



COMPARATIVE ANALYSIS OF
**REDDIT AND
SUBREDDITS**

TEAM M

PRESENTATION OUTLINE

- What's Reddit?
- Data Cleaning and Tools
- Trends for 2014
- Trends for 2016
- Comparison of Metrics
- Metrics using NetworkX



DEFINITION

Reddit is a social news website and forum where content is socially curated and promoted by members through voting.

COMPOSED BY SUBREDDITS OR SUBCOMMUNITIES WITH DIFFERENT TOPICS.

OUR TOOLS

GEPHI

for network and communities visualization, and
having a sense of the trends

NETWORKX

for advanced metrics and analysis

WORDCLOUD

for fancy word plotting

PLOTLY

for advanced plots

RESEARCH QUESTION

- HOW REDDIT IS USED BY THE USERS?
- HOW THE SUBREDDIT TOPICS CHANGED IN YEARS?
- HOW THE NEGATIVE TOPICS HAS CHANGED?

SUBREDDIT DATASET

	2014	2016
Nodes	827	642
Edges	1208	899
Average degree	1,461	1,4
Network diameter	4	3
Graph density	0.0002	0.0002
Modularity	0.548	0.557
Av. Clustering Coeff	0.072	0.039
Av. Path Length	2,047	1,483
Proportion of Pos Sentiment	92%	89%

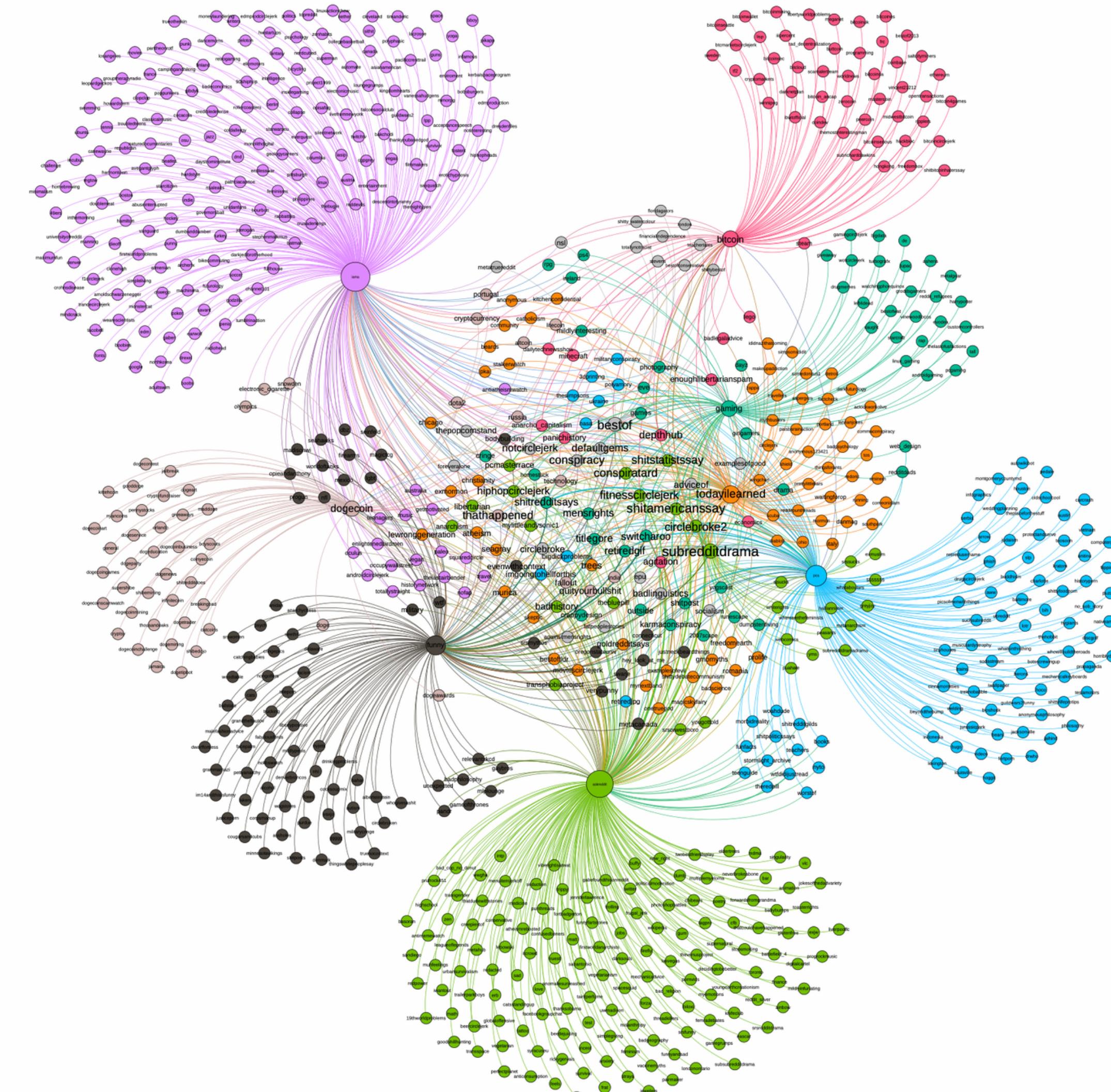
ATTRIBUTES:

