



Historia i Polityka
Nr 49(56)/2024, ss. 9–25

www.hip.umk.pl



ISSN 1899-5160, e-ISSN 2391-7652

DOI: <http://dx.doi.org/10.12775/HiP.2024.019>

Zbigniew WILCZYŃSKI

Instytut Pamięci Narodowej, Szczecin, Polska

Wiktor WERNER

Uniwersytet im. Adama Mickiewicza, Wydział Historii, Poznań, Polska

Portret tajnego współpracownika „bezpieki”. Badania źródeł archiwalnych SB (Wydziału VI w Szczecinie) z wykorzystaniem algorytmów uczenia maszynowego oraz analizy prozopograficznej

Portrait of a Secret Collaborator of the “Security Service” (SB) of the Polish People’s Republic:
Research on Archival Sources of the SB (Department VI in Szczecin) Using Machine Learning
Algorithms and Prosopographic Analysis

• Abstrakt •

Celem artykułu jest charakterystyka tajnych współpracowników służby bezpieczeństwa PRL dokonana metodami tradycyjnymi (analiza prozopograficzna) i cyfrowymi (zastosowanie algorytmu uczenia maszynowego w formie drzewa decyzyjnego) oraz porównanie skuteczności obu metod dla wskazania, co charakteryzowało współpracowników składających donosy personalne.

Słowa kluczowe: Służba Bezpieczeństwa (SB); prozopografia; tajni współpracownicy; drzewo decyzyjne

• Abstract •

The purpose of the article is to characterize the secret collaborators of the Security Service of the Polish People’s Republic by traditional (prosopographic analysis) and digital (application of the decision tree machine learning algorithm) methods, and to compare the effectiveness of the two methods for indicating what characterized the collaborators making personal denunciations.

Keywords: Security Service (SB); prosopography; secret collaborators; decision tree

Wstęp – cele badawcze, dyskusja

Niezwykle istotną częścią pracy operacyjnej Służby Bezpieczeństwa było wykorzystywanie tajnych współpracowników (TW) do zbierania informacji oraz wykonywania zleconych zadań. Ten temat doczekał się wielu opracowań oraz był przedmiotem głośniejszych i emocjonalnych dyskusji obecnych w mediach, w których znacznie częściej zadawano pytania dotyczące konkretnych osób, których nazwiska pojawiały się w archiwach SB (Dudek, 1998; Skibiński, 2003; Wiślicki, 2005; Buryła, 2009; Łaszczyński, 2010; Zarzycki, 2004), niż próbowano zobiektywizować obraz „konfidenta” z wykorzystaniem metod ilościowych i wspomaganych komputerowo. Taka perspektywa nie budzi zdziwienia w kontekście gorących emocji, jakie wywoływały (i nadal wywołują) odkrywane fakty dotyczące współpracy z SB osób znanych z życia publicznego. Jednak wraz z upływem czasu pojawia się również miejsce na „odpersonalizowanie” dyskusji o tajnych współpracownikach SB. Widać taką, paralelną, tendencję na przykładzie badań lingwistycznych Małgorzaty Hołdy (Hołda, 2009) czy psychologicznych Agnieszki Gałkowskiej (Gałkowska, 2007), jednak złożoność i ważkość zagadnienia zostawiają nadal pole dla wyszukiwania nowych metod i perspektyw badawczych, na co zwraca uwagę w swoim artykule z 2022 r. Daniel Wicenty. Autor twierdzi, że badacze zwykle poprzestają na wykorzystaniu akt tajnych współpracowników do opisu samej współpracy, ewentualnie do opisu danego środowiska, nie poszukując nowych sposobów ich zgłębiania (Wicenty, 2022). Warto zatem zastanowić się nad nowymi kierunkami i perspektywami badawczymi możliwymi do zastosowania w tej specyficznej dziedzinie.

Nasza propozycja dotyczy aplikacji metod statystycznych, prozopograficznych oraz metod algorytmicznych uczenia maszynowego do zbadania źródeł archiwalnych SB powiązanych z działalnością Wydziału VI WUSW w Szczecinie. Zastosowanie metod ilościowych daje możliwość agregacji danych i wyjścia poza problem wiarygodności pojedynczych teczek oraz informacji, na które zwracana jest uwaga podczas metodycznych dyskusji na temat pracy z tego rodzaju źródłami (Michalski, 2006; Brzechczyn, 2012), wprowadzenie metod algorytmicznych stwarza natomiast (nadal rzadko wykorzystywaną) możliwość znajdowania wśród badanych zjawisk powiązań nieoczywistych z poziomu pracy nad pojedynczymi teczkami, a zauważalnych dopiero z perspektywy zagregowanego zbioru danych. Na możliwości kryjące się w nowych, wspomaganych komputerowo metodach zwracają coraz częściej uwagę badacze z głównych ośrodków naukowych na świecie, takich jak Stanford University czy Princeton University (Grimmer, Roberts, Stewart, 2021; Cranmer, 2019), a wyniki uzyskiwane z zastosowania tych metod, jak spektakularne badania nad nieoczywistymi przyczynami wybuchu amerykańskiej wojny domowej

przeprowadzone przez zespół wykorzystujący algorytmy uczenia maszynowego dla historii polityczno-militarnej (Basuchoudhary i in., 2021), zachęcają do implementacji tych metod do analiz rodzimych źródeł i zagadnień.

Cele przyświecające powstaniu tego artykułu mają zatem zarówno charakter metodologiczny (porównanie metod statystycznych, prozopograficznych i metody uczenia maszynowego zastosowanych do zbadania danych z archiwaliów SB), jak i poznawczy: znalezienie odpowiedzi na pytanie, jacy ludzie współpracowali z niesławną Służbą Bezpieczeństwa i składali donosy na swoich rodaków, często kolegów czy koleżanki. W naszych badaniach nie chodzi nam jednak o personalia konfidentów, lecz określenie typu osoby skłonnej nie tylko do ogólnie rozumianej współpracy, ale także do składania bardziej zaangażowanych, personalnych donosów.

