ENIGMA 2023, SECURITY AND PRIVACY IDEAS THAT MATTER
JAN 24-26, 2023, SANTA CLARA, CA

Why Is Our Security Research Failing?

Five Practices to Change!

Marcus Botacin
Texas A&M University, USA
@MarcusBotacin

Everybody Complains About Security Research

Outside Academia Complaints

1. *“The scientific way of defining requirements is too strict for real world use”*

2. *“Early-stage research is useless in the sense of not being close to transitioning to practical use.”*

3. *“Cybersecurity is failing due to ineffective technology”*

4. *“Universities failing at cybersecurity education”*

Academic Complaints

Herley and P. C. van Oorschot about the JASON report:

> ***"The science seems under-developed in reporting experimental results, and consequently in the ability to use them. The research community does not seem to have developed a generally accepted way of reporting empirical studies so that people could reproduce the work"***

Let’s Suppose Security Research is Broken

Study: Challenges & Pitfalls in Malware Research

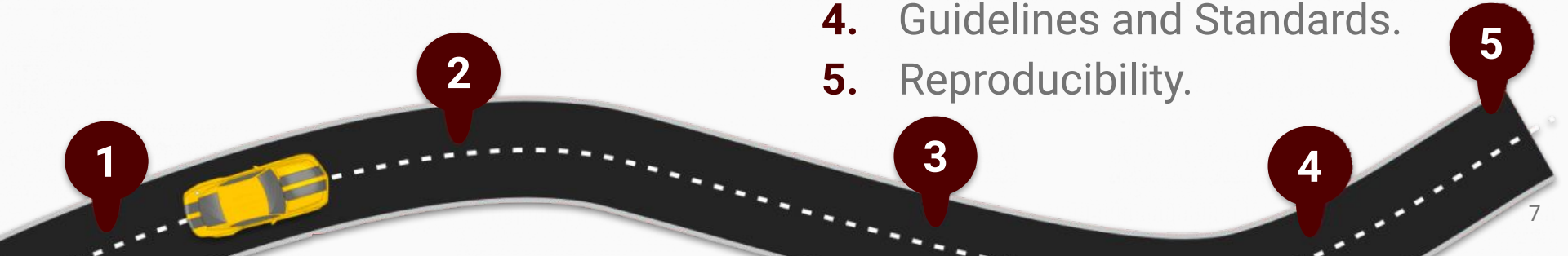
Table 1: **Selected Papers**. Distribution per year (2000 – 2018) and per venue.

Venue/Year	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
USENIX (Security, LEET & WOOT)	1	0	0	0	0	1	1	6	2	3	7	8	10	12	9	7	9	13	6	95
CCS	0	0	0	0	0	0	0	2	4	6	6	7	11	9	11	14	2	11	6	89
ACSAC	0	0	0	0	2	3	2	4	4	1	3	8	10	7	10	6	3	7	8	78
IEEE S&P	0	1	0	0	0	1	3	2	1	0	0	10	17	12	3	6	4	5	3	68
DIMVA	0	0	0	0	0	4	4	3	8	2	3	0	8	4	8	7	7	5	4	67
NDSS	0	0	0	0	1	0	2	0	3	3	3	3	2	4	5	4	9	7	3	49
RAID	0	0	1	0	0	1	3	0	0	0	0	0	3	5	5	3	4	3	3	31
ESORICS	0	0	0	0	0	1	0	0	2	1	0	0	2	3	3	0	1	1	0	14
Total	1	1	1	0	3	11	15	17	24	16	22	36	63	56	54	47	39	52	33	491

Goals and Roadmap

- Not to point fingers.
- I also make mistakes.
- Teach some lessons.
- Learn from others' mistakes.
- Learn from our own mistakes.