- SOURCE
- TARGET
- POST TIME
- SENTIMENT SCORE

DATA PREPARATION:

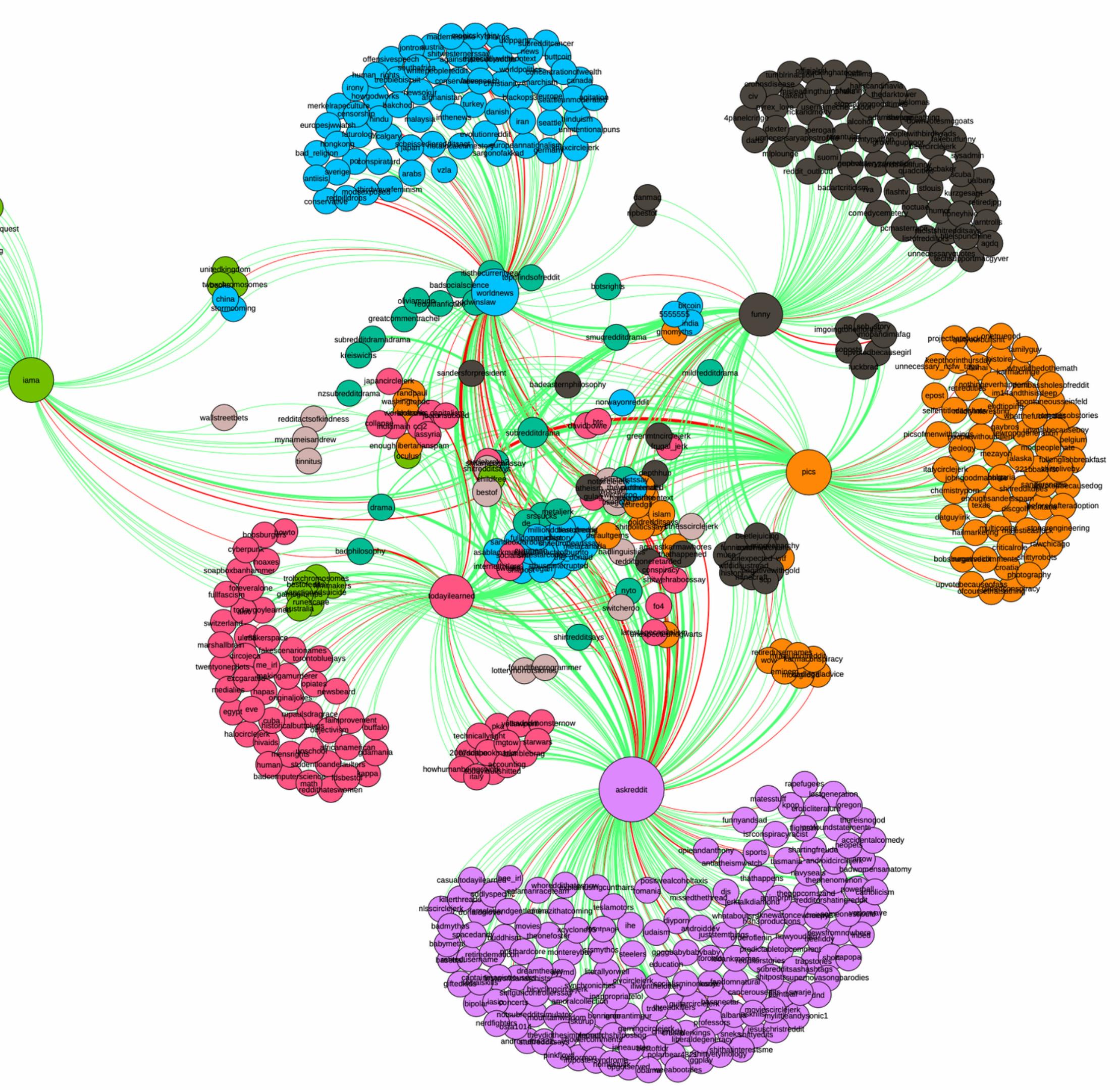
- FILTERING DATASETS PER YEAR
- 10% RANDOM SAMPLE
- QUERYING FOR PAGERANK TARGET > 0.005