Przygotowanie bazy empirycznej

Wydział VI WUSW w Szczecinie został przez nas wybrany jako empiryczna baza do badań z powodu relatywnie wysokiego odsetka zachowanej dokumentacji tajnych współpracowników. Na dzień 1 stycznia 1989 r. liczba akt TW VI oddziału liczyła 171 (AIPN, Sz 0012/375 t. 175, s. 113). Na potrzeby niniejszego artykułu udało się pozyskać i zeskanować ponad 130 jednostek archiwalnych (j.a.) łącznie liczących ponad sześć tysięcy stron. Niektóre wytypowane j.a. trzeba było odrzucić (np. okazały się aktami kandydatów na TW, a nie czynnych TW), więc ostatecznie baza danych objęła 123 teczki¹ stanowiące 73% spośród akt TW tworzących sieć agenturalną VI Wydziału.

Dane zawarte w zgromadzonych aktach posłużyły do wykonania bazy danych, w której wszystkie poszczególne przypadki były opisane za pomocą zbioru kategorii, dla badań prezentowanych w prezentowanym artykule wykorzystywaliśmy poniższe kategorie:

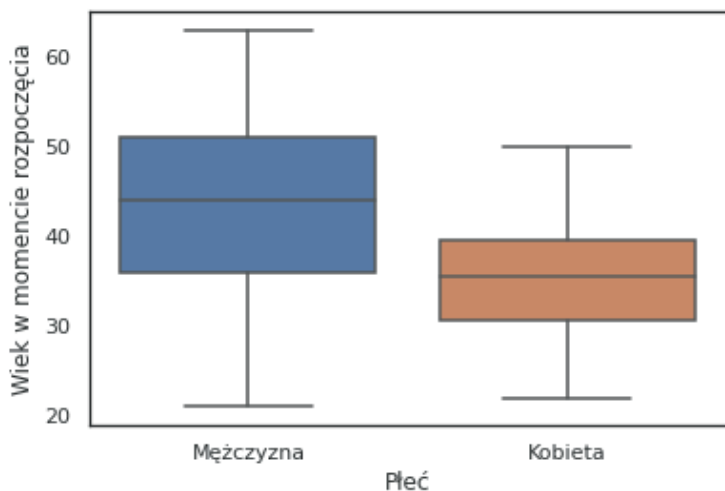
- kategorie identyfikacyjne: sygnatura akt i pseudonim TW,
- rok rozpoczęcia współpracy oraz rok zakończenia współpracy, a także długość współpracy w latach,
- wcześniejsze rozpracowanie,
- karalność,

¹ Przez (używane również w publikacjach naukowych) potoczne wyrażenie „teczki” należy rozumieć akta personalne i akta pracy tajnego współpracownika. Często mamy jednak do czynienia tylko z jedną, połączoną teczką (akta personalne i pracy TW) zawierającą zarówno kwestionariusz o współpracowniku i opis przebiegu samego pozyskania, jak i sporządzone w toku współpracy informacje (przyp. aut.).

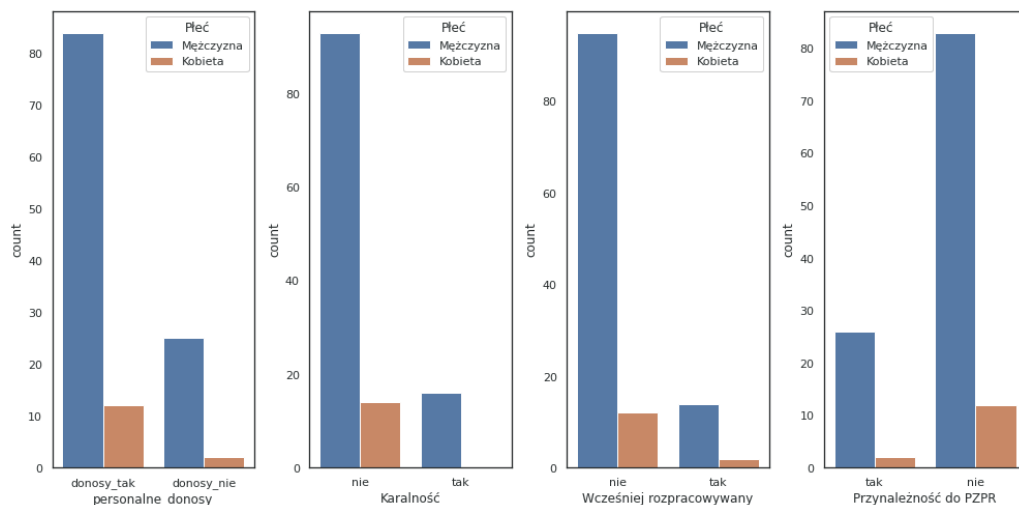
- wiek w momencie pozyskania,
- płeć,
- wykształcenie,
- charakter pracy,
- czy TW był nagradzany,
- łączna wysokość wynagrodzenia (ta kategoria nie mogła być jednak wykorzystana do analizy ilościowej ze względu na dużą ilość brakujących danych),
- przynależność do PZPR,
- czy udzielane informacje odnosiły się personalnie do innych osób (miały charakter personalnych donosów),
- informacja o zniszczeniu teczki pracy,
- liczba zachowanych informacji.

Analiza statystyczna danych ilościowych

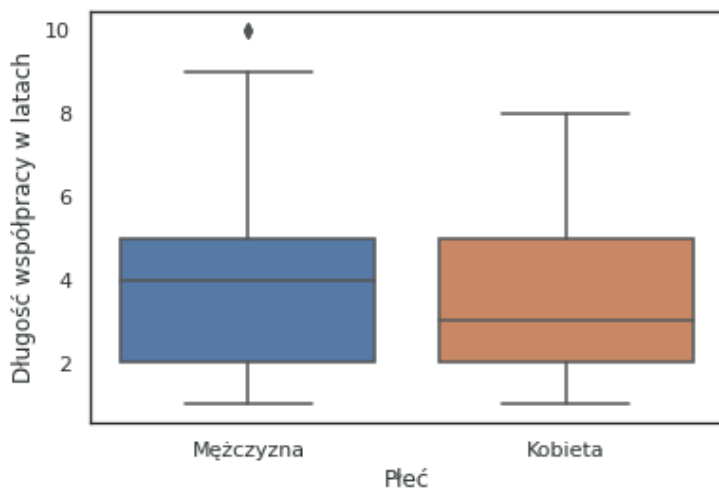
Pierwszym etapem naszych badań po opracowaniu bazy danych (opisanej we wcześniejszej sekcji) było wykonanie analizy statystycznej mającej ukazać portret typowego (przeciętnego) tajnego współpracownika Wydziału VI. Zebrane i zagregowane dane zostały zilustrowane poniższymi grafikami 1–3.



