

Measuring Online Feedback Loops

Antone Christianson-Galina

Abstract

My research looks at the struggle between greater simplicity and greater complexity online conversations. I lay out two different feedback loops that shape online conversations, the Simplicity loop and the Complexity Loop. By online conversations, I am referring to the process by which people come together and decide what means what. Are GMOs good or bad? Is the president a sinner or a rogue? Was the film genius or mad? The first feedback loop, which I will call *the Simplicity Loop* has three parts that generate each other. 1) A simple conversation becomes 2) popular and 3) generates a consensus which leads to greater simplification. The second loop, which I will call the *Complexity Loop* also consists of three parts. 1) An intricate and complex conversation 2) breaks into diverging positions 3) generating new ideas and positions. The Complexity Loop generates and recombines ideas, but does not make them popular. I then lay out how to study the two loops using Yule I measure. The measure is a statistical index of linguistic complexity and can be used to study the degree to which online conversations are becoming more simple or complex. I then illustrated the utility of the theory and used the Yule I measure to study conversations on Reddit, an online messaging board.

Theoretical Background

In 1960, Karl Popper posed a question.

“The problem with the validity of empiricism may be roughly put as follows: is observation the ultimate source of our knowledge of nature? And if not, what are the sources of our knowledge? ”(Popper, 1960)

Popper would later go on to argue that what separated the reason of a madman from that of a scientist was that a scientist can test his assertions. (Popper, 1974) But what happens when he cannot? What happens

when we live in a world that changes so fast that by the time papers are written on issues, they are already obsolete?

Ludwig Wittgenstein, in his blue and brown books, argued that we decide what means what through “language games” (Wittgenstein, 1958).

While one may speak english or Spanish or french, Wittgenstein refers to more specific “languages” such as geometry, chemical symbols, and the use of charts and diagrams. However, his examples largely dealt with simplified one-to-one interactions. What happens when hundreds, or even millions of people are speaking at once?

Another german philosopher, Niklas Luhman, created a robust model to understand how groups decide on meaning. He called it ” the social system”. He argued that that the primary goal of a social system is to reduce complexity(Luhmann, 1967). As humans, we do not have the time and ability to process all of the information on the world around us.

For Luhmann, *meaning is the product of the different choices that a system makes to deal with complexity* (Luhmann, 1990). Meanings spread across and through social systems through communication. Social systems can have sub-systems within them and they can be reflexive, modifying themselves (Luhmann & Baecker, 2014). To rephrase, Meaning is the product of the self-regulation that a system undergoes in order to cope with novel complexities: pending the preexisting mechanisms -and, hence, limitations and capabilities - of said system. In other words, meaning is an emergent effect of a system’s autopoetic self-maintenance.

According to Luhmann, systems begin with a minimal ability to reduce complexity, but evolve to be able to deal with more complexity, (Luhmann, 2000). By reproducing itself through communication, the system constitutes and reconstitutes itself (Bausch, 1997).

What I will do in the rest of this paper is use Luhman’s conception of a social system as the foundation of a model that can be used towards understanding how people decide what things mean on the internet. A millennial, online forums, wikis, and blogs, are some of the first places I look when I want to know what something means. This paper will focus specifically on Reddit, an online forum, but hopefully, the model and the tools to use it can be applied to other sites.

I am not the first to apply Luhmann’s social systems theories to digital media. Federico Farini in “Media Theory and Web-Based Groups as Social Systems”, explores how social media fits into Luhmann’s systems theory, and how a desire for social inclusion can limit communication on Facebook (Farini, 2017).

If meaning is the product of different choices that a system makes, how the choices are made becomes the critical point to consider. I now present two different way that these choices might be made- the Simplicity Loop and the Complexity Loop.

Why We Need a New Method

Traditional social psychology and behavioural science experiments have focused on individuals stripped away of context. An example would be Bartlett's work ([Bartlett, 1932](#)) on serial reproduction. In his famous "War of the Ghosts" experiment, he would ask people to read a passage and then copy it without looking. Today, online, we have the cut and paste function. We can "remember" anything we want by taking a picture of it or making a note of it. With our cellphones, we'll always have our memories in our pocket.

More recent work, on heuristics and biases like Kahneman and Tversky's work on Prospect Theory ([Kahneman & Tversky, 1979](#)) or Thaler's work on Behavioural biases ([Kahneman, Knetsch, & Thaler, 1991](#)) describes how individuals process information. Simply put, the idea is that people's biases stop them from looking at objective information rationally.

Outside of laboratories and thought experiments, the notion of "Objective Information" becomes more dubious. What country is Crimea in? Who is the President of Libya? Of Yemen? The answer is -it depends on who you ask.

With different narratives around the world competing on the same internet, who do you trust?