TRENDS FOR 2014

- Funny environment
- Place for entertainment
- Early niches (dogecoin!)
- 8% of negative posts
- Some components act like gigant ones

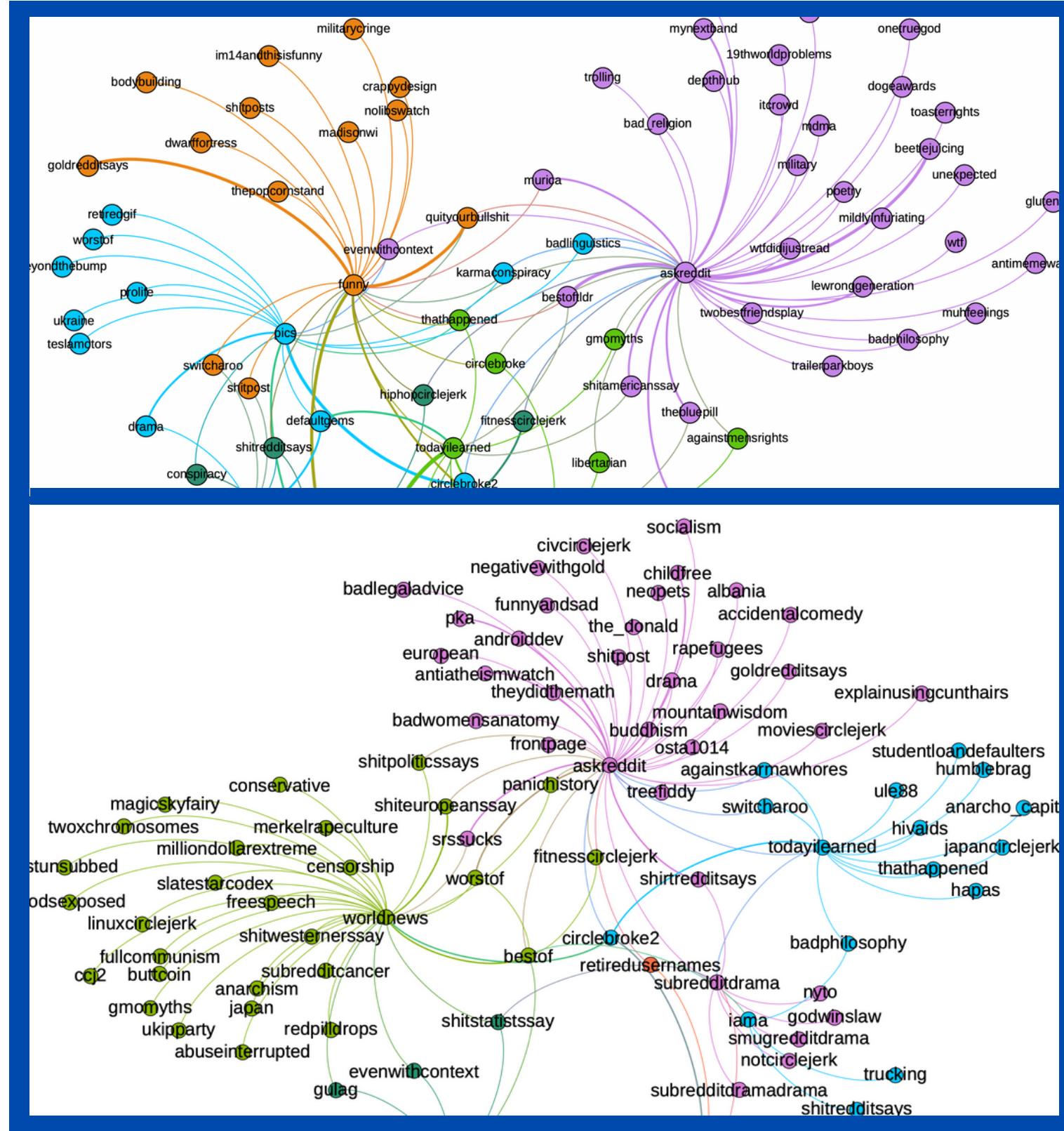


