

Interactive, web-based visualization of scientific collaboration networks

Ben Leone, Ryuichi Ohhata, Lukas Rosario

Jinyu Tian, Angela Vellante

Mentors:

- Eric Kolaczyk (Hariri Institute, Director)
- Arezoo Sadeghi (SAIL)
- Margot Menestrot from Red Hat UX team
- Joe Farmer (BU Office of Research, Program Manager & Data Analyst)
- Jonathan Chamberlin



Demo 4 Feedback

- Where is the D3 code for the visualization?
- For the docker file, it looks like you are just running things on your local machines, have you started working on openshift environment at all?
- Integration of visualization in WordPress

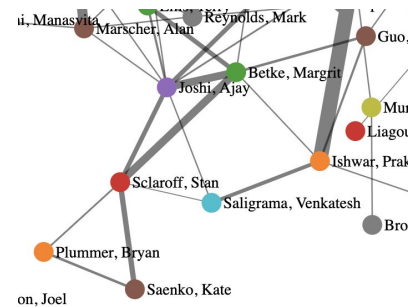
Visualization features

- Each **node** represents an individual person of the Hariri Institute.
- Hovering** your mouse over the **node** will increase the size
- A **link** between two nodes represents a direct collaboration; the **thicker** the link, the greater the number of collaborations between them.

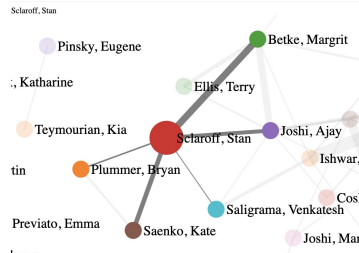
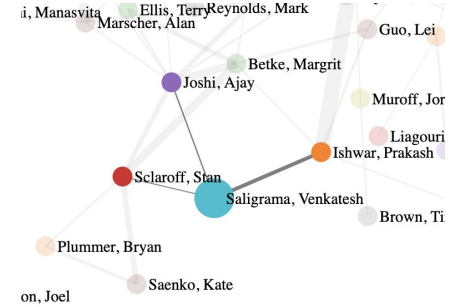


- Clicking** on a **node** will display the person's name and isolate direct links.

Before click:



After click:



Setbacks

- Creating a persistent storage database in case it goes down
- Connecting the PosgreSQL database to the hosted API on OpenShift

APPLICATION

postgresql-ephemeral <http://postgresql-ece-528-interactive-web-visualization.k-apps.osh.massopen.cloud>

> | DEPLOYMENT CONFIG **postgresql, #1** 1 pod ⋮

APPLICATION

postgresql-persistent <http://testsql-ece-528-interactive-web-visualization.k-apps.osh.massopen.cloud>