The study of feedback loops to see how people process information online is not a study in objective truth- it is a study in subjective truth, why people believe what they believe. In this study, I am following in the work of the Father of Sociology, Emile Durkheim. He studied "the collective consciousness" and "conscious effervescence" over a hundred years ago. The study of how the group thinks has been continued by Pareto (Social Sentiment), Le Bon (The Madness of the Crowds), Moscovici (Social Representations), and others. However, instead of conducting hundreds of interviews, the social media age has produced mountains of data to analyse. Therefore, a robust tool to study how online conversations change is more valuable and useful than ever.

The Simplicity Loop

The Simplicity Loop arises when ideas are judged by popularity. A simple idea becomes popular, and generates a consensus which leads to further simplification. Why a simple idea? For a message to spread easily, it must be short and filled with redundancy (Shannon, 1948). For example, this tweet, “broken. from the bottom of my heart, i am so so sorry. i don’t have words” spread to 2.7 million people. In it, singer Ariana Grande post repeats a message of remorse in each short sentence. Meanwhile, intricate, meaning-rich posts such as this documentary on systems theory <https://archive.org/details/Luhmann1973>, remain obscure. People are lazy and when given a choice between simplicity and complexity will tend towards simplicity (Fiske, 2013). Simple spreads better. Simple becomes more popular.

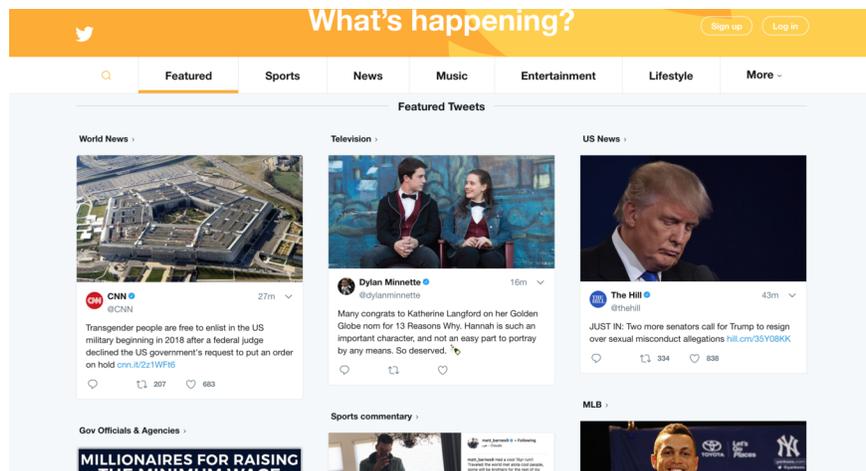


Figure 1: <https://twitter.com/?lang=en>

A simple, popular idea can lead to less new ideas through a process called cascading. An information cascade is when someone observes the actions of others and then engages in similar actions or takes up similar positions. This can spread even when people have information that signals that they should take other actions or positions. An information cascade occurs when people make decisions sequentially rather than simultaneously. Later people watch the actions of earlier people, and base their actions on the earlier ones. An example would be someone who searches for a new song by looking for the most popular songs. The cascade can be based on little information and people could ignore new information. This cascade can easily be based on faulty information, and the credibility of crowds could quickly surpass that provided by traditional experts and authorities. Yet, cascades are fragile and can be overturned by people with more compelling information causing a counter-cascade.

If you use Luhmann’s definition of meaning ”*meaning is the product of the different choices that a system makes to deal with complexity*” the simpler message holds less choices and is less able to accurately deal with complexity. Thus Feedback loops based driven by simplicity are poor at dealing with the complexity of the world we live in, instead painting a simpler, compelling picture.

One of the sites where this loop dominates is Facebook. This is no accident, it is a product of Facebook’s design.

According to its mission statement “. . . Facebook’s mission is to give people the power to share and make the world more open and connected. . .” In this aim, Facebook’s engineers manage people’s home page feed (newsfeed) so that voters see things they like and stay longer. For example, if a voter is interested in politics will show posts by their friends that support their political views but repress those that would clash (Keegan, 2016). Why is this a problem? By only showing views that users are likely to agree with, they are sheltered from new ideas. They end up trapped in an echo chamber listening only to the similar.

The Complexity Loop

Not all people are conformists. Punks, goths, dissidents, freethinkers. People alienated by simplicity loops, or who find some of their assumptions dodgy, often come together to create feedback loops on their own. Although less popular than sites that follow simplicity loops, there exists a thriving ecosystem of sources of alternative ideas online. And these sites are petri dishes for new ideas.

