

GRONKIEWICZ PATRYK

KRAWIEC PIOTR

MORSKYI VITALII

# DETEKCJA PODEJRZANYCH ADRESÓW URL

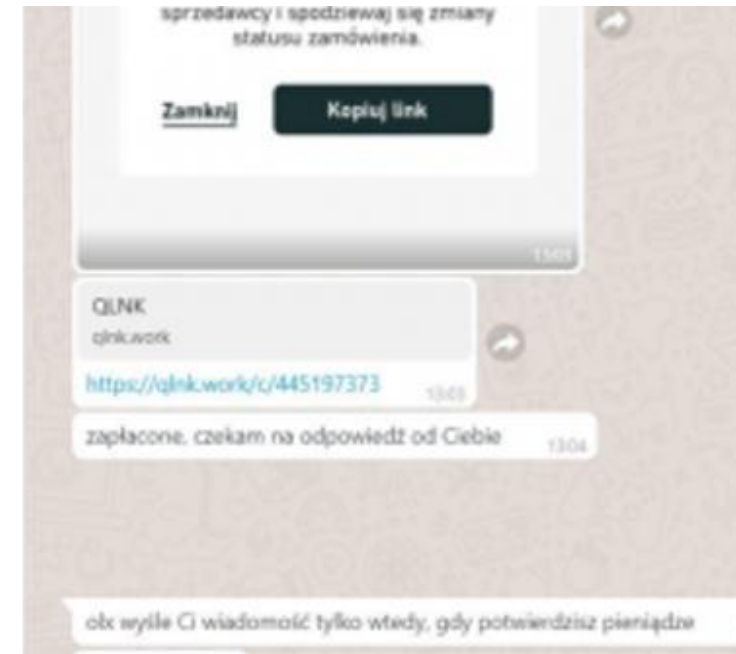
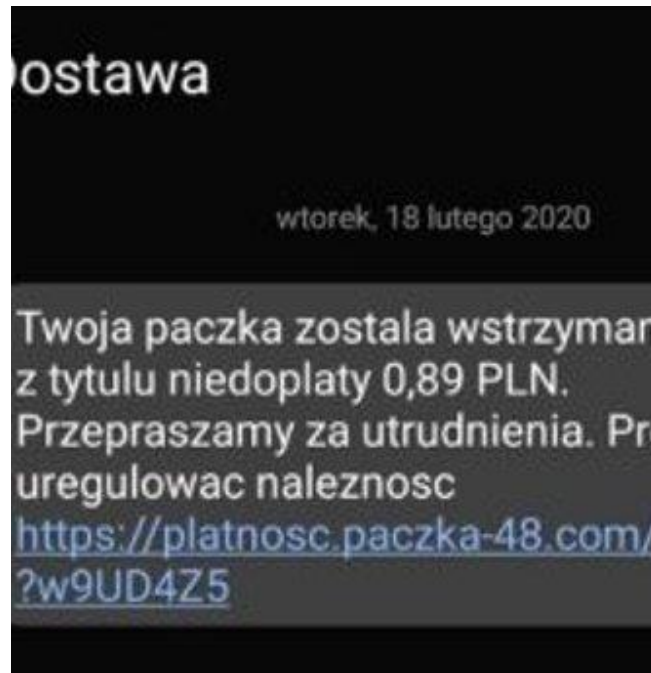
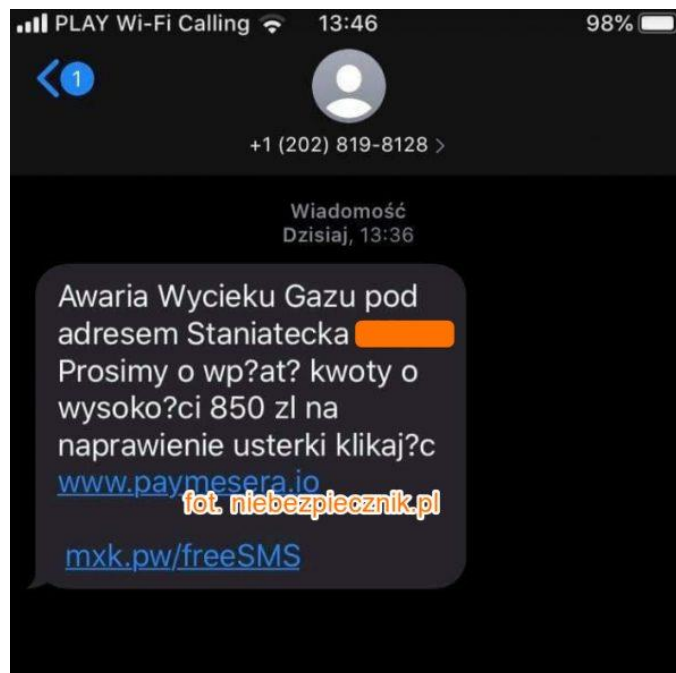