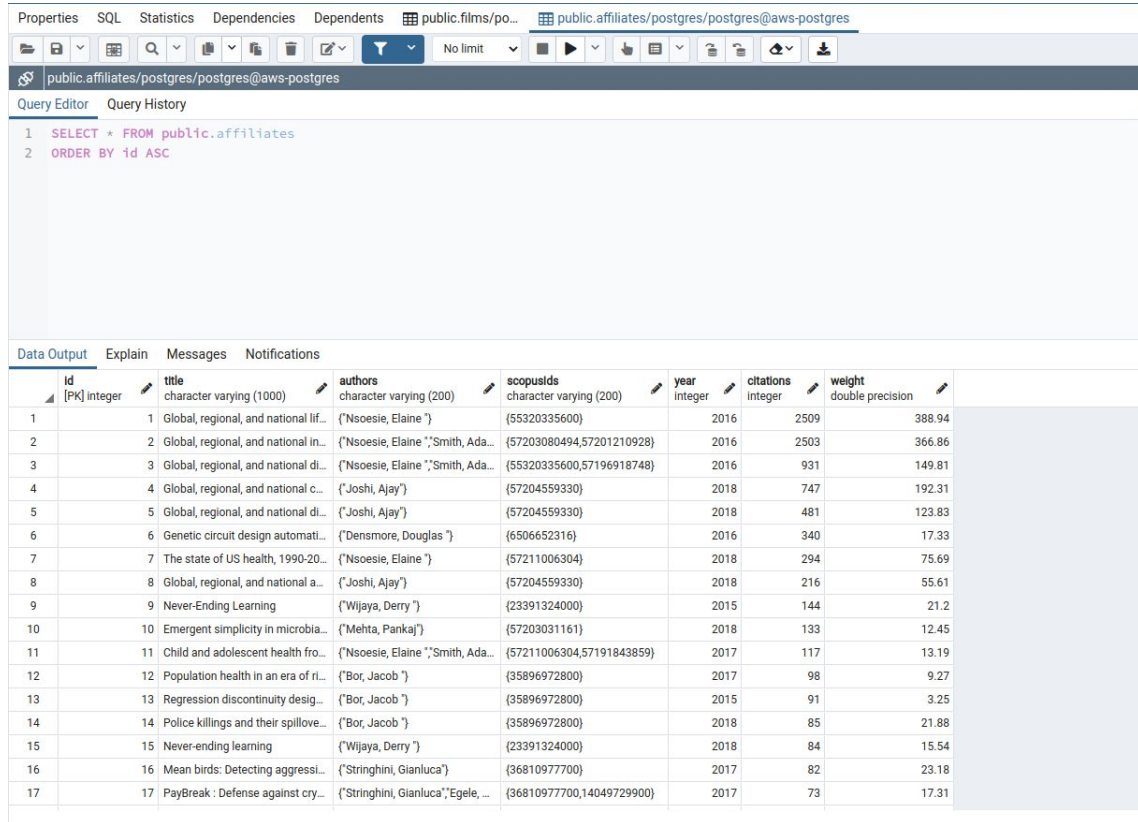
> | DEPLOYMENT CONFIG **testsql** ✖ 1 Error No deployments for testsql ⋮

Setbacks

- Connecting to database using GUI (i.e. pgAdmin) to upload the data to be used by the API
- [Port forwarding](#) did not expose the database to the service

Current Obstacles

- Using AWS RDS free tier as interim solution
- Slow for data uploads (db.t2.micro)
 - 1 vCPU
 - 1 GiB RAM
 - 20 GiB Disk



The screenshot shows a database management interface with a query editor and a results table. The query editor contains the following SQL query:

```
1 SELECT * FROM public.affiliates
2 ORDER BY id ASC
```

The results table displays 17 rows of data with the following columns:

	id [PK] integer	title character varying (1000)	authors character varying (200)	scopusids character varying (200)	year integer	citations integer	weight double precision
1	1	Global, regional, and national lif...	{'Nsoesie, Elaine'}	{55320335600}	2016	2509	388.94
2	2	Global, regional, and national in...	{'Nsoesie, Elaine','Smith, Ada...	{57203080494,57201210928}	2016	2503	366.86
3	3	Global, regional, and national di...	{'Nsoesie, Elaine','Smith, Ada...	{55320335600,57196918748}	2016	931	149.81
4	4	Global, regional, and national c...	{'Joshi, Ajay'}	{57204559330}	2018	747	192.31
5	5	Global, regional, and national di...	{'Joshi, Ajay'}	{57204559330}	2018	481	123.83
6	6	Genetic circuit design automati...	{'Densmore, Douglas'}	{6506652316}	2016	340	17.33
7	7	The state of US health, 1990-20...	{'Nsoesie, Elaine'}	{57211006304}	2018	294	75.69
8	8	Global, regional, and national a...	{'Joshi, Ajay'}	{57204559330}	2018	216	55.61
9	9	Never-Ending Learning	{'Wijaya, Derry'}	{23391324000}	2015	144	21.2
10	10	Emergent simplicity in microbia...	{'Mehta, Pankaj'}	{57203031161}	2018	133	12.45
11	11	Child and adolescent health fro...	{'Nsoesie, Elaine','Smith, Ada...	{57211006304,57191843859}	2017	117	13.19
12	12	Population health in an era of ri...	{'Bor, Jacob'}	{35896972800}	2017	98	9.27
13	13	Regression discontinuity desig...	{'Bor, Jacob'}	{35896972800}	2015	91	3.25
14	14	Police killings and their spillove...	{'Bor, Jacob'}	{35896972800}	2018	85	21.88
15	15	Never-ending learning	{'Wijaya, Derry'}	{23391324000}	2018	84	15.54
16	16	Mean birds: Detecting aggressi...	{'Stringhini, Gianluca'}	{36810977700}	2017	82	23.18
17	17	PayBreak : Defense against cry...	{'Stringhini, Gianluca','Egele, ...	{36810977700,14049729900}	2017	73	17.31

