

SCREEN SCRYING

ALI RAZZAK

BRANCHES OF KNOWLEDGE

1

ABSTRACT

International affairs were thrown into chaos and global operations halted in the face of COVID-19. While times of crisis cast fear into an uncertain future they also act as reminders of past conquered hardships. Pathogens have devastated entire civilisations and put humans at the brink of extinction throughout history. Despite innumerable scientific advances the arms race between human and pathogen is as competitive as ever, highlighting the sophisticated processes that govern microorganism behaviour. Scientists have erected databases of molecular information to seek an understanding of the networks that animate pathogens. This has attracted the employment of computational tools to glean insight from that data, conjuring landscapes of projection maps. However the intelligence that underlies machine algorithms is often nebulous and can obfuscate meaning instead of elucidating it. Nevertheless, developments of the COVID-19 pandemic entices examina-

tion of budding pathogenic catastrophes and what knowledge exists to address them. In hopes of envisioning life after a crisis endeavours to trawl these libraries of research can employ tools which manifest a logic unto themselves. My project examines how using data driven techniques to address biomolecular catastrophes can recapitulate hysteria instead of resolving it.

INFORMATION

Ali

www.getgot.space

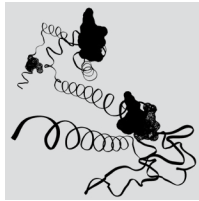
***University of Basel
Structural Bioinformatics***

OVERVIEW OF YOUR 45 VISUALS

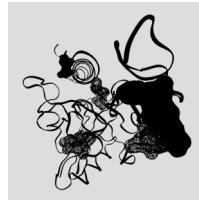
2



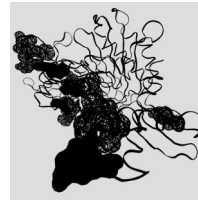
Enterobacter spp.



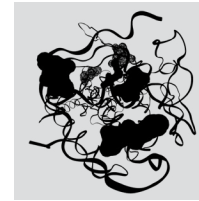
Corona virus



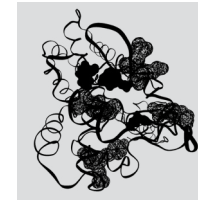
A. cholera



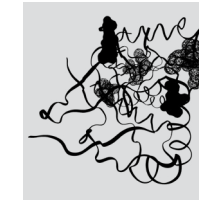
S. pneumoniae



Rabies virus



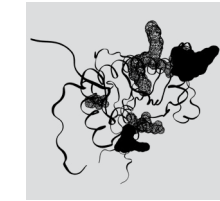
Variola virus



B. pseudomallei



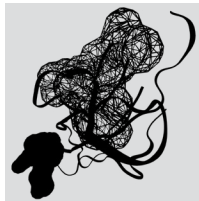
F. tularensis



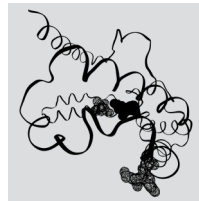
N. gonorrhoeae



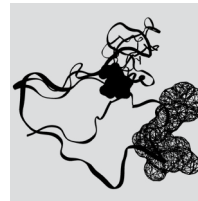
Rotavirus



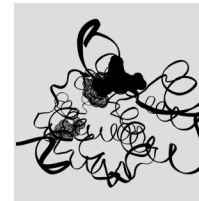
Dengue virus



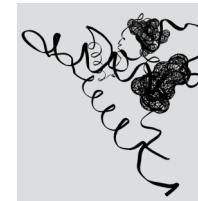
CCHF virus



L. monocytogenes



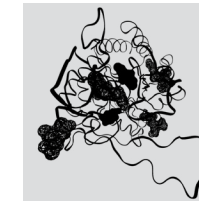
H. pylori



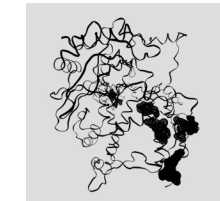
S. aureus



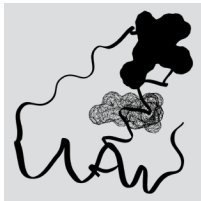
Monkeypox virus



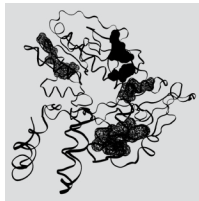
Shigella spp.