Grafika 1. Wiek TW w momencie rozpoczęcia współpracy, odrębnie dla kobiet i mężczyzn. Oś Y – wiek TW mierzony w latach

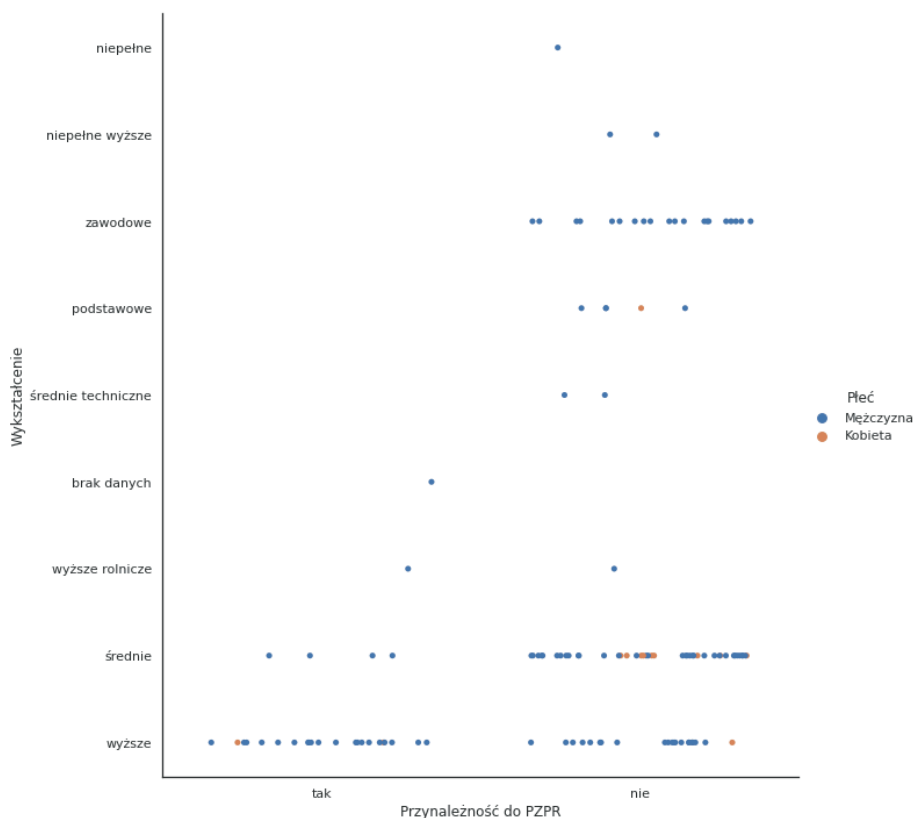


Grafika 2. Charakterystyka TW w kontekście kategorii: płeć (legenda), fakt składania personalnych donosów, karalność, wcześniejsze rozpracowywanie i przynależność do PZPR



Grafika 3. Długość współpracy w latach dla TW obu płci

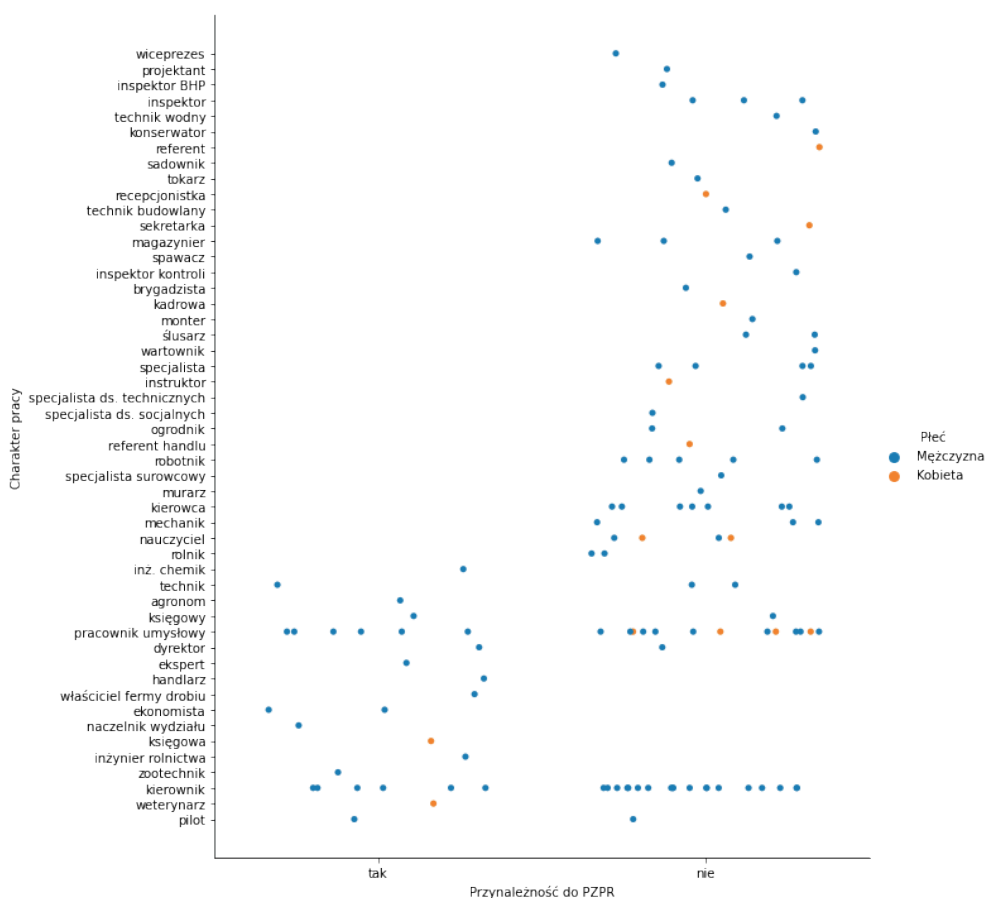
Z akt wynika, że tajny współpracownik to zdecydowanie częściej mężczyzna niż kobieta. Mężczyźni rozpoczynali współpracę w wieku powyżej czterdziestego roku życia, kobiety po trzydziestce. Dla mężczyzn możliwy jest fakt karalności (dla mniejszości TW), dla kobiet ten fakt nie występuje. Dla obu płci zdecydowana mniejszość była wcześniej rozpracowywana, w obu przypadkach osoby współpracujące znacznie częściej składały personalne donosy (obciążające konkretnych ludzi), niż się od tego skutecznie uchylały. Kwestia składania donosów będzie przez nas jeszcze analizowana odrębnie, gdyż wstępna analiza nie wykazała jednoznacznego powiązania tej kategorii z innymi. Skomplikowana jest również kwestia przynależności do PZPR, zasadniczo SB miała zakaz werbowania członków partii (Pastuszewski, 2019). Mniejsza liczba członków PZPR wśród TW świadczy o tym, że zakaz był do pewnego stopnia respektowany. Sam fakt, że niektórzy TW do partii należeli, wskazuje, że w praktyce bywał omijany.



