GitHub

GitHub Network Intelligence

Nittala Sai Venkata Aditya
Soumith Reddy Palreddy
Praneet Kumar Alamuri
Soumya Agrawal
Anudeep Kumar Akkana

OBJECTIVE

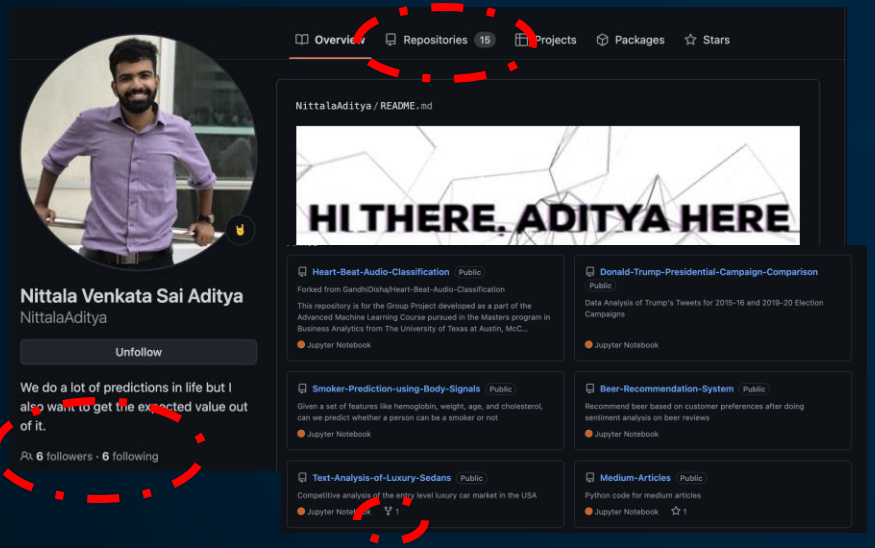
To analyse social network graphs on GitHub

Applications

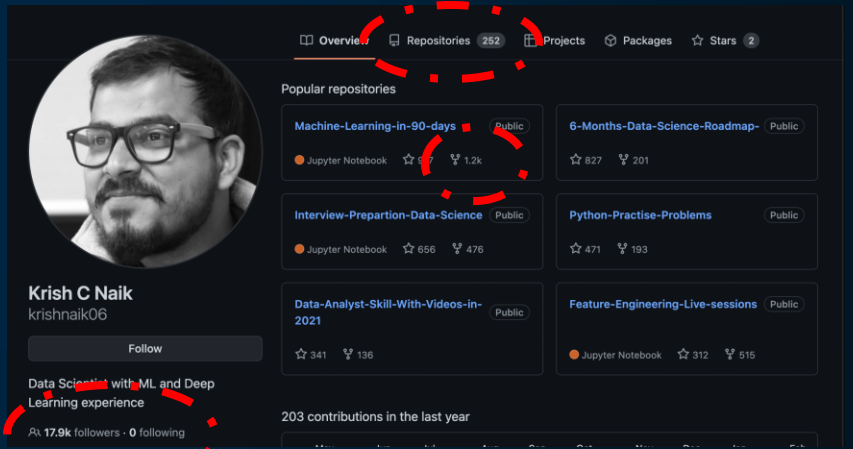
- *Helps companies in identifying the right talent*
- *Better communities on GitHub*
- *Recommendations for Developers*