TRENDS FOR 2016

- More mature userbase
 - Place to share ideas
 - Reddits related to NSFW

contents are usually
negative

 - Anarchy



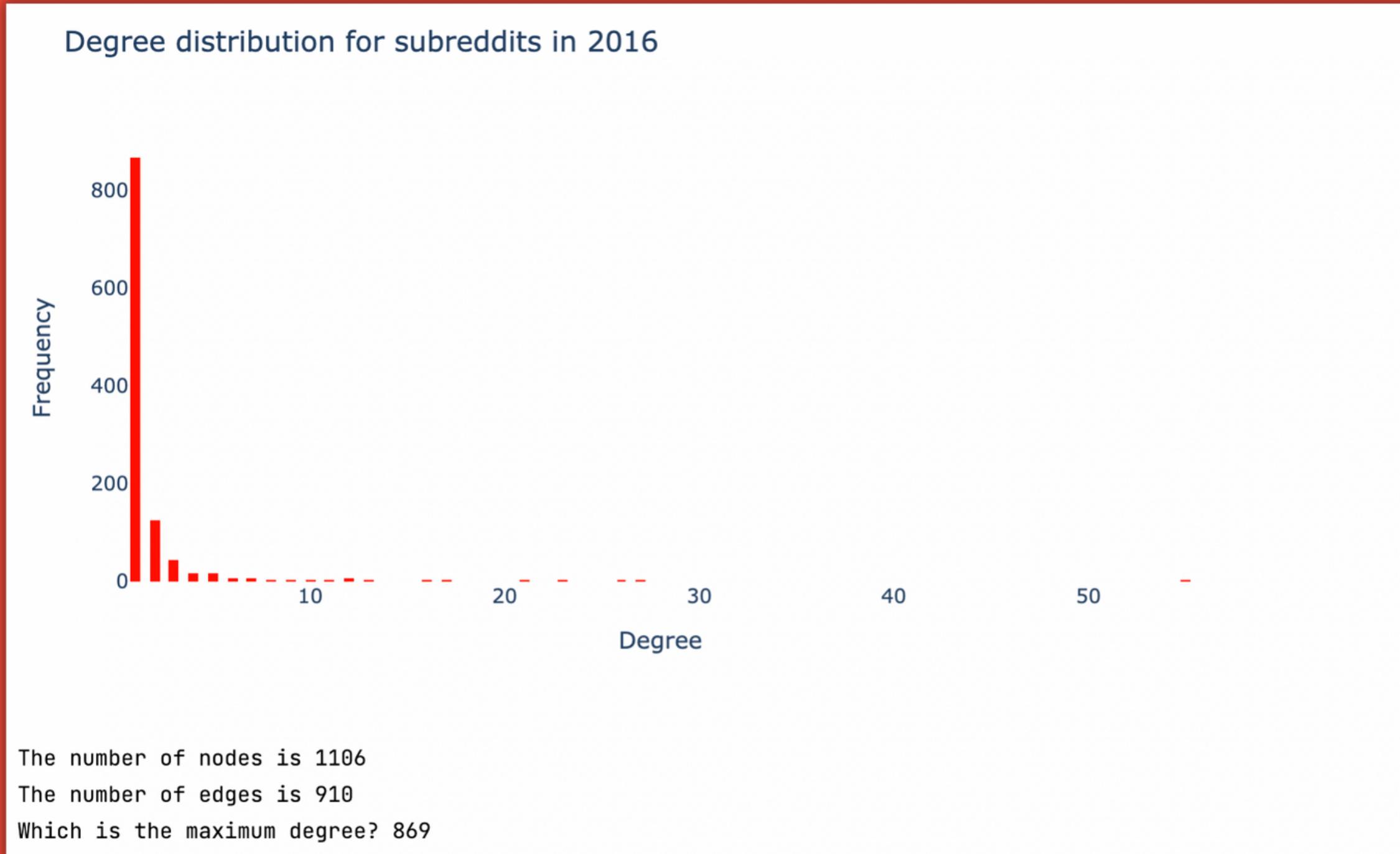
HOW NEGATIVE TRENDS EVOLVED?

