

SAS<sup>®</sup>  
GLOBAL  
FORUM  
2020

MARCH 29 - APRIL 1  
WASHINGTON, DC



USERS PROGRAM



# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer



**Krystian Matusz**

**Information** is a business's most valuable asset, as it provides a **competitive advantage** and supports **sustainable** growth. But information must be **securely** stored, shared, and processed. Currently, **this is the main challenge for every organization**. Any mistake can damage the entire business.

Worldwide, information security spending in 2019 exceeded \$124 billion (Gartner). The average cost of cybercrime has increased by 72% in the last five years to \$12 million and continues to grow. Estimates show **near-constant frequency of attacks every 39 seconds, affecting one in three Americans**.

**Malware plays the leading role in these attacks**. As a malicious program, intended to infiltrate or destroy PCs and networks without users' knowledge, malware is particularly dangerous today, when the penetration of the market by Windows-based systems is above 78%, creating an attractive target.

I will present the possibility of establishing the persistence of malware in operating systems based on behavioral analysis of the malware in the context of analyzing an internal Windows Registry. This approach is **reinforced** by using Machine Learning Predictive Models built in SAS® Viya® to evaluate **at a scale** how likely the malware will survive the restart of an infected operating system.

Malware might use **genetic** and **polymorphic** obfuscation or code packing, so the behavioral-based approach is an **effective way** to connect the Cybersecurity and Data Science domains in order **to increase** the overall level of security and awareness.

## Problem Statement and Research objective

**Predicting Malware Persistence**

through

Windows Registry

**Behavioral**

**Advanced**

**Cybersecurity**

**Analysis**



## 1. Topic & Abstract

2. Agenda

3. About me

4. Facts: KPIs/KRIs

5. Malware - def.

6. Malware - types

7. Research - goal

8. Methodology

9. Architecture

10. Insights

11. Conclusions



# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer



**Krystian Matusz**

**Information** is a business's most valuable asset, as it provides a **competitive advantage** and supports **sustainable** growth. But information must be **securely** stored, shared, and processed. Currently, **this is the main challenge for every organization**. Any mistake can damage the entire business.

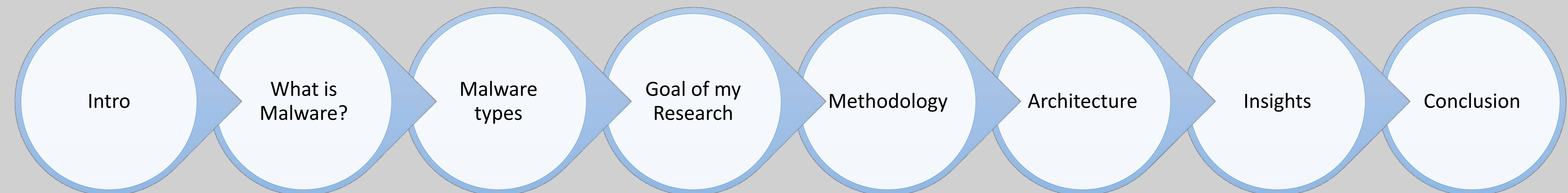
Worldwide, information security spending in 2019 exceeded \$124 billion (Gartner). The average cost of cybercrime has increased by 72% in the last five years to \$12 million and continues to grow. Estimates show **near-constant frequency of attacks every 39 seconds, affecting one in three Americans**.

**Malware plays the leading role in these attacks**. As a malicious program, intended to infiltrate or destroy PCs and networks without users' knowledge, malware is particularly dangerous today, when the penetration of the market by Windows-based systems is above 78%, creating an attractive target.

I will present the possibility of establishing the persistence of malware in operating systems based on behavioral analysis of the malware in the context of analyzing an internal Windows Registry. This approach is **reinforced** by using Machine Learning Predictive Models built in SAS® Viya® to evaluate **at a scale** how likely the malware will survive the restart of an infected operating system.

Malware might use **genetic** and **polymorphic** obfuscation or code packing, so the behavioral-based approach is an **effective way** to connect the Cybersecurity and Data Science domains in order **to increase** the overall level of security and awareness.

## What will be covered during my presentation in Washington D.C Convention Center



### Target Audience:

It is **technical** presentation delivered for **both: technical and business level experts** with the focus on:

- Analysts
- Engineers
- CxO / C-Suite
- Project, Program and Portfolio Managers

### Skills level

- Presented topic and research is advanced
- All skills level, abilities and your feedback are very welcomed

### Keywords

- Advanced Analytics
- Machine Learning
- Data Mining
- Cloud Computing

1. Topic & Abstract

2. Agenda

3. About me

4. Facts: KPIs/KRIs

5. Malware - def.

6. Malware - types

7. Research - goal

8. Methodology

9. Architecture

10. Insights

11. Conclusions



## Krystian Matusz, Passionate Data Scientist

Individual Contributor: Technology Evangelist and Business Inspiration Officer



Krystian Matusz

### About me

**Krystian** is very ambitious, well rounded, inspiring and visionary **Passionate** Data Scientist, Technology Evangelist and Business Inspiration Officer with a broad spectrum of technical and business domain expertise, and proven success in bringing **measurable added value** to companies ranging from startups to corporations.

Experienced across many

**industries:** Healthcare , Insurance , Banking , Security , Education

**domains:**

- Data Science, Market and Operations Research, Machine Learning, Computer Science, Business Intelligence
- Statistics, Product Development, Growth Hacking
- Customer Experience, IoT, AoT etc.

**and roles :** Data Scientist, Founder , Architect, Manager, Strategic Consulting Advisor.

**Highly certified** (all SAS certifications) and **dedicated** true positive Professional. The first person who passed all of SAS certifications in the world (2017/2018).

**Enabler**, who support people to make **brilliant** decisions. Hard-working personality, who believes that to dream BIG is a matter of choice but to reach that dream is a matter of discipline. He works smart and extremely hard to give people the true inspiration & power-to-know through the right actionable insights and Advanced Analytics.