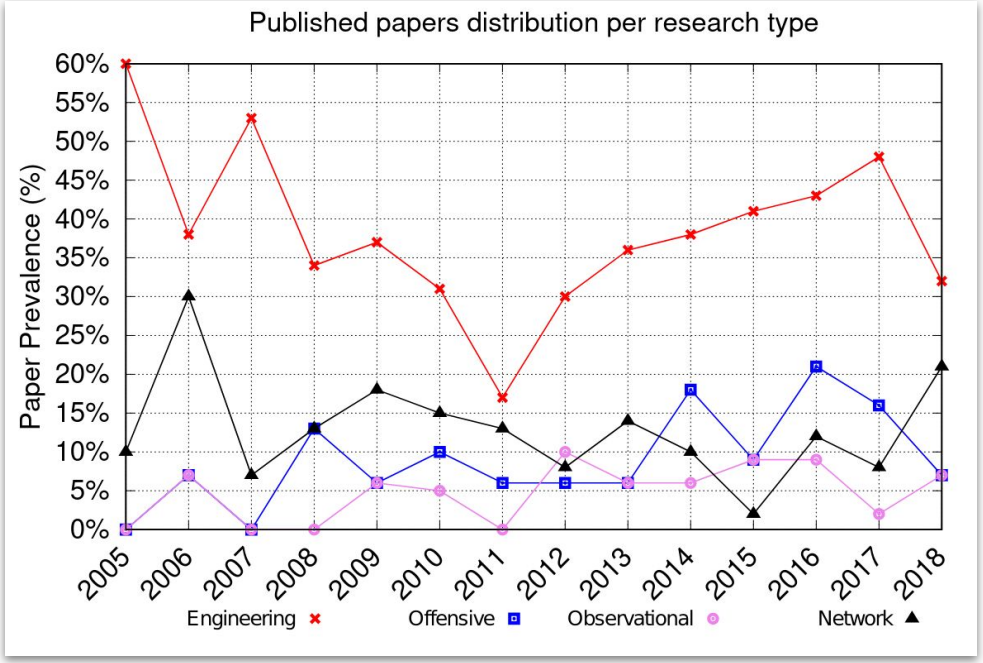
1. Study Types.
2. When to start looking to the industry.
3. When to stop looking to the industry.
4. Guidelines and Standards.
5. Reproducibility.



1. Focusing too much on a single study type

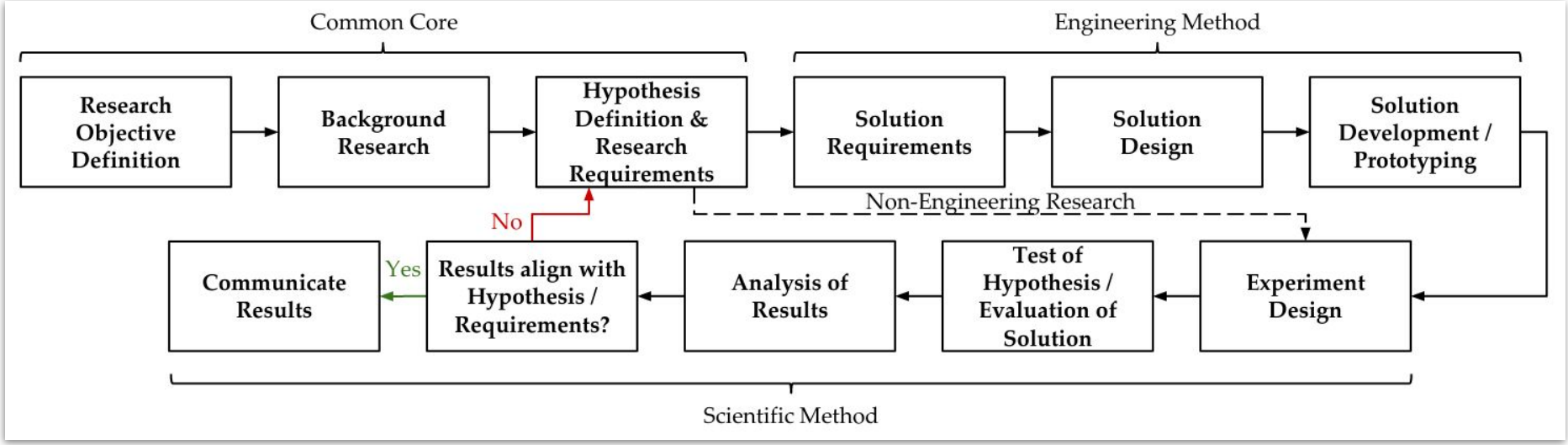
The Most Common Research Types

- Engineering Solutions are more than 50% of all published papers.
- Only a few papers relied on previous measurements and observational studies.
- It suggests researchers have been taking ad-hoc project decisions.