REPOSITORY AND USER EXAMPLE

Upcoming Popular User



Popular User



METHODOLOGY

Network Analysis on Soumith Palreddy's GitHub Profile

Levels of GitHub Followers

LEVEL 1

Soumith Palreddy's github with 6 followers

LEVEL 2

Soumith Pareddy's followers' followers

LEVEL 3

Soumith Palreddy's followers' followers' followers

LEVEL 1 GRAPH

Nodes

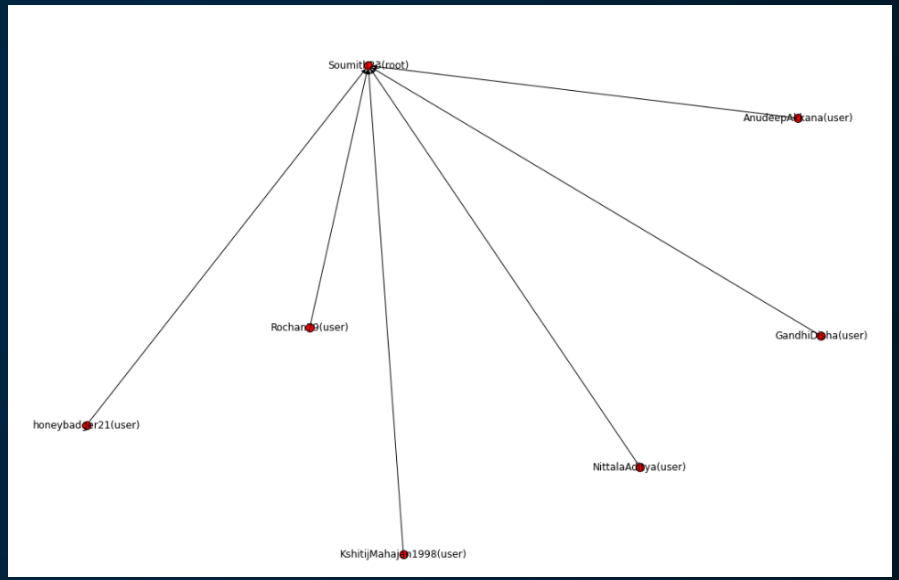
Soumith + 6 Followers

Edges

Denotes the following relation between 2 users

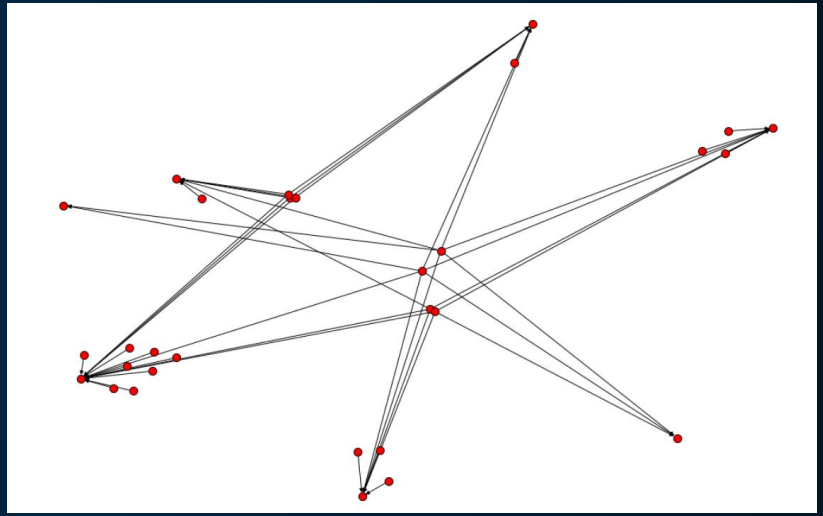
Examples

Aditya follows Soumith, so both Aditya and Soumith will have nodes and there will be an edge between the two showing that relation

