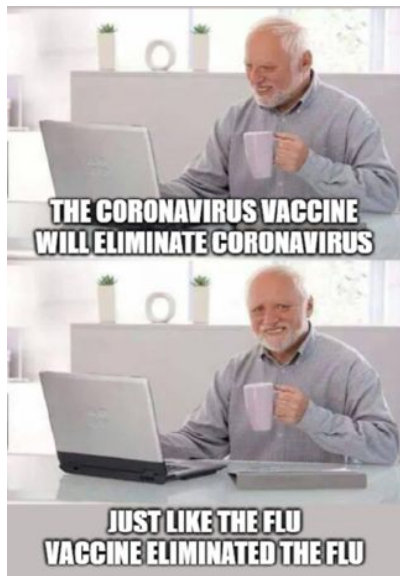


Understanding the Provenance of Visual Disinformation Targeting Science

Walter J. Scheirer

Department of Computer Science and Engineering
University of Notre Dame

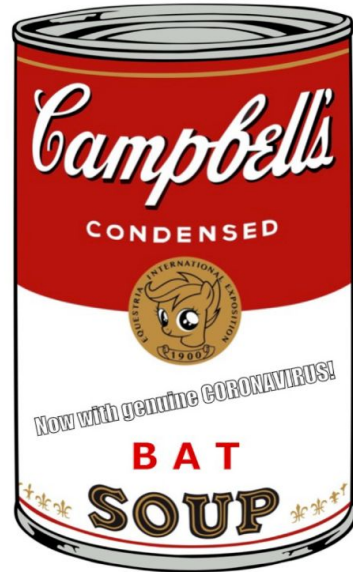
Public health counter-messaging from an unlikely source



When they're about to give you
Coronavirus Vaccine



Memetic warfare

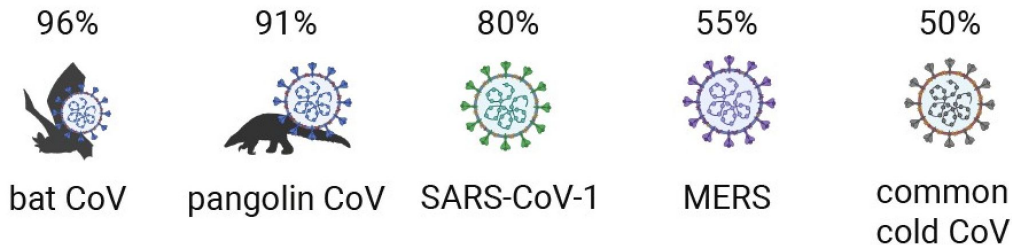


W. J. Scheirer, "A Pandemic of Bad Science," Bulletin of the Atomic Scientists, July 2020

Origin of the bat soup myth: genomic link

Genome

Nucleotide identity to SARS-CoV-2



Y. M Bar-On, A. Flamholz, R. Phillips and R. Milo, "SARS-CoV-2 (COVID-19) by the numbers," eLife, March 2020

Origins of the bat soup myth: the Huanan Seafood Wholesale Market



Huanan Seafood Wholesale Market © BY 3.0 China News Service

Origin of the bat soup myth: tabloid news



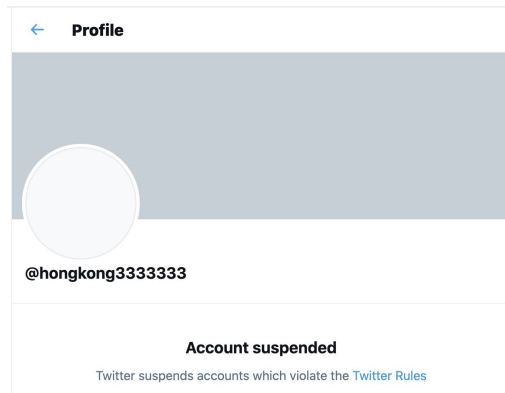
<https://video.foxnews.com/v/6145888115001#sp=show-clips>



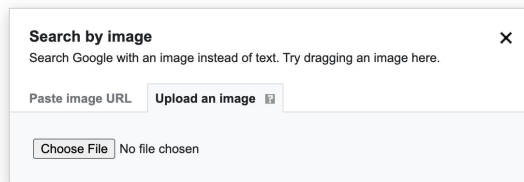
Origin of the bat soup myth: Twitter disinfo.

