

## Stacja 3 — anonimizacja (RODO)

**Plik źródłowy:** ta sama relacja co w Stacji 2 (przykład-korpus/poznan-1956/relacje-swiadkow/relacja-01-sur

Najlepiej zrób to **w tym samym czacie** zaraz po Stacji 2 — Bielik ma już tekst w pamięci, możesz napisać prompt bez ponownego wklejania.

### Prompt

Zanonimizuj powyższy tekst zgodnie z RODO. Zamień wszystkie dane osobowe na neutralne placeholderzy: imiona i nazwiska → [ŚWIADEK\_1], [OSOBA\_2]; adresy → [ULICA], [MIASTO]; nazwy zakładów pracy → [ZAKŁAD]. Zachowaj sens i czytelność tekstu.

Na końcu wypisz tabelę zmian: co zostało zamienione i na co.

(Jeśli zaczynasz od nowa: wklej najpierw prompt, a pod nim tekst z pliku.)

**Czego się spodziewać:** dane osobowe (Tadeusz Kowalczyk, ul. Dąbrowskiego, żona Halina, ZISPO) zastąpione placeholderami + lista zmian do weryfikacji.

### Wariant trudniejszy — sama detekcja (NER)

Nie zmieniaj tekstu. Tylko wypisz wszystkie dane osobowe, jakie w nim występują, pogrupowane w kategorie: imiona/nazwiska, adresy, miejsca pracy, daty, inne.

**Po co to w pracy naukowej:** warunek zgody komisji etyki i publikacji danych w modelu open data. Zawsze sprawdź wynik — model bywa nadgorliwy (anonimizuje za dużo) albo coś przeoczy.