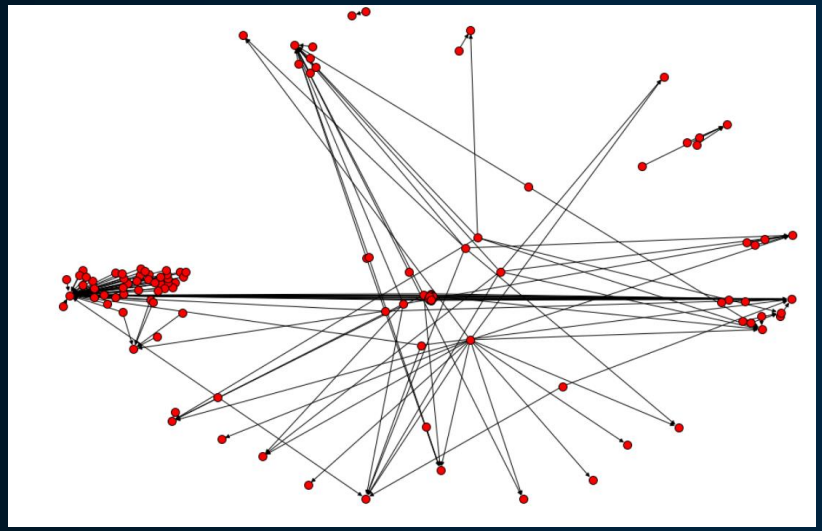


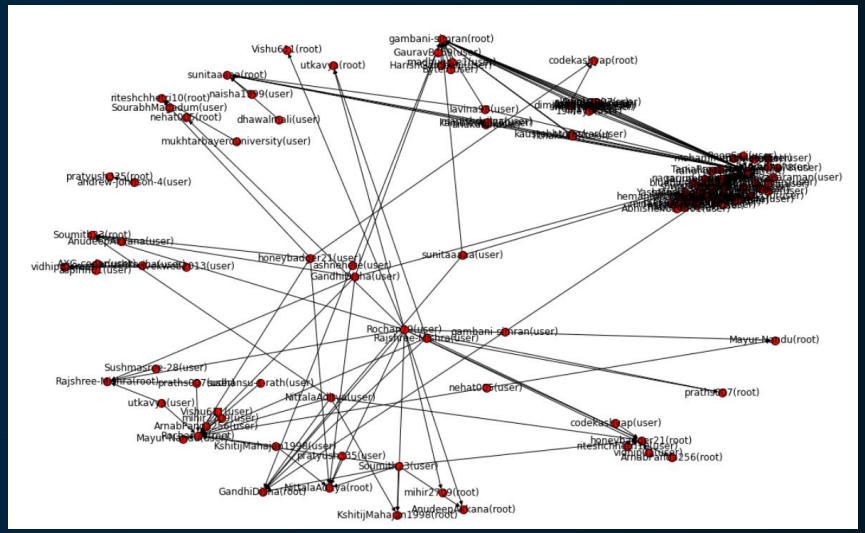
LEVEL 2 GRAPH

Nodes	Soumith + 6 Followers + Followers of 6 Followers
Edges	Denotes following relation between 2 users
Examples	Aditya follows Soumith, so both Aditya and Soumith will have nodes and there will be an edge between the two showing that relation