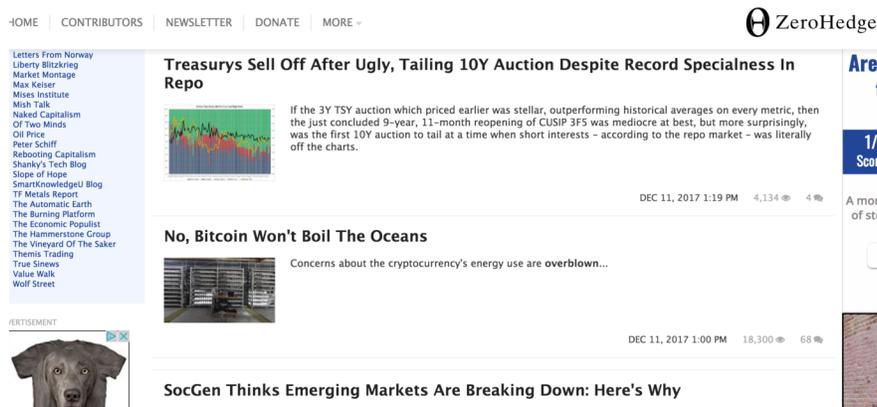


Figure 2: <http://www.zerohedge.com/>

The Philosophy of Rick & Morty: Szechuan Ricksauce

leave a comment »



Recent events with McDonald's and Mulan Szechuan sauce have steered me down the philosophical path and I feel that a post on these events might be in order. I myself have been feeling like Rick Sanchez C-137 lately and I think a lot of that has to do with the state of the world today. Another reason is really kind of encapsulated in Rick's comments to Beth in SE03E09 of The Adventures of Rick & Morty. To wit my "I drink and know things, because I know things I drink" moment.

Beth: - *Am I evil?*

Rick: - *Worse. - You're smart. - When you know nothing matters, the universe is yours. - And I've never met a universe that was into it. - The universe is basically an animal. - It grazes on the ordinary. - It creates infinite idiots just to eat them, not unlike your friend Timmy.*

Beth: - *Tommy.*

Blog Stats

951,452 hits

October 2017

M	T	W	T	F	S	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

« Sep Nov »

Categories

Select Category

Blogroll

[America In Context](#)
[Cleartext](#)
[Cognitive Dissidents](#)
[cub3tcrim3](#)
[Cyber Strategies for a World at War](#)
[eSecurity Intelligence](#)
[Galliumfree](#)
[IThreats](#)
[mediapsychology101](#)
[Michael555x](#)
[Red Teams](#)
[Scam Of The Day Central](#)
[Scot A. Terban LinkedIn](#)
[Scot Terban at Smugmug](#)
[Sunfell](#)
[Tcrweb's Blog](#)
[Team Cymru](#)

Figure 3: <https://krypt3ia.wordpress.com/>

Arguably, the Complexity Loop preceded the Simplicity Loop in digital preeminence. There was a time when the internet was a rebellious, countercultural place (Dana Boyd, 2017). However, times have changed. Hossein Derakshan, an activist blogger who spent 6 years in an Iranian prison after blogger-fuelled protests against the government tells the story poignantly:

Blogs gave form to that spirit of decentralization: They were *windows* into lives you'd rarely know much about; *bridges* that connected different lives to each other and thereby changed them. Blogs were *cafes* where people exchanged diverse ideas on any and every topic you could possibly be interested in. They were Tehran's taxicabs writ large.

Since I got out of jail, though, I've realized how much the hyperlink has been devalued, almost made obsolete.

Nearly every social network now treats a link as just the same as it treats any other object — the same as a photo, or a piece of text — instead of seeing it as a way to make that text richer. You're encouraged to post one single hyperlink and expose it to a quasi-democratic process of liking and plussing and hearting: Adding several links to a piece of text is usually not allowed. Hyperlinks are objectivized, isolated, stripped of their powers. ([Derakhshan, 2015](#))

A hyperlink allows people to spread ideas irrespective of their popularity, avoiding the dangers of the Simplicity Loop. In the Complexity Loops, intricate and complex conversations spew new ideas and criticism. However, the intricacy and diversity of discussions in this feedback loop prevents them from growing more popular

How to Measure the Differences Between Feedback Loops

How could you tell a Simplicity Loop from an Complexity Loop?

My solution: by looking at the increase or decrease of linguistic richness.

Measuring linguistic richness has been called the Gordian Knot of literary studies ([Miranda-García & Calle-Martín, 2005](#)). I explored many different ways of measuring linguistic richness on online text., Hapax Legomena, the Type-Token ratio, Adjectives and Adverbs, and Yule's I characteristic. I chose Yule's I characteristic because the previous three measures suffered from two key problems, jargon sensitivity and length dependence. Online comments vary tremendously in length and often feature thread-specific jargon and non-standard grammar. The variable length made Hapax Legomena and the Type Token ratio ineffective tools to measure richness (and from there meaning). A study of Adjectives and Adverbs quickly proved

futile as the threads used non-standard grammar.

