Analiza wycieku danych osobowych z ALAB Laboratoria

10 grudnia 2023

Niniejszy artykuł jest publikowany na prośbę autorki, która podpisała się pseudonimem Valkyria. Autorka artykułu nie chce publikować tej analizy pod swoim imieniem i nazwiskiem z uwagi na fakt, że pracodawcy negatywnie oceniają osoby, które zwracają uwagę na kwestie RODO oraz na bezpieczeństwo systemów informatycznych. Z powodu takiego, a nie innego podejścia pracodawców, do tej pory żądają oni nadmiernych danych osobowych, które później wykorzystują do węszenia w życiu prywatnym pracowników.

Stan taki prowadzi do przyjmowania do pracy osób niekompetentnych i niewykwalifikowanych, bo przecież te nie będą zadawały zbędnych pytań. System bez niezbędnych zabezpieczeń kryptograficznych powstanie też szybciej niż taki, gdzie na etapie jego projektowania zastosowano odpowiednie zabezpieczenia. Tego się później nie da nadrobić, ponieważ całość trzeba by projektować od nowa – często również z przebudową front-endu. Na to zazwyczaj później nie będzie już czasu.

Wyciek danych z 19 listopada 2023 roku

W dniu 19 listopada 2023 roku, a właściwie trochę wcześniej, doszło do wycieku danych medycznych prawdopodobnie wszystkich Polaków, którzy w latach 2017-2023 robili badania laboratoryjne, cytologiczne i inne w warszawskiej firmie ALAB Laboratoria. Wyciek ten objął minimum 246 gigabajtów danych. Z niektórych źródeł można się dowiedzieć, że prawdopodobnie doszło do wykradnięcia danych też z innych podległych placówek. Tych dodatkowych danych przestępcy jednak nie

opublikowali. Pobranie części danych było możliwe na darknetowej stronie http://pa32ymaeu62yo5th5mraikgw5fcvznnsiiwti42carjliarodltmqcqd.onion/post/9836.html. Tam będzie też możliwe pobranie całej bazy danych ALAB Laboratoria, gdy spółka nie opłaci okupu. Oczywiście, jeżeli prawdą jest to co pisze «RA World» na swojej stronie. «RA World» opublikował już wyniki około 55 tys. badań, które należały do 12 tysięcy różnych osób (różnych numerów PESEL).

W danych tych można znaleźć imię, nazwisko, PESEL, adres zamieszkania pacjenta oraz dane lekarza zlecającego. W wynikach można znaleźć dane medyczne takie jak wyniki badań laboratoryjnych. Są to dane medyczne szczególnie chronione. Część z nich może być krytyczna dla pacjenta, bowiem wyniki takich badań jak badanie na obecność przeciwciał wirusa HIV, czy też wysoki poziom cholesterolu mogą sprawić, że ta osoba będzie dyskryminowana przy udzielaniu pożyczki, jak też na rynku pracy.

Co zawiodło w ALAB Laboratoria?

W ALAB Laboratoria zawiodło wszystko, co tylko mogło zawieść. 246 gigabajtów danych to jest bardzo dużo danych. Ich pobieranie musiało zatem spowodować duże nienaturalne obciążenie na serwerach. Gdyby istniał zespół do monitorowanie nienaturalnego ruchu na serwerze w zakresie pobierania danych medycznych, zdarzenie byłoby wykryte i skończyłoby się jedynie na wycieku małej porcji danych.

Skoro pobrano wszystkie dane medyczne z 6 ostatnich lat, dane te nie były szyforwane indywidualnymi kluczami algorytmów asymetrycznych. Oznacza to, że dane te najprawdopodobniej nie były w ogóle szyfrowane. Wskazuje na to fakt, że podczas odbioru wyników wymagany jest tylko numer badania i PESEL lub data urodzenia klienta.

