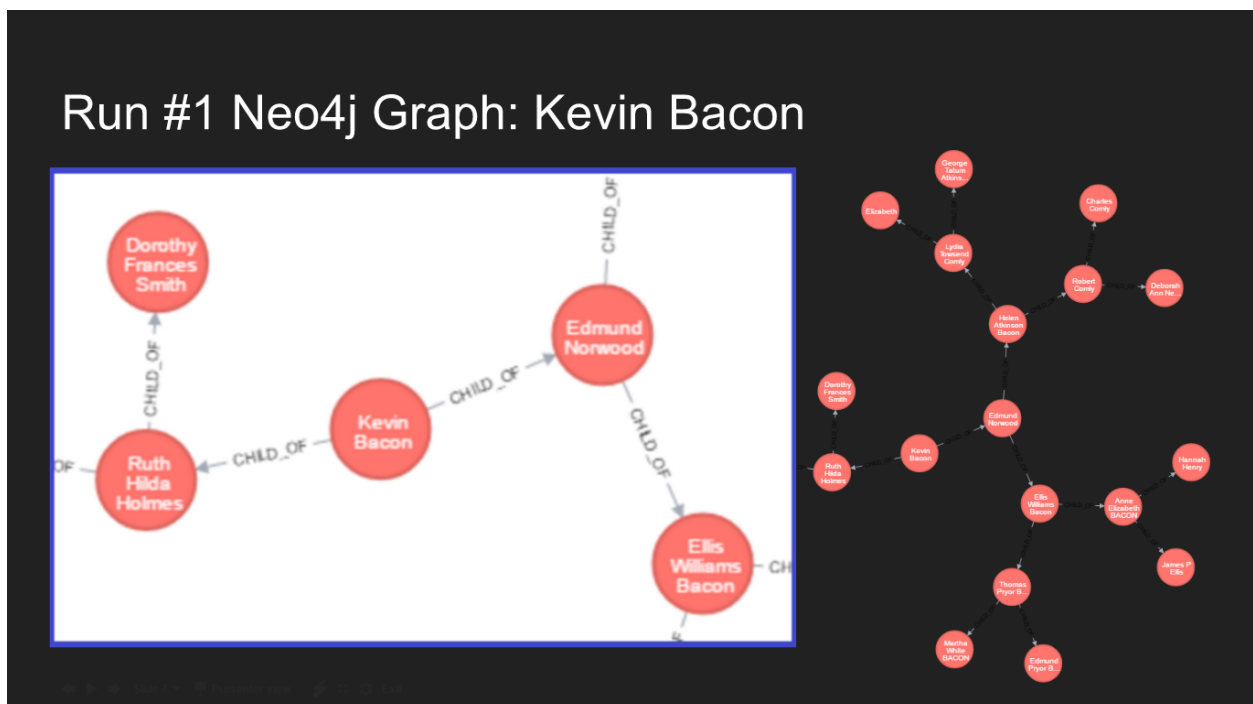


Figure 3:

Start Node: JFK



John Fitzgerald Kennedy, Sr. MP 100

Gender: Male

Birth: May 29, 1917
Brookline, Norfolk, Massachusetts, United States

Death: November 22, 1963 (46)
Dallas, Texas, United States (murdered)

Place of Burial: Arlington National Cemetery, Arlington, VA, USA

Immediate Family: Son of [Joeseeph P. Kennedy, Sr.](#) and [Rose Elizabeth Kennedy](#)
Husband of [Jacqueline Kennedy Onassis](#)
Father of [Arabella Kennedy](#); [Caroline Bouvier Schlossberg](#); [John F. Kennedy, Jr.](#) and [Patrick Bouvier Kennedy](#)
Brother of [Joseph P. Kennedy, Jr.](#); [Rosemary Kennedy](#); [Kathleen, Marchioness of Hartington](#); [Eunice Mary Shriver](#); [Patricia Helen Lawford](#) and 3 others

Added by: [Rick](#) on June 2, 2007

Managed by: [Tina](#) and 51 others

Curated by: [Daniel Dupree Walton](#)

[Overview](#) [Media](#) [Timeline](#) [Discussions \(3\)](#) [Sources \(2\)](#) [Revisions](#) [DNA](#)

