# Social Networks in the Marvel Universe

**ISSS606 Social Analytics and Applications**

**Group 8**

Chen Jin Chuan Brian

Clara Chua Kiah Hwii

Jin Xi Yuan

Lee Kern Choong

# Agenda

- Project Objective
- Data Collection & Preparation
- Analysis & Results
- Limitations and Further Work
- Q&A


# Project Objectives

- **21 MCU movies grossed $17.9B worldwide over the past 12 years**

- **To discover:**
  - **How importance of MCU characters have changed over time**
  - **If future characters in the MCU can be predicted**

- **Ideas:**
  - **Comics dataset size >> MCU dataset size**
  - **Comics serve as storyline for MCU movies**

# Data Collection

- **Primary Datasets**
  - MCU Movie Cast (IMDB): spanning 2008 – 2019
  - Marvel Universe Social Graph: spanning 1939 - 2018

- **Secondary Dataset**
  - Marvel Superheroes Dataset (incomplete)

- **Assumptions**
  - Character co-occurrence is equivalent to an edge

# The Infinity Saga: 3 phases of MCU movies

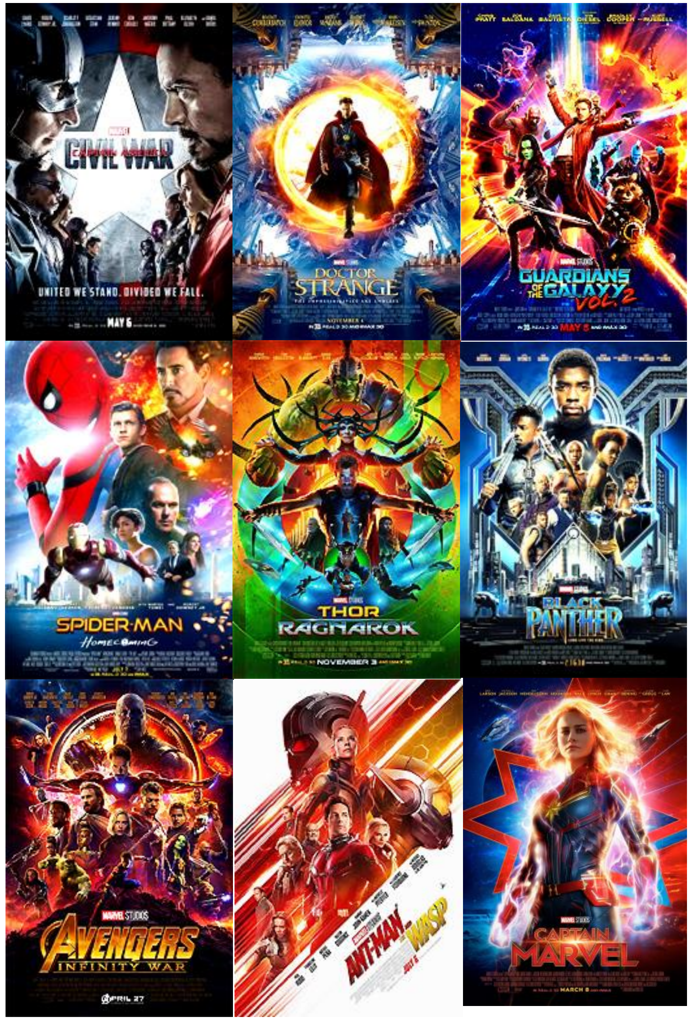
## Phase 1 (2008 – 2012)


- Iron Man
- The Incredible Hulk
- Iron Man 2
- Thor
- Captain America: The First Avenger
- The Avengers

## Phase 2 (2013 – 2015)


- Iron Man 3
- Thor: The Dark World
- Captain America: The Winter Soldier
- Guardians of the Galaxy
- Avengers: Age of Ultron
- Ant-Man

## Phase 3 (2016 – 2019)




- Captain America: Civil War
- Doctor Strange
- Guardians of the Galaxy Vol 2
- Spiderman – Homecoming
- Thor: Ragnarok
- Black Panther
- Avengers: Infinity War
- Ant-Man & The Wasp
- Captain Marvel

# Movie Data Collection & Preparation


Data from IMDB

| Movie | Year | Character | Time | Phase |
|---|---|---|---|---|
| Iron Man | 2008 | Tony Stark / Iron Man | 77:15:00 | 1 |
| Iron Man | 2008 | Pepper Potts | 23:15 | 1 |
| Iron Man | 2008 | Obadiah Stane / Iron Monger | 22:00 | 1 |
| Iron Man | 2008 | Professor Ho Yinsen | 10:45 | 1 |
| Iron Man | 2008 | Col. James "Rhodey" Rhodes / Wa | 8:15 | 1 |
| Iron Man | 2008 | Raza | 6:00 | 1 |
| Iron Man | 2008 | Agent Phil Coulson | 3:45 | 1 |
| Iron Man | 2008 | Christine Everhart | 3:45 | 1 |
| Iron Man | 2008 | Abu Bakaar | 1:45 | 1 |
| Iron Man | 2008 | Harold "Happy" Hogan | 1:15 | 1 |
| Iron Man | 2008 | Nick Fury | 0:15 | 1 |
| Iron Man | 2008 | J.A.R.V.I.S. | | 1 |
| X-Men Origins: Wolverine | 2009 | Logan / Wolverine / Weapon X | 50:30:00 | |
| X-Men Origins: Wolverine | 2009 | Victor Creed / Sabretooth | 17:00 | |
| X-Men Origins: Wolverine | 2009 | William Stryker | 11:15 | |

**Clean data**

Nick Fury = Director Nick Fury = Agent Nick Fury

# Movie Data Collection & Preparation

| Movie | Year | Character | Time | Phase |
|---|---|---|---|---|
| Iron Man | 2008 | Tony Stark / Iron Man | 77:15:00 | 1 |
| Iron Man | 2008 | Pepper Potts | 23:15 | 1 |
| Iron Man | 2008 | Obadiah Stane / Iron Monger | 22:00 | 1 |
| Iron Man | 2008 | Professor Ho Yinsen | 10:45 | 1 |
| Iron Man | 2008 | Col. James "Rhodey" Rhodes / Wa | 8:15 | 1 |
| Iron Man | 2008 | Raza | 6:00 | 1 |
| Iron Man | 2008 | Agent Phil Coulson | 3:45 | 1 |
| Iron Man | 2008 | Christine Everhart | 3:45 | 1 |
| Iron Man | 2008 | Abu Bakaar | 1:45 | 1 |
| Iron Man | 2008 | Harold "Happy" Hogan | 1:15 | 1 |
| Iron Man | 2008 | Nick Fury | 0:15 | 1 |
| Iron Man | 2008 | J.A.R.V.I.S. | | 1 |