ALAB nie ogranicza danych medycznych dostępnych online do

ostatniego tygodnia lub ostatniego miesiąca, tak aby w razie wycieku danych był to mały zbiór danych. Dane z 2017 roku wciąż były dostępne do pobrania.

Co można było zrobić, aby tego wycieku nie było?

Wystarczyło zachować odpowiednie środki bezpieczeństwa. Choć jeden z wymienionych poniżej.

1. Nie udostępniać i nie zbierać adresu zamieszkania w ogóle. Nie zapisywać imienia, nazwiska i PESEL-u pacjenta. Zamiast tych danych zapisać online jedynie hash tych danych.

2. Zastosować PKI – infrastrukturę klucza prywatnego. Każde badanie szyfrowane indywidualnym kluczem.

W rozwiązaniu pierwszym sól stosowana podczas tworzenia skrótu (hashu) tych wiadomości powinna zawierać hasło wymagane do odebrania badania. PDF powinien być generowany po stronie front-endu (JavaScript przeglądarki) po wpisaniu przez pacjenta swojego PESEL-u, imienia, nazwiska, numeru badania i hasła zabezpieczającego. System sprawdziłby czy skróty tych danych są identyczne – wygenerowany i wcześniej zapisany – i wtedy wygenerowałby plik PDF oraz XML. W razie, gdy się coś nie zgadza, zaprosiłby do odbioru wyniku osobiście.

W tym miejscu należy zaznaczyć, że wyniki badań są bezwartościowe dla przestępcy, gdy może je odczytać, ale nie może się dowiedzieć do kogo one należą. Są one przydatne do dyskryminacji i zastraszania, gdy można wskazać ofiarę. Gdy na podstawie danych nie da się wskazać osoby, do której one należą – zazwyczaj przestają być danymi osobowymi chronionymi w myśl RODO. Danymi osobowymi chronionymi nie są bowiem dane zanonimizowane.

Ten scenariusz nie dotyczy jednak wyników zawierających dane genetyczne – tu na podstawie samego wyniku możliwe jest

określenie do kogo one należą – trzeba zastosować inną metodę. Tylko ile takich badań wykonuje się dziennie? Zawsze w przypadku danych genetycznych można zaprosić pacjenta obligatoryjnie po odbiór osobisty.

W drugim rozwiązaniu na etapie zlecania badania należy wygenerować klientowi losowe hasło jednorazowe o długości minimum 32 znaków. Jednocześnie trzeba wygenerować klucz algorytmu Rabina, AES lub ElGamal i hasłem tym zaszyfrować wygenerowany klucz prywatny. To umożliwia klientowi odzyskanie klucza prywatnego (zna hasło). ALAB byłby w stanie zaszyfrować wyniki badań i publicznie umieścić je na serwerze wraz z zaszyfrowanym kluczem prywatnym. Do przygotowania wyniku ALAB użyłby znanego klucza publicznego.

W takiej sytuacji szyfrogram wyników laboratoryjnych i zaszyfrowany klucz prywatny pozostają publicznie dostępne na serwerze. Przestępca może je pobrać bez potrzeby włamywania się. Nie może jednak uzyskać dostępu do danych, ponieważ musiałby znać klucz prywatny. Ten pozostaje zaszyfrowany długim hasłem niemożliwym do złamania w krótkim czasie. W sytuacji, gdy klucze publiczne nie są publikowane online i spółka kasowałaby je po przygotowaniu wyników – nie jest możliwe też podrobienie danych na serwerze, bowiem przestępca nie zna klucza publicznego. Ataki mające na celu podrobienie danych są jednak rzadkością w stosunku do ataków mających na celu pozyskanie danych osobowych.

Po stronie JavaScriptu przeglądarki musiałoby być możliwe:

– poproszenie klienta o podanie numeru badania i hasła do badania (bez numeru PESEL),

– odszyfrowanie klucza prywatnego przy użyciu wprowadzonego hasła,

– pobranie i odszyfrowanie danych po uzyskaniu dostępu do jednorazowego klucza prywatnego,

– wygenerowanie pliku PDF i XML lub zwrócenie komunikatu błędu (klucz jest nieprawidłowy lub dane są nieczytelne po deszyfrowaniu).