Grafika 4. Rozkład przynależności TW do następujących kategorii: płeć (kolor), przynależność do PZPR (oś X), wykształcenie (oś Y)

Pytanie, kim byli członkowie PZPR, którzy zdecydowali się na współpracę z naszym wydziałem SB, wprowadza nas do zagadnienia pozycji zawodowej i wykształcenia TW, które są bardziej złożonymi (podzielonymi na wiele zmiennych) kategoriami i wymagają bardziej skomplikowanych wizualizacji w postaci wykresów typu „stripplot” ukazujących rozkład danych na poszczególne kategorie, gdzie punkty oznaczają poszczególnych TW rozdzielonych na trzy kategorie (oś X i Y na wykresie oraz wyróżnienie kolorystyczne).

Jak można zobaczyć na wykresach 4 i 5, TW należący do PZPR mieli zazwyczaj wyższe wykształcenie i odpowiednią do niego pozycję zawodową kierowników, pracowników umysłowych i specjalistów.



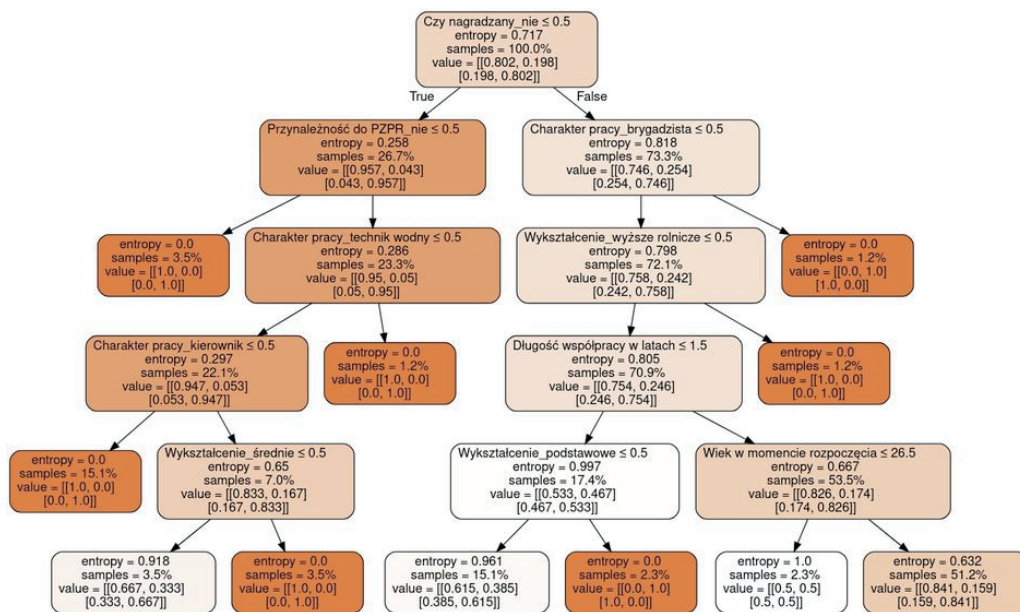
Grafika 5. Rozkład przynależności TW do następujących kategorii: płeć (kolor), przynależność do PZPR (oś X), pozycja zawodowa (oś Y)

Wśród TW nie wszyscy składali donosy obciążające konkretne osoby (personalne donosy), jednak fakt składania takich donosów nie koreluje jednoznacznie (przynajmniej na pierwszy rzut oka) z żadną z kategorii przypisanych do osób współpracujących z SB. W celu znalezienia korelacji, które są mniej oczywiste, postanowiliśmy zastosować wspieraną komputerowo metodę „drzewa decyzyjnego” (*decision tree*).

