Warszawa, 08.07.2024 r.

Tworzenie baz danych na podstawie profilów osób na portalach branżowych typu LinkedIn w kontekście przetwarzania danych osobowych i zgodności z RODO

W dobie cyfryzacji i globalizacji, portale branżowe takie jak LinkedIn stały się integralną częścią nowoczesnego rynku pracy i kluczowymi narzędziami w rozwoju kariery zawodowej oraz w zarządzaniu zasobami ludzkimi. Platformy te, w szczególności, umożliwiają profesjonalistom nie tylko prezentowanie swojego doświadczenia zawodowego, umiejętności i osiągnięć, ale także nawiązywanie i utrzymywanie cennych kontaktów biznesowych. Firmy z kolei wykorzystują dane w nich publikowane również w innych celach - poniżej najbardziej popularne:

- Rekrutacja: Firmy tworzą bazy danych kandydatów w celu usprawnienia procesów rekrutacyjnych.
- Networking biznesowy: Tworzenie a następnie wykorzystywanie baz do budowania i utrzymywania relacji biznesowych. Informacje dostępne w profilach mogą pomóc w identyfikacji kluczowych osób w branży oraz nawiązywaniu nowych kontaktów.
- Badania rynku i analiza konkurencji: Firmy mogą zbierać dane np.: z profili LinkedIn w celu analizowania trendów rynkowych, identyfikowania kluczowych graczy oraz monitorowania działalności konkurencji.
- Marketing i sprzedaż: Informacje zebrane z LinkedIn mogą być wykorzystywane do tworzenia listy potencjalnych klientów lub partnerów biznesowych. Taka baza danych może być przydatna w celach marketingowych i sprzedażowych

W ramach realizacji wyżej wymienionych celów dochodzi najczęściej do gromadzenia danych, w tym danych osobowych użytkowników portali branżowych. Na rynku istnieje wiele praktyk dotyczących gromadzenia danych z LinkedIn, w tym danych osobowych, do najczęściej spotykanych należą:

- Scraping danych: polega na automatycznym zbieraniu danych z profili użytkowników za pomocą narzędzi, systemów IT = technicznie bardzo skuteczny ale często spotyka się z problemami natury prawnej tak w zakresie naruszania warunków użytkowania (regulaminów) portali branżowych jak i przepisów dotyczące ochrony danych;
- „Ręczne" zbieranie danych: polega na ręcznym przeklikiwaniu, przeszukiwaniu profili i kopiowaniu interesujących informacji lub całych profili = bardziej czasochłonne niż scaping, ale może być bardziej dokładne a przede wszystkim zgodne z przepisami, pod warunkiem prawidłowego zaprojektowania samego procesu pozyskiwania takich danych;
- Używanie dedykowanych narzędzi rekrutacyjnych: Platformy takie jak LinkedIn oferują narzędzia rekrutacyjne, które umożliwiają profesjonalne zarządzanie kandydatami i ułatwiają przeszukiwanie profili. Te narzędzia są zazwyczaj zgodne z regulacjami prawnymi, gdyż są udostępniane przez samą platformę.
- Zewnętrzne bazy danych i agregatory: inaczej firmy oferujące bazy danych i agregatory informacji, które gromadzą dane z różnych źródeł, w tym z LinkedIn. Korzystanie z takich

usług wymaga jednak upewnienia się, że dane są zbierane i przetwarzane zgodnie z przepisami RODO

Pomimo korzyści, jakie niesie za sobą korzystanie z danych publikowanych na portalach branżowych, należy mieć na uwadze, że istnieje także szereg potencjalnych ryzyk, w szczególności w sferze ochrony danych osobowych na które podmiot gromadzący dane musi zwrócić uwagę, przeanalizować i wdrożyć odpowiednie procedury jeżeli chce zachować tzw. „Zgodność z RODO". Aby zapobiec materializacji ryzyk związanych z gromadzeniem i przetwarzaniem danych z portali branżowych, firmy powinny stosować, przede wszystkim zasady "privacy by design" i "privacy by default". Oznacza to wprowadzenie ochrony danych osobowych już na etapie projektowania procesów oraz zapewnienie, że domyślnie dane są chronione. W dalszej części publikacji omówione zostaną kroki, jakie firmy mogą podjąć, aby zminimalizować ryzyko naruszeń w kontekście gromadzenia danych z LinkedIn i innych podobnych platform na podstawie zidentyfikowanych ogólnych ryzyk.