### Main objective of my Research presented here in Washington D.C

**Predicting Malware Persistence**

through

Windows Registry

**Behavioral Advanced Cybersecurity Analysis**



1. Topic & Abstract

2. Agenda

3. About me

4. Facts: KPIs/KRIs

5. Malware - def.

6. Malware - types

7. Research - goal

8. Methodology

9. Architecture

10. Insights

11. Conclusions



# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

- 1. Topic & Abstract
- 2. Agenda
- 3. About me
- 4. Facts: KPIs/KRIs**
- 5. Malware - def.
- 6. Malware - types
- 7. Research - goal
- 8. Methodology
- 9. Architecture
- 10. Insights
- 11. Conclusions

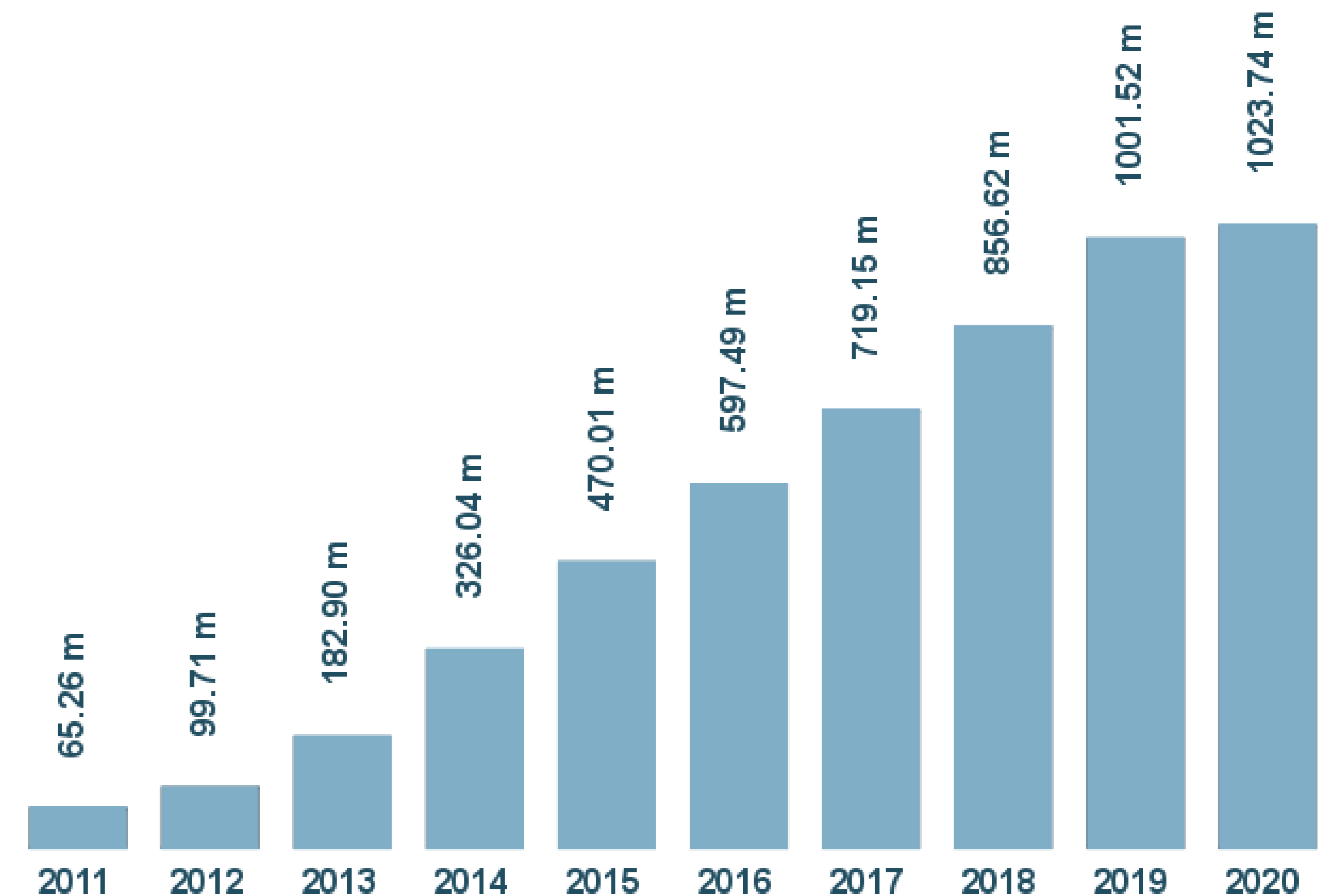
- **Worldwide**, information security spending in 2019 exceeded \$124 billion (Gartner). The average cost of cybercrime has increased by 72% in the last five years to \$12 million and continues to grow.
- Estimates show **near-constant frequency of attacks every 39 seconds, affecting one in three Americans.**
- **Malware plays the leading role in these attacks.** As a malicious program, intended to infiltrate or destroy PCs and networks without users' knowledge, malware is particularly dangerous today, when the penetration of the market by Windows-based systems is **above 78%** creating an attractive target.

## A few frightening statistics:

- **More than 70% of Americans** concerned about having their personal data stolen
- Approximately ~2 billion (1,769,185,063) records have been **stolen** in January 2019
- The average cost of breach(es) is **rapidly increasing.**
- Malware is a **serious threat** to your company's bottom line and is taking **consistently and increasingly** large Economic toll. At the current trajectory, the total cost can reach \$6 trillion by 2021.
- **One out of every thirteen Internet requests** like searches, links leads to the Malware
- The most common way for the malware to be delivered is through **emails** (usualy by phishing emails)
- Threat landscape is becoming increasingly divided between **consumer** and **business** targets. Over 70% of them are **unprepared** to face down even the most basic attempt at a security breach.
- From a business standpoint: much more diverse and sophisticated Malware is coming out. **Your phone** is at now a major target.
- **Awareness** is the Best Defense Against Malware