Zastosowanie algorytmu drzewa decyzyjnego

Algorytm drzewa decyzyjnego jest stosowany w naukach społecznych od lat 60. ubiegłego stulecia w celu grupowania badanych zjawisk w zbiory jednorodne (klasyfikacja) i wyszukiwania nieoczywistych powiązań między zjawiskami (Loh, 2011). Jego podstawową zaletą z perspektywy nauk społecznych (oprócz relatywnej prostoty) jest fakt, że może on analizować każdy typ danych – zarówno ilościowy, jak i jakościowy. Algorytm działa na zasadzie dokonywania podziałów binarnych tak długo, aż doprowadzą one do powstania zbiorów jednorodnych (w przypadku klasyfikacji). Badana jest również korelacja między zmienną zdefiniowaną jako „zależna” a zbiorem zmiennych zdefiniowanych jako „niezależne”, w naszym przypadku oznacza to badanie korelacji między faktem składania donosów personalnych a cechami TW jako osoby (płcią, przynależnością do PZPR, wiekiem, długością współpracy i pozycją zawodową). Algorytm działa na zasadzie zadawania pytań zbiorowi danych, które powodują podział zbioru na dwie części (przykładowe pytanie: „Czy TW należał do PZPR?”), a następnie sprawdzania, jak wpływa odpowiedź na dane pytanie na wiedzę o tym, czy dany TW składał donosy personalne. Szczególną zaletą algorytmu jest zdolność testowania zależności zachodzących między pojedynczą zmienną, którą możemy określić jako skutek, a całym zbiorem zmiennych, z których chcemy wyodrębnić najbardziej prawdopodobne przyczyny zaistnienia konkretnego skutku (Younas i in., 2022). W celu „uruchomienia” algorytmu musimy zatem podzielić nasze dane na zmienne określone jako „zależne” i „niezależne”, a następnie na część treningową (w której algorytm będzie doszukiwał się zależności między zmiennymi niezależnymi a zależnymi) oraz testową, w której będzie testował, na ile określone zależności dają możliwość „zgadnięcia”, kiedy występuje określona wartość zmiennej zależnej.

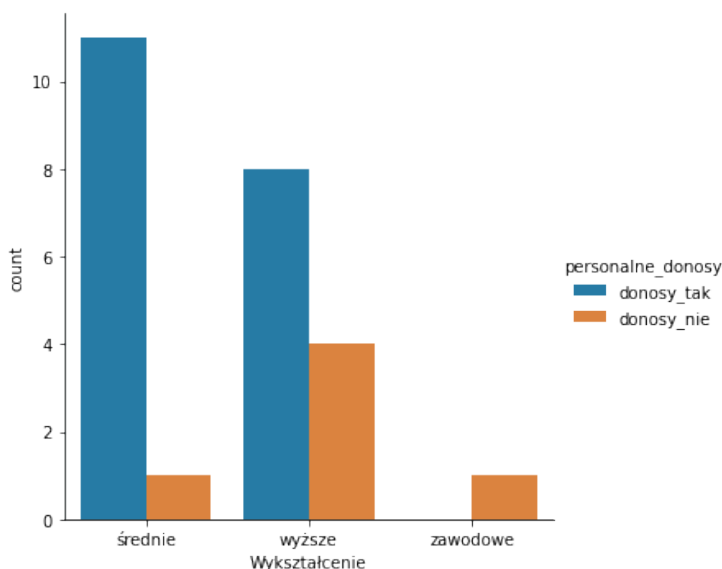
Na potrzeby naszych badań dokonaliśmy implementacji algorytmu drzewa decyzyjnego, wykorzystując moduł do badań naukowych języka programowania python: scikit-learn 1.3.2. W wyniku przeprowadzonych działań uzyskaliśmy następujące efekty zwizualizowane w postaci wykresu drzewiastego:



Grafika 6. Wykres drzewa decyzyjnego obrazujący, w jaki sposób dzielenie zbioru danych pytaniami o odpowiedzi binarnej (tak lub nie) wpływa na wiedzę o fakcie składania donosów. Strzałki odchodzące w lewo (*true*) wskazują nam kierunek pozytywnej korelacji z faktem składania donosów. Efektem zadawania pytań jest zmiana w poziomie informacji mierzona entropią. Wartość entropii (*entropy*) ukazana jest dla całej alternatywy (liścia) i dla poszczególnych składowych pytań (*value*). Pytania dzielą zbiór na podzbiory zawierające określoną część (ujętą procentowo) elementów (*samples*)

Miarą zdobywania wiedzy na temat zmiennej zależnej, który to proces jest zobrazowany na powyższym wykresie drzewiastym, jest entropia stanowiąca pomiar zmiany w informacji obecnej w zbiorze danych – konkretnie entropia jest logarytmem naturalnym odwrotności prawdopodobieństwa zaistnienia danego faktu (Maszczyk, Duch, 2008). Parametr entropii pokazuje nam, jak wraz z zadaniem pytaniem wzrasta lub spada nasza pewność na dany temat, przy czym ważna jest dynamika jej spadku. Wartość entropii równa 1 oznacza pełną niepewność co do faktu, wartość równa 0 to całkowita pewność co do faktu. Zmiana w wartości (*value*) entropii dla zbioru danych dokonuje się w wyniku zadawanych pytań i obliczana jest odrębnie dla każdej odpowiedzi oraz łącznie dla całej alternatywy. Jeżeli w wyniku zadania pytania nie uzyskujemy pewności na temat faktu, sumaryczna wartość entropii (*entropy*) obecnej na danym poziomie przekracza jeden i konieczne jest zadanie kolejnego pytania i zejście poziom niżej.

W interpretacji wyników uzyskanych z analizy wykorzystującej algorytmy uczenia maszynowego nie należy zapominać, że zależności wychwytywane przez algorytm mają charakter statystyczny, zadaniem badacza-interpretatora jest umieszczenie ich w odpowiednim, ludzkim, politycznym i historycznym kontekście. Z oglądu wykresu drzewiastego możemy się dowiedzieć, że na pierwszym (licząc od góry) poziomie wyodrębniania tych zmiennych, które zwiększają naszą wiedzę na temat tego, czy dany TW donosił, czy nie, należy umieścić fakt nieotrzymywania nagród – co wynika z tego, że nagrody wypłacane były bardzo rzadko w grupie badanych TW, zatem nienagradzani stanowią statystyczną większość. Także brak przynależności do PZPR jest taką większością kategorią, założenia o nienagradzaniu i nieprzynależności do partii wszelako zwiększają naszą wiedzę (zachodzi redukcja entropii w algorytmie) na temat faktu donoszenia. Na kolejnych (trzecim i czwartym od góry) poziomach drzewa dowiadujemy się, że przynależność do określonych grup zawodowych wywiera wpływ na naszą wiedzę o donoszeniu. Algorytm wyłania tu dwie grupy – jedna bardzo nieliczna, lecz jednoznaczna pod interesującym nas względem (jeden przykład z faktem składania donosów), druga znacznie bardziej liczna (kierownik), w której znajdowało się o wiele więcej donosicieli niż ludzi wstrzymujących się od tej czynności (19 : 6). Kolejną kategorią zwiększającą wiedzę na temat faktu donosicielstwa jest wykształcenie średnie, które w przypadku podgrupy kierowników częściej wiązało się z donosicielstwem niż wyższe bądź zawodowe, co sprawdziliśmy odrębnie.