itempool.com/  
pgronkievitz/live



EKSPERYMENT

**PROBLEM**



# PROBLEM

Ź R Ó D Ł O :  
N I E B E Z P I E C Z N I K . P L



## Klawiatura Apple Magic Keyboard A1314

Zdobądź fundusze

**Podaj swoją kartę, aby otrzymać środki:**

Numer karty\*

XXXX XXXX XXXX XXXX

Data Ważności Karty\*

mn/ yy

Kod CVV\*

\*\*\*

**W sumie:**

Wartość towaru:

**160 PLN**

**Otrzymać:**

**160 PLN**

Bezpieczeństwo gwarantowane



Klawiatura Apple Magic Keyboard A1314

**160 PLN**

**OLX-PAYCORE.CLUB/PAYMENT/SELL/**

**PROBLEM**

**ŹRÓDŁO:**  
**NIEBEZPIECZNIK.PL**





olx paycore



Wszystko

Zakupy

Wiadomości

Grafika

Mapy

Więcej

Narzędzia

Okolo 2 160 wyników (0,58 s)

<https://pomoc.olx.pl> › articles › 360016098540-Płatności-...

## Płatności OLX - OLX - Centrum Pomocy

W przypadku tej strony informacje nie są dostępne.

Dowiedz się dlaczego

<https://blog.olx.pl> › 2020/01/20 › uwaga-na-falszywe-li... ▼

## Uwaga na SMSy z fałszywymi linkami do usługi wysyłki przez ...

20 sty 2020 — Ten wpis ma już kilka miesięcy – sprawdź nasze na bieżąco uaktualniane ostrzeżenie przed linkami do fałszywych stron Przesyłek OLX – "Uwaga ...

<https://www.olx.pl> › oferty › q-pay-za ▼

## Ogłoszenia - OLX.pl - pay za

Gry PS3 Far Cry Resident Call of Duty Pay Day Sniper Uncharted Driver. Gry i Konsole » Gry.

30 zł. Żory wczoraj 16:54 ...

# PROBLEM

Ź R Ó D Ł O :  
N I E B E Z P I E C Z N I K . P L

**CZY MOŻE TO ZROBIĆ  
KOMPUTER?**

**TAK!\***



CERT Polska



Wykaz stron internetowych  
podmiotów publicznych

**OTWARTE ŹRÓDŁA DANYCH**



## > Lista ostrzeżeń przed niebezpiecznymi stronami \_

23 marca 2020 | CERT Polska | #phishing

Wspólnie z operatorami telekomunikacyjnymi rozpoczynamy walkę ze stronami wyludzającymi dane osobowe, dane uwierzytelniające do kont bankowych i serwisów społecznościowych. Lista ostrzeżeń rozpoczyna funkcjonowanie w odpowiedzi na znaczący wzrost liczby wyludzeń danych w związku z treściami dotyczącymi epidemii koronawirusa. W ramach współpracy będziemy publikować listę domen wykorzystywanych do nadużyć – każdy może ją pobrać i wdrożyć w swoich systemach bezpieczeństwa do blokady złośliwych treści. Operatorzy, którzy przystąpili do współpracy, będą uniemożliwiali dostęp do zidentyfikowanych i przeanalizowanych przez nas domen.

Strony wyludzające dane osobowe oraz dane uwierzytelniające są obecnie zjawiskiem masowym, dotyczącym różne grupy



### Uwaga! Ta strona stanowi zagrożenie

Może ona wyludzać dane osobowe, dane uwierzytelniające do kont bankowych lub serwisów społecznościowych. W trosce o Twoje bezpieczeństwo dostawca internetu powstrzymał próbę ataku poprzez tę stronę.

Przypominamy:

- Dokładnie sprawdzaj** adres i wygląd strony, na której podajesz dane logowania, dane osobowe czy karty płatniczej.
- Nie działaj pod presją czasu**, uważaj na wszelkie wiadomości, które skłaniają do działania natychmiast.
- Weryfikuj źródło** informacji zanim podejmiesz działania na jej podstawie lub ją powielisz.
- Nie jesteś pewien czy dana wiadomość jest prawdziwa? **Skontaktuj się** z rzekomym nadawcą innym znanym kanałem i/lub poszukaj potwierdzenia informacji w innych źródłach.
- Zgłaszaj do CERT Polska** każdą podejrzaną stronę, a także wiadomości email i SMSy, które mogą wyludzać dane. Formularz znajdziesz na stronie <https://incydent.cert.pl>

Oficjalne informacje i komunikaty na temat koronawirusa znajdziesz na stronie: <https://gov.pl/koronawirus>.