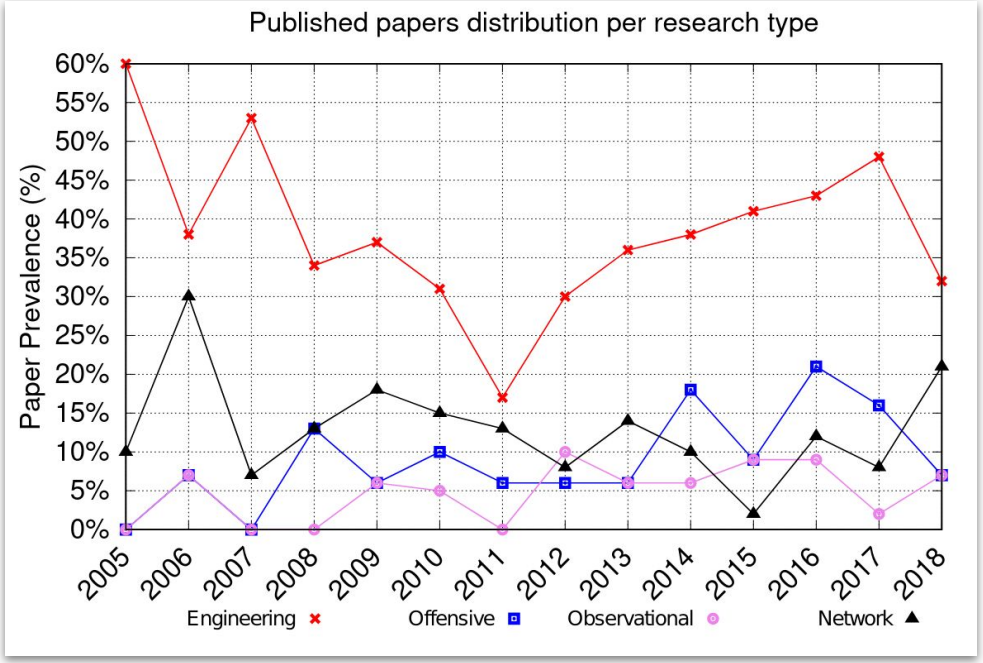
A new malware research method

Solution Proposal: Integrate Science and Engineering methods



The Most Common Research Types

- Engineering Solutions are more than 50% of all published papers.
- Only a few papers relied on previous measurements and observational studies.
- It suggests researchers have been taking ad-hoc project decisions.