Banned Anti-CPC Account

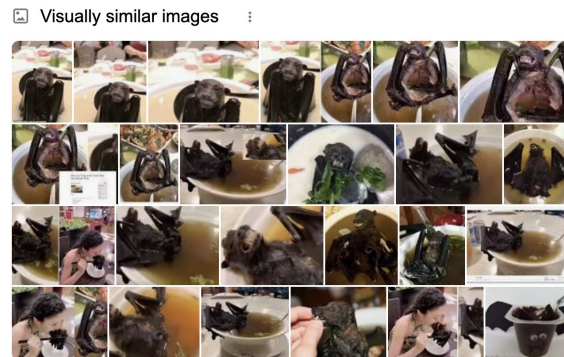
Alleged Chinese Delicacy



Reverse Image Search



Turns up links about Palau and somebody named Wang Mengyun?



Report images

Origin of the bat soup myth: Wang Mengyun



Wang Meng Yun

This is a public statement on the bat video
It's also a letter of apology

Sorry everyone, I shouldn't eat bats

"You should go to hell. You should be killed in the evening. You're abnormal. You're disgusting. Why haven't you died?" These are the messages I just received today.

It's all because in 2016, when I was screening a tour program in Palau, a south pacific island, I ate a soup of local people's daily food.

Back in May 2016, I didn't know what the virus was at that time. When the video was released, I only want to introduce the lifestyle of the local people. I don't know bat will become the host. Although it's a local fruitbat, I didn't refer to expert info and fail to clarify the danger before eating the bat

Since Decr 2019, the pneumonia of unknown cause has occurred in Wuhan. There are numerous confirmed cases infected with new coronavirus, and the victims are increasing everyday. Experts say that the source of the new coronavirus is wild animals illegally sold in a seafood market in Wuhan.

W. J. Scheirer, "A Pandemic of Bad Science," Bulletin of the Atomic Scientists, July 2020

Origin of the bat soup myth: Palau

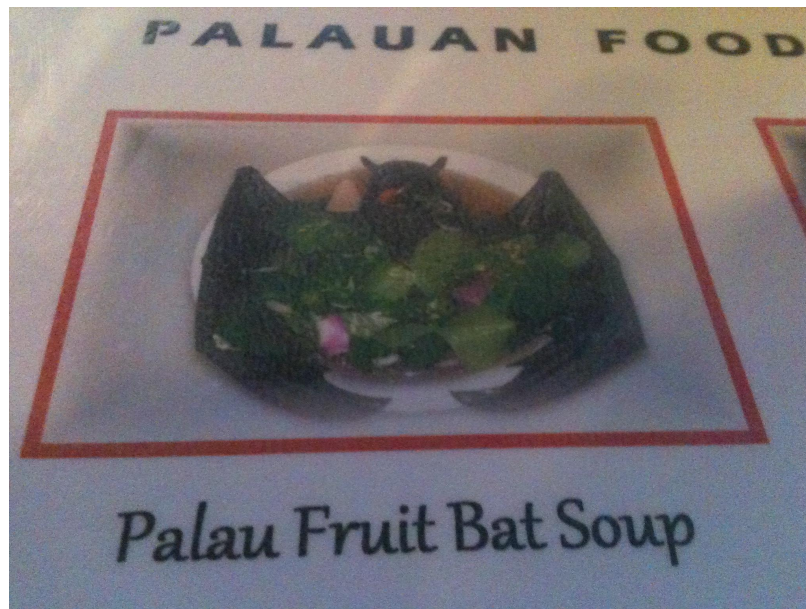


Photo Credit, Elizabeth Warren, Taken in 2012.