## Did you know? Presented total numer of Malware

### Total malware



Last update: February 15, 2020

Copyright © AV-TEST GmbH, www.av-test.org



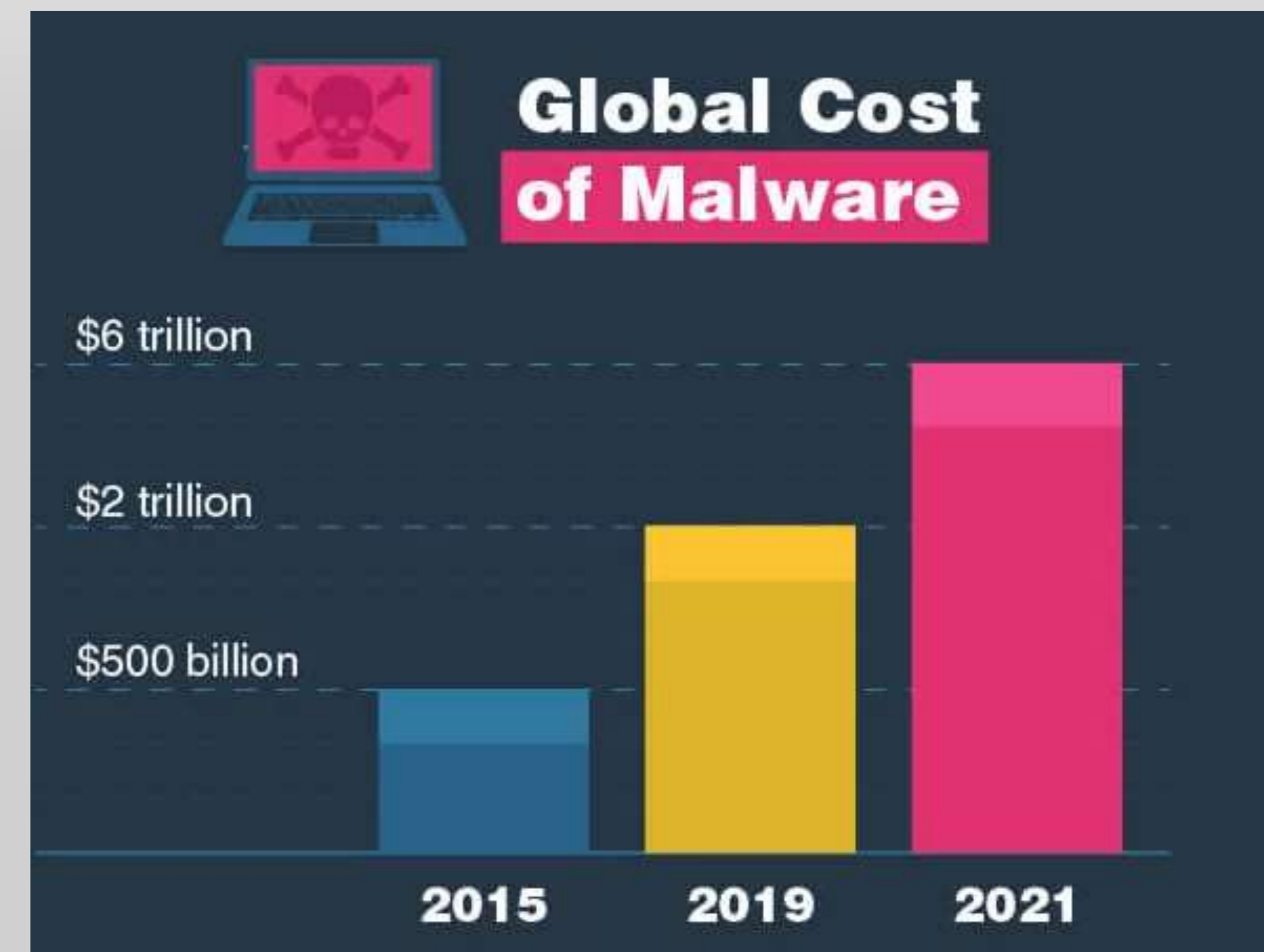
# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

**Malware** is a **malicious software** which is designed specifically with the purpose of infiltrate, gaining access, cause damage or disruption of computer systems and services with no knowledge and consent of the owners. **It is one of the most dangerous threats we currently face.**

Malwarebytes defines Malware, or "malicious software," as an umbrella term that describes **any malicious program or code that is harmful** to systems. Hostile, intrusive, and **intentionally** nasty, malware seeks to invade, damage, or disable computers, computer systems, networks, tablets, and mobile devices, often by taking partial control over a device's operations. Like the human flu, it interferes with normal functioning.



## A little bit more about what Malware is and what it can do

Malware is **intended to infiltrate or destroy PCs and networks** without users' knowledge. It is particularly dangerous today, where the penetration of the market by Windows-based systems is above 78%, creating this OS as an attractive target.

Malware might use **genetic** and **polymorphic** obfuscation or code packing, so the behavioral-based approach is an **effective way** to connect the Cybersecurity and Data Science domains in order **to increase** the overall level of security and awareness.

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
5. Malware - def.
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
10. Insights
11. Conclusions



**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
- 5. Malware - def.**
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
10. Insights
11. Conclusions