2. Not looking at the industry when needed.

Systems Architectures

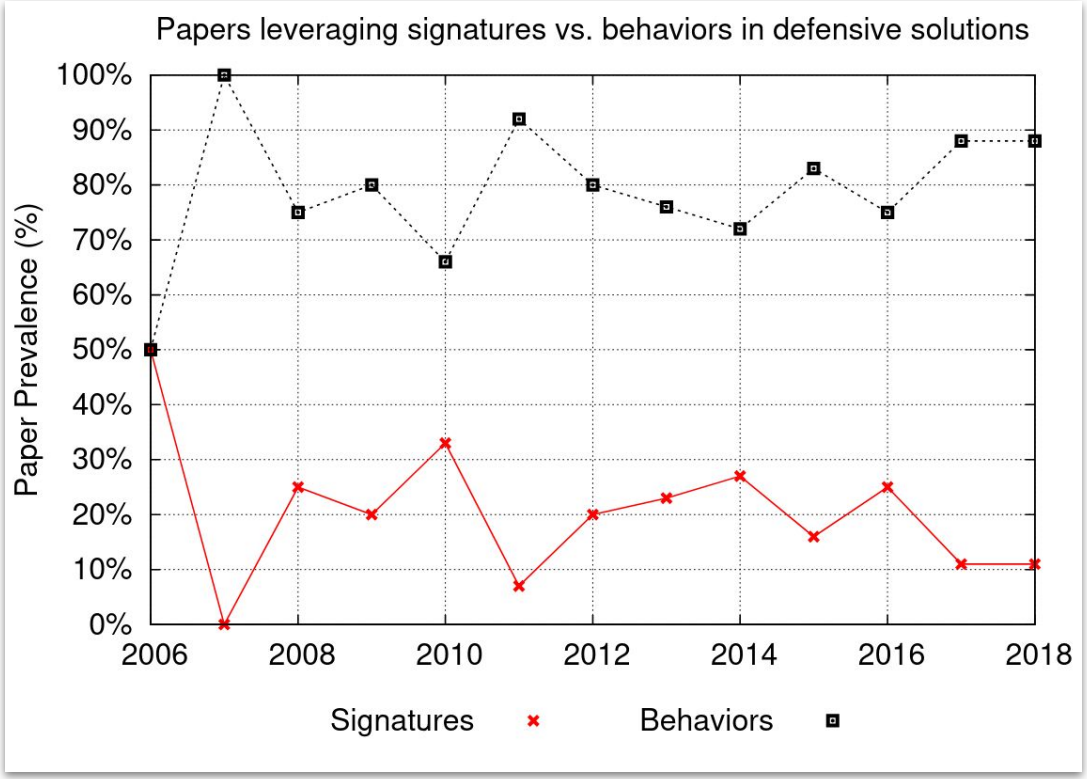
- Multi-core processors are prevalent (>99%)
- Literature is limited in multi-core examples (<1%)
- Prototypes can be single-core.
- Multi-core threats must be researched.

- Distributed Threadless Malware

- VANILLA: Layers and Layers of Attacks

Malware Detection Approaches

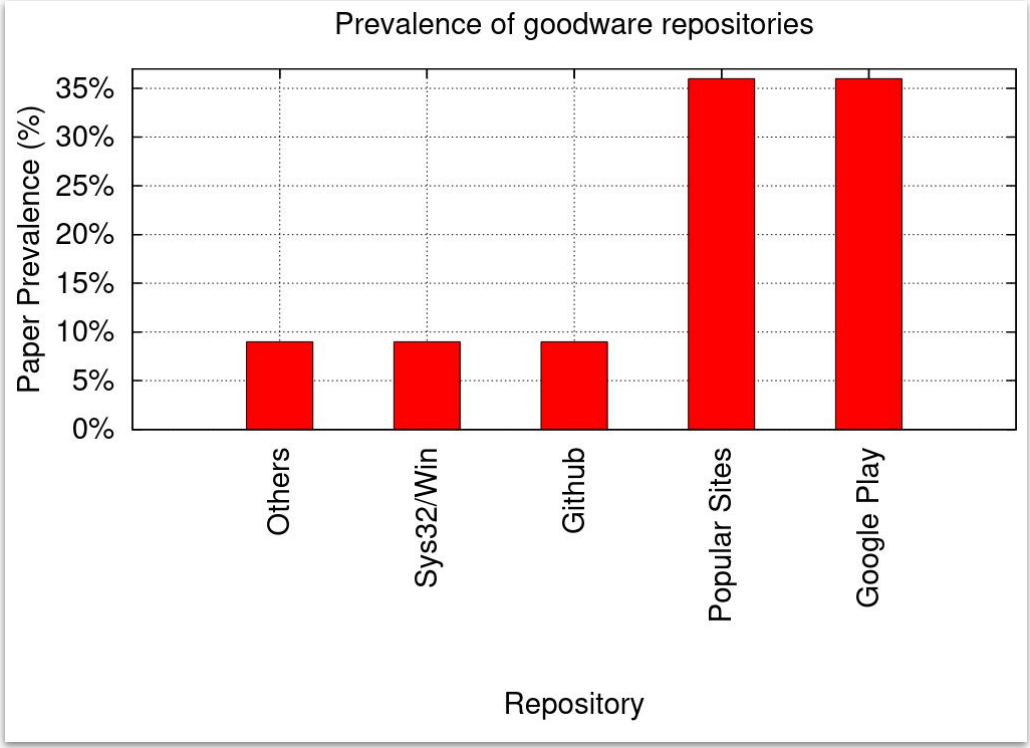
- Majority of academic studies using ML rather than signatures.
- Industry uses both.
- Signatures still relevant.
- Should we stop researching signatures?