Drugie rozwiązanie może być stosowane również do danych genetycznych. Badanie pozostaje zaszyfrowane do czasu wprowadzenia hasła. Na serwerze nigdy nie przebywa w postaci niezaszyfrowanej. Zaszyfrowane pliki i zaszyfrowane klucze można przechowywać w publicznej chmurze.

Podatność jaka tu pozostaje to możliwość podmiany JavaScriptu tak, by przesyłał przestępcy klucze lub odszyfrowane wyniki badań. Jest to jednak proces długotrwały i ograniczony jedynie do ofiar, które pobrały zmodyfikowany skrypt. Ryzyko jest więc małe o ile istnieje przegląd zawartości witryny służącej do odbioru wyników.

Jak się teraz ochronić?

Skoro dane są w rękach przestępców (co zostanie potwierdzone ostatecznie 31 grudnia 2023 roku, gdy okup nie zostanie opłacony przez ALAB lub kogoś innego), niewiele da się zrobić, bowiem w Polsce nie istnieje instytucja zmiany numeru PESEL. Każdy komu zależy na prywatności, powinien napisać do MSWiA wniosek o umożliwienie wymiany numeru PESEL na żądanie. Zmiana samego imienia i nazwiska nie ma sensu, bowiem każdy wciąż z opublikowanych archiwów może pobrać dane po niezmienialnym numerze PESEL. Są dwie sytuacje, w których można wraz z imieniem i nazwiskiem zmienić numer PESEL – jest to zmiana oznaczenia płci w dokumentach oraz przypadek świadka koronnego. Korzystanie z tych opcji jest jednak szkodliwe i nieodpowiednie dla obywateli, których te sytuacje nie dotyczą. Oznacza więc to niedostępność faktyczną zmiany numeru PESEL.

Aby na problem ten zwrócono w końcu uwagę, należy pisać pozwy do sądu o odszkodowanie za upublicznienie danych. Skutki wycieku są dalekie, a już samo prawo do prywatności jeszcze jest prawem chronionym konstytucyjnie. Bankructwo kilku firm z

sektora medycznego po pozytywnych wyrokach sądów powinno pomóc przetrzeć im oczy.

Dyskryminacja w pracy

W pracy trzeba podać numer PESEL już po podpisaniu umowy. Umowę i skierowanie na badania wstępne sporządza się obecnie bez numeru PESEL, o ile czyni się to papierowo. Elektronicznie wciąż nie można zawrzeć umowy o pracę bez podania numeru PESEL (oprogramowanie podpisu cyfrowego zawsze wstawia numer PESEL).

W pracy nie tylko pracodawca może dyskryminować obywatela. Mogą to robić inni pracownicy, szczególnie ci wyżsi w hierarchii, np. menedżerowie. Aby temu zapobiec należy w pracy posługiwać się pseudonimem, który budową przypomina imię i nazwisko. W tym celu należy napisać i rozesłać CV podając zamiast imienia i nazwiska ustalony pseudonim (stały w określonym dłuższym czasie). Wypełniając kwestionariusz osobowy należy podać «prawdziwe» imię i nazwisko i zaznaczyć w nim, że używamy innego imienia i nazwiska. Można więc w formularzu tym napisać „Imię: Danuta, używane imię: Alicja”.

Można też napisać na końcu, aby w systemach pozakadrowych (e-mail, nazwa konta Windows/Google itp.) i wśród kolegów firma używała podanych używanych danych. Sprawę tą zawsze można wyjaśnić tym, że jest się w procesie zmiany danych, ale aby je zmienić, trzeba je najpierw używać.