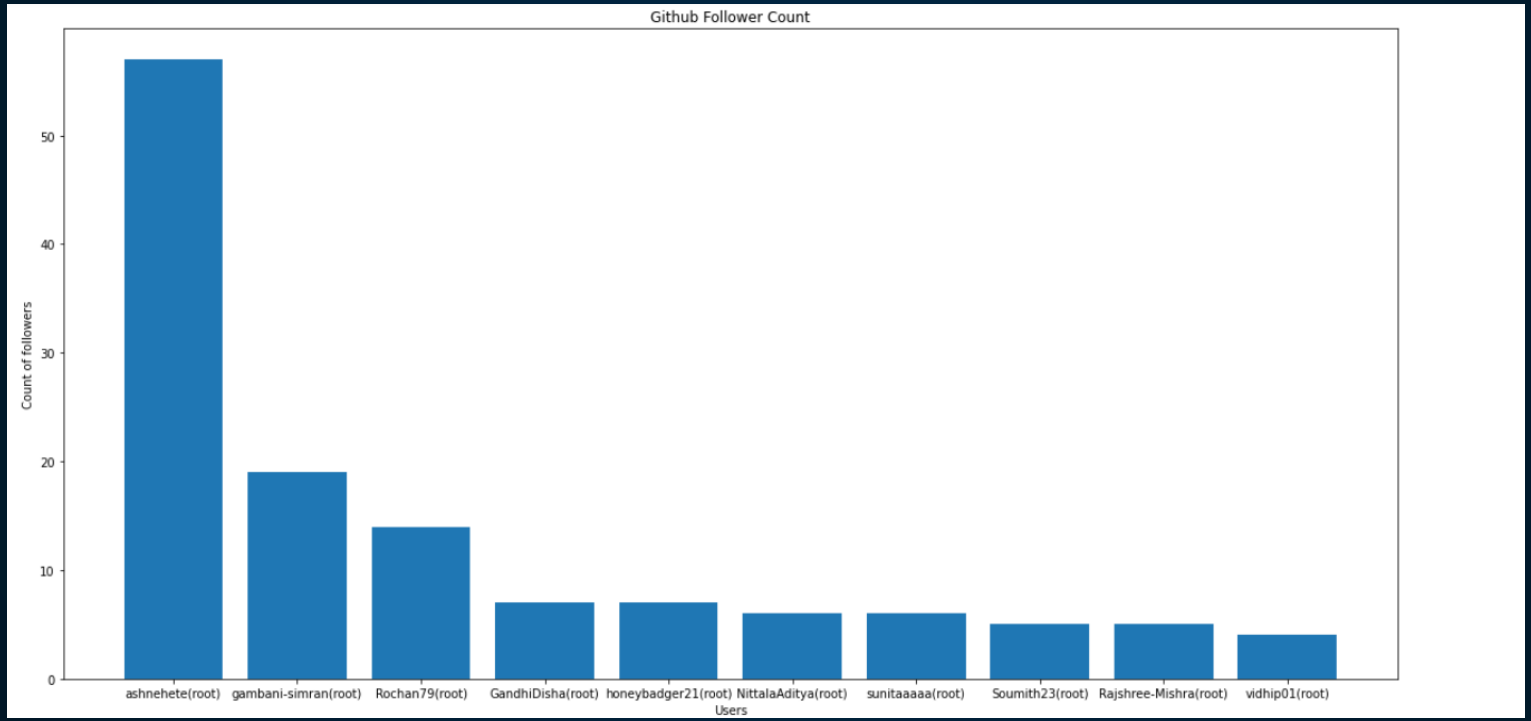


LEVEL 3 GRAPH





WHO'S FAMOUS AMONG SOUMITH'S FOLLOWERS



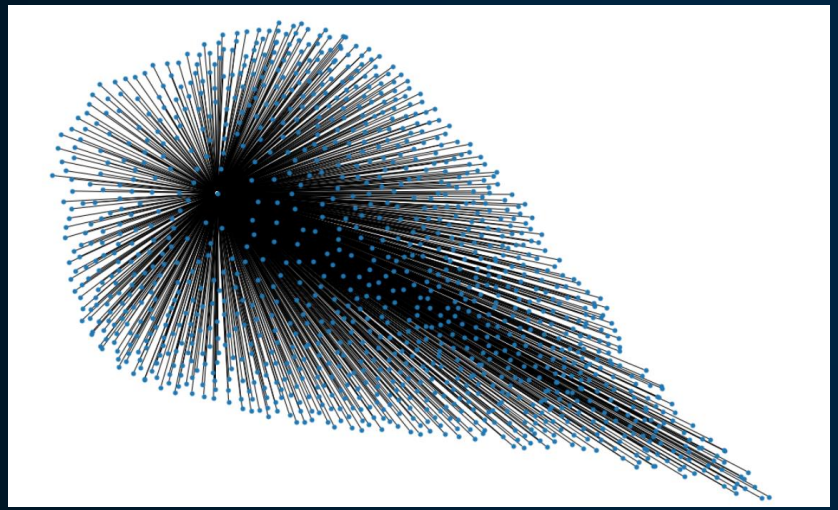
USER NETWORK ANALYSIS

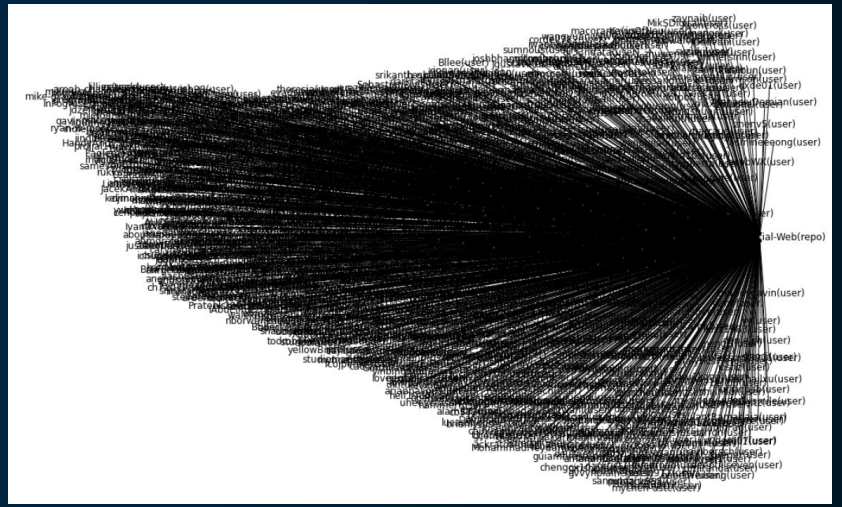
Levels	Nodes	Edges	In Degree	Out Degree