<https://threatmap.checkpoint.com/>

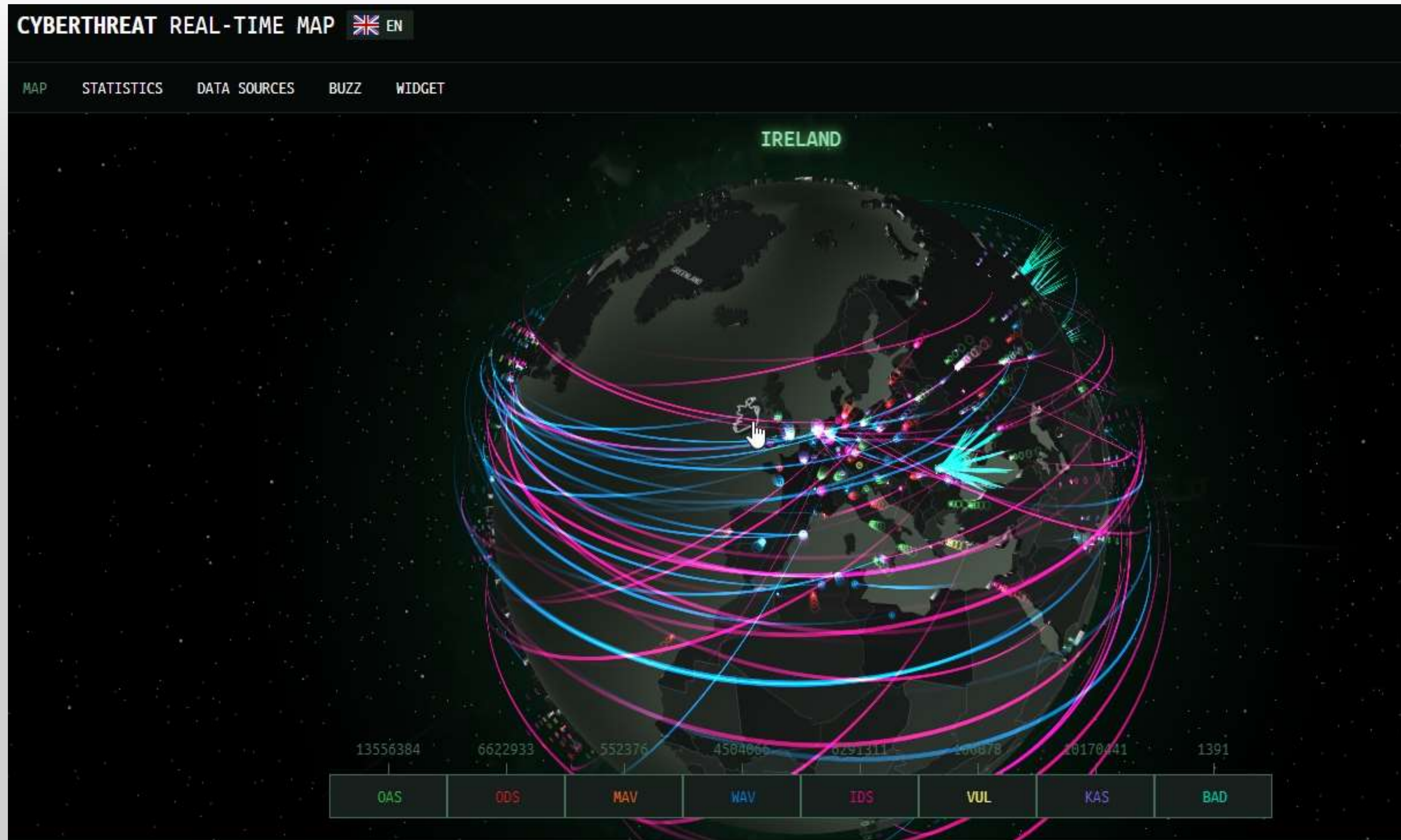


# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
- 5. Malware - def.**
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
10. Insights
11. Conclusions



<https://cybermap.kaspersky.com/>



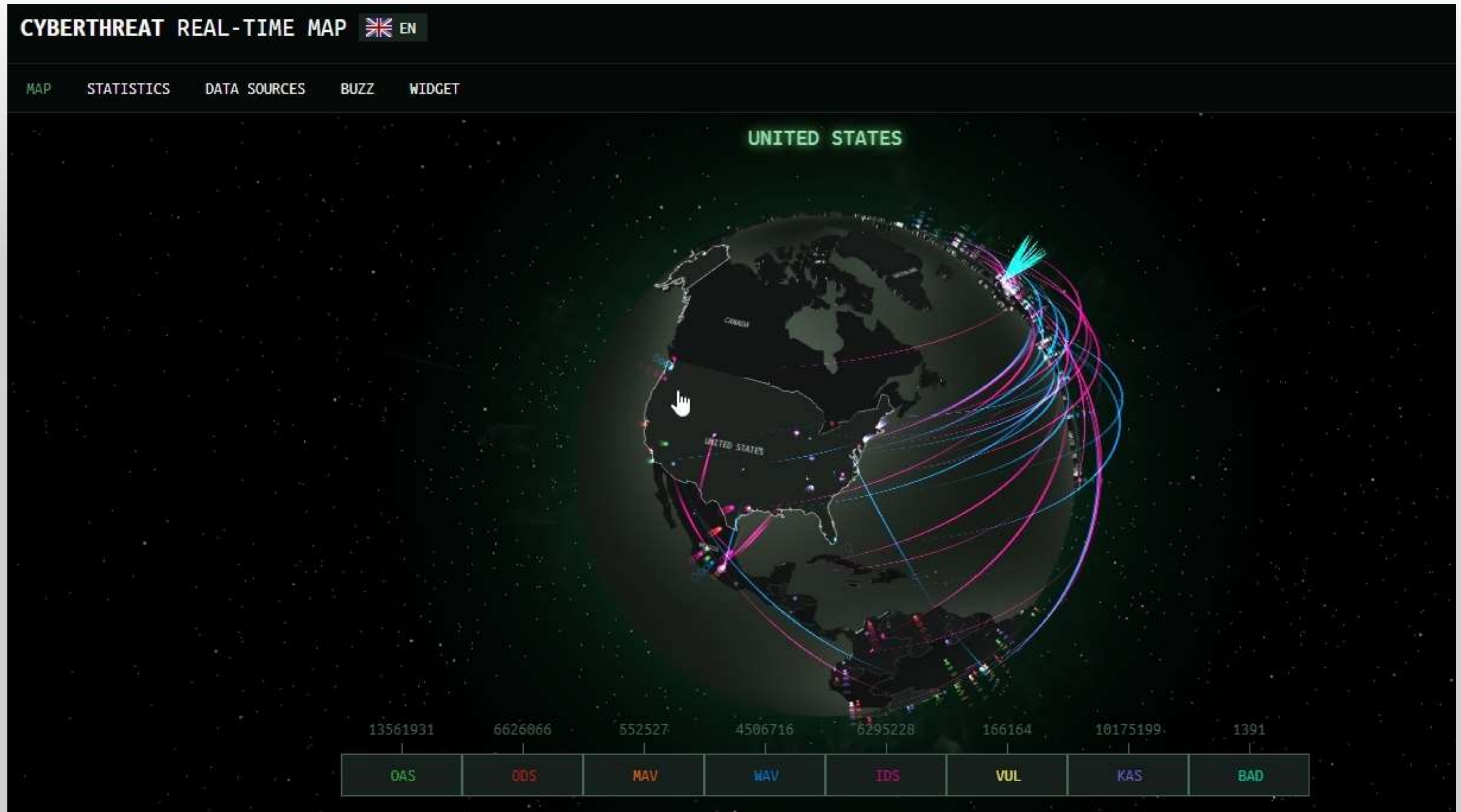


# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
- 5. Malware - def.**
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
10. Insights
11. Conclusions