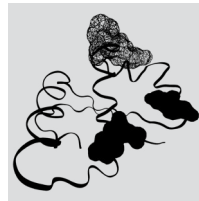
Influenza A virus



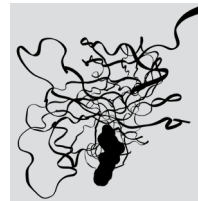
C. auris



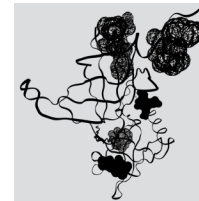
Providencia spp.



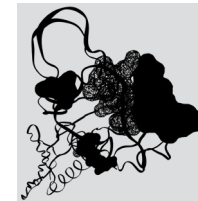
Ebola virus



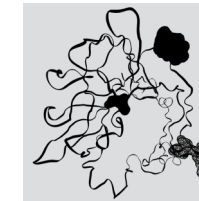
Hantavirus



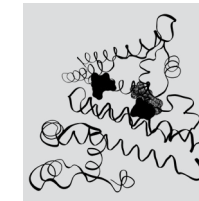
K. pneumonia



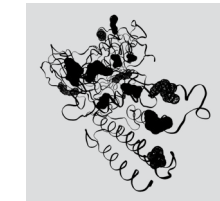
M. tuberculosis



E. coli



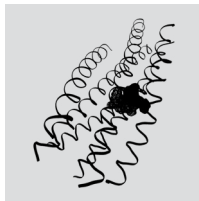
B. burgdorferi



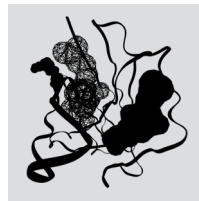
L. pneumophila



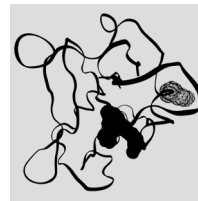
MERS



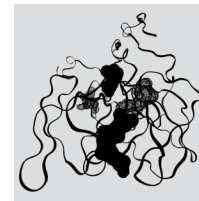
HIV



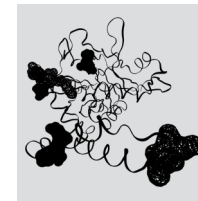
P. falciparum



E. Faecium



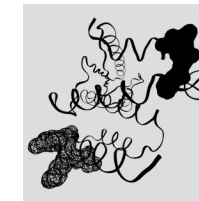
E. cloacae



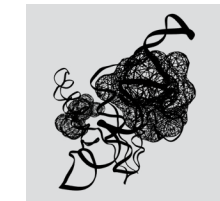
S. enterica



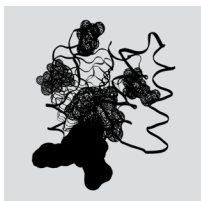
A. baumannii



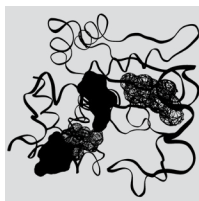
Mammarenavirus



H. influenzae



Proteus spp.



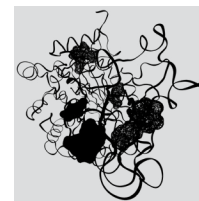
C. difficile



Serratia spp.



Morganella spp.



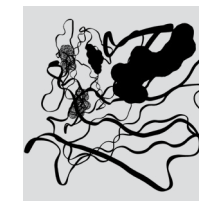
Campylobacter spp.