Yule's I Measure of Richness

Udny Yule's I characteristic comes from his 1944 work, "A Statistical Study of Literary Vocabulary" (Yule, 1944) called the result of this formula Characteristic I. Basically, it measures the ratio of total words and unique words in a text while at the same time being length dependent. The phrase CAT CAT CAT would have a lower score than the phrase CAT CAT DOG even though both had the same amount of words. What separates Yule's I score from a simple type-token ratio is the ability to deal with texts of varying lengths. This is achieved by the use of the following formula:

$$\frac{(M1 \cdot M1)}{M2 - M1}$$

Where M1 is the total number of words (in the above examples both would be 3), and M2 is the sum of the product of the number of words at each frequency and the frequency squared. For example, for the post "Dogs eat cats. Cats eat pigs. Cats eat cats" there are 4 instances of the word "cats", three instances of the word "eat", and one instance of the words "dogs" and "pigs". Therefore M2 would be: $1 \cdot 4^2 + 1 \cdot 3^2 + 2 \cdot 1^2 = 27$

CAT CAT CAT would have Yule's I score of $\frac{(3 \cdot 3)}{(9 - 3)} = 1.5$

CAT CAT DOG would have a Yule's I score of $\frac{(3 \cdot 3)}{(5 - 3)} = 4.5$

"Dogs eat cats. Cats eat pigs. Cats eat cats" would also have a Yule's I score of $\frac{(9 \cdot 9)}{(27 - 9)} = 4.5$

This formula should roughly hold true no matter the size of the sample (Yule, 1944), unlike Hapax measures of type token measures. Furthermore, coding this measure in Python was straightforward and had been done before by web engineer Swizek Teller (Teller, 2017).

Returning to Luhmann's definition of meaning- 'the product of the different choices that a system makes to deal with complexity'- How does Yule's I score fit in?

Like the type token ratio, the higher the score, the more information and different choices that a post could hold.

For example, "Take a left turn" has a higher score than "Cat Cat Cat Cat," and has more information. To further develop this connection, I will use a concept from Chaos Theory, Kolmogorov Complexity. Laid out

by Soviet Mathematician Andrey Kolmogorov in “Three Approaches to the Quantitative Definition of Information” (Kolmogorov, 1968). Kolmogorov Complexity is the length of the shortest description it would take to produce an output. For example, the string “aaa” (50 Characters Long) can be described by the description “a 50 times” (8 Characters). However, the string “esfjbkskjbgslfdnapsdkngirngvlsjrnvsjadb” can only be completely described as “esfjbkskjbgslfdnapsdkngirngvlsjrnvsjadb”. even though it has less characters than the a string, it has a higher Kolmogorov Complexity.

Posts with higher Yule I scores have higher Kolmogorov complexities, as there are less repeated words. The higher the Yule I score, the higher the Kolmogorov complexity. The more complex a comment, the more potential choices, and therefore meaning it can store. Therefore, Yule’s I score measures the possible meaning in a comment.

By looking at the Yule’s I score of each comment, we can fit a line to the increase or decrease of Yule’s I over time. We can then compare the rate of Yule’s I increase/decrease with the number of comments on a post, thread or community. By using rate of increase of a measure instead of average Yule I score, we ignore the line intercept. This allows for a standardised measure across all categories, as each categories might have different baseline mean Yule I scores shaped by what the commenters knew before entering the thread.

Therefore, we can see which feedback loop is at work on an online conversation.

A Simplicity Loop is characterised by a decreasing Yule Score while a Complexity Loop is characterised by an increasing Yule I score.

An Example Case: Reddit

I decided to test this theory and method on Reddit, looking at polemic science threads. My hypothesis: Polemic Science issues of Reddit should be sites of

Reddit is an online forum divided into subreddits, threads, and comments. Each subreddit has its own unique rules and moderators and is built around set themes. As of June 2017, the most popular subreddits are AskReddit, Funny, TodayILearned, and Science according to <http://redditlist.com>. Each subreddit has different threads, which can be a question, a statement, or even a link to an article or video. People comment on these threads and then comment on those comments. They can then comment on those comments-

commenting indefinitely. Furthermore, they can upvote or downvote both comments and threads.

I wanted to explore Reddit because it has elements of both the Simplicity Loop and the Complexity Loop. Reddit is a place where fringe ideas can thrive and new ideas arise. However, there is a process of liking that could allow for simple ideas to push out more complex ideas.

Why Polemic Science Issues?