Level 1	7	6	0.86	0.86
Level 2	30	45	1.5	1.5
Level 3	114	152	1.33	1.33

Network Analysis on "Mining the Social Web" Repo

REPO EGO GRAPH (LEVEL 1)

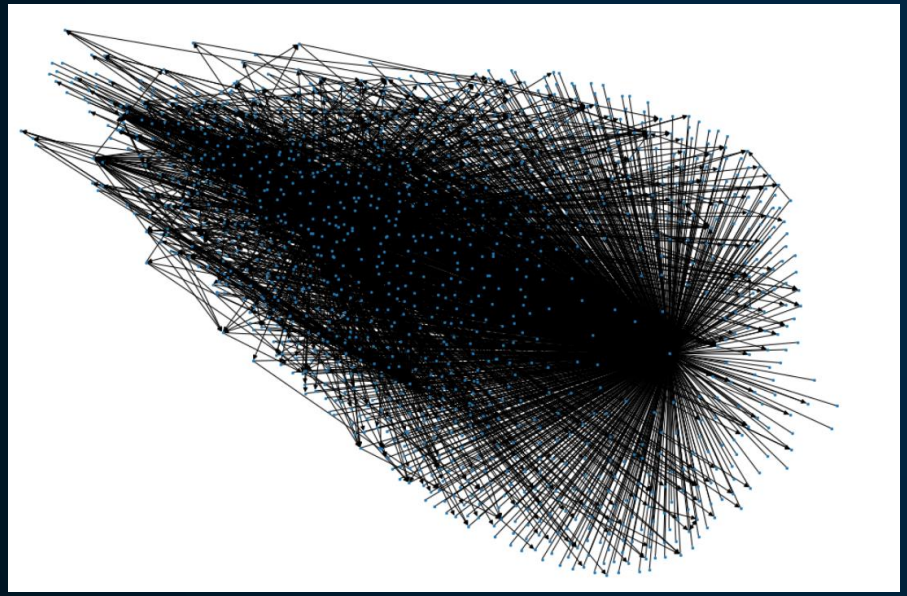
Edges indicate ‘gazing’