<https://elizabethinpalau.wordpress.com/2012/05/20/exploring-the-island/>

Limitations of determining provenance by hand

- Reverse image search only works well for duplicate and near duplicate images
- Staggering number of memes and source images
 - Every minute: 243,000 images are uploaded to Facebook
 - Every day: 500 million people use Instagram and over two billion users watch or post to YouTube
 - Every year: hundreds of billions of photos are uploaded to social media

Watching the meme landscape with AI

- Real-time processing pipeline with content flagging
- Assess meme landscape with respect to events and actors
- Scale to the order of billions of images per day
- Match composite images to other composite and donor images



Automatic Image Provenance Analysis



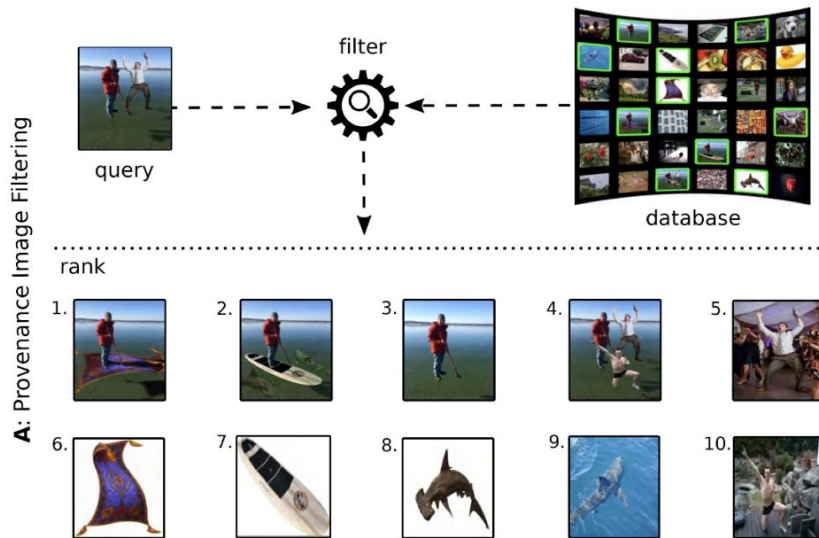
Aparna Bharati



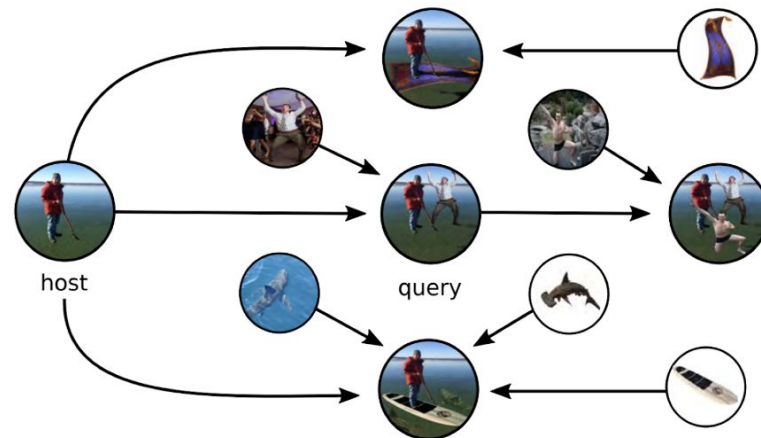
Joel Brogan



Daniel Moreira

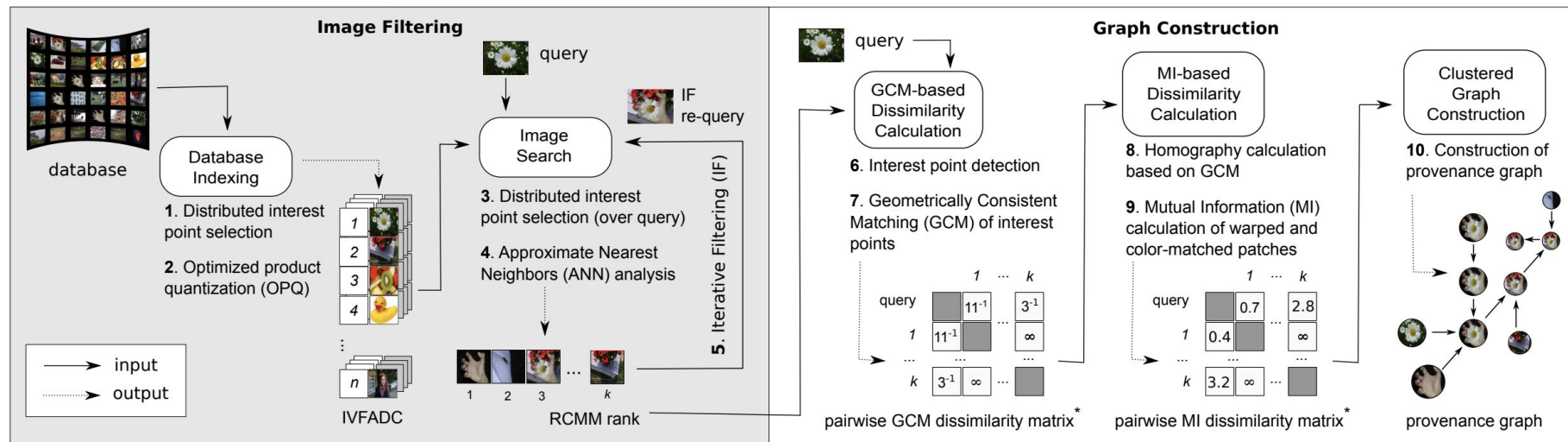


B: Provenance Graph Construction

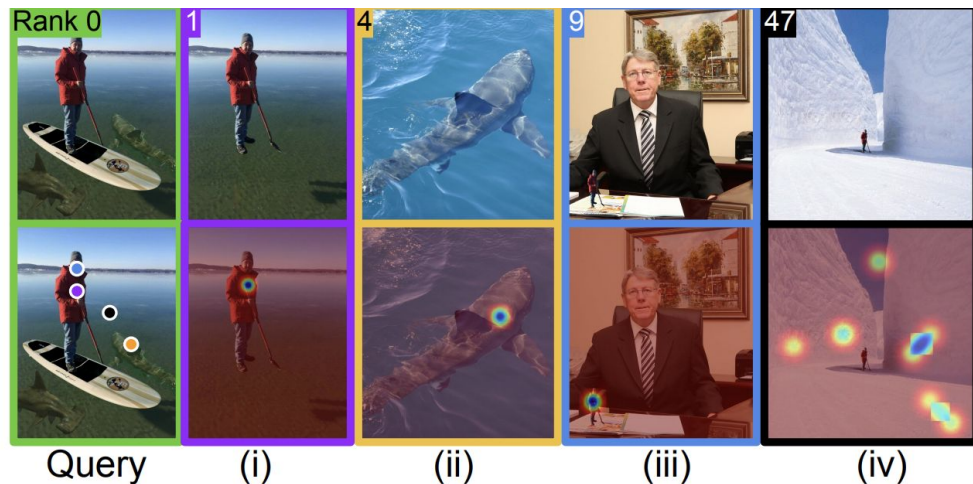
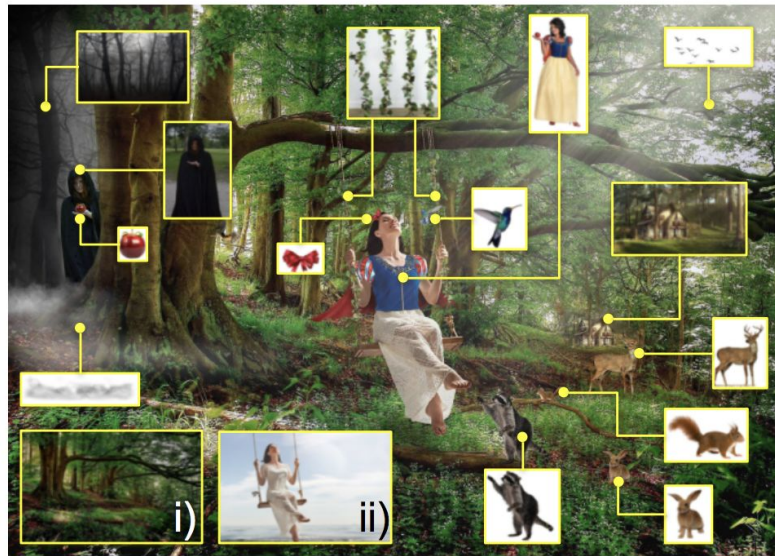