More scientific papers than ever are being published ([Larsen & von Ins, 2017](#)), and in cases of public interest, often contradict each other. This preponderance of conflicting information means that there are too many studies for a layman to easily judge. Thus, according to Luhman's theory, we rely on social systems to process complexity and give us a simpler picture of the world. By looking at polemical science debates, we can analyse a dynamic and changing system. For these reasons, I chose to look at science issues.

The categories I chose, Artificial Intelligence, Global Warming, Genetically Modified Organisms, the CRISPR gene editing tool, and the debate over Vaccines were categories that interested me. I have been building a crude Artificial Intelligence throughout the year and was introduced to debates over Artificial Intelligence. The debate over Global Warming and Vaccines has come to the forefront in my home country, the United States, after a recent election and seemed pertinent. An interest in the debate over Genetically Modified Organisms and CRISPR stems from growing up in a family of geneticists.

Instruments

To collect and analyze Reddit comments, I wrote a script in the Python computing language. Reddit, which is open source, offers a service called PRAW written in Python, which allows for the easy mining of threads and comments. I, therefore, used Python for this project. Another computer language- R, provides stronger statistical analysis tools, and a more straightforward implementation of Yule's I score, but seeing that I could measure it on Python and also use the PRAW kit, I decided to use Python when writing my program "SuperYule".

```
##### Here I import the necessary packages
import praw
import numpy as np
import matplotlib.pyplot as plt
```

```

import nltk
import pylab
from scipy import stats
from nltk.stem.porter import PorterStemmer
from itertools import groupby
import sys
import os

##### Here I define Yule's I indicator, using code from the blog of Swizec Teller.
# taken from: http://swizec.com/blog/measuring-vocabulary-richness-with-python/swizec/2528
def words(entry):
    return filter(lambda w: len(w) > 0,
                  [w.strip("0123456789!:,.?(){}[]") for w in entry.split()])
# take from: http://swizec.com/blog/measuring-vocabulary-richness-with-python/swizec/2528
def yule(entry):
    d = {}
    stemmer = PorterStemmer()
    for w in words(entry):
        w = stemmer.stem(w).lower()
        try:
            d[w] += 1
        except KeyError:
            d[w] = 1

    M1 = float(len(d))
    M2 = sum([len(list(g))*(freq**2) for freq,g in groupby(sorted(d.values()))])

    try:
        return (M1*M1)/(M2-M1)
    except ZeroDivisionError:
        return 0

```

```
slopelist= []
```

```
##### Here, I use PRAW, a python package used to pull Reddit threads using the official Reddit applica
```

```
while True:
```

```
    person = input('Enter code: ')
```

```
    reddit = praw.Reddit(client_id='QiF3f5b59NCviQ',
```

```
                        client_secret='hdn8vB1obETJkFC0xumHv1jDEWU',
```

```
                        user_agent='Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_4) AppleWebKit/603.1.1
```

```
    ltime =[]
```

```
    lyulei = []
```

```
    submission = reddit.submission(id=person)
```

```
    submission.comment_sort = 'old'
```

```
    submission.comments.replace_more(limit=0)
```

```
    for comment in submission.comments.list():
```

```
        sentences = nltk.tokenize.sent_tokenize(comment.body)
```

```
        time = comment.created
```

```
        ltime.append(comment.created_utc)
```

```
        text= str(comment.body)
```

```
        gettingyules = yule(text)
```

```
        lyulei.append(gettingyules)
```

```
#### Here, I use the arrays generated from the lists to fit  
a regression line for my points with the Yule score on the Y axis and  
time on the X Axis. Computers generally use Epoch Time, the number of  
seconds since January 1, 1970, to calculate time. It is linear and  
avoids time zones, making it easier to use in a program. By dividing  
this number by 86400, I used days rather than seconds as a time
```

interval, giving me a less unwieldy number.

```
    atime = np.asarray(ltime)
    ay = np.asarray(lyulei)

    x = (atime)/86400
    y = (ay)

    finiteYmask = np.isfinite(y)
    yclean = y[finiteYmask]
    xclean = x[finiteYmask]

    x = xclean
    y = yclean

    slope, intercept, r_value, p_value, std_err = stats.linregress(x,y)
    print ("slope",slope)
    slopelist.append(slope)
    saver=(person + ".png")
    predict_y = intercept + slope * x

#### Here, I create a scatterplot with the regression line, and the points with the Yule score on the
    pylab.plot(x, y, 'o')
    pylab.plot(x, predict_y, 'k-')
    pylab.xlabel("Time in Days")
    pylab.ylabel("Yule's I")
    plt.savefig(saver)
    plt.clf()

    continue
```

Method of analysis:

On Microsoft EXCEL, I created a column with the slope of the line fitting the Yule I scores for each of the 50 threads. I then created a column with the number of comments for each thread. The number of comments

can be seen on the front page of the thread, and I had gathered it for each thread during earlier analysis.