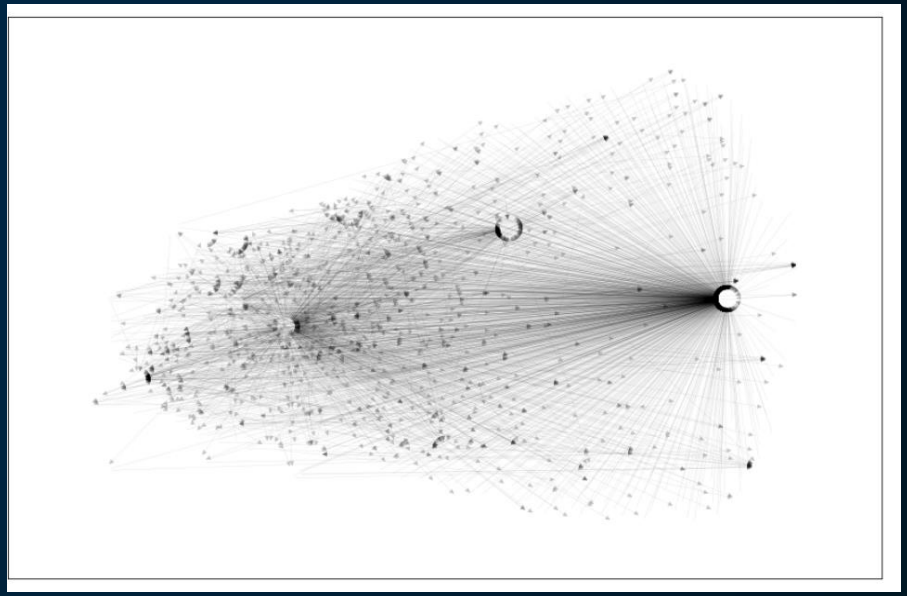




CONNECTIONS AMONG STARGAZERS (LEVEL 2)

Edges indicate ‘repo gazing’ and ‘within gazers following’





LEVEL - 2 OBSERVATIONS

Most popular users based on degree & followers among stargazers

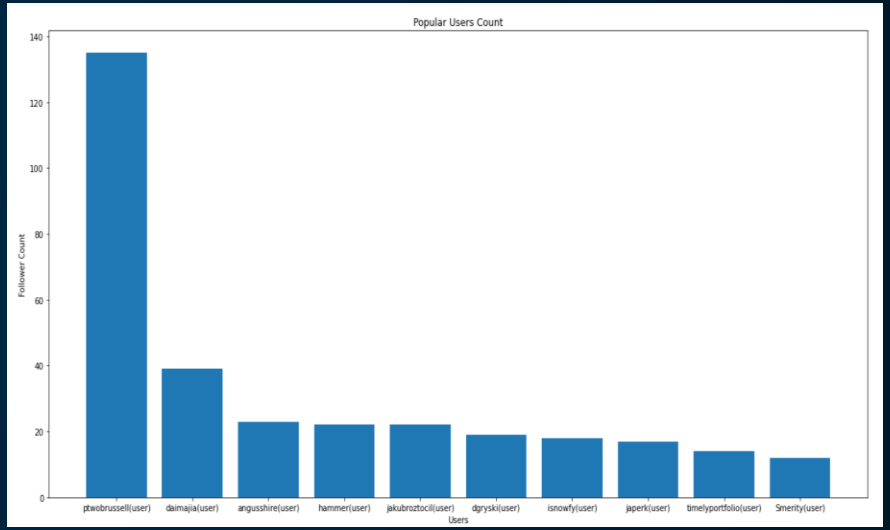
Angus Hung - PhD UC Berkeley, 11.k followers & 213k stars