Poniżej niektóre z ryzyk związane z gromadzeniem danych z portali branżowych oraz praktyczne rady jak im zasadzić:

1. Automatyczne zbieranie danych (scraping):

Ryzyka:

- Naruszenie regulaminu portalu branżowego (np.: LinkedIn wyraźnie zabrania scraping'u swoich danych w warunkach użytkowania). Użycie tego typu technologii może prowadzić do zablokowania konta użytkownika lub podjęcia działań prawnych przez właściciela portalu;
- Naruszenie podstawowych zasad RODO - automatyczne pobieranie danych bez zgody użytkownika (użytkownik nie ma możliwości wyrażenia zgody ponieważ nie ma w ogóle informacji, że jego dane są gromadzone) może naruszać jedne z podstawowych zasad RODO tj. RODO wymaga, aby dane były zbierane zgodnie z prawem, uczciwie i w sposób przejrzysty dla osoby, której dotyczą. Ponadto, scraping może prowadzić do gromadzenia nadmiarowych lub wrażliwych danych, które nie są niezbędne dla celu przetwarzania, co jest sprzeczne z zasadą minimalizacji danych.

Zapobieganie:

- Zrezygnuj z scrapingu / zamiast scrapingu, korzystaj z oficjalnych API dostarczanych przez platformy, które oferują zgodne z prawem sposoby na dostęp do danych. Na przykład LinkedIn oferuje LinkedIn Recruiter i inne narzędzia zgodne z regulaminem, umożliwiające dostęp do danych kandydatów.
- Uzyskaj wyraźną zgodę jeśli scraping jest nieunikniony - informuj użytkowników o tym procesie i uzyskaj ich wyraźną zgodę. Podczas procesu rekrutacji, należy poinformować kandydatów, że ich profil na LinkedIn będzie analizowany w sposób automatyczny i poprosić o zgodę na przetwarzanie tych danych (z wykorzystaniem czat lub adresu e-mail)

2. Korzystanie z zewnętrznych baz danych i agregatorów informacji.

Ryzyka:

- Brak transparentności: Korzystanie z zewnętrznych baz danych może utrudniać firmom zapewnienie pełnej przejrzystości wobec osób, których dane są przetwarzane. Trudno jest śledzić, jak i skąd dokładnie dane zostały zebrane.
- Nielegalne źródła danych: Firmy muszą upewnić się, że dane dostarczane przez zewnętrzne bazy zostały zebrane zgodnie z przepisami RODO. Nielegalne źródła danych lub nieprzestrzeganie procedur zbierania danych mogą prowadzić do poważnych naruszeń.
- Utrzymanie zgodności: Firmy muszą sprawdzić i utrzymywać zgodność z RODO w całym łańcuchu przetwarzania danych (od och pozyskania „u źródła" aż do zmaterializowania okresu retencji i ich usunięcia/anonimizacji) co może być znacznie utrudnione w przypadku korzystania z zewnętrznych dostawców.

Zapobieganie:

- Audyt dostawców przed skorzystaniem z ich usług w zakresie dostarczenia zewnętrznej bazy danych. Firma powinna przeprowadzić audyt dostawcy (weryfikację procesora), sprawdzając źródła pochodzenia danych i metody zbierania. Można np. zażądać dowodów zgodności z RODO oraz procedur ochrony danych stosowanych przez dostawcę.
- Zawarcie umów o przetwarzanie danych (DPA). Firma powinna zawierać szczegółowe umowy z dostawcami danych, które określają, jakie dane są przetwarzane, na jakiej podstawie prawnej, jak długo będą przechowywane i jakie są obowiązki związane z ochroną tych danych a także odpowiedzialność w przypadku wystąpienia naruszeń oraz jakie procedury bezpieczeństwa są / mają być stosowane

3. Gromadzenie informacji zawartych na profilu bez informowania użytkowników.

Niestety powszechną praktyką rynkową jest, że firmy przetwarzają dane z portali branżowych bez informowania użytkowników o tym, co się dzieje z ich danymi – brak przekazania obowiązku informacyjnego (art. 13 lub 14 RODO) w związku z gromadzeniem danych. Może to dotyczyć zarówno „ręcznego", jak i automatycznego (scraping) gromadzenia informacji z profili.