Lista ostrzeżeń zawierająca wykaz witryn stanowiących zagrożenie znajduje się na stronie [https://cert.pl/ostrzezenia\\_idosifing](https://cert.pl/ostrzezenia_idosifing)

# OTWARTE ŹRÓDŁA DANYCH

ŹRÓDŁO: CERT.PL

Jesteś w sekcji: [Strona główna](#) / [Dane](#) / [Statystyki rejestru domen .gov....](#) / [Lista nazw domen .gov.pl z usług...](#)

## Lista nazw domen .gov.pl z usługą www

Lista nazw zarejestrowanych w domenie .gov.pl, które posiadają skonfigurowaną usługę www na dzień 2021-07-14.

Poziom otwartości danych: ★★☆☆☆

Data udostępnienia: 15 lipca 2021, 10:34

Dane na dzień: 14 lipca 2021

Typ zasobu: plik

Warunki wykorzystywania: [Rozwiń](#)

Pobierz: [CSV \(23kB\) >>](#)

Zobacz: [JSON >>](#)

http://hans:geheim@www.example.org:8080/demo/example.cgi?land=de&stadt=aa

\\_/\_ \\_/\_ \\_/\_ \\_/\_ \\_/\_ \\_/\_ \\_/\_

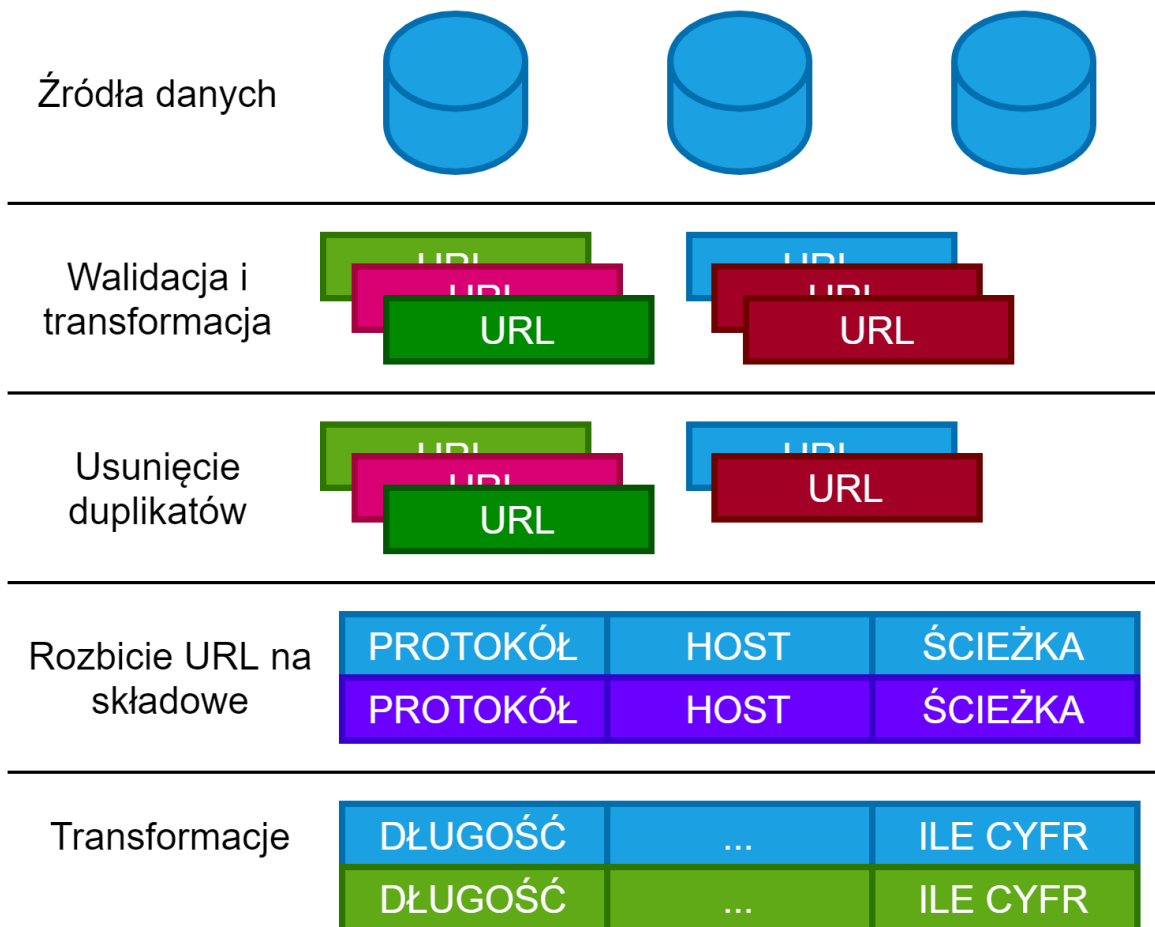
| | | | | | |

login | host port ścieżka ścieżka

protokół hasło (adres serwera) do zasobu wyszukiwania

DANE

ŹRÓDŁO: WIKIPEDIA

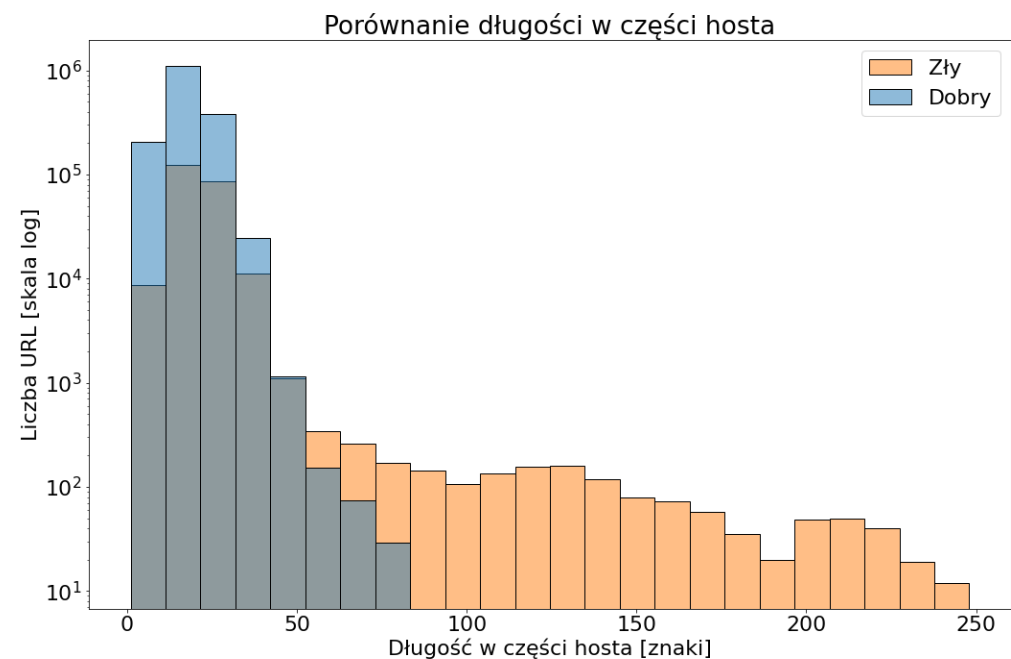
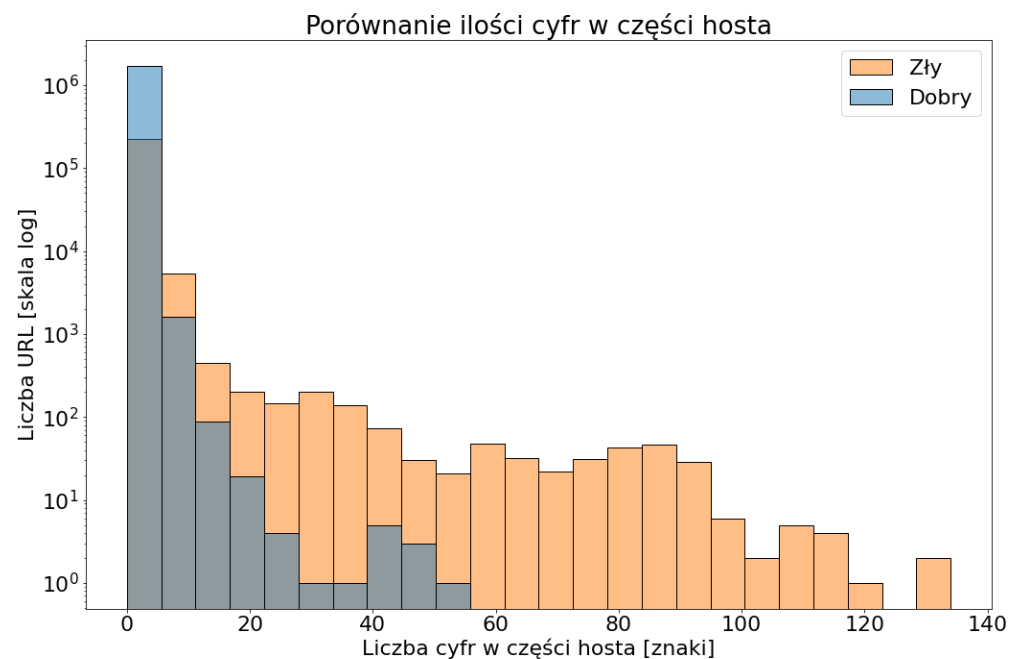