D. Moreira, A. Bharati, J. Brogan, A. Pinto, M. Parowski, K. Bowyer, P. Flynn, A. Rocha and W. J. Scheirer, "Image Provenance Analysis at Scale," IEEE Transactions on Image Processing (T-IP), December 2018

Provenance Analysis Pipeline



Dynamic Spatial Verification for Large-Scale Object-Level Image Retrieval



J. Brogan, A. Bharati, D. Moreira, K. Bowyer, P. Flynn, A. Rocha and W. J. Scheirer, "Dynamic Spatial Verification for Large-Scale Object-Level Image Retrieval," arXiv 2019

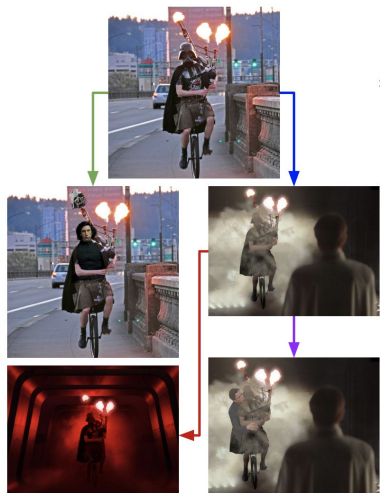
How well does it work?

NIST Media Forensics Challenge (<https://mfc.nist.gov/>)

Order millions of images

	Recall@200	Node Recall	Link Overlap	Node Overlap	Node + Link Overlap
MFC 2018	0.9	0.798	0.272	0.798	0.543
MFC 2019	0.819	0.693	0.295	0.703	0.519
MFC 2020	0.855	0.753	0.250	0.490	0.714

How about the real data?



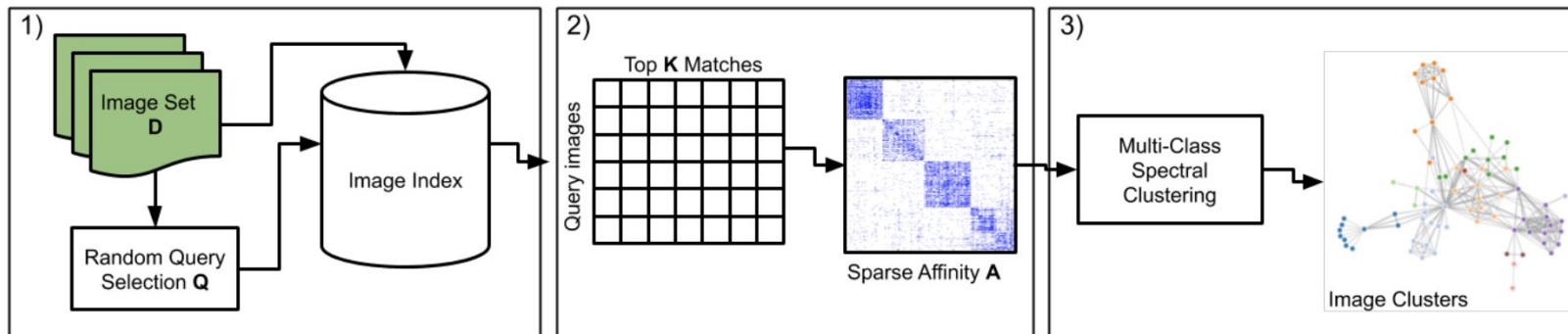
Images from /r/photoshopbattles

Solution	VO	EO	VEO
GCM-SURF [38]	0.884	0.156 [†]	0.523 [†]
GCM-MSER	0.924	0.121[†]	0.526[†]
MI-SURF	0.757	0.037	0.401
MI-MSER	0.509	0.027	0.271

[†]: Values for undirected edges. In bold, the solution with the best VEO.