Rohithadassanayake - 11.8k followers

VagrantStory - Developer Program Member, 17.3k followers

Mathew Russell - Owner & Author of the 'Mining the social web' repo

Daimajia - Student, 22.1k followers and 2.9k stars




```
[('ptwobrussell(user)', 136), ('VagrantStory(user)', 102), ('trietptm(user)', 71),
('rohithadassanayake(user)', 68), ('daimajia(user)', 44), ('mcanthony(user)', 34), ('JT5D(user)',
32), ('andrewwxy(user)', 31)]
```

LEVEL-2 CENTRALITY STATS

```
Degree Centrality
[('angusshire(user)', 0.41288191577208916), ('ptwobrussell(user)', 0.11147811725846407), ('VagrantStory(user)', 0.08340214698596202), ('trietptm(user)', 0.057803468208092484), ('rohithadassanayake(user)', 0.05532617671345995), ('daimajia(user)', 0.03550784475639967), ('mcanthony(user)', 0.027250206440957884), ('JT5D(user)', 0.02559867877786953), ('andrewwxy(user)', 0.02477291494632535), ('hammer(user)', 0.023121387283236993)]

Betweenness Centrality
[('angusshire(user)', 0.014194880776259482), ('trietptm(user)', 0.0026759413566630895), ('samholt(user)', 0.0011850734656830297), ('rohithadassanayake(user)', 0.0010883959338893773), ('VagrantStory(user)', 0.0008561830456452001), ('daimajia(user)', 0.0007513012516790038), ('uetchy(user)', 0.0007052652120248737), ('hammer(user)', 0.0005859658894985315), ('JT5D(user)', 0.0005794514477137015), ('miku(user)', 0.000548444589876096)]

Closeness Centrality
[('ptwobrussell(user)', 0.11366949912540067), ('dgryski(user)', 0.03967540943114316), ('odewahn(user)', 0.037430069834464424), ('samuel(user)', 0.036335725931101646), ('daimajia(user)', 0.03464341639762599), ('japerk(user)', 0.033993252804069896), ('mcroydon(user)', 0.03385091994408553), ('miku(user)', 0.03206929257860735), ('albertsun(user)', 0.03193457380571307), ('acdha(user)', 0.03148713395545476)]
```

STARRED REPOS OF STARGAZERS (LEVEL 3)

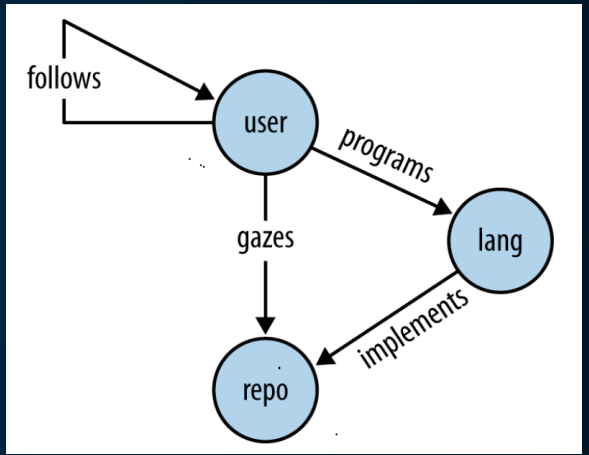
Nodes - Seed Repository + 1176 Stargazers + Starred Repositories of stargazers Edges - ‘Stargazing’

Repo Name	# Stargazers
Mining-the-Social-Web	1212
dotfiles	178
bootstrap	166
tensorflow	166
d3	165
System-design-repo	151
Free-programming-books	136