S. aureus



P. aeruginosa



Nipah virus



SARS

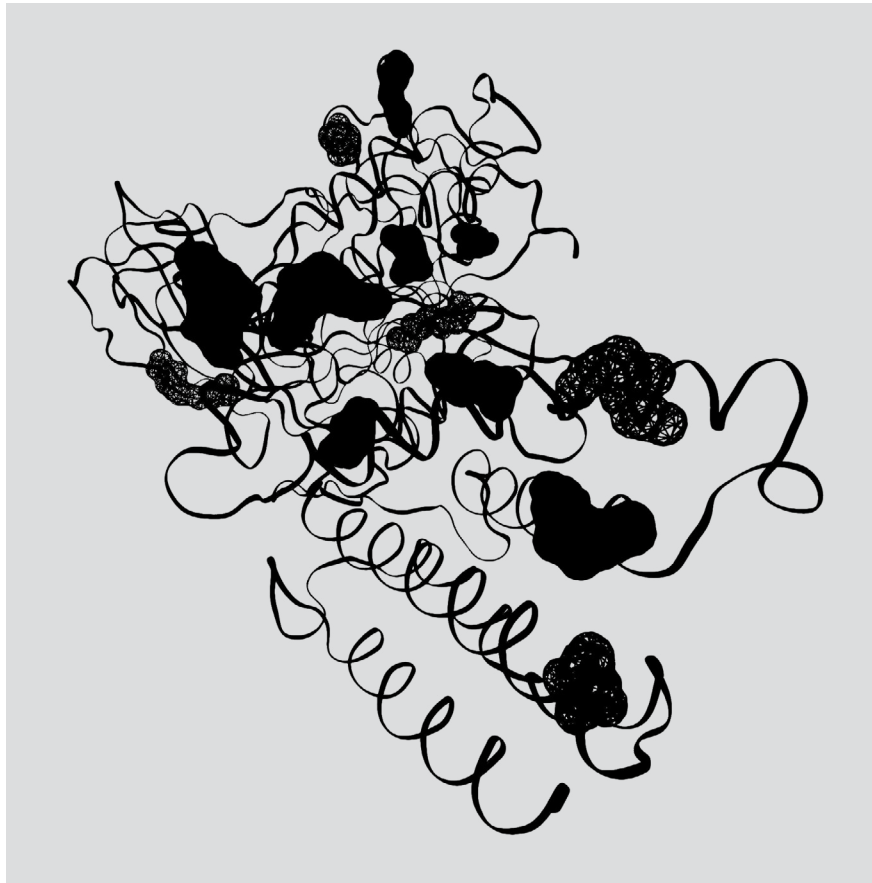
SELECTED VISUALS

Select 1-5 visuals and arrange them in different sizes on this page.