Grafika 7. Rozkład danych dotyczących kategorii „kierowników” na zmienne dotyczące wykształcenia (oś X) i faktu składania donosów (legenda)

Zaprezentowane powyżej działanie algorytmu wzbogaca naszą wiedzę, wskazując na potencjalnie ważne zależności, lecz dostarczanych wyników nie powinniśmy traktować bezkrytycznie. Algorytm drzewa klasyfikacji i regresji działa w oparciu o wartość „czystości” podziałów kreowanych pytaniami (zdolność pytań do redukcji entropii) i nie zawsze zwraca nam uwagę na kategorie istotne z szerszej perspektywy, z drugiej strony jego zdolność do zestawiania i testowania dużej ilości danych bywa przydatna tam, gdzie „nieuzbrojony” w perspektywę umysł badacza męczy się ogładem niejednoznacznych faktów.

Obraz TW – analiza prozopograficzna

Metoda prozopograficzna koncentruje się na zbiorczej analizie danych opisujących określoną grupę ludzką. Zazwyczaj jest to grupa dość jednorodna: członkowie jednej klasy, organizacji, grupy zawodowej (Carnegie, McBride, 2023), w naszym przypadku zdecydowaliśmy się jednak na dokonanie opisu prozopograficznego osób powiązanych ze sobą (poza zbliżonym miejscem zamieszkania) faktem znalezienia się w orbicie wpływów SB jako tajni współpracownicy. Nie jest to oczywiście grupa monolityczna, niemniej metoda prozopograficzna dopuszczana jest również wobec grup heterogenicznych (Ankoud, 2020), jeżeli występują one we wspólnym systemie (który fakt tu zachodzi). Pytania badawcze o elementy różnicujące członków grupy są także ważnymi pytaniami badawczymi (Øland, 2013), nas zatem interesuje przede wszystkim to, dlaczego niektóre osoby były bardziej skłonne do donoszenia, inne mniej? W poprzedniej części artykułu usiłowaliśmy odpowiedzieć na to pytanie, posługując się metodami statystycznymi i algorytmami uczenia maszynowego, teraz przyjrzymy się zagadnieniu w bardziej tradycyjny dla badań historycznych sposób.

Z rozkładu prawdopodobieństwa wynika, że najczęściej nasz TW przed rozmową z funkcjonariuszem SB nie był wcześniej rozpracowywany; możliwe, że dawno temu został skazany na karę grzywny lub popełnił jakieś wykroczenie. SB nie dysponuje jednak materiałami obciążającymi ani kompromitującymi, a on sam w rozmowie deklaruje, że dba o ład i porządek, wyraża pozytywny stosunek do władzy. Takie zapewnienia zadowolają funkcjonariusza. W swojej pracy przyszły TW zajmuje stanowisko średniego szczebla lub kierownicze, ma rozeznanie w swoim środowisku i może cieszyć się pewnym zaufaniem ze strony innych osób. Podczas rozmowy jest uprzejmy i nie stawia warunków, rozumie potrzebę pomocy, a funkcjonariusz zapewnia go, że wszelkie ewentualne poniesione koszty zostaną mu zwrócone. Funkcjonariusz przyjmuje pisemne zobowiązanie do współpracy, a chwilę później tajny współpracownik otrzymuje pierwsze zadanie o niewielkim stopniu trudności.

To zadanie kontrolne i bardzo możliwe, że SB już dysponuje odpowiednią wiedzą, a jeżeli nie, to prawdomówność i rzetelność informującego zostaną skontrolowane poprzez rozmowę z innym TW. Zresztą podczas tej rozmowy przekazał już pewne ogólne informacje o zakładzie pracy, o nastrojach wśród załogi lub opowiadał o nieprawidłowościach, drobnych oszustwach i przekrętach innych pracowników. Za informacje zazwyczaj nie było pobierane wynagrodzenie, a jeżeli tak, to rzadko i nie były to duże pieniądze.

Jeśli chodzi o wykonywaną pracę i jej charakter, można zauważyć, że aż 29 TW zajmowało stanowiska kierownicze (od kierownika średniego szczebla do dyrektora/prezesa), a kolejnych 28 było w miejscu pracy albo pracownikami umysłowymi, albo specjalistami i ekspertami. Do grona TW wykonujących bardziej prestiżowe zawody należy też zaliczyć weterynarza, inżyniera chemii oraz pilota. Robotnikami było 13 TW, zawodowymi kierowcami – 7, nauczycielami – 4, rolnikami i ogrodnikami – 6.

Kategoria *Długość współpracy* może zaskakiwać, ponieważ średnia długość współpracy TW wynosi aż 3,7 lat, przy czym negatywnie koreluje ze skłonnością do składania donosów (algorytm wskazuje, że czas współpracy poniżej 1,5 roku sprzyjał uchylaniu się od tej czynności).

Przedstawione wyżej sytuacje nie świadczą o tym, że prezentowane dane należy z góry odrzucić jako niewiarygodne, lecz pokazują, że długość współpracy staje się użyteczną kategorią dopiero po jej zestawieniu z innymi kategoriami: liczbą zachowanych informacji (czyli ile zachowało się dokumentacji z odbytych spotkań), liczbą stron akt (zachowanych stron) oraz liczbą przeprowadzonych kontroli. W każdym razie najczęstsze trzy wartości prezentują się tak: 4 i 2 lata współpracy występują *ex aequo* po 22 razy, rok współpracy wystąpił 21 razy, współpraca trwająca 3 lata wystąpiła 19 razy. Co może wydać się nietypowe, naprawdę długotrwałą, sześciolletnią współpracą wystąpiła aż 13 razy. Dwóch TW współpracowało z SB po 10 lat (współpracę zaczęli przed powstaniem Wydziału VI i po 1985 r. zostali przekazani kolejnym funkcjonariuszom).