3. Looking too much at the industry and market.

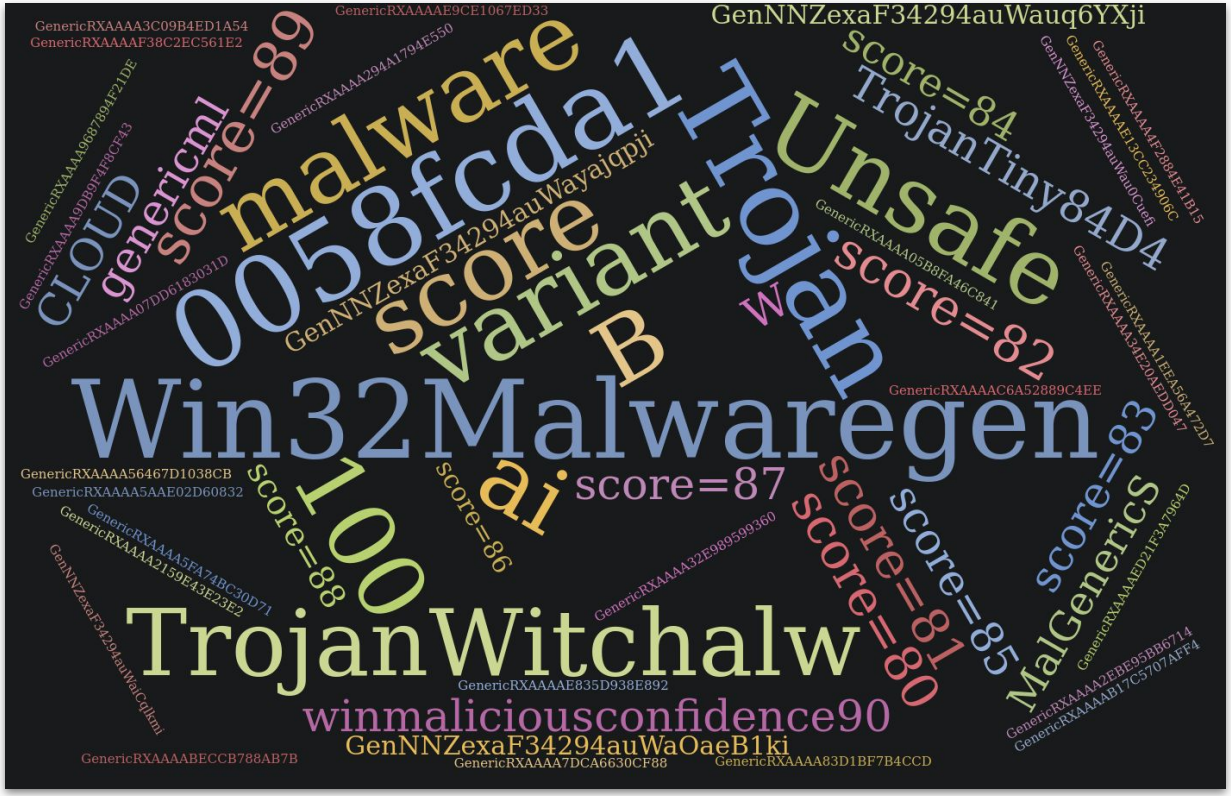
Goodware Sources

- Goodware samples are used as ground truth for ML models.
- Crawling software repositories allows getting popular software.
- Some binaries might be trojanized.
- A few studies filter out trojanized binaries.
- ML models might be biased!