Po takim zabiegu gdy Danuta Nowak jest znana jako Alicja Kowalska wśród kolegów z pracy – nie mogą się oni dowiedzieć jaki jest stan zdrowia Danuty i czy jest ona zakażona HIV. Nie da się też wtedy dyskryminować z innych przyczyn medycznych, których nie widać na zewnątrz (wysoki cholesterol czasami potrafi się przedstawić zmianami na twarzy).

Kradzież tożsamości

Dane podane w wynikach ALAB-u pozwalają na kradzież tożsamości. Brakuje w nich numeru dowodu osobistego, gdy ma się nadany numer PESEL. Należy nie podawać nikomu numeru dowodu osobistego. Pytanie o tą daną podczas rekrutacji do pracy powinno od razu wzbudzić alarm. Numer PESEL bowiem jest potrzebny dopiero, gdy pracownik zostanie dopuszczony do pracy. W umowie o pracę nie wpisuje się numeru PESEL.

Możliwe jest wciąż wyłudzanie większej ilości danych, bowiem podczas rejestracji w placówkach medycznych jedyną weryfikacją podczas rozmów telefonicznych jest podanie PESEL.

Należy też unikać publikacji biometrycznego wizerunku oraz odmawiać wszelkich aktywności, gdzie jedna osoba wymaga od drugiej włączenia kamery. Żądanie przekazania wizerunku podczas rozmowy kwalifikacyjnej jest złamaniem prawa polskiego. Danych biometrycznych nie wolno bowiem przetwarzać bez zapisanej w prawie podstawy prawnej. Podczas rekrutacji do pracy nigdzie nie występuje taka podstawa prawna. Nawet podczas rekrutacji na ochroniarza!

Odmawiać przekazania wizerunku należy dlatego, że wizerunek ten wraz z danymi z ALAB-u może zostać użyty do wyłudzenia kredytu oraz udowodnienia, że osoba ta była w danym miejscu, na co oszust przedstawi dowód w postaci fotografii lub też fotografii przerobionej komputerowo z użyciem AI. Taki dowód może być ciężki do obalenia w sądzie.

Do przemyślenia

Czy systemy informatyczne sądów są zbudowane w oparciu o PKI? Sprawy sądowe obywateli zarejestrowane na kamerze zawierają: wizerunek biometryczny, zapis głosu biometrycznego (wprost do szkolenia sztucznej inteligencji), imię, nazwisko, informacje o chorobach (zwolnienia L4 sądowe), adresy zamieszkania, dane

o skazaniu.

Czy oni też czekają na palec boży światowego RA?

Podsumowanie

Systemy medyczne należy budować tak, aby nie było potrzeby się do nich włamywać. Dane mogą być przechowywane w publicznych chmurach w formie zaszyfrowanej kluczem algorytmu asymetrycznego. Hasła i klucze tam gdzie się da, powinny być jednorazowe (jeden arkusz wyników = 1 klucz). Nie należy systemów budować tak, że dziura gdziekolwiek spowoduje dostęp do danych.

Architektura PKI jest odpowiednia do tego, aby więcej nie było przypadków typu „ALAB failed successfully" lub też „ALAB sucks A55 fully". Architektura PKI nie wymaga od klienta, żeby posiadał podpis cyfrowy lub jakikolwiek klucz. Wystarczające jest przekazanie długiego hasła do każdego arkusza wyników.

Obywatele powinni również robić protesty przeciwko przymusowej cyfryzacji rejestrów medycznych oraz pisać wnioski do Ministerstwa Zdrowia o cofnięcie przechowywania danych w formie cyfrowej, jako podatnej na wykradnięcie. Wykradnięcie realne, nie tylko teoretyczne, jak pokazało życie. Cyfryzacja przynosi więcej szkód niż korzyści.

Autorstwo: Valkyria
Źródło: Wolne-Forum-Transowe.plhttps://wolne-forum-transowe.pl/art-analiza-wycieku-danych-osobowych-z-alab-laboratoria