<https://cybermap.kaspersky.com/>

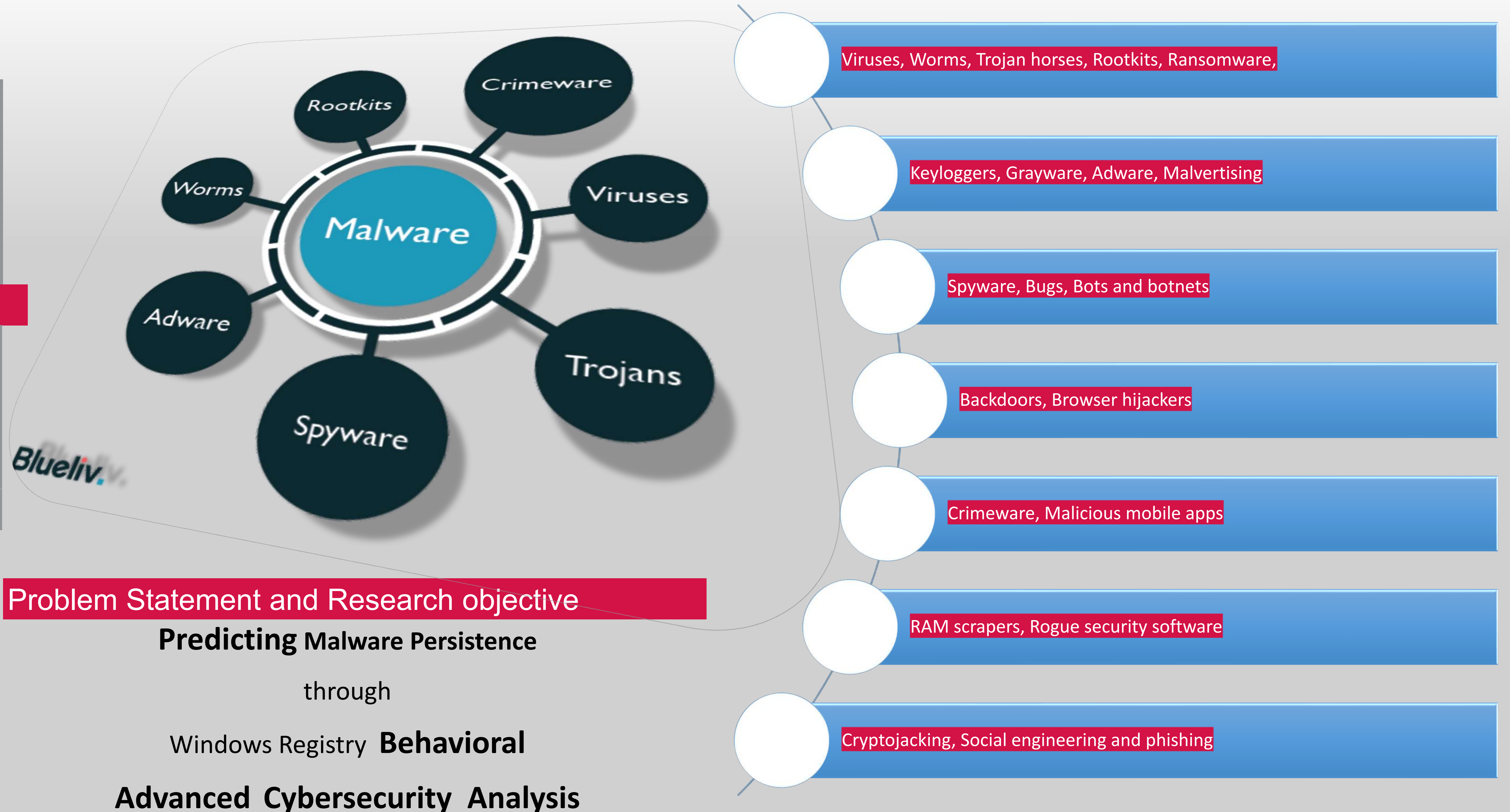


# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

- 1. Topic & Abstract
- 2. Agenda
- 3. About me
- 4. Facts: KPIs/KRIs
- 5. Malware - def.
- 6. Malware - types**
- 7. Research - goal
- 8. Methodology
- 9. Architecture
- 10. Insights
- 11. Conclusions





**Krystian Matusz, Passionate Data Scientist**  
Individual Contributor: Technology Evangelist and Business Inspiration Officer

- 1. Topic & Abstract
- 2. Agenda
- 3. About me
- 4. Facts: KPIs/KRIs
- 5. Malware - def.
- 6. Malware - types
- 7. Research - goal**
- 8. Methodology
- 9. Architecture
- 10. Insights
- 11. Conclusions

## >>Research objective statement

**Predicting Malware Persistence**

through

Windows Registry

**Behavioral**

**Advanced**

**Cybersecurity**

**Analysis**





**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

## CRISP-DM and SEMMA as my leading methodologies

- 1. Topic & Abstract
- 2. Agenda
- 3. About me
- 4. Facts: KPIs/KRIs
- 5. Malware - def.
- 6. Malware - types
- 7. Research - goal
- 8. Methodology**
- 9. Architecture
- 10. Insights
- 11. Conclusions