# PRZYGOTOWANIE I ZEBRANIE DANYCH

	host	label
	smartshares.co.nz	0.0
	joyproject.it	1.0
	www.mtrwestern.com	0.0
	phonegap.matainja.com	1.0
	eng.infootball.ru	0.0
	chevrolet.com.ar	0.0
	ventures.pl.ru	0.0
	vinashare.vn	0.0
	qwick.com	0.0
	mulliganformayor.blogspot.com	0.0

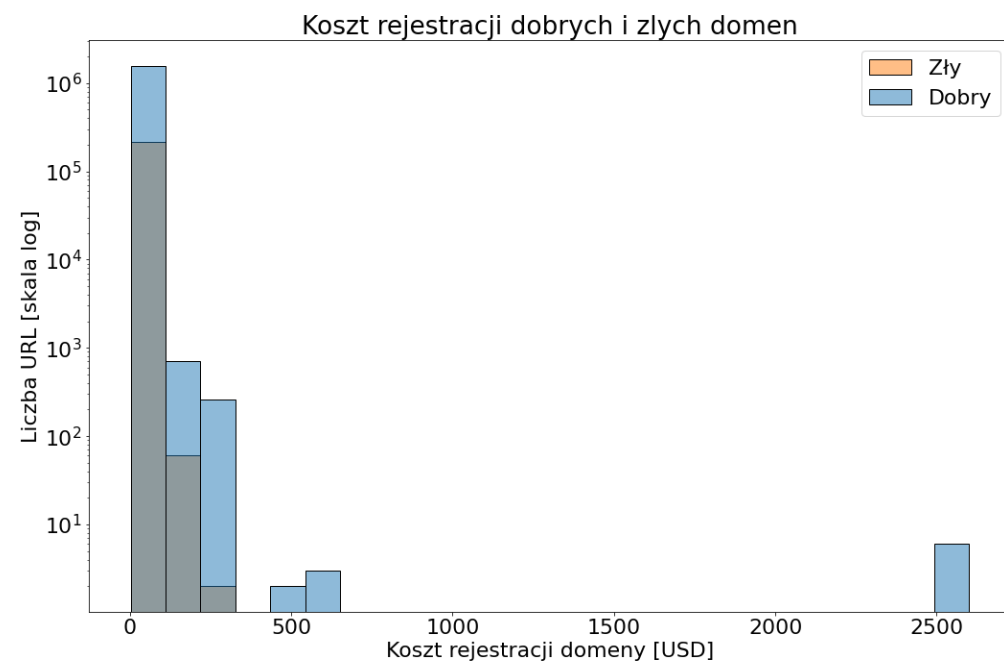
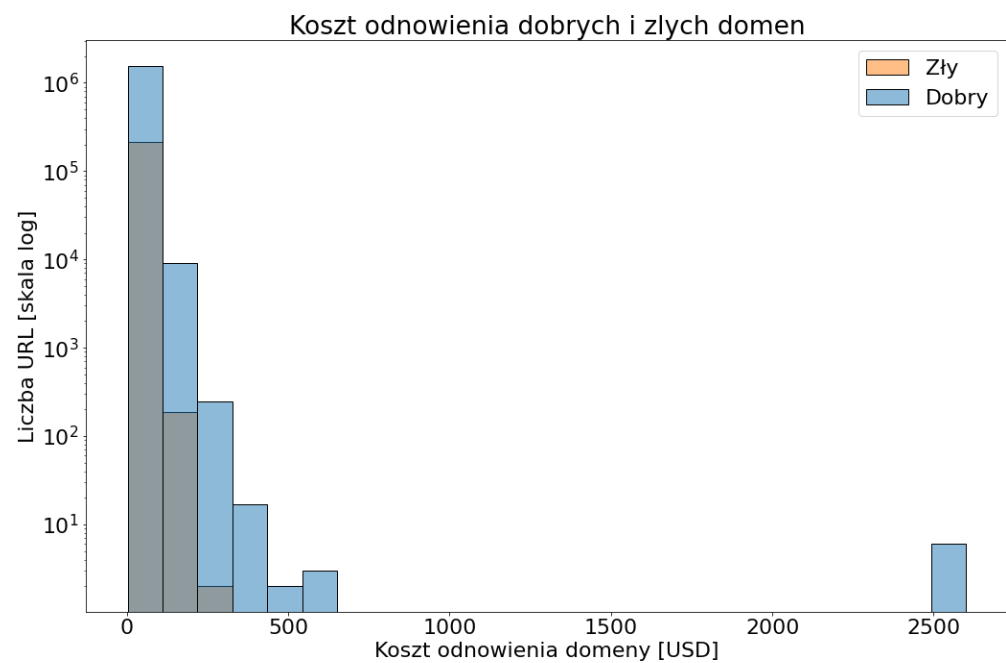
	host	label	host_len	host_digit	Renewal Price
	'www.njofficeol.com	1.0	19.0	0.0	13.99
	restaurantdepot.com	0.0	19.0	0.0	13.99
	www.wrma.org	0.0	12.0	0.0	14.99
	www.pdinfo.com	0.0	14.0	0.0	13.99
	www.kofc8600.org	0.0	16.0	4.0	14.99
	enterpleasure.tripod.com	0.0	24.0	0.0	13.99
	backyardboxing.ru	0.0	17.0	0.0	28.99
	www.stuffed-animals.com	0.0	23.0	0.0	13.99
	www.gfbkonline.com	1.0	18.0	0.0	13.99
	www.fbbmilwaukee.com	0.0	20.0	0.0	13.99