Pewną trudność w badaniu omawianych materiałów stanowi fakt, że wiele zachowanych teczek jest niekompletnych lub że teczki pracy zostały zniszczone. Skalę tego zjawiska można dość dobrze oszacować dzięki określeniu liczby stron w każdej teście oraz poprzez informacje o zniszczeniu teczki pracy. To, żeteczka personalna liczy 40, 60 czy 80 stron, nie mówi nam zbyt wiele, ale jeżeli przykładowo znajdziemy w środku wykaz spotkań lub wykaz wypłat liczący 20 pozycji – a przekazywanych informacji brakuje – będzie jasne, że wspomniane i sporządzone podczas spotkań doniesienia zostały zniszczone. Doniesienia te, zwłaszcza w przypadku dłuższej współpracy, powinny trafiać do teczki pracy. Na szczęście, kiedy dochodziło do zakończenia współpracy, funkcjonariusze SB najczęściej robili adnotację o tym,

żeteczka pracy została zniszczona. W przypadku badanych akt 69 razy zapisano, że teczkę pracy zniszczono, 53 razy tego nie napisano (co nie znaczy wcale, że teczka przetrwała), raz funkcjonariusz napisał, że teczkę pracy zniszczono, co zostało następnie przekreślone (teczka pracy nie zachowała się mimo skreślenia) („Stary”, IPN, Sz 0010/1988). Ponadto w co najmniej czterech przypadkach teczka pracy została po prostu włączona do teczki personalnej. Statystycznie zachowana teczka liczy średnio 47 stron; najkrótsze liczą po 30, najdłuższa ma 245 stron.

Należy ocenić, że ze zdecydowanej większości badanych akt TW usunięto część przekazywanych informacji. Jedynie około dziesięciu teczek jest na tyle obszernych i zawiera na tyle liczne doniesienia, że można zaryzykować (znowu ostrożnie) stwierdzenie, że są „raczej kompletne”. Ten problem dość dokładnie opisuje kategoria *Liczba zachowanych informacji*. W przypadku przytłaczającej większości teczek (93 j.a. ze 123) zachowała się tylko pierwsza, pojedyncza informacja udzielona przez TW podczas spotkania, na którym został zwerbowany. Wiemy, że były kolejne spotkania, ale wiemy też, że sporządzone w ich trakcie doniesienia zostały później zniszczone.

Nie oznacza to jeszcze, że nie możemy w ogóle oceniać współpracy. W zdecydowanej większości akt zachowały się przecież pisane przez funkcjonariuszy plany wykorzystania TW lub analizy dotychczasowej współpracy, podobnie zresztą jak uwagi i zalecenia przełożonych. Omawiane materiały prezentują po prostu perspektywę SB, ponieważ taka jest nieodwracalna specyfika tych materiałów źródłowych.

Ważnym kryterium jest również to, czy dany TW był nagradzany za współpracę, a jeżeli tak, to w jaki sposób. Kategoria *Czy nagradzany* pokazuje, że większość TW – czyli 89 – w czasie współpracy nie była nagradzana (co wyraźnie wpływa na pracę naszego algorytmu). Pozostałe 34 nagradzane przez SB osoby mogły w trakcie współpracy liczyć na pieniądze i upominki. Wysokość wynagrodzenia została ujęta w kategorię *Łączna wysokość wynagrodzenia*, obejmującą sumy od 2 tys. złotych do 45 tys. złotych. Należy jednak pamiętać, że ta druga wartość stanowi sumę otrzymaną w pewnym okresie i na przestrzeni wielu spotkań. SB nie wypłacała swoim TW jednorazowo aż tak wielu pieniędzy. Dwóch najlepiej wynagradzanych TW, do których łącznie trafiło po 45 tys. złotych, otrzymało je na przestrzeni odpowiednio dwóch i czterech lat współpracy. Dlatego też dobrze jest zestawiać wynagrodzenia z długością współpracy. 28 osób wśród badanych TW stanowili członkowie PZPR. Teoretycznie bycie członkiem PZPR powinno stanowić poważny problem dla SB i tacy ludzie powinni współpracować jako KO (kontakt operacyjny), a nie TW (Musiał, 2019, s. 166). Jednak w przypadku badanej grupy można stwierdzić, że werbowano również członków PZPR, choć nie tak często jak inne osoby, co także stanowi czynnik wywierający wpływ na pracę algorytmu uczenia maszynowego.

Kim byli TW, którzy donosili na konkretne osoby? Trudno odpowiedzieć jednoznacznie na to pytanie. O wiele łatwiej określić, ilu prawdopodobnie tego nie robiło: to przede wszystkim konsultanci oraz wszyscy, którzy przy udzielaniu informacji koncentrowali się wyłącznie na ogólnych problemach typu „nieprawidłowości gospodarcze”. Ci pierwsi byli werbowani po to, żeby służyć SB swoją fachową wiedzą, pisać analizy ekonomiczne, ekspertyzy czy oceny inwestycji. „Zwykli” TW informowali zarówno o funkcjonowaniu miejsca pracy, jak i o zachowaniu innych. Trudno wyobrazić sobie, żeby w informacjach nie padały żadne nazwiska, zwłaszcza w przypadku osób na kierowniczych stanowiskach (a więc odpowiedzialnych za funkcjonowanie „obiektu”), a jeszcze trudniej rozstrzygnąć, czy jeżeli dany TW pisze np. o prezie swojej spółdzielni, to czy robi to ze szlachetnych pobudek (poprawa działania spółdzielni), czy wręcz przeciwnie (np. osobista chęć zaszkodzenia nielubianemu przełożonemu). Dodajmy do tego, że bardzo często zachowała się jedynie pierwsza informacja, dlatego ocena intencji jest skomplikowaną, niepewną kwestią. Akta SB są źródłem, dzięki któremu zdobywamy wiedzę o wydarzeniach, lecz nie zawierają wszystkiego, gdyż zarówno funkcjonariusz, jak i TW dokonywali selekcji informacji (SB wyznaczała zadania na podstawie zainteresowań i potrzeb, TW sporządzał informację). Często nie ma w nich również śladów nacisków, zawołowanych grózb dotyczących nie tylko kariery, lecz także możliwości ułożenia sobie godziwego życia w ogarniętym kryzysem kraju. Większość TW to osoby w średnim wieku, mające rodziny i poza stałym zatrudnieniem niemające innych źródeł utrzymania. Fakt, że częściej donosili kierownicy niż szeregowi pracownicy, może nam powiedzieć co nieco o tym, że ich łatwiej można było tego stanowiska pozbawić. To, że wśród nich częściej personalne donosy składały osoby z wykształceniem średnim niż wyższym, może świadczyć o niepewności co do zajmowanego stanowiska i większej wrażliwości na naciski oficera prowadzącego.