### 1. CRISP-DM

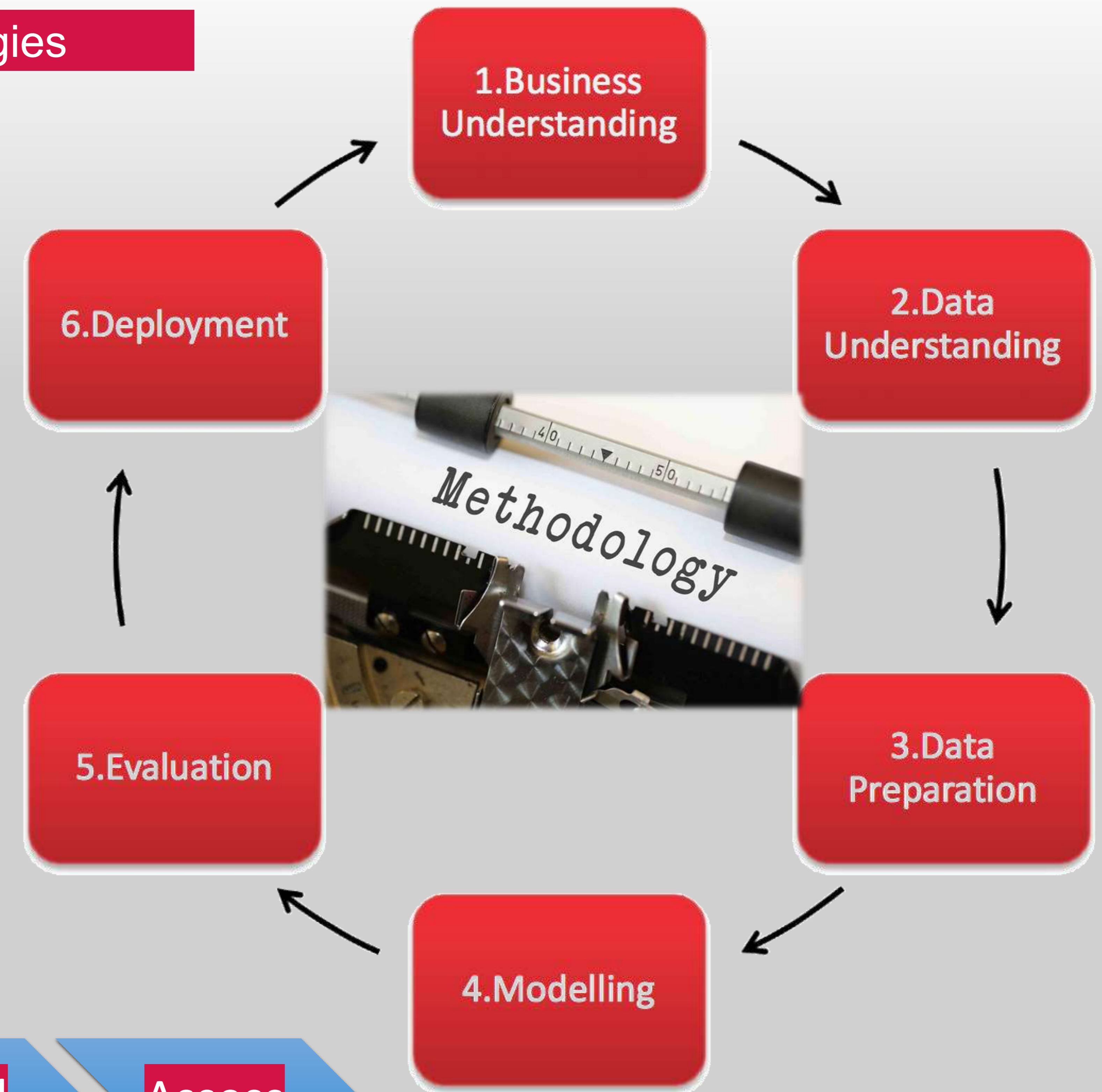
– focused on business understanding

- Cross-industry standard process for data mining, that standardizes a analytics process
- Currently it is the most tasks commonly used methodology for analytics, data mining and data science projects
- Six phases in the process
- These phases are **iterative**
- Each step has own deliverables and tasks

### 2. SEMMA

– focused on modelling

- The process is a list of best practices
- Five phases in the process
- logical organization of the functional tool set tools for carrying out the core tasks of data mining