Ryzyka:

- Brak informacji: Przetwarzanie danych użytkownika bez jego świadomości w tym względzie stanowi naruszenie w sferze ochrony danych osobowych. Użytkownicy muszą być informowani o przetwarzaniu ich danych oraz o celach tego przetwarzania w sposób oraz na zasadach określonych przepisami prawa.
- Naruszenie prawa do prywatności: Użytkownicy mają prawo do ochrony swojej prywatności. Przetwarzanie ich danych bez informowania ich o tym narusza ich prawo do prywatności.
- Pogorszenie reputacji.

Zapobieganie:

- Zapewnij transparentność, w szczególności informuj użytkowników o tym, że ich dane mają być/są gromadzone. Firma rekrutacyjna może na swojej stronie internetowej udostępnić politykę prywatności, która wyjaśnia, jak dane z LinkedIn są wykorzystywane

a następnie linkować ją w komunikacji z użytkownikami. Dodatkowo, kandydaci mogą być informowani w wiadomościach (czat, e-mail), że ich profil zostanie „przetworzony" w ramach procesu rekrutacyjnego.

- Przestrzegaj obowiązku informacyjnego. Przed rozpoczęciem gromadzenia danych, firma powinna wysyłać wiadomości e-mail do kandydatów lub poprzez czat informując ich o celach przetwarzania, danych kontaktowych administratora danych, przysługujących im prawach, celach i sposobach przetwarzania oraz innych wymienionych w art. 13 i 14 RODO. Potencjalny kandydat który ma być brany pod uwagę w konkretnej rekrutacji powinien otrzymać wiadomość z pełnym zakresem informacji, jakie dane będą przetwarzane, jak również przez jaki czas.

4. Zbieranie nadmiarowych danych

Zbieranie zbyt dużej ilości danych, które nie są konieczne do osiągnięcia określonego celu, jest powszechną praktyką, szczególnie gdy dane są zbierane automatycznie lub bez odpowiednich procesów weryfikacji.

Ryzyka:

- Naruszenie zasady minimalizacji danych: RODO nakłada obowiązek na firmy, aby zbierały tylko te dane, które są niezbędne do osiągnięcia konkretnego celu. Nadmiarowe zbieranie danych może prowadzić do niezgodności z tą zasadą.
- Zwiększenie skali przy ewentualnym naruszeniu danych:. nadmiarowe dane mogą stać się celem dla cyberataków, a ich nadmiarowy zakres może prowadzić do pogorszenia konsekwencji prawnych i finansowych.
- Wyzwania związane z zarządzaniem danymi: Gromadzenie dużej ilości danych może skomplikować procesy zarządzania nimi, w tym ich aktualizację, ochronę i usuwanie, co może być problematyczne w kontekście spełnienia wymogów RODO.

Zapobieganie:

- Ogranicz zbieranie danych do minimum. Firma powinna zbierać tylko te dane, które są absolutnie niezbędne do określonego celu. Na przykład, rekrutując na konkretne stanowisko, można zbierać tylko informacje o doświadczeniu zawodowym i umiejętnościach kandydatów, pomijając dane osobowe, które nie są bezpośrednio związane z kwalifikacjami do pracy.
- Przeprowadzaj regularne przeglądy danych. Dobrą praktyką jest, że Firma co kwartał przeprowadza przegląd zebranych danych kandydatów i usuwa te, które są zbędne lub nadmiarowe. Na przykład, jeżeli kandydat nie był rozpatrywany przez określony czas (np. sześć miesięcy), jego dane mogą być usuwane z systemu.

5. Przechowywanie danych przez dłuższy czas niż jest to konieczne

Firmy mogą przechowywać dane z LinkedIn przez dłuższy okres niż jest to faktycznie potrzebne, zwłaszcza gdy dane są gromadzone do ogólnych celów biznesowych bez jasno określonej polityki retencji.

Ryzyka:

- Naruszenie zasady ograniczenia przechowywania: RODO wymaga, aby dane były przechowywane jedynie przez okres niezbędny do realizacji celu, w jakim zostały zebrane. Przechowywanie danych przez dłuższy czas może prowadzić do naruszenia tej zasady.

- Zwiększenie ryzyka niezgodności: Długoterminowe przechowywanie danych zwiększa ryzyko, że dane te staną się przestarzałe lub będą nieaktualne, co może prowadzić do niezgodności z innymi zasadami RODO – zasada prawidłowości danych.
- Koszty związane z przechowywaniem i zarządzaniem danymi: Długotrwałe przechowywanie danych może wiązać się również z wyższymi kosztami.