DANE



WYKRESY WYBRANYCH METRYK





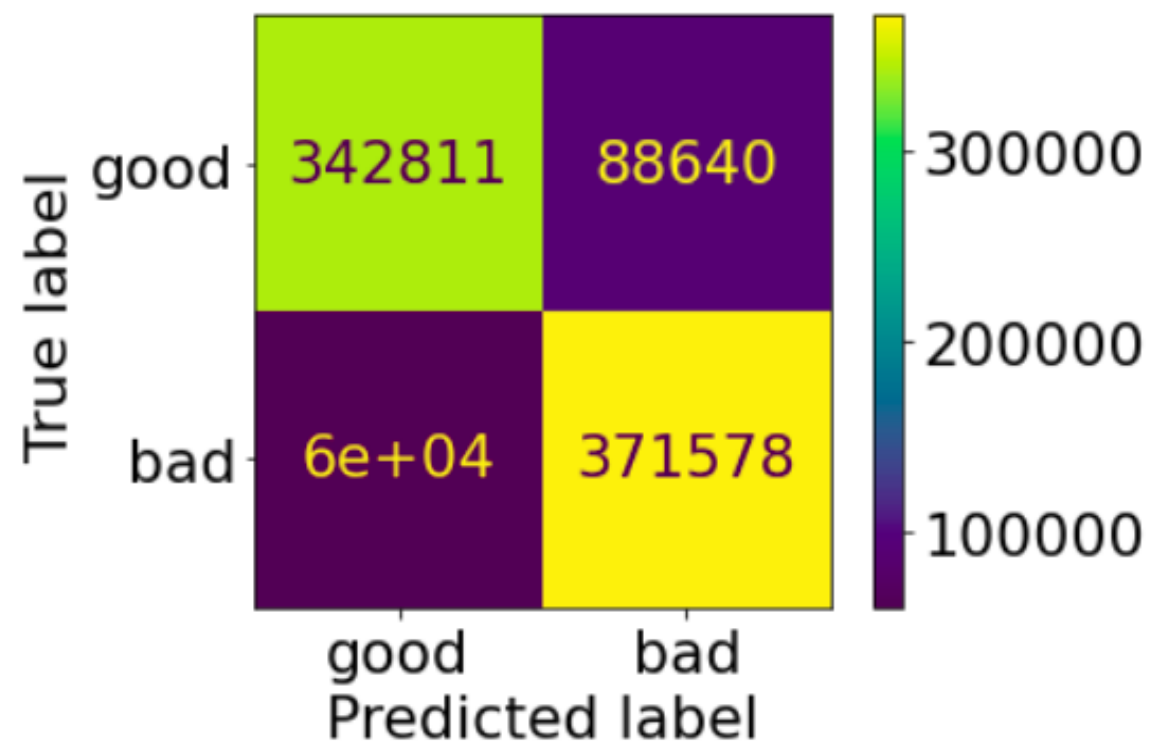
WYKRESY WYBRANYCH METRYK

**MODELE**

Bazujące na:

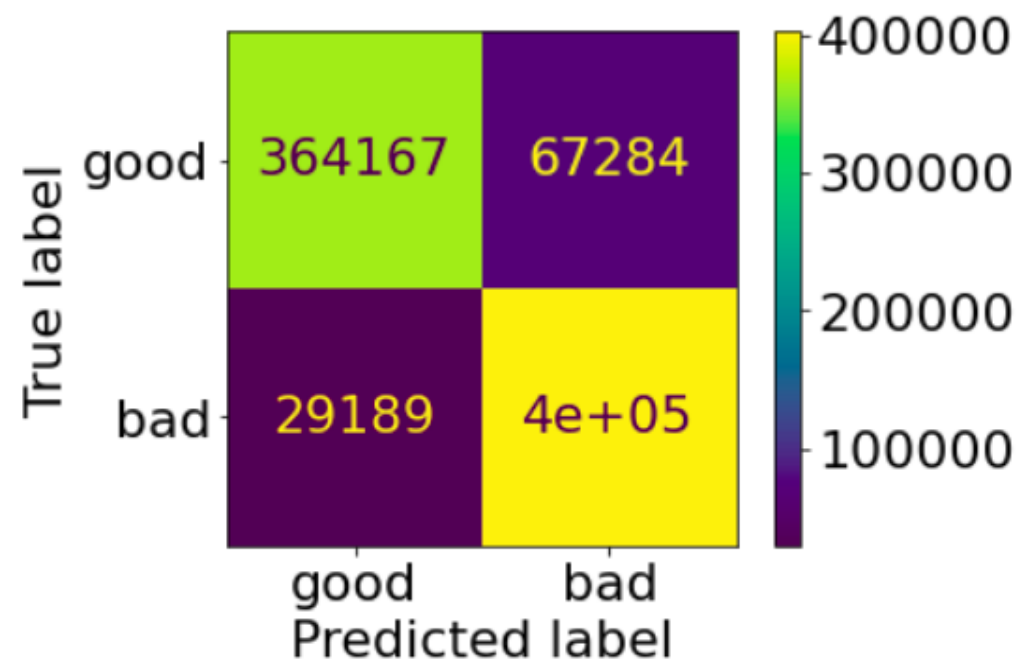
- Długości hosta
- Długości pierwszej subdomeny
- Długości całego URL

	precision	recall	f1-score	support
0	0.85	0.79	0.82	431451
1	0.81	0.86	0.83	431452
accuracy			0.83	862903
macro avg	0.83	0.83	0.83	862903
weighted avg	0.83	0.83	0.83	862903



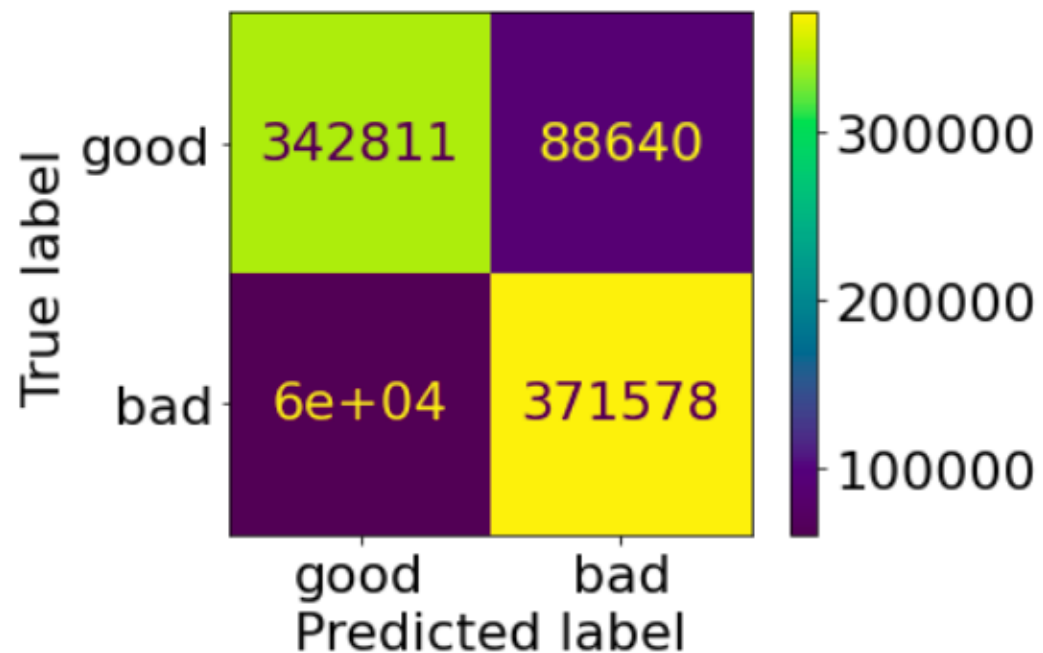
# MODELE NAIWNE

	precision	recall	f1-score	support
0	0.93	0.84	0.88	431451
1	0.86	0.93	0.89	431452
accuracy			0.89	862903
macro avg	0.89	0.89	0.89	862903
weighted avg	0.89	0.89	0.89	862903