Cleaned

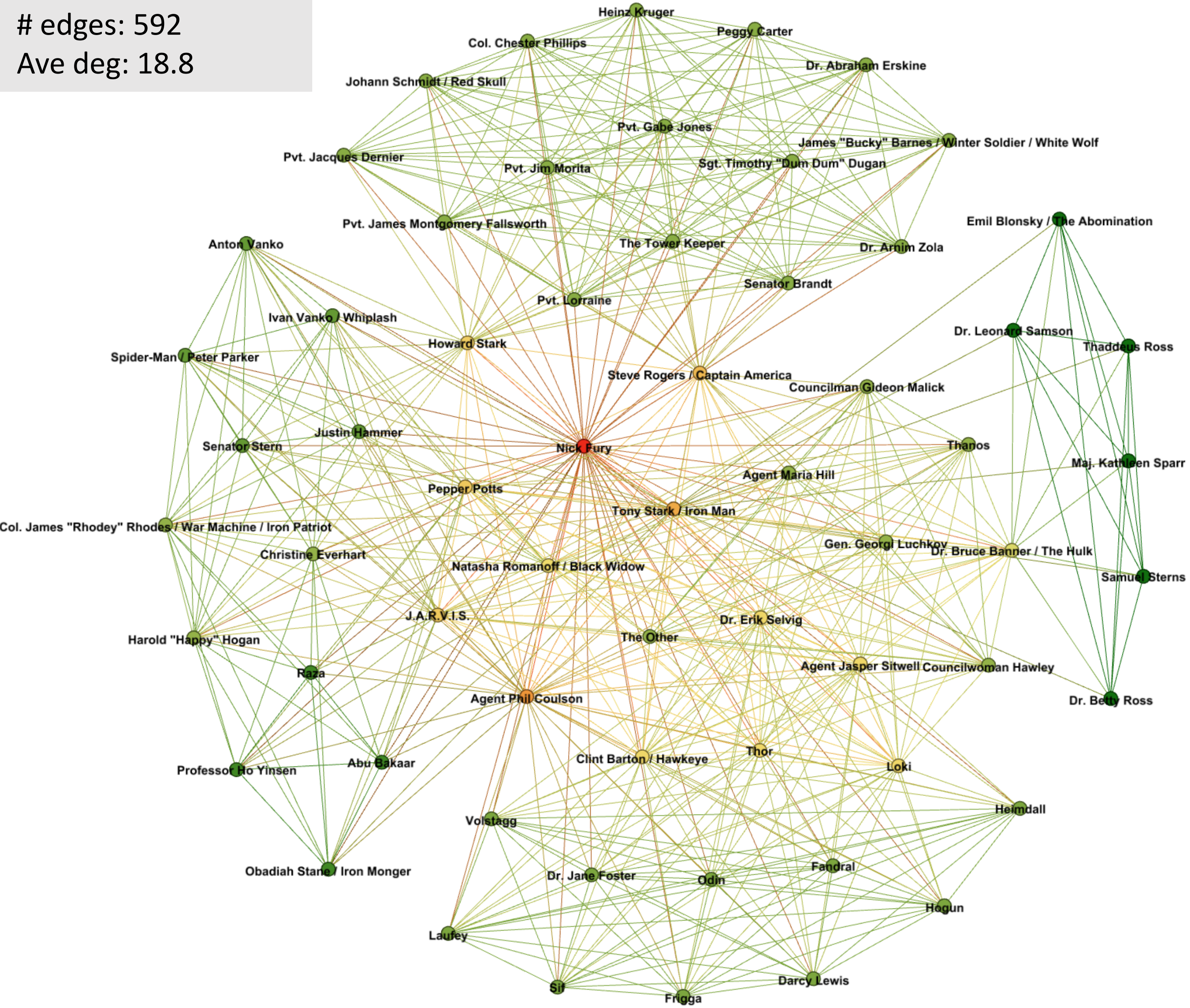
| | Movie | char1 | char2 | Phase |
|---|---|---|---|---|
| 3202 | Iron Man | Tony Stark / Iron Man | Pepper Potts | 1 |
| 3203 | Iron Man | Tony Stark / Iron Man | Obadiah Stane / Iron Monger | 1 |
| 3204 | Iron Man | Tony Stark / Iron Man | Professor Ho Yinsen | 1 |
| 3205 | Iron Man | Tony Stark / Iron Man | Col. James "Rhodey" Rhodes / War Machine / Iro... | 1 |
| 3206 | Iron Man | Tony Stark / Iron Man | Raza | 1 |
| 3207 | Iron Man | Tony Stark / Iron Man | Agent Phil Coulson | 1 |
| 3208 | Iron Man | Tony Stark / Iron Man | Christine Everhart | 1 |
| 3209 | Iron Man | Tony Stark / Iron Man | Abu Bakaar | 1 |
| 3210 | Iron Man | Tony Stark / Iron Man | Harold "Happy" Hogan | 1 |
| 3211 | Iron Man | Tony Stark / Iron Man | Nick Fury | 1 |
| 3212 | Iron Man | Tony Stark / Iron Man | J.A.R.V.I.S. | 1 |
| 3213 | Iron Man | Pepper Potts | Obadiah Stane / Iron Monger | 1 |
| 3214 | Iron Man | Pepper Potts | Professor Ho Yinsen | 1 |
| 3215 | Iron Man | Pepper Potts | Col. James "Rhodey" Rhodes / War Machine / Iro... | 1 |
| 3216 | Iron Man | Pepper Potts | Raza | 1 |
| 3217 | Iron Man | Pepper Potts | Agent Phil Coulson | 1 |
| 3218 | Iron Man | Pepper Potts | Christine Everhart | 1 |
| 3219 | Iron Man | Pepper Potts | Abu Bakaar | 1 |
| 3220 | Iron Man | Pepper Potts | Harold "Happy" Hogan | 1 |
| 3221 | Iron Man | Pepper Potts | Nick Fury | 1 |
| 3222 | Iron Man | Pepper Potts | J.A.R.V.I.S. | 1 |
| 3223 | Iron Man | Obadiah Stane / Iron Monger | Professor Ho Yinsen | 1 |

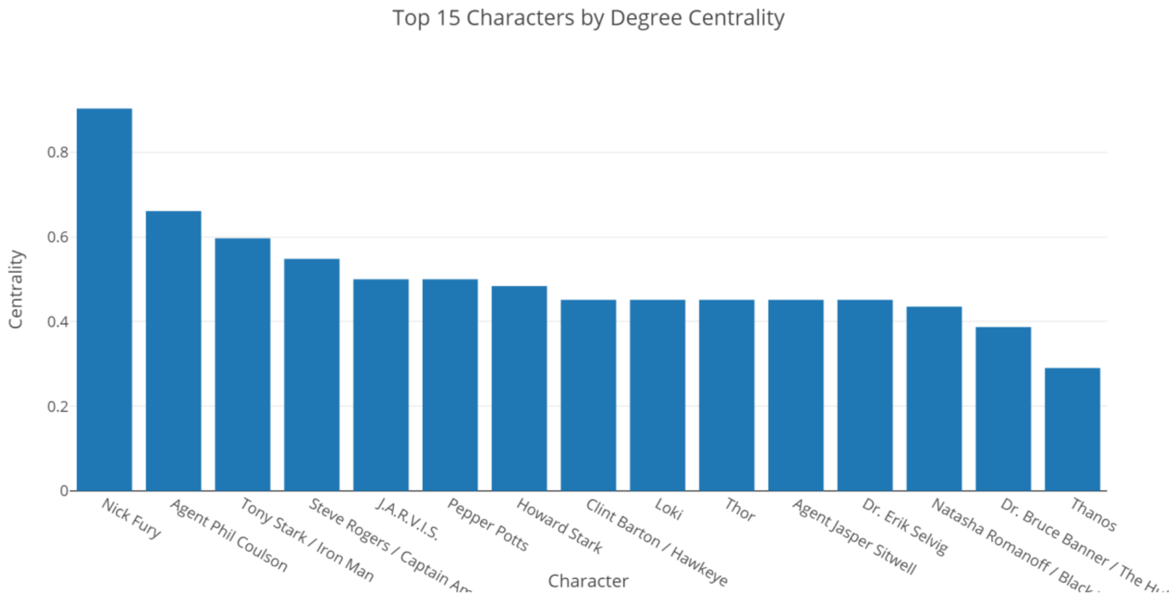
- **Itertools Combinations** function
- **pd.merge** with phases dataframe
- **pd.query** to slice into 3 different phases

```python
from itertools import combinations

mcu = (MCU_char.groupby('Movie')['Character'].apply(lambda x: pd.DataFrame(list(combinations(x,2))))
                                                                   .reset_index(level=1, drop = True)
                                                                   .reset_index())

#rename column names
mcu.rename(columns={0:'char1',
                    1:'char2'},
                 inplace=True)

#Add in phases column
#select phases from movie_info df
mcu_phase = movie_info[['Title', 'Phase']]
mcu_phase.rename(columns={'Title':'Movie'},
                 inplace = True)
#merge MCU with MCU_phases
mcu_new = pd.merge(mcu, mcu_phase,
                  left_on=['Movie'],
                  right_on=['Movie'],
                  how='inner')
```