Pipeline for Meme Genre Discovery

Replace Graph Building with Clustering



M. Yankoski, T. Weninger and W. J. Scheirer, "An AI Early Warning System to Monitor Online Disinformation, Stop Violence, and Protect Elections," Bulletin of the Atomic Scientists, March 2020

W. Theisen, J. Brogan, P. Bilo Thomas, D. Moreira, P. Phoa, T. Weninger, and W. J. Scheirer, "Automatic Discovery of Political Meme Genres with Diverse Appearances," AAAI Conference on Web and Social Media (ICWSM), June 2021.

Case Study: 2019 Indonesian National Election



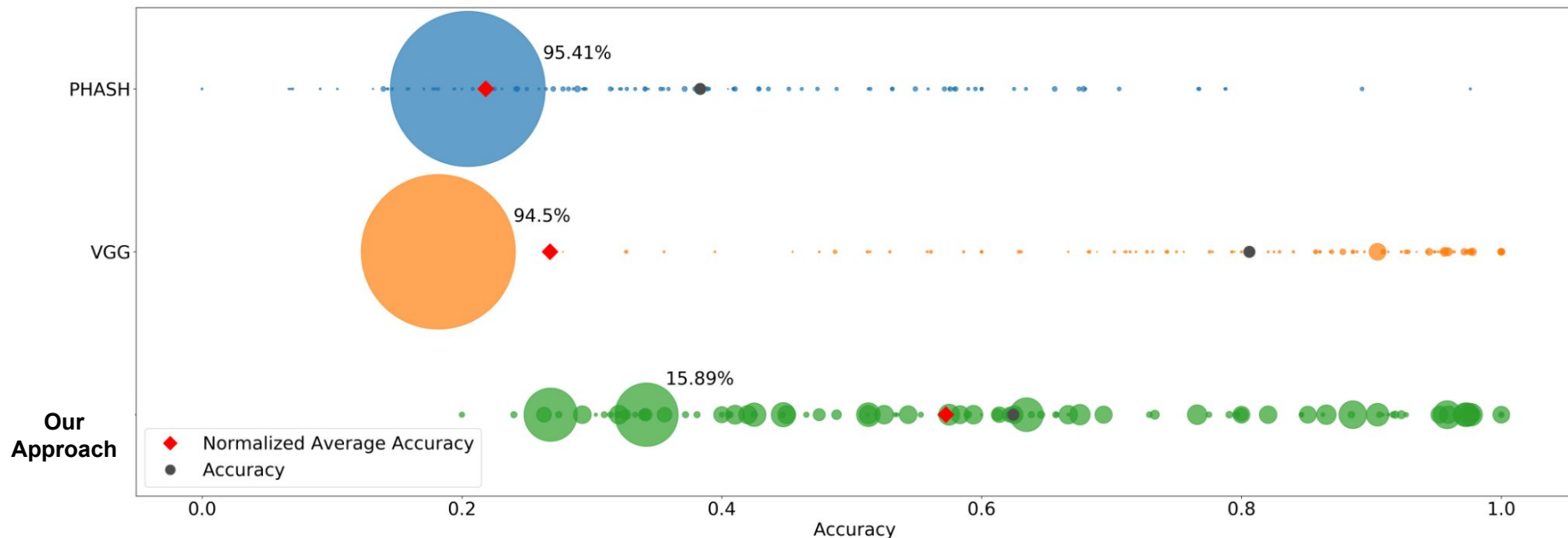
William Theisen



Pam Thomas



Tim Weninger



Unusual findings over 2 million images

A variety of messaging.
Not all of it disinformation!

Advertising regularly
piggybacks on significant
events like elections.



"PSA" Warning About Social Media



Meme Mocking Gaffe by Subianto



Ad for "Super Grow Up" Supplement

Bogus Medical Supplements

- “Super Grow Up” supplement
- Promises to increase the height of adults taking it
 - Only works under the age of 40
- Website went offline in early 2020



Who else is selling bogus supplements using elections as a vehicle?



<https://thecounter.org/would-you-buy-supplements-from-alex-jones/>

Coronavirus? There's a cheap cure for that too...



<https://www.washingtonpost.com/nation/2020/03/11/jim-bakker-coronavirus-cure/>

Outstanding Challenges

Scalability

Video Processing

Graph Building

Usability for Analysts



Questions?