Our assumption about the maturation of the userbase was correct: negative posts in 2014 were related to pop corn, bodybuilding and shitpost.

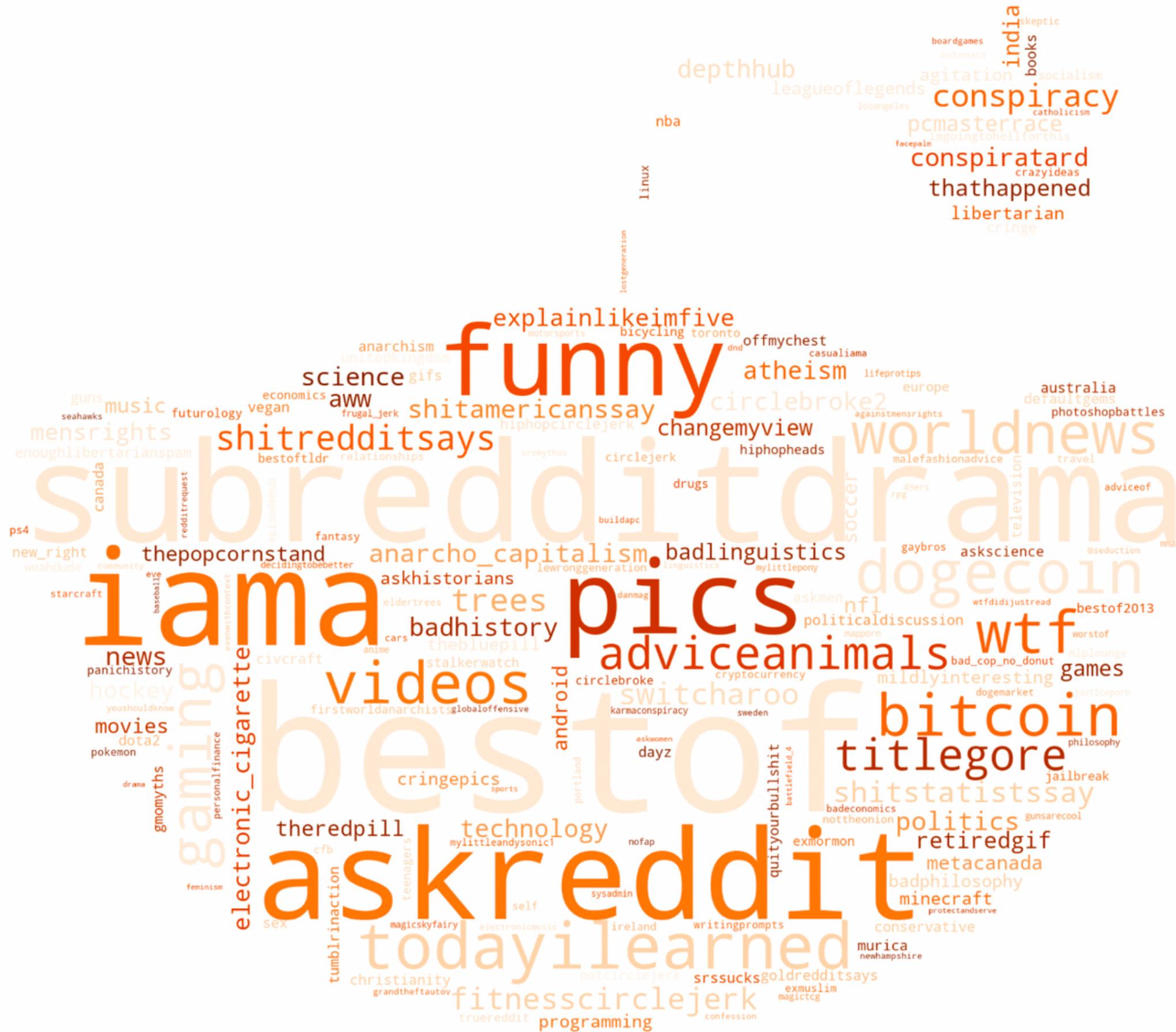
In 2016: socialism, conservative, censorship, buddism,
childefree etc.