# MCU Movies – Phase 1

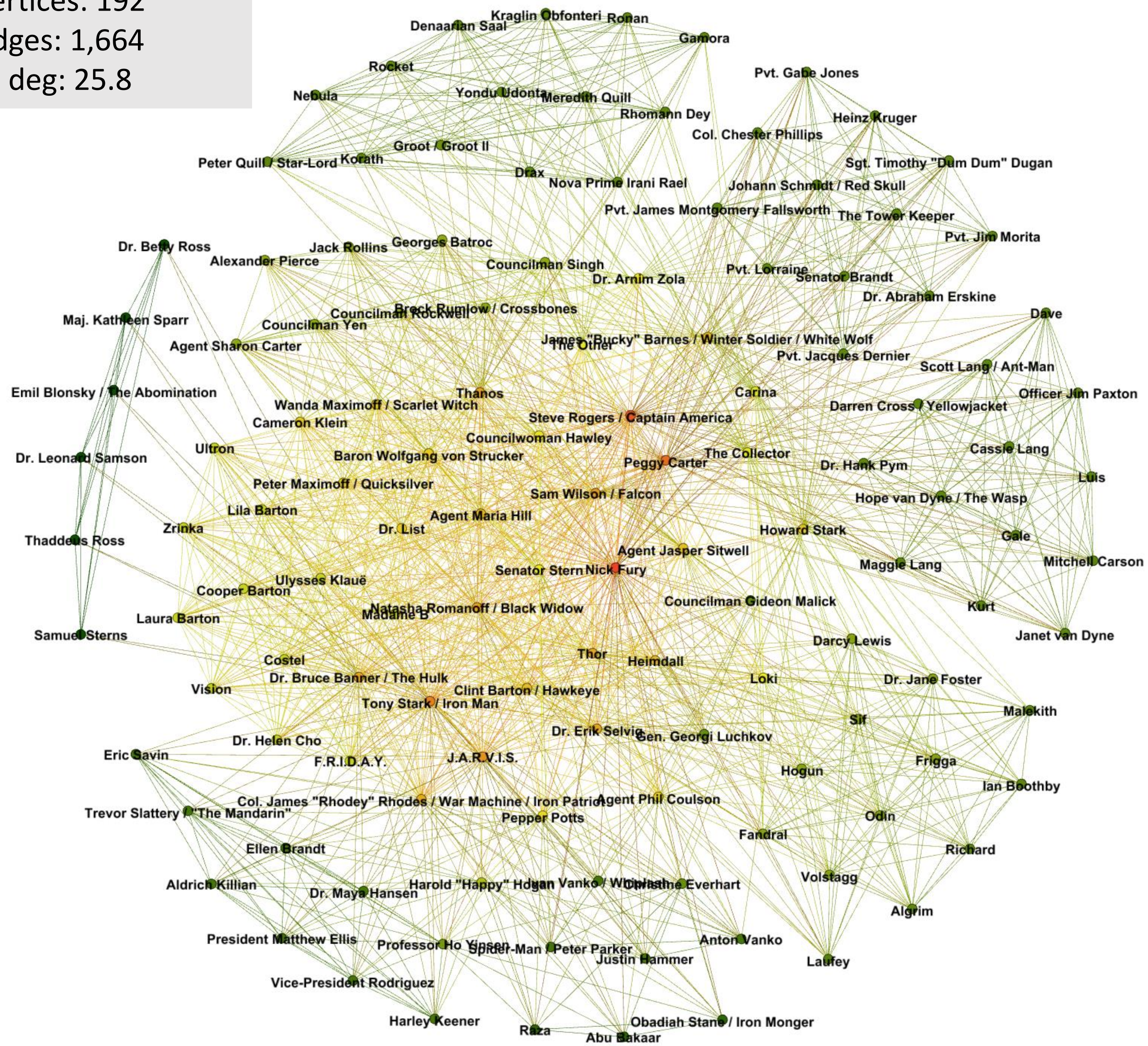
# vertices: 63
# edges: 592
Ave deg: 18.8

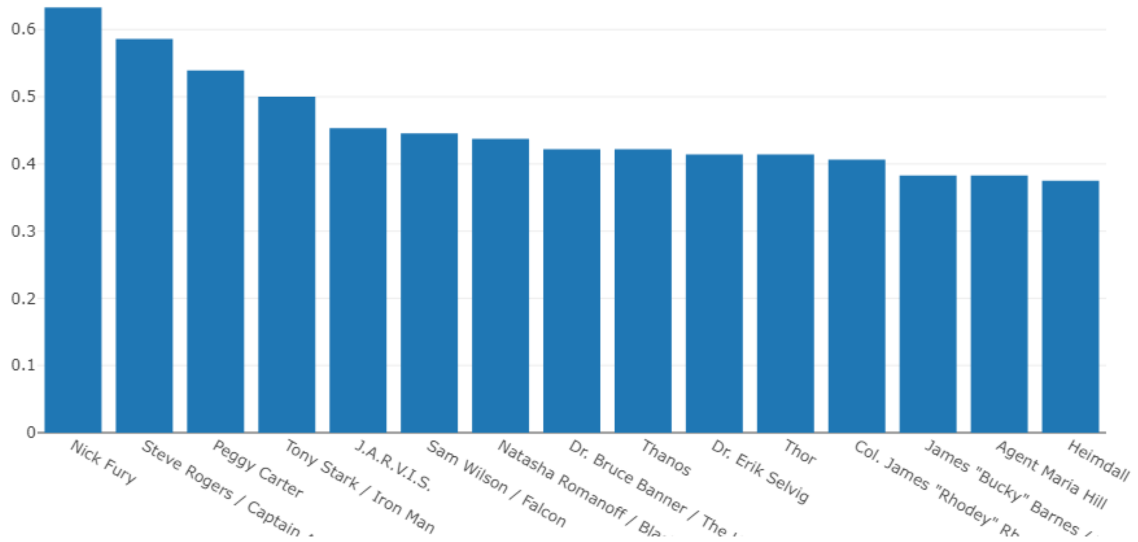



# MCU Movies – Phase 2

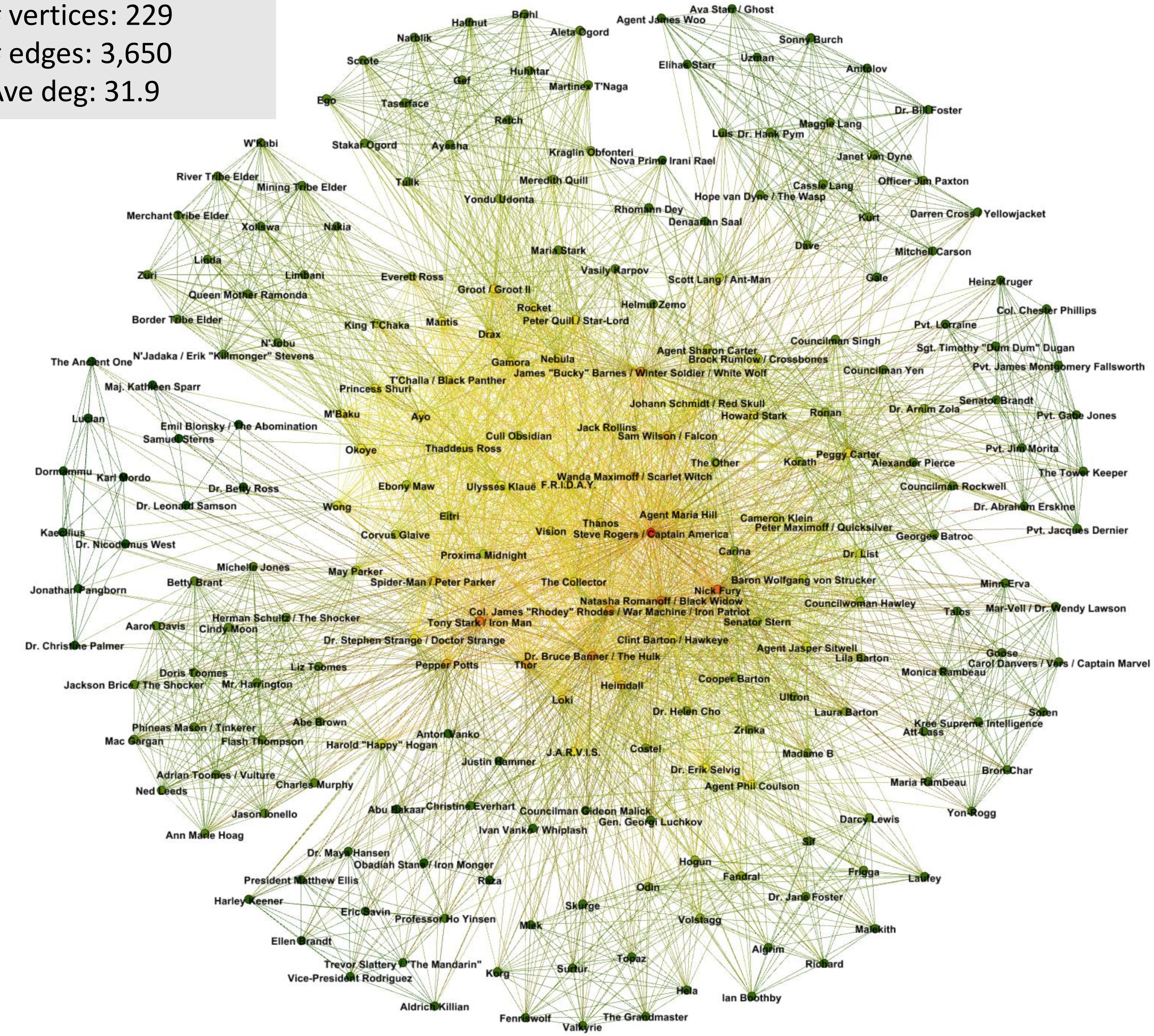
# vertices: 192
# edges: 1,664
Ave deg: 25.8