I then used Excel's in-built Regression tool with the slope of Yule I increase or decrease of each post as the X range and the number of comments as the Y range.

“SuperYule” Results

The Mean Yule I level was 17.01. The Yule I values ranged from -197.32 to 356.80, a range of 554.12.

The Average number of comments was 1719.4 The number of comments values ranged from -90 to 9636, a range of 9546. The debates over Global Warming and Genetically Modified Organisms had the highest increase in Yule's I score as the threads evolved, with an increase rate of 39.78 and 21.11 per day. Both are topics of widespread debate, and the threads likely drew in many people with diverse views. This brought in more past decisions and interpretations, leading to a meaning-rich social system. While Artificial intelligence threads had slightly more comments per thread than GMO threads, the GMO threads had a 54% higher increase rate. I do not know the reasons for this.

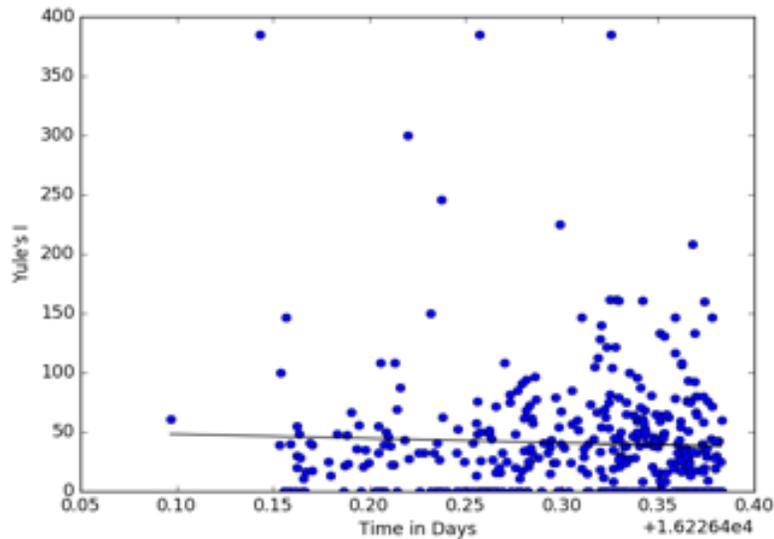


Figure 4: A map of Yule I scores on thread- Richard Tol accidentally confirms the global warming consensus.

The category with the least amount of comments per thread, CRISPR, had the lowest Yule I increase over time. The category with the second least amount of comments per thread, Vaccines, had the second least

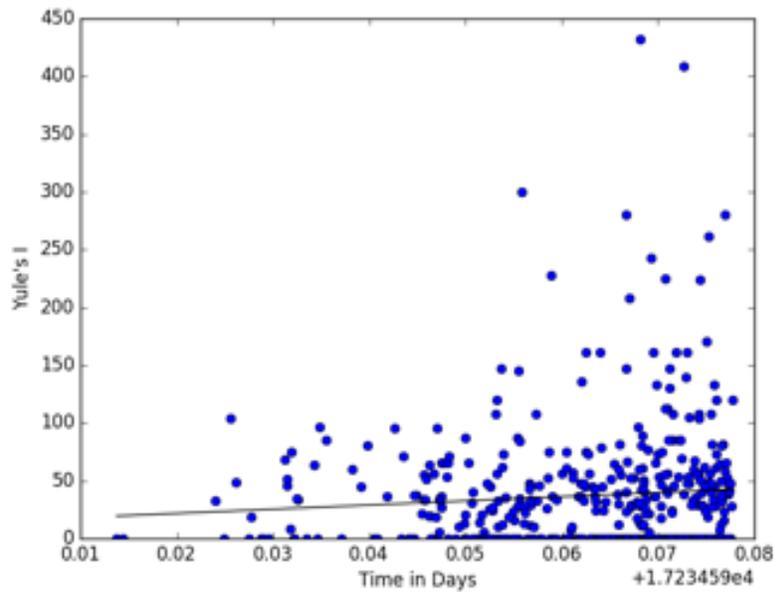


Figure 5: A map of Yule I scores on thread- EPA chief Scott Pruitt says carbon dioxide is not a primary contributor to global warming

increase. However, Vaccine threads had almost 300% higher comments per thread than CRISPR but only a 25% higher score increase rate. I do not know the reasons for this.