API

RestAPI hosted on OpenShift cloud using docker container

Python

Information

Configuration

Results

1

2

3

* Add to Project

Interactive, web-based visualization of scientific collaboration networks

Version

3.6 — latest

* Application Name

api-flask

* Git Repository

https://github.com/BU-CLOUD-F20/Interactive_Web_Visualization

[Try Sample Repository ↗](#)

If you have a private Git repository or need to change application defaults, view [advanced options](#).

Cancel

< Back

Create

API

API retrieves data from PostgreSQL by applying the credentials, and:

- Displays affiliate information
- Lists number of collaboration between each affiliates

URL:

<http://api-flask-ece-528-interactive-web-visualization.k-apps.osh.massopen.cloud/api/v1/relations>

API

```
▶ nodes:      [...]
▼ links:
  ▼ 0:
    source:    3
    target:    6
    value:     2
```

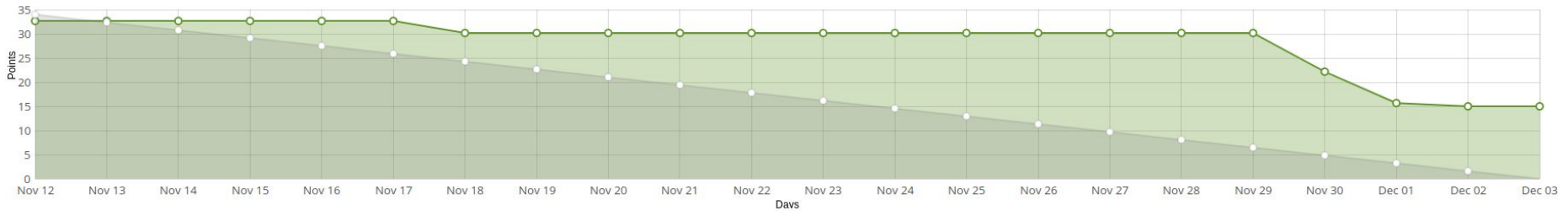
Source: Affiliate 1
Target: Affiliate 2
Value: How many collaborated papers

```
▼ nodes:
  ▶ 0:      {}
  ▶ 1:      {}
  ▶ 2:      {}
  ▼ 3:
    id:      "Smith, Adam"
    department: "Computer Science"
    college:  "College of Arts and Sciences"
    email:    "ads22@bu.edu"
    ▶ interests: "Computability and Comput...preserving Computation"
    ▶ domains:  "Computational Social and Society, Ethics, and Law"
  ▶ 4:      {}
  ▶ 5:      {}
```

```
▼ 6:
  id:      "Becker, Alexander"
  department: "Administrative Sciences"
  college:  "Metropolitan College"
  email:    "apbecker@bu.edu"
  ▶ interests: "Computational Logic and ...raphy and Cryptosystems"
  ▶ domains:  "Computational Finance an... and Political Sciences"
  ▶ 7:      {}
  ▶ 8:      {}
```

Demo

Taiga Burndown Chart



Flat holiday burndown but progress during our final week of development

Wrapping up

Dec 8:

- Resolve remote database issues with OpenShift and connect to API
- Refine visualization for best UI/UX
 - Search/filter
 - Displaying emails and other info.
 - Department key color
 - Legend
- Finalize documentation on how to upload data to our database after the end of the project

Thank you!
Questions?