Antivirus Labels

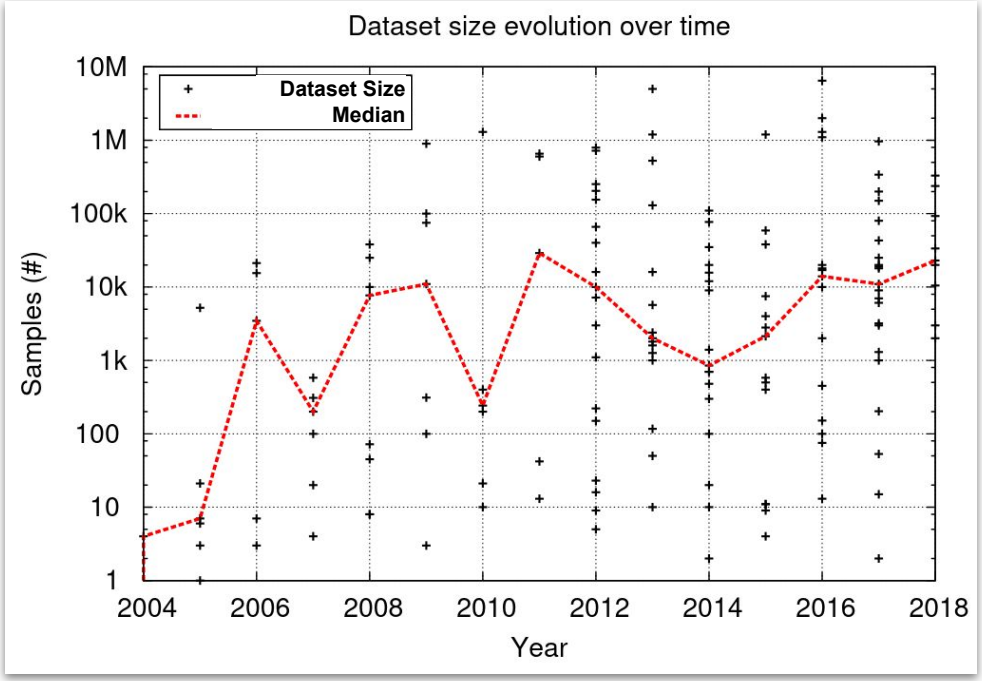
- The problem of heterogeneous AV labels has been known for a long.
- Comparisons are not fair because labels are not easily comparable.
- AVClass



4. Not developing standards and guidelines.

Dataset Sizes

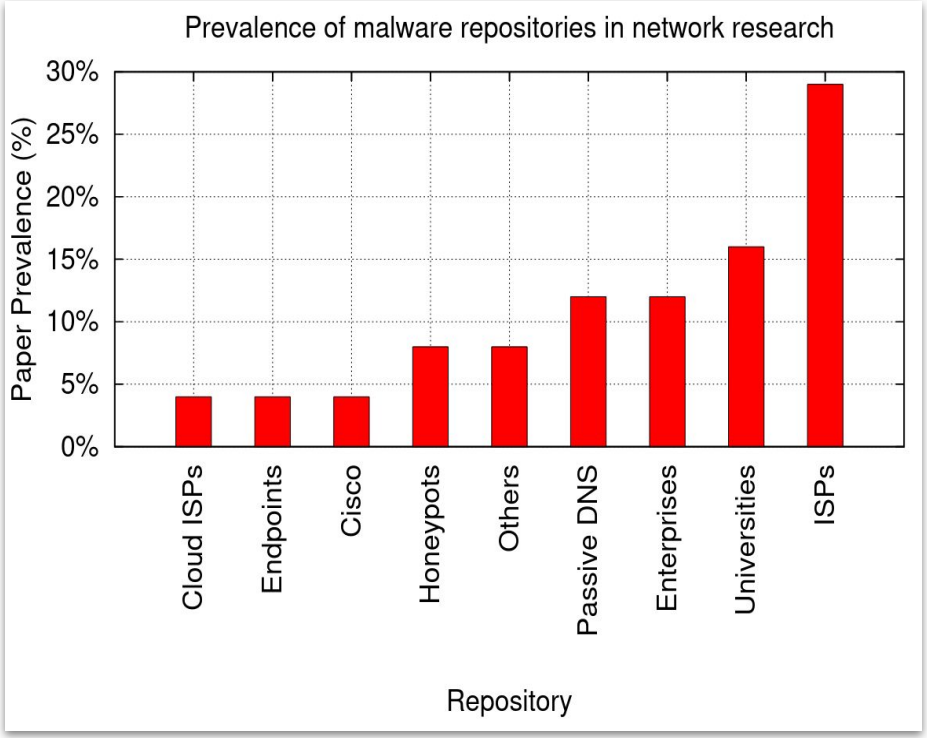
- No guideline and not practical standard.
- Researchers adopting ad-hoc decisions.
- **Anchor bias:** the median is ever-growing.
- How much is enough?
- **Contradictory verdicts:** 900K vs. 1M samples.