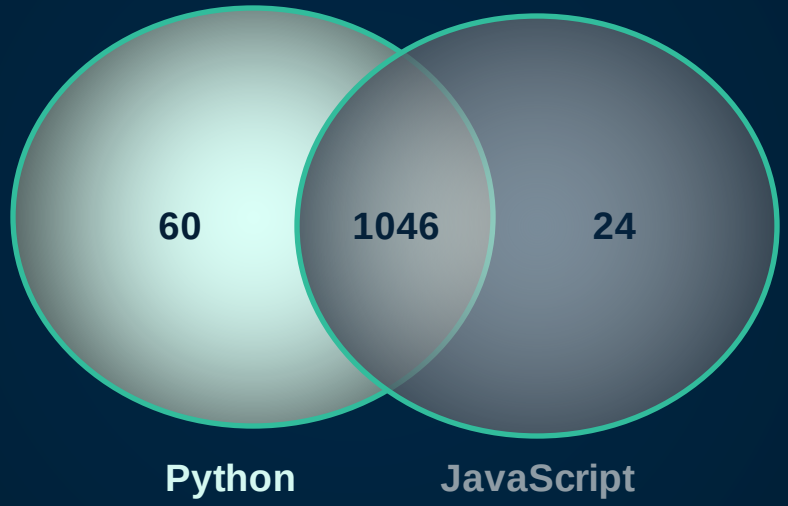
REPOSITORY NETWORK ANALYSIS

Number of Stargazers = 1212 (1213 nodes)

Levels	Nodes	Edges	In Degree	Out Degree
1	1213	1212	0.99	0.99
2	1213	2695	2.22	2.22
3	145898	309042	2.18	2.18



LANGUAGE OF THE STARRED REPOS (LEVEL 4)



The 2 most popular languages are Python and JavaScript

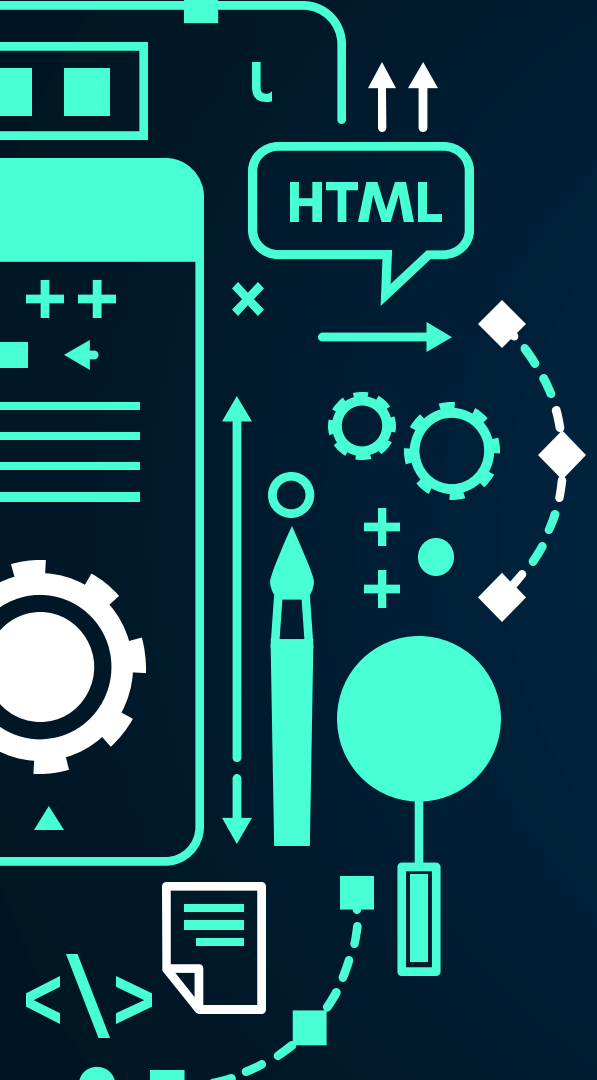
FUTURE SCOPE

Performance Tracking — *assess the number of commits and pull requests to define user importance and impact of individual contributions in a community*

Tech-stack Insights — *Identify the key trends and traction on the programming languages*

Additional Analytics — *centrality measures, we can use cliques or bipartite algorithms to derive insights*

Project Management — *can identify bottlenecks in workflows and improve project management by analyzing user-repository interactions*



THANKS!