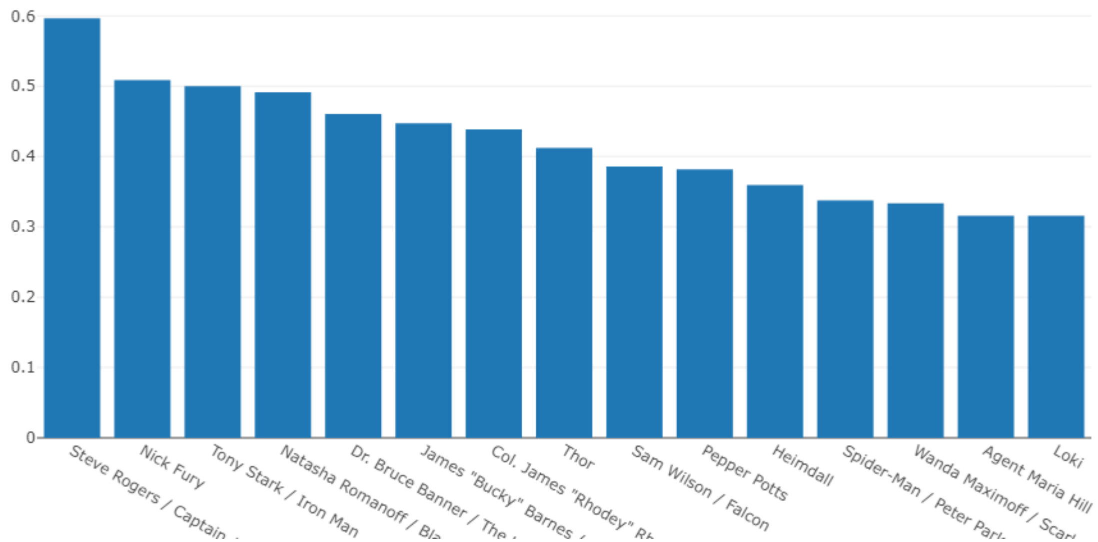


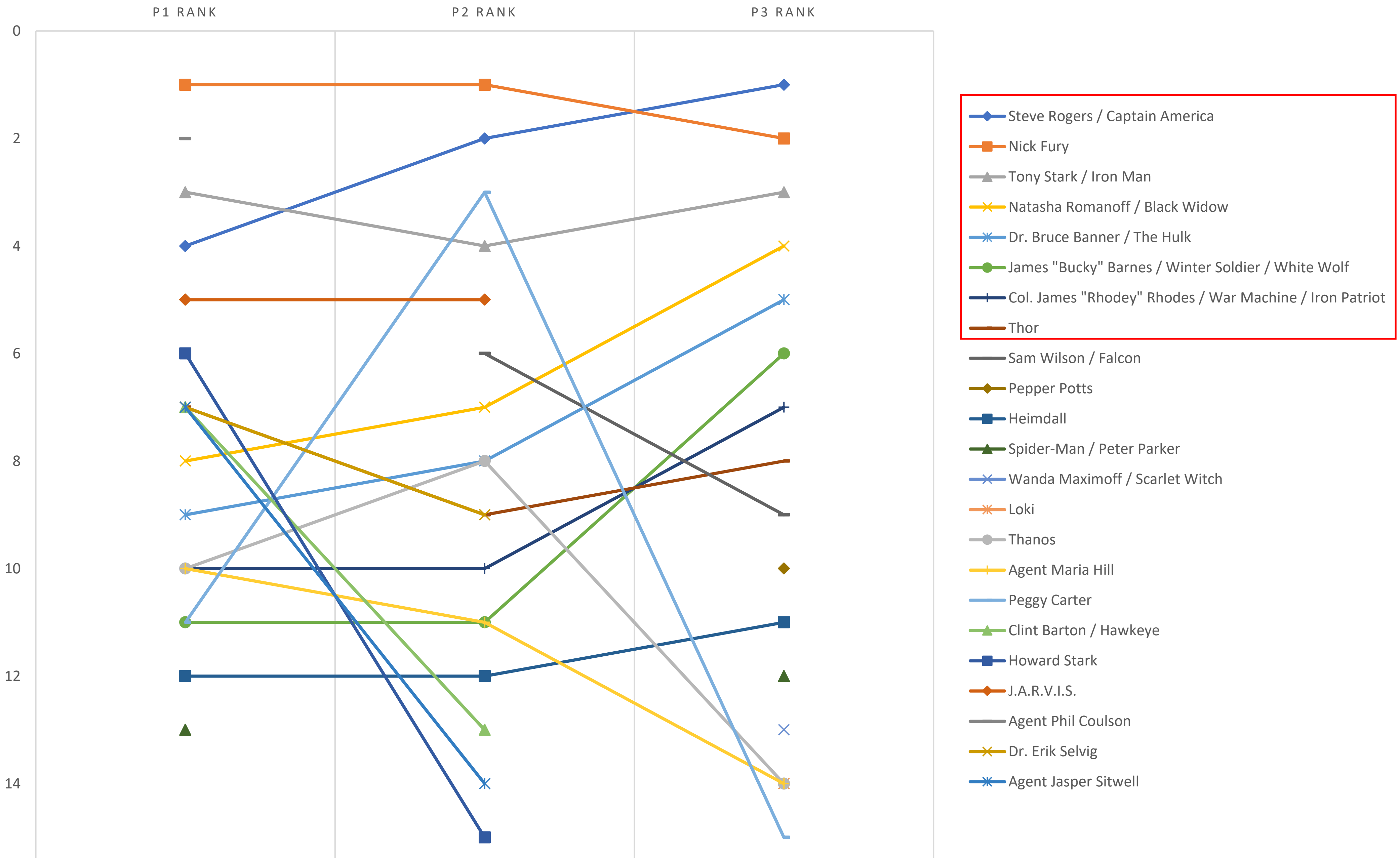

# MCU Movies – Phase 3

# vertices: 229
# edges: 3,650
Ave deg: 31.9






# How did characters' centrality fare in 3 phases?



# All 50 ‘Marvel’ related movies

# vertices: 495
# edges: 6,100
Ave deg: 24.6

**Interesting Findings:**

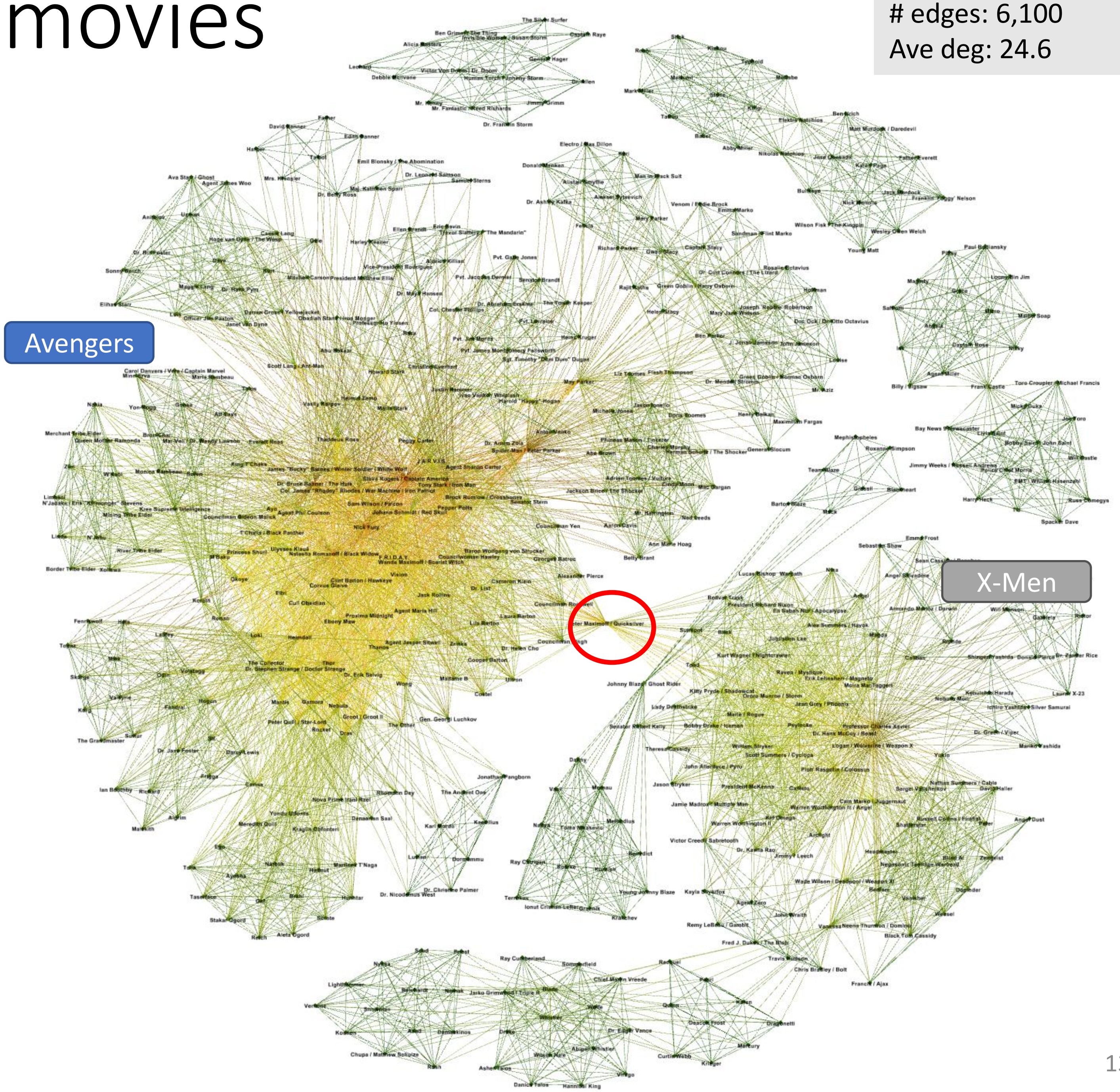
**Top 10 Central Characters**

| |
|---|
| Steve Rogers / Captain America |
| Nick Fury |
| Tony Stark / Iron Man |
| Spider-Man / Peter Parker |
| Natasha Romanoff / Black Widow |
| Dr. Bruce Banner / The Hulk |
| James "Bucky" Barnes / Winter Soldier / White Wolf |
| Col. James "Rhodey" Rhodes / War Machine / Iron Patriot |
| **Logan / Wolverine / Weapon X** |
| **Professor Charles Xavier** |