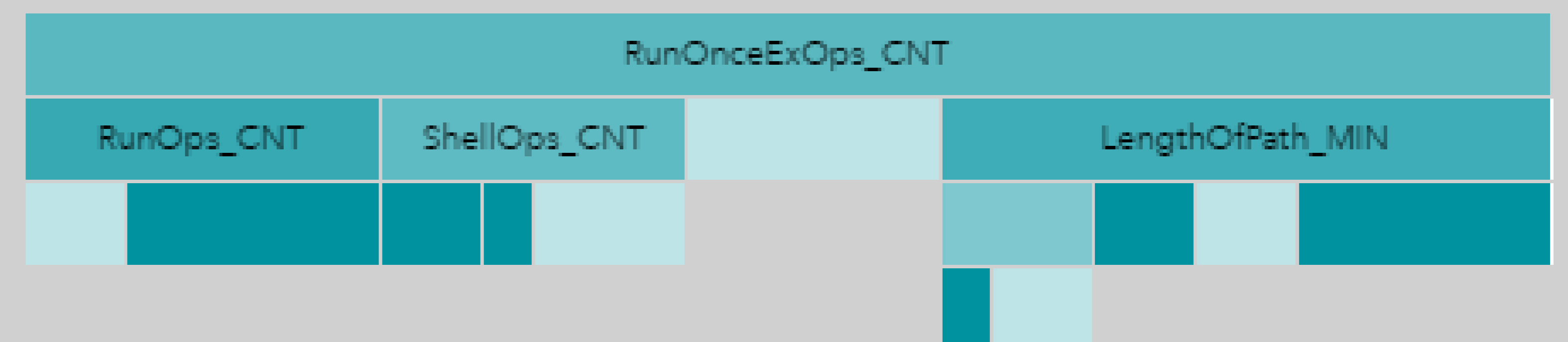
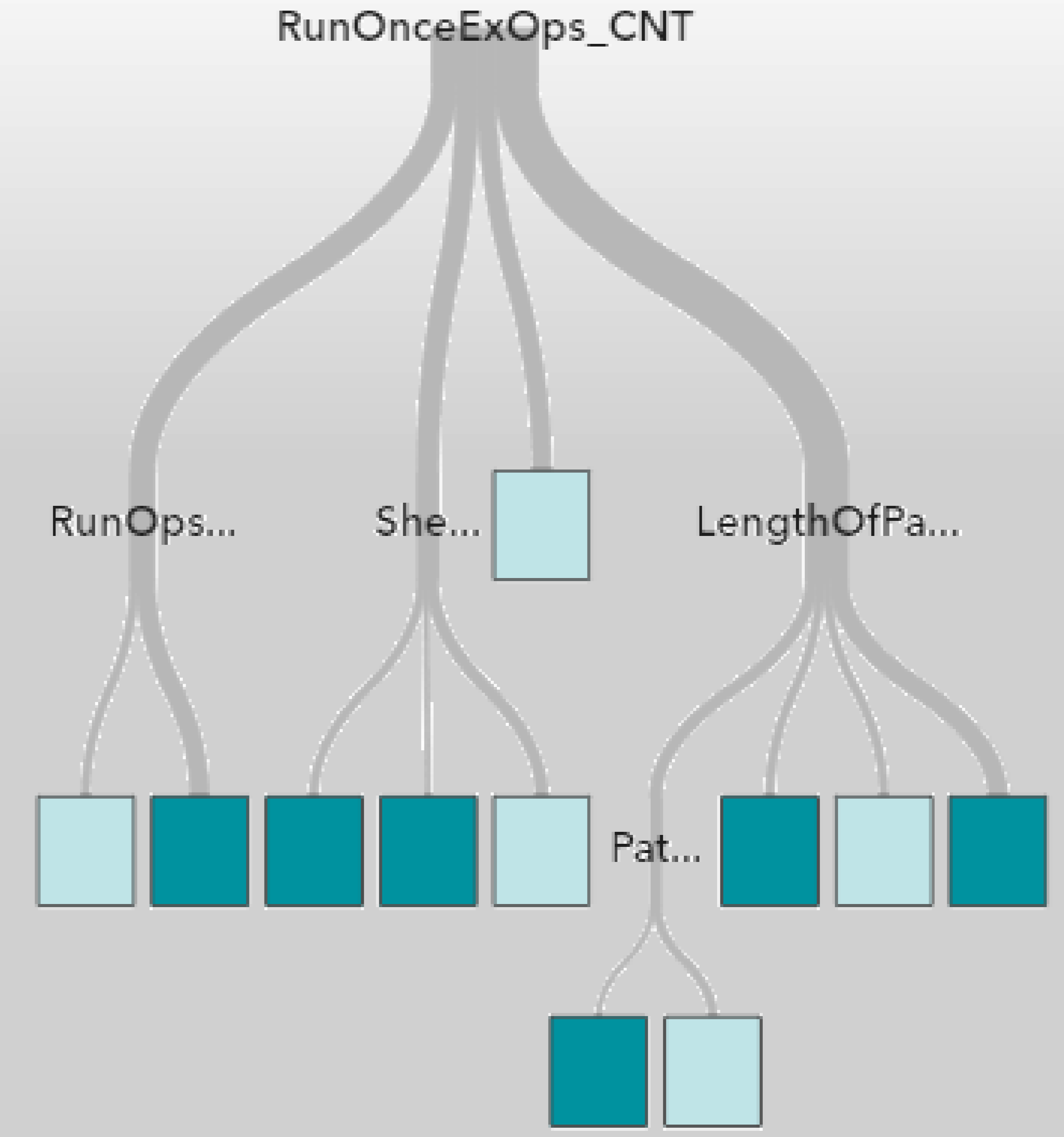
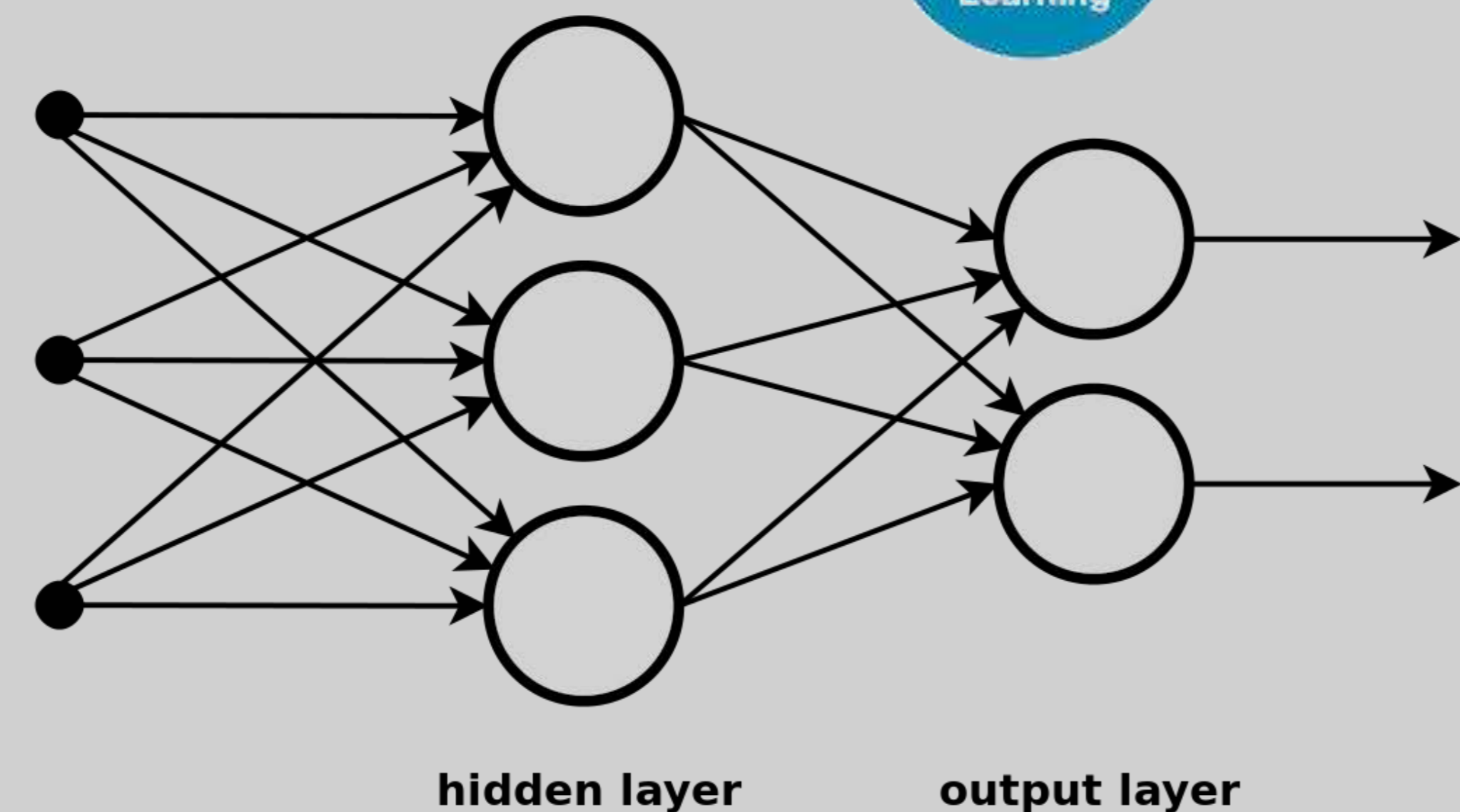
# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
5. Malware - def.
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
10. Insights
11. Conclusions

## Problem Statement and Research objective





**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

## How time-to-insight Is Driving Business Innovation and Security?

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
5. Malware - def.
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
- 10. Insights**
11. Conclusions