When not looking at aggregates, there was a large range of both comment levels and Yule I score levels. For one of the Vaccine threads, the Yule I score decreased by a rate of 197.32 per day while another thread in the same category showed a 101.47 increase per day. The thread with the most comments and highest Yule I score, “EPA chief Scott Pruitt says carbon dioxide is not a primary contributor to global warming” had a whopping 9636 comments and an average daily Yule I increase of 356.80. The link between comments and Yule I rates is visible after a cursory glance of the results table (below), but to find out these results are significant, I ran a significance test. The p-value for the null hypothesis that the level of commenters would not be associated with the rate of Yule I increase is less than .01 ($p < 0.00635$). As this is below all standard levels of significance, we reject the null hypothesis in favour of the alternative hypothesis that there is an association between comments and a Yule I increase. The result of the experiment produces this premise:

There is a positive relationship between the number of comments in a thread and the rate at which the Yule's I characteristic increases.

We can also use this equation to draw a line predicting the meaning generated in a thread using the coefficient

of the X Variable1 (0.02) and the intercept (-9.09).

$$Yule's\ I\ Score = -9.09 + Comments \times 0.02$$

This means what with each extra comment in the threads studied, the Yule's I score increased by 0.02.

Why?

- Strength in diversity: diverse ideas bounce off each other and generate new diverse ideas.
- The anonymity of crowds: because Reddit emphasizes the text of the comment rather than the user, users might feel more confident saying their ideas in public than under a smaller thread with more scrutiny.
- All the simple thoughts were taken: people are pushed further and further off the beaten path to generate more meaning rich posts because all of the simplest ideas have already been said.

Therefore, the Polemic Science issues on Reddit studied were sites of the Complexity Loop. Consensus does not arise, but this is a rich environment for the generation of new ideas.

Conclusion

Yule's I indicator provides a powerful, automated method to analyse online conversations. Using the Simplicity Loop and the Complexity Loop one can determine if new ideas are being developed or if a coherent, hegemonic narrative is developing. Using Luhmann's definition of meaning, *the product of the different choices that a system makes to deal with complexity*, this tool allows us to see if those choices are accumulating or simplifying.

Looking at Polemic science issues in reddit, I found that many new ideas were generated- issues became more complex. However, this was just a pilot study opening up a new area of study. Yule's I and feedback loops could allow for analysis of more threads- even the entirety of Reddit. Furthermore, the tool would be far more useful if the Yule I scores were compared across different websites. I hope, in the future, to research the feedback loops at work on social media sites like Twitter, VK and Facebook, and also online news sources like The Huffington Post, Business Insider, and Breitbart.

Feedback loops are a useful tool to understand how people decide what things mean online, and I hope that this paper opens the door to more research.

Category	Reddit ID	Yule I Rate	Comments	Subreddit
GLOBAL WARMING	5hwhlp	-20.22	1562	WORLDNEWS
GLOBAL WARMING	1umk34	55.07	2084	The NEW REDDIT JOURNAL OF SCIENCE
GLOBAL WARMING	5yfc4y	356.8	9636	news
GLOBAL WARMING	27di3k	-34.59	1495	The NEW REDDIT JOURNAL OF SCIENCE
GLOBAL WARMING	616x8p	15.39	363	TumblrComedyGold
GLOBAL WARMING	5e2gq7	3.03	671	Art
GLOBAL WARMING	4ktef6	-0.04	492	Politics
GLOBAL WARMING	3a3s9y	-37.07	2953	WORLDNEWS
GLOBAL WARMING	1nq1kt	36.37	718	Politics
GLOBAL WARMING	5m11tu	23.12	1408	The NEW REDDIT JOURNAL OF SCIENCE
GMO	62e4cm	17.53	4149	Videos
GMO	3gcwgh	-10.2	2818	TumblrComedyGold
GMO	3ggz6z	-0.56	254	UNEXPECTED
GMO	2j0gis	0.31	464	CRINGE
GMO	2vdrk1	96.89	6051	food
GMO	4qzxq1	-20.94	1268	Politics
GMO	34thm6	65.12	2448	Advice Animals
GMO	2af33s	-8.52	1460	Politics
GMO	4p16rf	1.76	443	GRINGEPICS
GMO	31kqoq	69.69	701	TumblrComedyGold
Vaccine	57mfgu	38.81	3237	WORLDNEWS
Vaccine	3lahv8	-197.32	3527	news
Vaccine	2lou8v	8.88	576	TIL
Vaccine	3lahwa	33.54	835	WORLDNEWS
Vaccine	31akrv	4.09	193	The NEW REDDIT JOURNAL OF SCIENCE
Vaccine	5mlz2h	101.47	6047	news
Vaccine	5oxpqt	27.88	708	UNEXPECTED
Vaccine	1z35k1	-0.57	438	VACCINES
Vaccine	48m2wt	0.97	377	DATA IS BEAUTIFUL
Vaccine	5066yp	57.58	1924	WORLDNEWS
AI	5g3ezx	-58.68	4048	Futurology
AI	4brgvm	-80.65	3911	Technology
AI	5xdou5	33.92	1225	TIL
AI	5jyy0k	26.82	3336	Futurology
AI	5ay137	59.94	2628	Futurology
AI	45l03x	-55.81	3298	Technology
AI	59thrj	-50.66	1274	Futurology
AI	6cfm24	1.59	486	Android
AI	61rmzt	-2.28	760	Technology
AI	4j5q9d	261.66	1235	news
CRISPR	4fpvqv	58.32	63418	The NEW REDDIT JOURNAL OF SCIENCE
CRISPR	5rd5f1	12.39	527	The NEW REDDIT JOURNAL OF SCIENCE
CRISPR	5ymfbu	-1.6	350	The NEW REDDIT JOURNAL OF SCIENCE
CRISPR	6cral1	-34.12	1953	Futurology