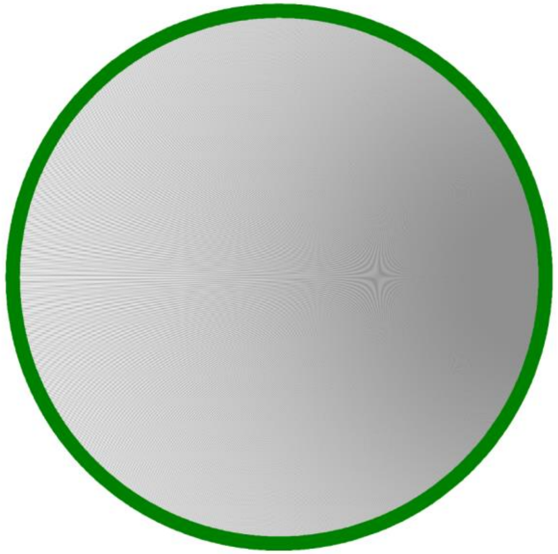
**Bridge between “communities”:**

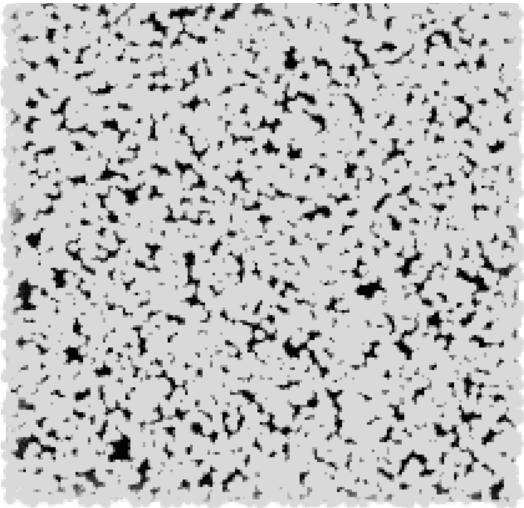
- **Quicksilver** (Avengers to X-Men)

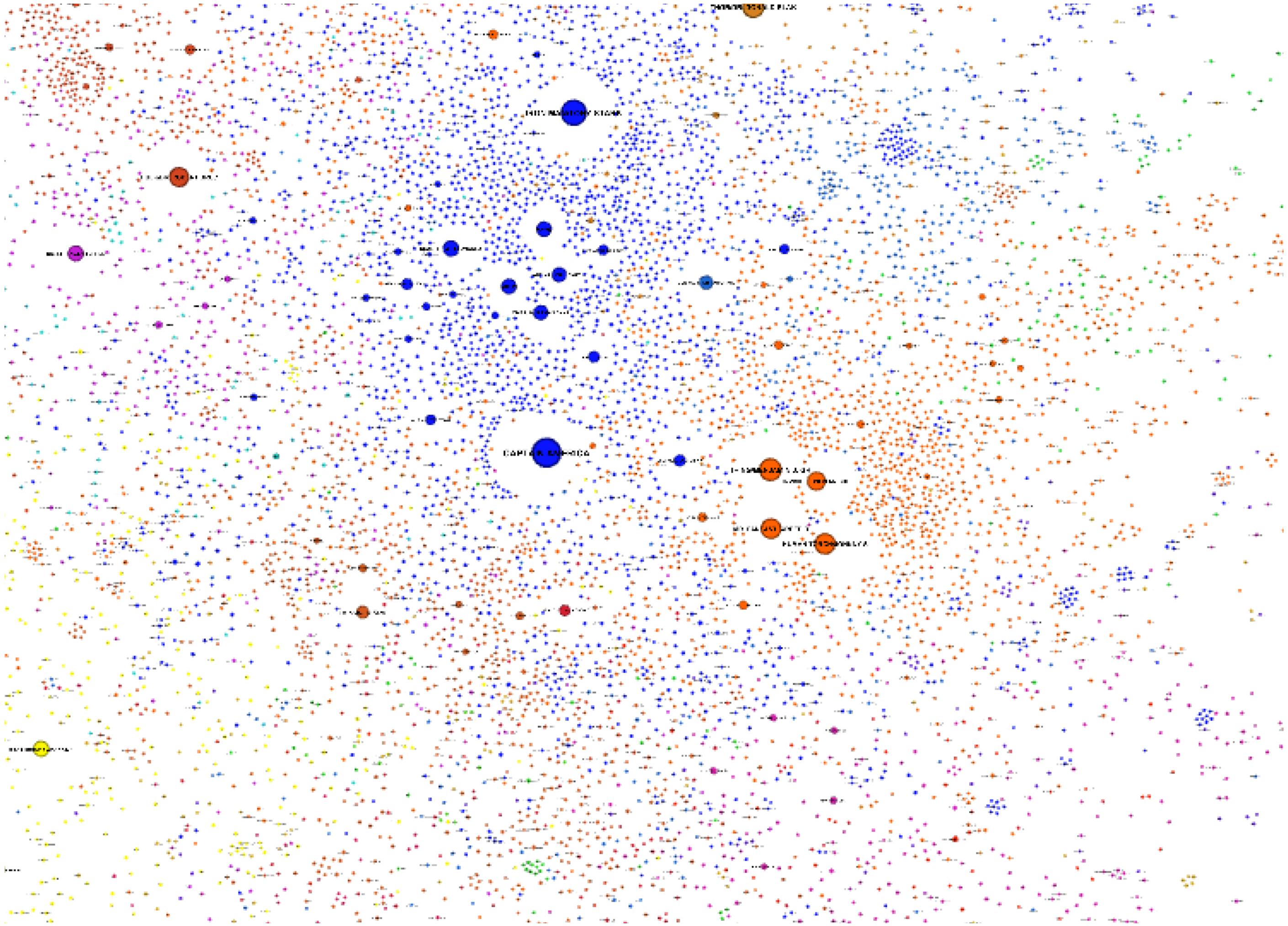
| |
|---|
| Spider-Man / Peter Parker |
| Logan / Wolverine / Weapon X |
| Professor Charles Xavier |
| Steve Rogers / Captain America |



# Comic Data Preparation







# Comic Data Preparation

| | hero1 | hero2 |
|---|---|---|
| 213085 | PATRIOT/JEFF MACE | PATRIOT/JEFF MACE |
| 186598 | MISS AMERICA/MADELIN | PATRIOT/JEFF MACE |
| 213083 | PATRIOT/JEFF MACE | MISS AMERICA/MADELIN |
| 186594 | MISS AMERICA/MADELIN | MISS AMERICA/MADELIN |
| 292584 | THING/BENJAMIN J. GR | HUMAN TORCH/JOHNNY S |
| 128312 | HUMAN TORCH/JOHNNY S | THING/BENJAMIN J. GR |
| 193246 | MR. FANTASTIC/REED R | HUMAN TORCH/JOHNNY S |
| 127900 | HUMAN TORCH/JOHNNY S | MR. FANTASTIC/REED R |
| 292879 | THING/BENJAMIN J. GR | MR. FANTASTIC/REED R |

| hero1_hero2 | weight |
|---|---|
| A--B | 3 |
| B--C | 2 |
| D--E | 1 |

| hero1_hero2 |
|---|
| B -- A |
| C -- B |
| E -- D |

| | hero1 | hero2 | weight |
|---|---|---|---|
| 213085 | PATRIOT/JEFF MACE | PATRIOT/JEFF MACE | 2550 |
| 186598 | MISS AMERICA/MADELIN | PATRIOT/JEFF MACE | 1894 |
| 213083 | PATRIOT/JEFF MACE | MISS AMERICA/MADELIN | 1894 |
| 186594 | MISS AMERICA/MADELIN | MISS AMERICA/MADELIN | 1344 |
| 292584 | THING/BENJAMIN J. GR | HUMAN TORCH/JOHNNY S | 744 |
| 128312 | HUMAN TORCH/JOHNNY S | THING/BENJAMIN J. GR | 744 |
| 193246 | MR. FANTASTIC/REED R | HUMAN TORCH/JOHNNY S | 713 |