3



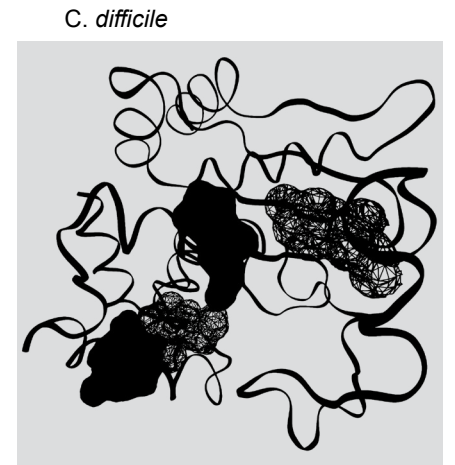
Rotavirus



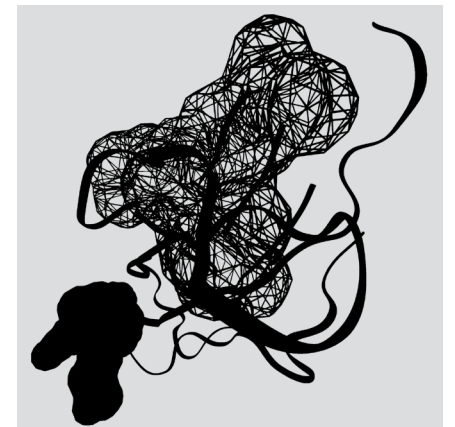
L. pneumophila



Hantavirus



C. difficile



Dengue virus

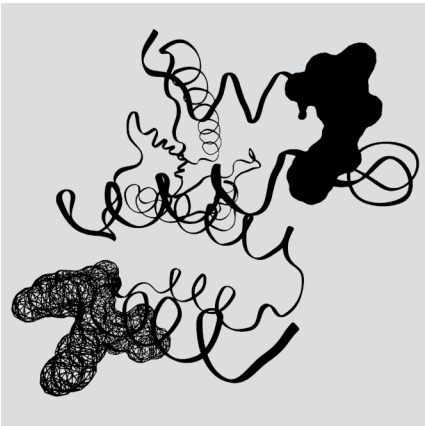
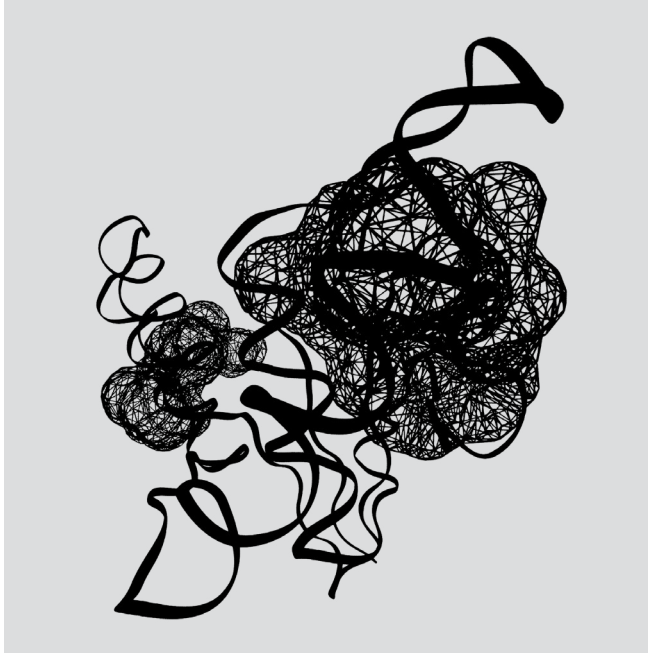
SELECTED VISUALS

Select 1-5 visuals and arrange them in different sizes on this page.

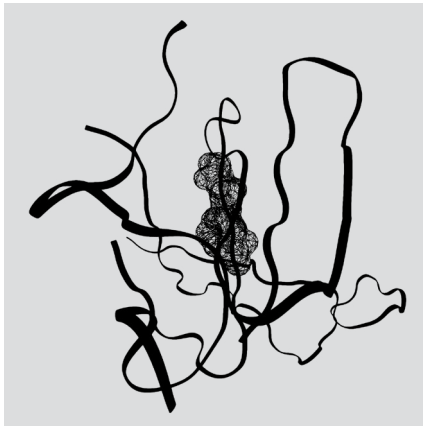
S. aureus



H. Influenzae



Mammarenavirus



SARS



Proteus spp.

DOCUMENTATION OF THE PROCESS

SHORT CV

5

PROCESS

1) I researched the 45 most dangerous pathogens as of 2020 and compiled them into a list.

2) I wrote a script that searched a protein database (rcsb.org/) for the top 25 proteins related to each pathogen in the list.

3) The script then scanned each proteins code sequence and returned a list of english words found in the sequence.

4) Each word was then used to search a sentiment analysis database (data.world/crowdfunder/sentiment-analysis-single-word) which returns whether the word has a positive (“+”) or negative (“-”) sentiment and to what degree (score).

5) The protein with the highest number of sentiment analysis matches and highest cumulative score was selected to represent that pathogen.