SUMMARY
OUTPUT

Regression Statistics	
Multiple R	0.38
R Square	0.15
Adjusted R Square	0.13
Standard Error	71.87
Observations	50

ANOVA

	df	SS	MS	F	Significance F
Regression	1	42075.39	42075.39	8.15	0.01
Residual	48	247924.15	5165.09		
Total	49	289999.54			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-9.09	13.81	-0.66	0.51	-36.86	18.68	-36.86	18.68
X Variable 1	0.02	0.01	2.85	0.01	0.0	0.03	0.0	0.03

Table 2: Regression 1 with the slope of Yule I increase or decrease of each post as the Y range and the number of comments as the X range

References

- Bartlett, F. (1932). *Remembering*. Cambridge.
- Bausch, K. C. (1997). The Habermas/Luhmann Debate and Subsequent Habermasian Perspectives on Systems Theory. In *Systems research and behavioral science 14* (p. 315-339). Wiley-Blackwell. doi: 10.1002/(sici)1099-1743(199709/10)14:5<315::aid-sres173>3.3.co;2-q
- Dana Boyd, E. A. (2017). *The kids are (still) alright*. Retrieved from <http://news.microsoft.com/stories/people/danah-boyd.html> (Accessed on Mon, December 11, 2017)
- Derakhshan, H. (2015). *The Web We Have to Save*. <https://medium.com/matter/the-web-we-have-to-save-2eb1fe15a426>. Retrieved from <https://medium.com/matter/the-web-we-have-to-save-2eb1fe15a426> (Accessed on Mon, December 11, 2017)
- Farini, F. (2017). *Media theory and web-based groups as social systems*. Mapping Systemic Knowledge: In G. Mancini & M. Angrisani.
- Fiske, S. (2013). *Social Cognition*. SAGE Publications Ltd. Retrieved from <https://doi.org/10.4135/2F9781446286395> doi: 10.4135/9781446286395
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991, feb). Anomalies: The Endowment Effect Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives*, 5(1), 193–206. Retrieved from <https://doi.org/10.1257/2Fjep.5.1.193> doi: 10.1257/jep.5.1.193
- Kahneman, D., & Tversky, A. (1979, mar). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263. Retrieved from <https://doi.org/10.2307/2F1914185> doi: 10.2307/1914185
- Keegan, J. (2016, August). Red Feed Blue Feed. *Wall Street Journal*, 9, 2017. Retrieved from <https://graphics.wsj.com/blue-feed-red-feed/>
- Kolmogorov, A. N. (1968). Three Approaches to the Quantitative Definition of Information. *International Journal of Computer Mathematics*, 2, 157-68. doi: 10.1080/00207166808803030
- Larsen, P., & von Ins, M. (2017, August). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Retrieved*, 9.
- Luhmann, N. (1967). Soziologische Aufklärung. *Soziale Welt*, 18.
- Luhmann, N. (1990). *Meaning as sociology's basic concept*. In Essays on self-reference. New York: Colombia University Press.
- Luhmann, N. (2000). *Art as a social system*. Stanford, Calif.: Stanford University Press.
- Luhmann, N., & Baecker, D. (2014). *Introduction to Systems Theory*. Polity.

- Miranda-García, A., & Calle-Martín, J. (2005, dec). Yule's Characteristic K Revisited. *Language Resources and Evaluation*, 39(4), 287–294. Retrieved from <https://doi.org/10.1007/s10579-005-8622-8>
doi: 10.1007/s10579-005-8622-8
- Popper, K. (1960). Knowledge without Authority.
- Popper, K. (1974). The Problem of Induction.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell Systems Technical Journal*, 27.
- Teller, S. (2017, August). Measuring Vocabulary Richness with Python. *A Geek With a Hat*, 9, 2017. Retrieved from <https://swizec.com/blog/measuring-vocabulary-richness-with-python/swizec/2528>
- Wittgenstein, L. (1958). *The Blue and Brown Books*. Harper and Row.
- Yule, G. U. (1944). *The Statistical Study of Literary Vocabulary*. Cambridge University Press.