# How do Movies Compare to the Comics?

Degree Centrality:

| Rank | Movies | Comics |
|---|---|---|
| 1 | **Steve Rogers / Captain America** | **Captain America / Steve Rogers** |
| 2 | Nick Fury | Spiderman / Peter Parker |
| 3 | **Tony Stark / Iron Man** | **Iron Man / Tony Stark** |
| 4 | Natasha Romanoff / Black Widow | Thing / Benjamin Grimm |
| 5 | Dr. Bruce Banner / The Hulk | Mr. Fantastic / Reed Richards |
| 6 | James "Bucky" Barnes / Winter Soldier / White Wolf | Wolverine / Logan |
| 7 | Col. James "Rhodey" Rhodes / War Machine / Iron Patriot | Human Torch / Johnny Storm |
| 8 | Thor | Scarlet Witch / Wanda Maximoff |
| 9 | Sam Wilson / Falcon | Thor / Dr. Donald Blake |
| 10 | Pepper Potts | Beast / Hank McCoy |

# Predicting Future Characters

1. Pick a complete 'trilogy'

2. Apply different measures of centrality to select best neighbours from comic dataset

3. Determine the best measure based on accuracy rate

4. Apply chosen measure to predict characters in upcoming movies

- There are 2 complete trilogies in the entire MCU – Iron Man & Thor
- Identify all unique characters in the trilogy to use for verification of our measures

- Apply the following measures to the comic dataset to predict characters that will appear in the Thor movie series:
  - Top weighted neighbours
  - Degree Centrality
  - Similarity

- Compare accuracy rate of three measures to determine which one is the best

- Apply the chosen measure to upcoming movies:
  - Black Widow
  - Black Panther 2
  - Doctor Strange 2
  - Gambit

# Identify Unique Characters of a Series




```
for char in g.adj['Thor']:
    print(char)
```


```
for char in g.adj['Thor: The Dark World']:
    print(char)
```


```
for char in g.adj['Thor: Ragnarok']:
    print(char)
```

| 1. Thor | |
|---|---|
| **Thor** | **Frigga** |
| **Loki** | **Heimdall** |
| **Dr. Jane Foster** | **Agent Phil Coulson** |
| **Dr. Erik Selvig** | **Laufey** |
| **Odin** | **Agent Jasper Sitwell** |
| **Darcy Lewis** | **Nick Fury** |
| **Sif** | **Clint Barton / Hawkeye** |
| **Volstagg** | **Doctor Strange** |
| **Fandral** | |
| **Hogun** | |

| 2. Thor: The Dark World | |
|---|---|
| *Thor* | *Frigga* |
| *Loki* | *Heimdall* |
| *Dr. Jane Foster* | **Malekith** |
| *Dr. Erik Selvig* | **Algrim** |
| *Odin* | **Ian Boothby** |
| *Darcy Lewis* | **The Collector** |
| *Sif* | **Richard** |
| *Volstagg* | **Carina** |
| *Fandral* | |
| *Hogun* | |

| 3. Thor: Ragnarok | |
|---|---|
| *Thor* | **Surtur** |
| *Loki* | **Dr. Stephen Strange** |
| **Dr. Bruce Banner / Hulk** | **Topaz** |
| **Valkyrie** | **Fenriswolf** |
| *Odin* | **Miek** |
| **Hela** | *Hogun* |
| **Skurge** | *Volstagg* |
| **The Grandmaster** | *Fandral* |
| *Heimdall* | **Natasha Romanoff / Black Widow** |
| **Korg** | |

**There are 35 unique main characters in the Thor Trilogy (excluding Thor)**

# Measures to identify potential neighbours



Top Weighted Neighbors

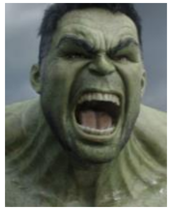
- **Create a weighted graph**
- **Select all neighbours of key character**
- **Rank neighbours by weight**
- **Select top neighbors**





| Character | Weight |
|---|---|
| Captain America | 386 |
| Iron Man/Tony Stark | 344 |
| Odin | 266 |
| Vision | 255 |
| Scarlet Witch/Wanda Maximoff | 254 |
| Wasp/Janet Van Dyne | 238 |
| Hawk | 210 |
| Balder | 209 |
| Sif | 204 |
| Ant-Man/Dr. Henry J. Pym | 189 |
| Volstagg | 187 |
| Fandral | 186 |
| Hogun | 186 |
| Loki | 182 |
| Jarvis, Edwin | 160 |
| Mr. Fantastic/Reed Richards | 129 |
| Thing/Benjamin J. Grimm | 126 |
| Wonder Man/Simon Williams | 125 |




| Character | Weight |
|---|---|
| Human Torch/Johnny Storm | 124 |
| She-Hulk/Jennifer Walters | 123 |
| Hercules | 114 |
| Kincaid, Dr. Jane Foster | 112 |
| Heimdall | 111 |
| Invisible Woman/Sue Storm | 111 |
| Black Panther/T'Challa | 103 |
| Quicksilver/Pietro Maximoff | 99 |
| Spider-Man/Peter Parker | 95 |
| Beast/Henry Hank Mccoy | 90 |
| Captain Marvel II/Monica Rambeau | 84 |
| Enchantress/Amora | 78 |
| Karnilla | 76 |
| Hulk/Dr. Robert Bruce Banner | 72 |
| Vizier | 72 |
| Sub-Mariner/Namor Mackenzie | 62 |
| Black Widow/Natasha Romanoff | 61 |

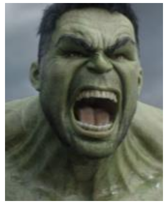
# Neighbours with top centrality scores



Centrality

- **Use the weighted graph previously created**
- **Get all the neighbors of one character**
- **Rank with degree centrality**
- **Select top neighbors**

| Character | Degree Centrality |
|---|---|
| Captain America | 0.296965 |
| Spider-Man/Peter Parker | 0.27035 |
| Iron Man/Tony Stark | 0.236887 |
| Thing/Benjamin J. Grimm | 0.220389 |
| Mr. Fantastic/Reed Richards | 0.21463 |
| Wolverine/Logan | 0.213385 |
| Human Torch/Johnny Storm | 0.211829 |
| Scarlet Witch/Wanda Maximoff | 0.206226 |
| Beast/Henry Hank Mccoy | 0.197198 |
| Vision | 0.193152 |
| Invisible Woman/Sue Storm | 0.192374 |
| Hawk | 0.182879 |
| Wasp/Janet Van Dyne | 0.169805 |
| Ant-Man/Dr. Henry J. Pym | 0.168405 |
| Cyclops/Scott Summers | 0.168249 |
| Angel/Warren Kenneth Worthington III | 0.167004 |
| Storm/Ororo Munroe | 0.166848 |
| She-Hulk/Jennifer Walters | 0.166693 |



| Character | Degree Centrality |
|---|---|
| Dr. Strange/Stephen Strange | 0.16607 |
| Hulk/Dr. Robert Bruce Banner | 0.164202 |
| Wonder Man/Simon Williams | 0.160778 |
| Professor X/Charles Xavier | 0.160623 |
| Colossus II/Peter Rasputin | 0.159533 |
| Marvel Girl/Jean Grey | 0.15642 |
| Hercules | 0.154241 |
| Jarvis, Edwin | 0.153463 |
| Sub-Mariner/Namor Mackenzie | 0.152374 |
| Daredevil/Matt Murdock | 0.150506 |
| Iceman/Robert Bobby Drake | 0.147082 |
| Black Widow/Natasha Romanoff | 0.143502 |
| Fury, Col. Nicholas | 0.143502 |
| Jameson, J. Jonah | 0.143191 |
| Quicksilver/Pietro Maximoff | 0.136187 |
| Nightcrawler/Kurt Wagner | 0.133385 |
| Rogue | 0.129339 |