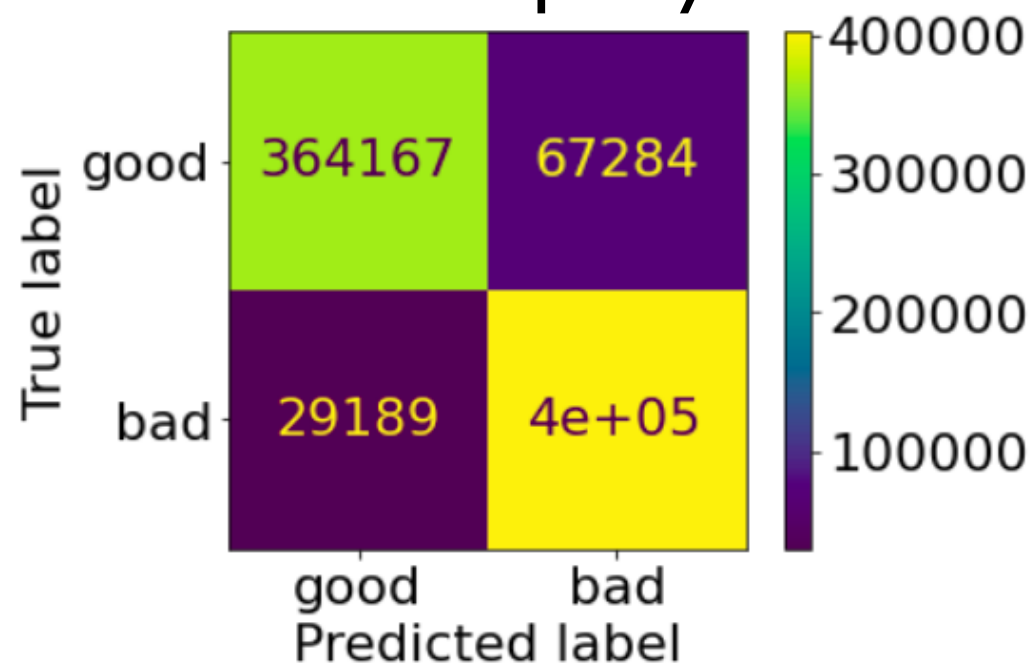


MODEL PEŁNY

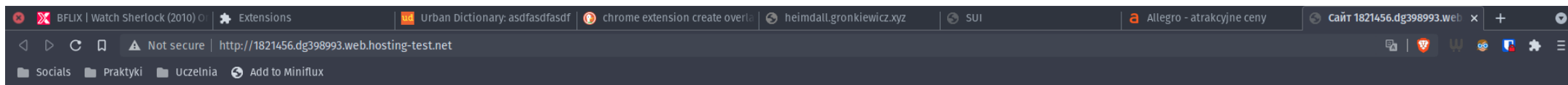
### Model naiwny



### Model pełny



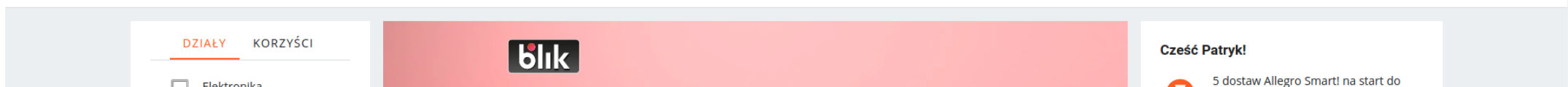
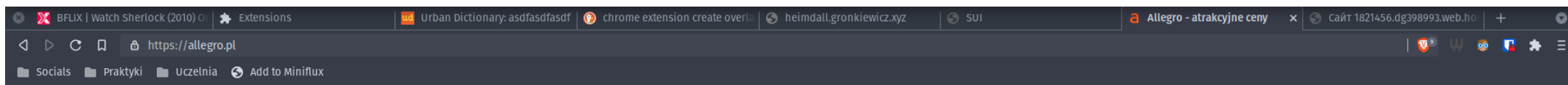
# PORÓWNANIE MODELI



**This domain probably is malicious!**

Чтобы здесь оказался Ваш сайт, необходимо добавить его в [панель управления](#) в разделе «Мои сайты».

Если Вы недавно добавили сайт в панель управления - подождите 15 минут и ваш сайт начнет работать.



# ZASTOSOWANIE - WTYCZKA



URL	Rodzaj url	Przewidywania
summervilleminiatureworkshop.com	dobry	dobry
skullsecurity.org	dobry	dobry
magnol.ru	dobry	dobry
knml.prz.edu.pl	dobry	dobry
olx.pl-informacje.info	zły	zły
pl-przelew.club	zły	zły
castitu.com	zły	dobry
tracklostdevice.com	zły	zły
vodesh.com	zły	dobry

Precyzja: 77%

# ESKPERYMENT - WYNIKI

```
In[8]: model.predict("https://w.prz.edu.pl/")  
Out[8]: 'good'
```