Zapobieganie:
- Opracuj i wdróż politykę retencji danych, Firma powinna określić, jak długo dane będą przechowywane, oraz wprowadzić automatyczne procedury usuwania danych po określonym czasie lub, chociażby automatyczne tagowanie w celu dalszego ręcznego usuwania w celu lepszej kontroli nad danymi.
- Regularne czyszczenie baz danych. Firma może ustawić harmonogram przeglądów i czyszczenia bazy danych, aby upewnić się, że niepotrzebne dane są regularnie usuwane. Procesy mogą być zautomatyzowane – na rynku dostępne są rozwiązania IT posiadające niezbędne i użyteczne funkcjonalności w tym zakresie.

Podsumowując, w ramach gromadzenia informacji z portali branżowych nieodłącznym elementem procesu jest również przetwarzanie danych osobowych. W kontekście RODO, każda organizacja musi podejść do tego zadania z pełnym zrozumieniem swojej odpowiedzialności, świadomością procesów i procedur biznesowych funkcjonujących wewnątrz organizacji i jakie cele mają być realizowane w związku z gromadzeniem danych (jakie dane i do czego są nam, de facto, potrzebne). Gromadzenie i wykorzystywanie danych to nie tylko kwestia techniczna, ale przede wszystkim działanie wymagające szacunku dla prywatności użytkowników i ich praw do ochrony danych. Aby zapewnić zgodność z RODO przy przetwarzaniu danych z portali branżowych, firmy powinny skupić się w pierwszej kolejności na kilku strategicznych obszarach:

1. **Budowanie kultury ochrony danych**: Ochrona danych powinna być integralną częścią projektowania systemów i procedur, co oznacza wprowadzenie zasad "privacy by design" i "privacy by default". Oznacza to również, że wszystkie systemy i procesy muszą być zaprojektowane z myślą o maksymalnej ochronie prywatności użytkowników od samego początku.
2. **Edukacja i szkolenia**: Pracownicy muszą być regularnie szkoleni w zakresie zasad ochrony danych i zgodności z RODO. Świadomość ryzyk związanych z przetwarzaniem danych osobowych oraz znajomość najlepszych praktyk w tym zakresie są kluczowe dla utrzymania zgodności i uniknięcia naruszeń.
3. **Transparentność wobec użytkowników**: Użytkownicy muszą być jasno informowani o tym, jak ich dane są zbierane, przetwarzane i przechowywane. Transparentność nie tylko buduje zaufanie, ale jest również kluczowym wymogiem RODO. Informowanie użytkowników o celu przetwarzania ich danych oraz o ich prawach jest nie tylko zgodne z prawem, ale także dobrym standardem biznesowym.
4. **Uzyskiwanie zgody podmiotu danych**: W każdym przypadku, gdy przetwarzanie danych wymaga zgody użytkowników, należy ją uzyskać w sposób świadomy i dobrowolny. Zgoda powinna być wyrażona jednoznacznie, po uprzednim poinformowaniu o celach

przetwarzania danych oraz prawach użytkownika. Praktyka ta nie tylko zapewnia zgodność z RODO, ale także wzmacnia zaufanie i przejrzystość.

5. **Regularne audyty i przeglądy**: Firmy powinny regularnie przeprowadzać audyty oraz przeglądy swoich procedur ochrony danych, aby identyfikować potencjalne problemy i zagrożenia. Ciągłe doskonalenie praktyk ochrony danych jest kluczowe dla utrzymania zgodności z przepisami.
6. **Odpowiedzialność za cały cykl życia danych**: Firmy muszą dbać o zgodność z RODO na każdym etapie przetwarzania danych – od momentu ich zebrania, poprzez przechowywanie, aż po usunięcie. Obejmuje to minimalizację zbieranych danych, właściwą politykę retencji oraz bezpieczne usuwanie danych, gdy nie są już potrzebne.

Wnioskiem z tych rozważań jest to, że odpowiedzialne zarządzanie danymi osobowymi z portali branżowych jest kluczowe zarówno z perspektywy prawnej, jak i dla utrzymania odpowiedniego poziomu reputacji. Stosowanie dobrych praktyk w zakresie ochrony danych osobowych powinno być traktowane jako strategiczny priorytet, a nie tylko obowiązek regulacyjny. W dynamicznie rozwijającym się świecie cyfrowym, zgodność z RODO i etyczne podejście do przetwarzania danych osobowych mogą stać się istotnym elementem przewagi konkurencyjnej oraz fundamentem zrównoważonego rozwoju biznesu.

Paweł Wojciechowski - adwokat, specjalista ds. ochrony danych osobowych w iSecure Sp. z o.o.