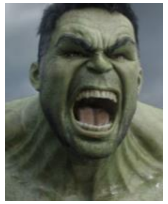
# Neighbours with top centrality scores



Centrality

- Use the weighted graph previously created
- Get all the neighbors of one character
- Rank with eigenvector centrality
- Select top neighbors

| Character | Eigenvector Centrality |
|---|---|
| Captain America | 0.291174519 |
| Thing/Benjamin J. Gr | 0.232840152 |
| Human Torch/Johnny Storm | 0.228728903 |
| Iron Man/Tony Stark | 0.225055946 |
| Mr. Fantastic/Reed Richards | 0.224791665 |
| Invisible Woman/Sue Storm | 0.216340253 |
| Scarlet Witch/Wanda Maximoff | 0.212751467 |
| Vision | 0.209606757 |
| Thor/Dr. Donald Blake | 0.199501378 |
| Wasp/Janet Van Dyne | 0.197125043 |
| Hawk | 0.176999356 |
| Ant-Man/Dr. Henry J. Pym | 0.171095374 |
| Beast/Henry Hank Mccoy | 0.157246619 |
| Wonder Man/Simon Williams | 0.143439689 |
| Cyclops/Scott Summers | 0.135707624 |
| Wolverine/Logan | 0.131486003 |
| Spider-Man/Peter Parker | 0.129055006 |
| Storm/Ororo Munroe | 0.122635321 |



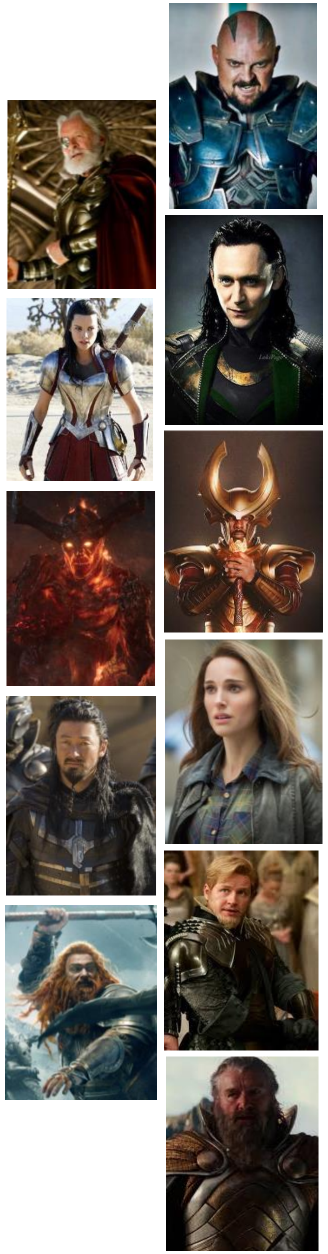
| Character | Eigenvector Centrality |
|---|---|
| Professor X/Charles Xavier | 0.116516396 |
| She-Hulk/Jennifer Walters | 0.114708977 |
| Jarvis, Edwin | 0.109228344 |
| Quicksilver/Pietro Maximoff | 0.107505041 |
| Colossus II/Peter Rasputin | 0.10650563 |
| Marvel Girl/Jean Grey | 0.105099105 |
| Angel/Warren Kenneth Worthington III | 0.104304004 |
| Iceman/Robert Bobby Drake | 0.102478167 |
| Hulk/Dr. Robert Bruce Banner | 0.100038363 |
| Hercules | 0.096838714 |
| Patriot/Jeff Mace | 0.095058577 |
| Sub-Mariner/Namor Mackenzie | 0.089066117 |
| Black Widow/Natasha Romanoff | 0.085247196 |
| Rogue | 0.083888824 |
| Nightcrawler/Kurt Wagner | 0.077362686 |
| Black Panther/T'Challa | 0.077100814 |
| Richards, Franklin B | 0.076174818 |

# Neighbours most similar to Thor

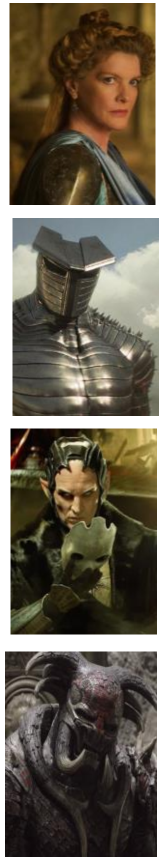
1. 2. 3. 4.

Similarity

- Create a node vector model
- Get Similarity with main character in the movies
- Rank with similarity
- Select top characters



| Character | Similarity |
|---|---|
| Executioner II/Skurge | 0.928727 |
| Odin | 0.907808 |
| Loki | 0.8981 |
| Enchantress/Amora | 0.889757 |
| Sif | 0.877101 |
| Heimdall | 0.873244 |
| Balder | 0.870802 |
| Surtur | 0.86834 |
| Kincaid, Dr. Jane Foster | 0.862963 |
| Krista | 0.862055 |
| Harokin | 0.859799 |
| Hogun | 0.85884 |
| Fandral | 0.855293 |
| Utgard-Loki | 0.851918 |
| Volstagg | 0.847318 |
| Tyr | 0.844327 |
| Hobbs, Harris | 0.837497 |
| Vizier | 0.837075 |



| Character | Similarity |
|---|---|
| Frigga | 0.836512 |
| Volla | 0.828432 |
| Destroyer III | 0.825879 |
| Seth II | 0.824532 |
| Hildegarde | 0.823125 |
| Lorelei II/Melodi | 0.818588 |
| Karnilla | 0.813871 |
| Neffethesk | 0.813761 |
| Pentigaar | 0.812846 |
| Designate/Tarene | 0.805783 |
| Nichols, Lorna | 0.788973 |
| Horus | 0.78628 |
| Toothgnasher | 0.784995 |
| Malekith/Malcolm Keith | 0.784638 |
| Case, Col. Preston | 0.782062 |
| Kurse/Algrim | 0.781911 |

# How did our measures do?



| | Weight of Next Neighbour | Degree Centrality | Eigenvector Centrality | Node2Vec Similarity |
|---|---|---|---|---|
| **Predicted number** | 35 | 35 | 35 | 35 |
| **Matched number** | 10 | 4 | 2 | 15 |
| **Accuracy** | 28.6% | 11.4% | 5.7% | 42.9% |

# Predicting Future Characters of these Movie




## Black Widow (2020)

| Character | Similarity |
|---|---|
| DEATHCRY | 0.827196 |
| *ANT-MAN/DR. HENRY J. PYM* | 0.778099 |
| HERCULES | 0.763727 |
| MANTIS | 0.759803 |
| **IVAN PETROVITCH** | **0.757092** |
| NEUT | 0.755555 |
| ZA'KEN | 0.734836 |
| MADAME MASQUE III | 0.734648 |
| TUC | 0.733714 |
| SWORDSMAN III/PHILIP JAVERT | 0.730123 |
| CARINA/CARINA WALTER | 0.730029 |
| JOCASTA II | 0.728925 |
| *VISION* | 0.726362 |
| T'KYLL ALABAR | 0.718949 |
| TABULA RASA | 0.710955 |
| STORM, CHILI | 0.710499 |
| WATCHLORD | 0.706668 |
| WASP/JANET VAN DYNE | 0.705857 |
| PROCTOR | 0.698965 |
| NELIT | 0.698442 |


## Black Panther 2 (TBD)

| Character | Similarity |
|---|---|
| LYNNE, MONICA | 0.821174 |
| NECRODAMUS | 0.789341 |
| ADAMS, NICOLE NIKKI | 0.775762 |
| SWORDSMAN/JACQUES DUQUESNE | 0.77544 |
| ROSS, EVERETT KENNET | 0.76188 |
| WHITE WOLF/HUNTER | 0.753689 |
| *KILLMONGER, ERIC/N'JADAKA* | 0.751231 |
| RED GUARDIAN II/ALEXEI SHOSTAKOV | 0.750105 |
| TAKU | 0.743949 |
| PRESTER JOHN | 0.742969 |
| STINGER II | 0.741208 |
| *ANT-MAN/DR. HENRY J. PYM* | 0.740225 |
| MASTER PANDEMONIUM/MARTIN PRESTON | 0.731731 |
| BLACK KNIGHT III/EOB | 0.728046 |
| GRIM REAPER/ERIC WILLIAMS | 0.727991 |
| *VISION* | 0.72677 |
| ZURI | 0.726121 |
| WASP/JANET VAN DYNE | 0.725275 |
| AMENHOTEP | 0.723408 |
| LLOIGOROTH | 0.721435 |

# Predicting Future Characters of these Movie




## Gambit (TBD)

| Character | Similarity |
|---|---|
| **PSYLOCKE/ELISABETH BRADDOCK** | 0.884617 |
| CHROME | 0.884347 |
| *ROGUE* | 0.874492 |
| DELGADO | 0.863572 |
| MACTAGGART, JOE | 0.859872 |
| HAZARD/CARTER RYKING | 0.853572 |
| *RASPUTIN, MIKHAIL* | 0.849118 |
| EJULP | 0.847835 |
| *STORM/ORORO MUNROE* | 0.843426 |
| TRION | 0.842097 |
| MARROW/SARAH | 0.841992 |
| MR. SINISTER/NATHAN | 0.838692 |
| SISTER MARIA | 0.82992 |
| BRAIN CELL | 0.82423 |
| WANDERER | 0.821231 |
| MEME | 0.818377 |
| *WOLVERINE/LOGAN* | 0.816266 |
| RYKING, ALEXANDER | 0.813096 |
| PAM | 0.812597 |
| AVATAR | 0.81259 |