6) The words in the protein code sequence were orientated and rendered in Visual Molecular Dynamics depending on whether they possessed a negative (solid) or positive (wire) sentiment and their score (resolution) while nonsense was left as a ribbon.

7) The rendering of each protein was saved using Tachyon and prepared in Adobe Photoshop and Adobe Illustrator.

SHORT CV

Ali Razzak is a structural bioinformatician and data scientist who was born in Iraq, raised in New Zealand and is now residing in Switzerland. Outside of science his interests lie in graphic and web design. He utilises digital mediums to traverse migration, data mining, and fashion culture discourse. He finds transdisciplinary discourse and manipulating unconventional tools to introduce alternative perspectives most exciting.

DOCUMENTATION OF THE PROCESS VISUAL IMPRESSIONS

Link to github project files:

https://github.com/ternlef11/PDB_protein_sentiment_analysis.git

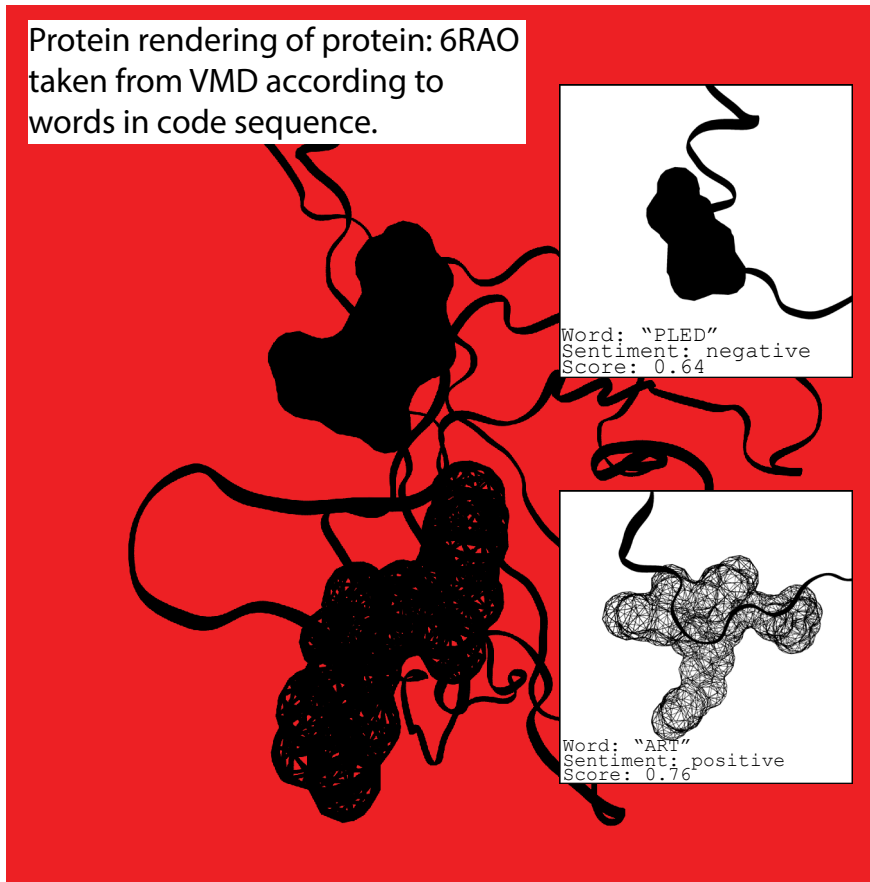
Pathogens	Sequences	Code
Influenza A virus	1SMN.fasta	<pre>>6RAO_2 Chain C Afp2 Serratia entomophila (42906) MTVTTTYPGVYLSE- DAVSSFSVNSAAT- AVPLFAYDSENTNTINKPI QVFRNWAFTVEYPTPLED AFYTSLSLWFMHGGGKCYL VNEANIADAVAQYDDITLI VAAGTDTTTYTAFTTVVGQ GYRIFGLFDGPKEKIAGTA KPDEVMEEYPTSPFGAVFY PWGTLASGAAPPSAIAAA SITQTDRTRGVWKAPANQA VNGVTPAFVSDDFQGGYN QGKALNMIRTFSGQGTVMW GARTLESDSNWRYIPVRR FNAVERDIQKSLNKLVEF NSQPTWQVRVKAADVSYLH LWQQGALAGNTPADANFVQ VGKDLTMTQEEINQGMII KIGLAAVRPAEFILQFSQ</pre>
Rotavirus	1QL0.fasta	
SARS	1SRP.fasta	
<i>P. aeruginosa</i>	6DRE.fasta	
<i>K. pneumonia</i>	5WUL.fasta	
Ebola virus	3DDR.fasta	
<i>Serratia spp.</i>	6RAO.fasta	
Corona virus	5WVM.fasta	
<i>H. pylori</i>	1I7Q.fasta	
<i>E. coli</i>	1B04.fasta	
Rabies virus	3ZFI.fasta	<pre>1QAE score: 1 5WVM score: 1 - HIT 0.79 + DEAL 0.76</pre>
<i>M. tuberculosis</i>	6RAP.fasta	
HIV	1G8T.fasta	
<i>E. faecium</i>	4B6I.fasta	
<i>Morganella spp.</i>	5NEN.fasta	