5. We have a reproducibility crisis!

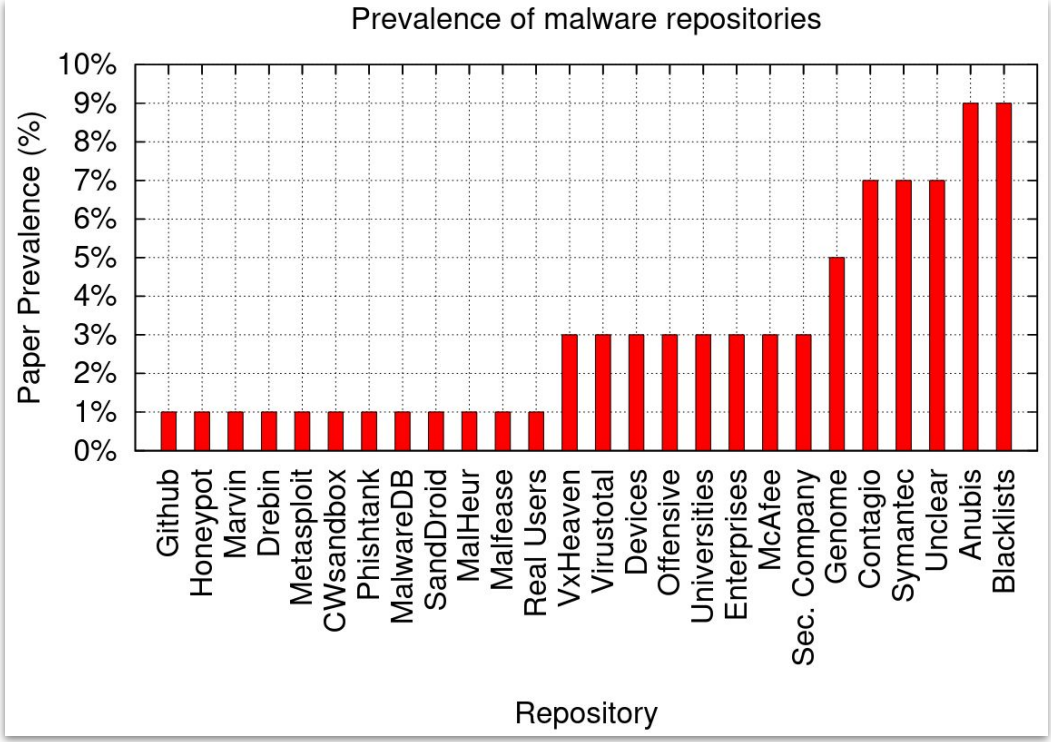
Networks: Where Datasets Come From?

- **The Human Aspect:**
  Datasets are made private via Non-Disclosure Agreements (NDAs)



Networks: Where Datasets Come From?

- **The Technological Aspect:** Malware analyses might not be reproducible due to payloads and C&Cs being sinkholed at any time.



Moving Forward

Call to Action

- Researchers:
  - Diversify the types of conducted studies.
- Reviewers and Program Committees:
  - Develop evaluation guidelines.
- The Field:
  - Focus on representativity rather than quantity.
- Venues and CFPs:
  - Ask for more diversified studies.
  - Be clear on requesting representative datasets.

Why Is Our Security Research Failing?
Five Practices to Change!

Thank you!

Contact: botacin@tamu.edu or @MarcusBotacin
My Website: marcusbotacin.github.io

Marcus Botacin
Texas A&M University, USA
@MarcusBotacin