**Data Science + Machine Learning + Advanced Analytics**

and

**Human Intelligence**

**can** successfully

**Predict Malware Persistence**

through

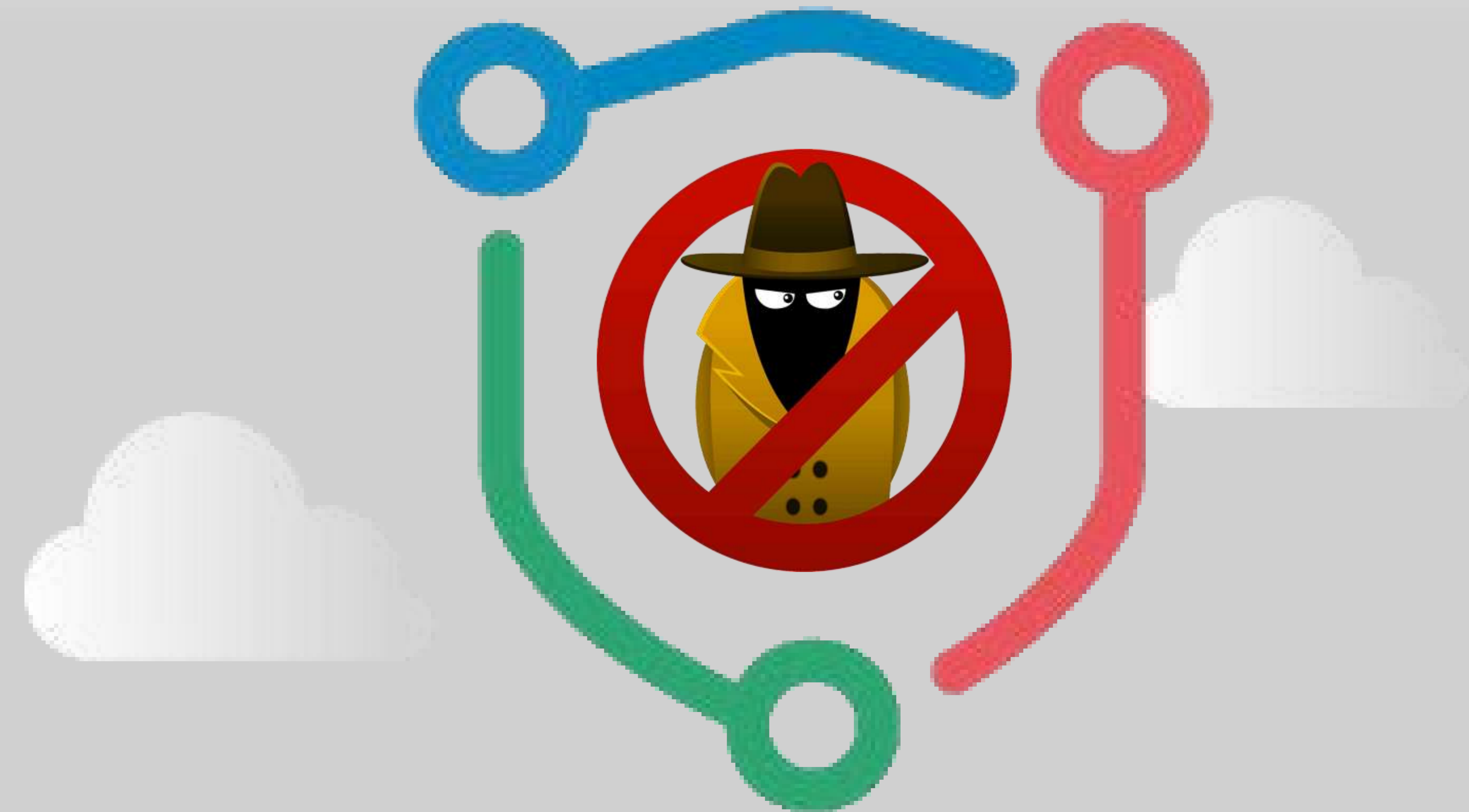
Windows Registry

**Behavioral**

**Advanced**

**Cybersecurity**

**Analysis.**



**However do not forget about security by design principles.**

**It is not the product – it is the process!**



# Predicting Malware Persistence through Windows Registry Behavioral Advanced Cybersecurity Analysis

**Krystian Matusz, Passionate Data Scientist**

Individual Contributor: Technology Evangelist and Business Inspiration Officer

1. Topic & Abstract
2. Agenda
3. About me
4. Facts: KPIs/KRIs
5. Malware - def.
6. Malware - types
7. Research - goal
8. Methodology
9. Architecture
10. Insights
- 11. Conclusions**

- **Malware** plays the leading role in most of the attacks causing significant damages.
- Technology is not intended to entirely replace human intelligence - but the human **always** will be the weakest element of the system.
- Security principles and user education **is crucial!**
- I have presented the possibility of establishing the persistence of malware in operating systems based on behavioral analysis of the malware in the context of analyzing an internal Windows Registry. This approach was been **reinforced** by using Machine Learning Predictive Models built in SAS® Viya® to evaluate **at a scale** how likely the malware will survive the restart of an infected operating system.
- Malware might use **genetic** and **polymorphic** obfuscation or code packing, so the behavioral-based approach is an **effective way** to connect the Cybersecurity and Data Science domains in order **to increase** the overall level of security and awareness.
- Despite of the fact that this is fascinating and challenging area, do not forget that your security and your organization starts from **YOU!**

## Useful Resources

- <https://www.av-test.org/en/statistics/malware/>
- <https://www.tigermobiles.com/blog/malware-statistics/>
- <https://threatmap.checkpoint.com/>
- <https://cybermap.kaspersky.com/>
- <https://www.malwarebytes.com/malware/>
- <https://www.upguard.com/blog/types-of-malware>
- [https://resources.malwarebytes.com/files/2020/02/2020\\_State-of-Malware-Report.pdf](https://resources.malwarebytes.com/files/2020/02/2020_State-of-Malware-Report.pdf)

## Acknowledgements

"I Have A Dream.." - I would like to express my sincere thanks to Dina and Speakers Committee for positive response to my ambitious technical Call for Papers submission and for giving me opportunity to share my research, knowledge and **passion (!)** in such important and challenging topic, being here in exceptional place: Washington D.C. and time during SAS Global Forum 2020 with you all across 6500 delegates and experts around the globe.

Go n-éirí an bóthar leat.

## Inspirational Quotes

- *The difference between a dreamer and a visionary is that a dreamer has his eyes closed and a visionary has his eyes open.*
- *The best things that capture your imagination are ones you hadn't thought of before and that aren't talked about in the news all the time.*



## Contact details

Your comments and questions are valued and encouraged.  
You can contact me at:

- [MatuszKrystian@Gmail.com](mailto:MatuszKrystian@Gmail.com)
- <https://www.linkedin.com/in/krystianmatusz/>

The background of the banner is a scenic view of the Washington Monument at dusk, reflected in the water of the Tidal Basin. The sky is a mix of blue, purple, and pink. In the foreground, there are cherry blossom trees with pink and white flowers. A dark teal rectangular box is centered over the image, containing the event title in white and teal text.

# SAS<sup>®</sup> GLOBAL FORUM 2020

USERS PROGRAM

MARCH 29 - APRIL 1 | WASHINGTON, DC | #SASGF

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.