Figure 4:

```
## [ 1 ] John F. Kennedy, 35th President of the USA
## [ 2 ] Joeseeph P. Kennedy, Sr.
## [ 3 ] Patrick Joseph Kennedy, Jr.
## [ 4 ] Patrick Joseph Kennedy, Sr.
## [ 4 ] Bridget Kennedy
## [ 3 ] Mary Augusta Kennedy
## [ 4 ] James Hickey
## [ 4 ] Margaret Hickey
## [ 2 ] Rose Elizabeth Kennedy
## [ 3 ] John F. "Honey Fitz" Fitzgerald, Sr.
## [ 4 ] Thomas Fitzgerald
## [ 4 ] Rosanna Fitzgerald
## [ 3 ] Mary Josephine FitzGerald
## [ 4 ] Michael Hannon
## [ 4 ] Mary Ann Fitzgerald
## [1] "Finished JFK"
```

Run #2 Neo4j Graph: JFK

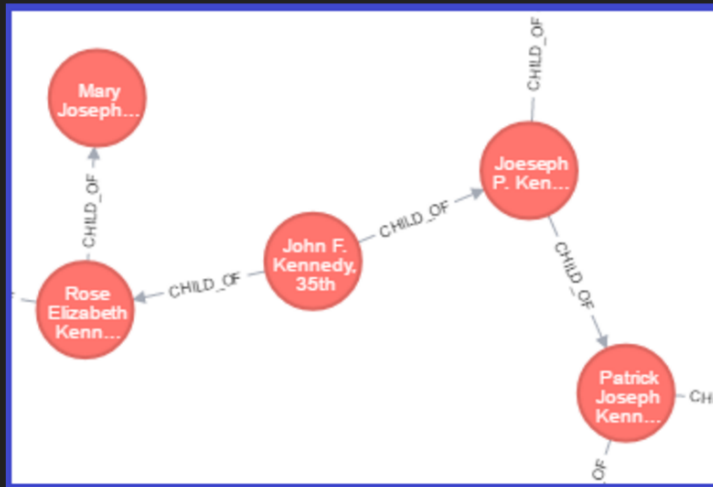


Figure 5:

All Code:

```
library(rvest)
library(knitr)
library(RNeo4j)

NEO4J_GRAPH_URL <- 'http://localhost:7474/db/data/'
graph = startGraph(NEO4J_GRAPH_URL, username="neo4j", password="neo5j")
clear(graph, FALSE)

addConstraint(graph, "Person", "name")
addConstraint(graph, "Person", "link")

NUM_GENERATIONS <- 3

get_geni <- function(source_person, this_name, this_link, gen_num){
  for(spaces in c(1:gen_num)){
    cat("\t")
  }
  cat("[",gen_num,"]",this_name,"\n")

  page_node <- html(this_link)
  fam_links <- page_node %>% html_nodes("#family_handprint a")
  for(i in c(1:2)){
    f <- fam_links[i]
    next_name <- f %>% html_text()
```

```

next_link <- f %>% html_attr("href")
if(!endsWith(next_link, "#")){
  curr_time <- format(Sys.time(), "%H:%M:%s")
  found_person = getOrCreateNode(graph, "Person", name=next_name, link=next_link, time=Sys.time(), g
  relative_rel = createRel(source_person, "CHILD_OF", found_person)
  if(gen_num <= NUM_GENERATIONS){
    gen_above <- gen_num + 1
    get_geni(found_person, next_name, next_link, gen_above)
  }
}
}
return (df)
}
start_name_1 <- "Kevin Bacon"
start_page_1 <- "https://www.geni.com/people/Kevin-Bacon/6000000009325127022"
the_person_1 = createNode(graph, "Person", name=start_name_1, link=start_page_1, time=Sys.time(), gen_al
geni_run_1 <- get_geni(the_person_1, start_name_1, start_page_1, 1)

sprintf('Finished Kevin Bacon')
start_name_2 <- "John F. Kennedy, 35th President of the USA"
start_page_2 <- "https://www.geni.com/people/John-F-Kennedy-35th-President-of-the-USA/60000000028809126
the_person_2 = createNode(graph, "Person", name=start_name_2, link=start_page_2, time=Sys.time(), gen_al
geni_run_2 <- get_geni(the_person_2, start_name_2, start_page_2, 1)
sprintf('Finished JFK')
##

```