1. List of most dangerous pathogens as of 2020.

2. Protein data base top hits for pathogen "*Serratia spp.*"

3. Amino acid code for protein 6RAO taken from sequence for "6RAO.fasta".

Protein rendering of protein: 6RAO taken from VMD according to words in code sequence.



Code

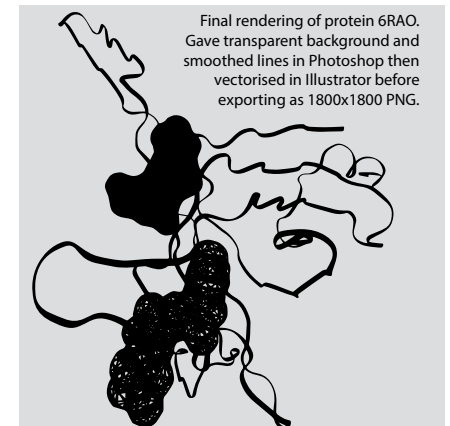
```
>6RAO_2|Chain C|Afp2|Serratia entomophila (42906)
MTVTTTYPGVYLSEDAVSSFSVNSAATAVPLFAYDSENTNTINK
PIQVFRNWAFTVEYPTPLEDAFYTSLSLWFMHGGGKCYLVNEA
NIADAVAQYDDITLIVAGTDTTTYTAFTTVVGQGYRIFGLFDG
PKEKIAGTAKPDEVMEYPTSPFGAVFYPWGTLASGAAPPSAI
AAASITQTDRTRGVWKAPANQAVNGVTPAFVSDDFQGGYNQKG
ALNMIRTFSGGGTVWVGARTLESDSNWRYIPVRRLFNAVERDIQ
KSLNKLVEFPNSQPTWQVRVKAADVSYLHSLWQQGALAGNTPADA
WFVQVGKDLTMTQEEINQGMIIKIGLAAVRPAEFILQFSQDI
AQ
```

Word "PLED"
Sentiment: negative
Score: 0.64

Word "ART"
Sentiment: positive
Score: 0.76

6RAO score: 2
- PLED 0.64
+ ART 0.76

Reads fasta file, scans words, analyses and scores words.



Final rendering of protein 6RAO. Gave transparent background and smoothed lines in Photoshop then vectorised in Illustrator before exporting as 1800x1800 PNG.

Code scores

6RAO score: 2 1SMN score: 1
- PLED 0.64 - HIT 0.79
+ ART 0.76

1QAE score: 1 5WVM score: 1
- HIT 0.79 + DEAL 0.76

5NEN score: 1 5NEN score: 1
- LATE 1.0 + PAY 0.75

6RAO had the highest score so using PDB to represent *Serratia spp.*

