**7. Podsumowanie**

Portale społecznościowe, dostępne dla każdego internauty, są to aplikacje ułatwiające komunikację ze znajomymi oraz znanymi użytkownikowi grupami osób. Dane zawarte na takich portalach, są to informacje pochodzące z różnych dziedzin życia, opisujące wydarzenia jakie miały miejsce czy opinie na dany temat. Dzięki tym systemom, można wyrobić sobie zdanie o konkretnej osobie, nigdy jej nie spotykając, co jest znaczące przy na przykład próbie otrzymania pracy. Złe treści udostępniane na publicznych profilach, mogą realnie wpłynąć na przyszłość użytkownika. Aplikacja stworzona w ramach tej pracy, pokazuje jak szczegółowe dane, może wyciągnąć osoba, łącząc informacje dostępne z wielu witryn społecznościowych. Przykładowy raport dotyczący lokalizacji, pokazuje szczegółowość zaimplementowanej analizy oraz zagrożenia płynące z nierozsądnego udostępniania treści związanych z życiem.

Biorąc pod uwagę fakt, że informacje zawarte na serwerach Facebooka, Instagrama czy Twittera są to często dane poufne, systemy te posiadają wiele mechanizmów zabezpieczających te dane. Pierwszym dostępnym mechanizmem jest profil prywatny, który użytkownik może włączyć wchodząc w ustawienia. Oznaczenie swojego profilu jako prywatny powoduje, że jedynie określona grupa osób może oglądać dane związane z użytkownikiem. Takie profile są niedostępne w aplikacjach, ale także w wystawionych przez portale społecznościowe serwisach API, więc analiza takich osób jest nieskuteczna. Jednym z wniosków, analizując utworzone raporty, jest słuszność w używaniu profilu prywatnego, dostępnego tylko dla znanych użytkowników. Pozwala to osiągnąć odpowiedni stopień bezpieczeństwa i anonimowości. Drugim mechanizmem zabezpieczającym przed analizą zbyt poufnych danych, są pozwolenia dla zewnętrznych aplikacji do pobierania odpowiednich danych. Najbardziej restrykcyjnym portalem społecznościowych z trzech analizowanych w ramach tej pracy, jest Facebook, który udostępnia jedynie podstawowe dane na temat użytkownika.

Pomimo wyżej opisanych mechanizmów bezpieczeństwa, ilość danych, które są pobierane, w szczególności dla znanych osób, powodują złe działanie prostych serwerów bazodanowych. Zaproponowana w ramach tej pracy struktura, pochodząca z dziedziny Business Intelligence, Architektura Lambda jest rozwiązaniem pozwalającym analizować te duże ilości danych oraz przedstawiać aktualne informacje bez oczekiwania. Rozwiązanie to jest rozwiązaniem ogólnym i może zostać zaimplementowane na wiele sposobów. W ramach tej pracy został użyty Spark oraz widoki postgresowe jako dwie najważniejsze warstwy tej struktury. W przypadku aplikacji tworzonej w ramach tej pracy, zaimplementowana architektura zadziałała prawidłowo, dostarczając przeliczone raporty przez Sparka oraz aktualne dane, dostarczone przez te widoki. Największym rozwiązanym problemem tworzenia tej architektury, było zdefiniowanie przeliczeń w języku Scala, w ramach Sparka.

Architektura Lambda pozwala na bardzo skuteczne skalowanie rozwiązania, poprzez możliwość dodania nowych widoków w odpowiednich warstwach. Dzięki temu tworzona aplikacja może być w łatwy sposób rozszerzona o nowe raporty, co jest działaniem planowanym w przyszłości.