Notably, according to Reddit, "Ukraine" was already a negative topic in 2014.

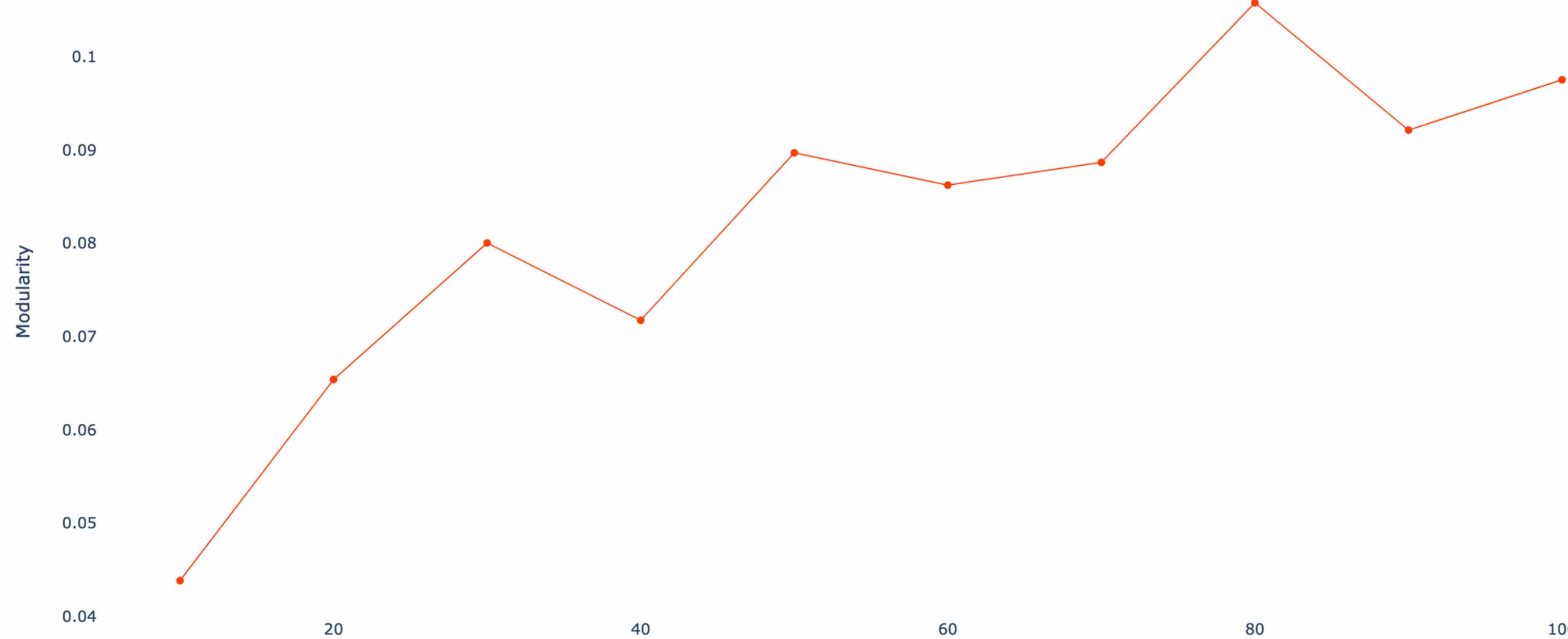
ADVANCED ANALYSIS NETWORKX



Most important subreddit according to their Degree Centrality in 2014



Modularity according to the number of K cluster, using K-Means



FINDINGS

1. Reddit became a place to share thoughts,
earlier was a place for entertaining
2. Reddit is adult oriented and a free platform
3. **Niches** are popular and increasing in
number