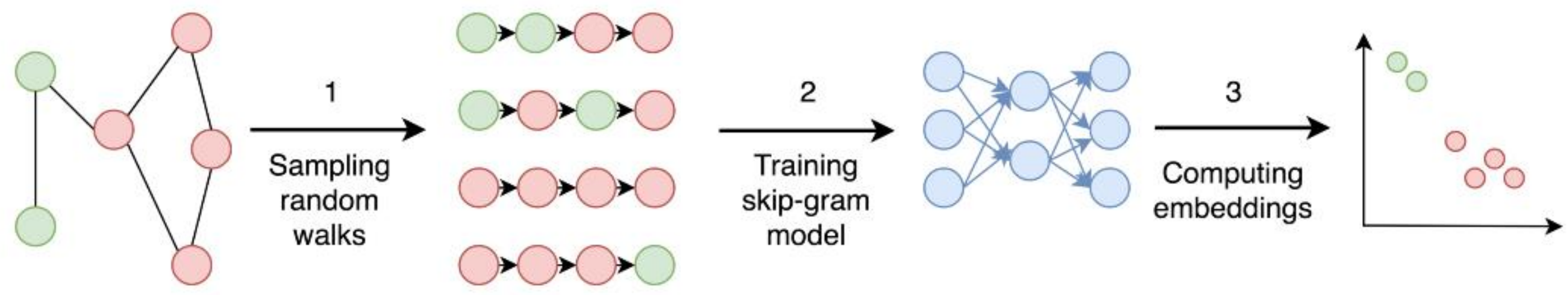

## Doctor Strange 2 (TBD)

| Character | Similarity |
|---|---|
| **CLEA** | 0.967789 |
| *WONG* | 0.959589 |
| CHANG, IMEI | 0.904088 |
| *ANCIENT ONE* | 0.899713 |
| RINTRAH | 0.890948 |
| WOLFE, SARA | 0.885755 |
| BLESSING, MORGANA | 0.859316 |
| *DORMAMMU* | 0.855099 |
| *BARON MORDO/KARL MORDO* | 0.834619 |
| BLACK, CYRUS | 0.828678 |
| ASMODEUS | 0.825324 |
| *AGAMOTTO* | 0.820484 |
| AZRAEL | 0.820051 |
| SHADOW QUEEN/SHIALMAR | 0.817132 |
| UMAR | 0.816197 |
| AGGAMON | 0.816153 |
| DR. STRANGER YET | 0.814923 |
| INTERLOPER | 0.813463 |
| SHUMA-GORATH | 0.81115 |
| APPALLA | 0.810538 |

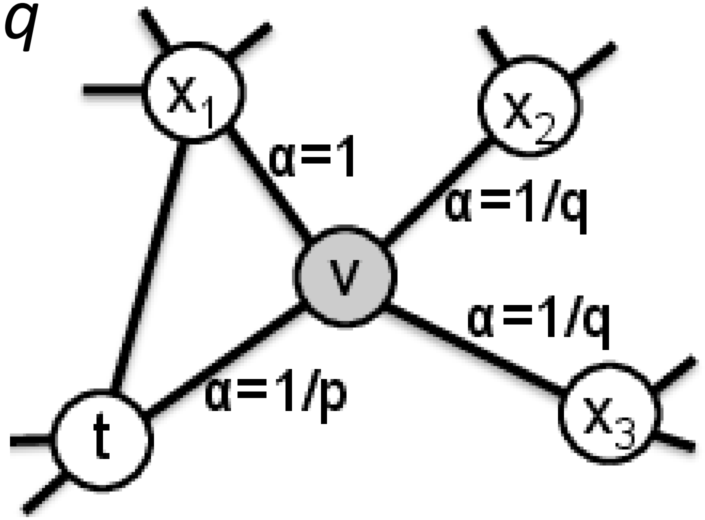
# Node2Vec Package – What is Node2Vec?

- Same principle as word2vec
- But!
  - Text => Linear, directed, contextual
  - Graphs => No natural order
- Can a network graph be made analogous to a corpus of documents?

# Node2Vec Package – What is Node2Vec?




- Select 1 node from the network
- Do *m* random walks from the node with *n* steps in each random walk, governed by parameters *p, q*
  - *p*: returning to the previous node
  - *q*: going to/"discovering" a part of the graph unconnected to the previous node
- Repeat for all nodes



# Node2Vec Package

```
In [5]: # Generate walks
        # will take a while
        node2vec = Node2Vec(g, dimensions=20, walk_length=25, num_walks=125)

Computing transition probabilities: 100%|██████████| 6426/6426 [05:01<00:00, 21.33it/s]
Generating walks (CPU: 1): 100%|██████████| 5/5 [00:06<00:00,  1.34s/it]
```

- Each random walk is analogous to a "sentence"
- All random walks together is analogous to a "corpus"
- Skip-Gram model to determine similarity

# Node2Vec Package

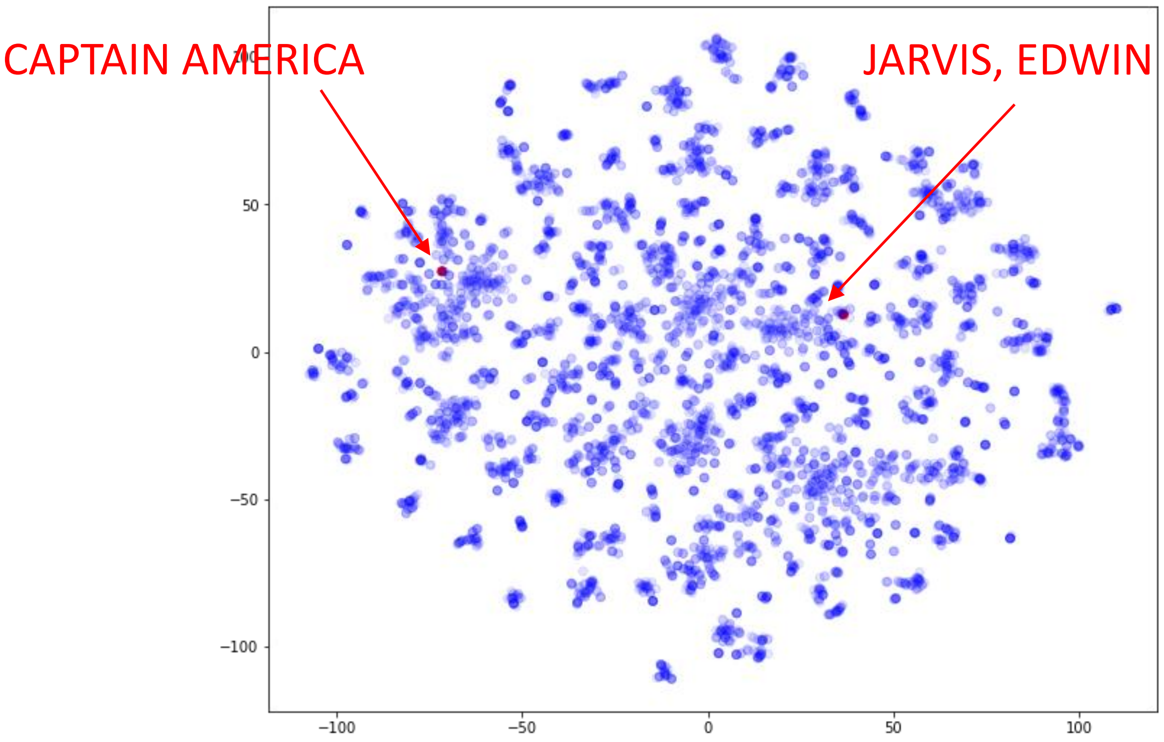
Perform the random "deep walk" based on set parameters

Learn the probabilities & embeddings to form model

Play with the characters in multi-dimension space (similarity etc) and make predictions

```
IDER-MAN/PETER PAR']

38346 ,  3.8081942 , -2.813464  ,  3.8015187 , -0.44214833,
15962 , -0.3458974 ,  2.7286322 , -4.449345  , -2.7225494 ,
778554,  2.3601246 ,  2.4234307 , -2.4428544 , -3.0922093 ,
21474 , -0.9120338 , -2.5853953 , -1.6498612 , -2.6722753 ],
float32)
```

Use TSNE to flatten multi-dimension dataset



# Limitations and Further Work

- Exploration of Node2vec with weighted edges. Experimenting with more parameters eg. Number of dimensions, length of walks
- Include other MCU properties, such as Marvel TV series, Agents of SHIELD, Jessica Jones, Daredevil, Iron Fist, Luke Cage, The Defenders, etc, and other MCU produced One-Shots.
- Inability to mine comics dataset for completely new superheroes

Questions?