Wnioski

W zagadnieniu o tak dużej złożoności i zarazem społecznej „wrażliwości” istotne jest zachowanie wielości perspektyw. Taka możliwość powstaje, gdy konfrontujemy ze sobą różne metody. Zimne spojrzenie „uczenia maszynowego” nieodróżniające dobra od zła ani prawdy od fałszu, a działające na zasadzie zależności statystycznych, możemy zestawić z refleksją historyka rozumiejącego kwestie etyczne. Sam fakt zależności określonych kategorii pochodzi „z komputera”, jej interpretacji może dokonać tylko człowiek. Współczesne, komputerowe, metody badawcze stwarzają

możliwości, które warto wykorzystać, jednak nie są i najprawdopodobniej nie będą autonomiczne w tworzeniu opisu przeszłości.

Bibliografia:

- Ankoud, R. (2020). An Approach to Studying Elites and Social Groups. *Almuntaqa*, 3(1), 70–85.
- Basuchoudhary, A., Bang, J. T., David, J., Sen, T. (2021). *Identifying the Complex Causes of Civil War: A Machine Learning Approach*. Palgrave: Springer.
- Brzechczyn, K. (2012). Problem wiarygodności teczek i opartej na nich narracji historycznej. Kilka uwag metodologicznych. *Pamięć i Sprawiedliwość*, 20(2), 53–77.
- Brzeziński, P. (2020). Prominenci i marionetki. Protopografia egzekutywy KW PZPR w Gdańsku w latach 1975–1990. *Rocznik Gdański*, 109–141.
- Buryła, S. (2009). Artyści i ich opiekunowie. *Znak*, 650–651, 138–143.
- Carnegie, G. D., McBride, K. M. (2023). Prosopography and Microhistory: Illuminating Historical Actors. *Handbook of Historical Methods for Management*, 245–263.
- Cranmer, S. J. (2019). Introduction to the Virtual Issue: Machine Learning in Political Science. *Political Analysis*, 1–9.
- Dudek, A. (1998). Sutanny w służbie Peerelu. *Karta: kwartalnik historyczny*, 8(25), 110–114.
- Gałkowska, A. (2007). Idealny TW. Niektóre psychologiczne aspekty werbunku tajnych współpracowników SB w świetle materiałów szkoleniowych MSW. *Politeja*, 7, 311–330.
- Grimmer, J., Roberts, M. E., Stewart, B. M. (2021). Machine Learning for Social Science: An Agnostic Approach. *Annual Review of Political Science*, 24, 395–419.
- Hołda, M. (2009). Pseudonimy tajnych współpracowników UB, SB – analiza lingwistyczna. *Język Polski*, 4–5, 308–321.
- Loh, W. Y. (2011). Classification and Regression Trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 14–23.
- Łaszczyński, M. (2010). „Krag” i jego krag. Wydawnictwo w świetle relacji. *Pamięć i Sprawiedliwość*, 16(2), 139–170.
- Maszczyk, T., Duch, W. (2008). Comparison of Shannon, Renyi and Tsallis Entropy Used in Decision Trees. *Proceedings*, 9, 643–651.
- Michalski, A. (2006) Dyskusja panelowa – Materiały IPN jako źródło historyczne. Szansa czy pułapka historyków. *Śląskie Studia Historyczne*, 12, 303–307.
- Musiak, F. (2019). *Podręcznik bezpieki. Teoria pracy operacyjnej Służby Bezpieczeństwa w świetle wydawnictw resortowych Ministerstwa Spraw Wewnętrznych PRL (1970–1989)*. Kraków–Warszawa: Wyd. IPN.
- Øland, T. (2013). The Diversity of ‘Progressive School Pedagogues’ 1929–1960: A Space of Opposites Making Society Making the Child. *Praktiske Grunde: Tidsskrift for kultur-og samfundsvidenskab*, 7, 5–24.
- Pastuszewski, S. (2019). Osobowe źródła informacji Służby Bezpieczeństwa w bydgoskiej przestrzeni kulturalnej lat siedemdziesiątych i osiemdziesiątych XX wieku. *Kronika Bydgoska*, 40, 261–296.

- Skibiński, P. (2003). Infiltracja komunistycznych służb specjalnych w polskim Kościele – co już wiemy? Informacja na temat stanu badań. *Teologia Polityczna*, 1, 57–70.
- Wicenty, D. (2022). O znaczeniu fałszywego świadectwa. Przypadek TW „Jakuba” a fikcje operacyjne SB dotyczące tajnych współpracowników. *Dzieje Najnowsze*, 54(3), 151–174.
- Wiścicki, T. (2005). Nie każdy musiał. *Więź*, 47(561), 5–8.
- Younas, N., Ali, A., Hina, H., Hamraz, M., Khan, Z., Aldahmani, S. (2022). Optimal Causal Decision Trees Ensemble for Improved Prediction and Causal Inference. *IEEE Access*, 10, 13000–13011.
- Zarzycki, M. (2004). Przypadek Henryka Karkoszy. *Karta*, 41, 143–145.

Źródła archiwalne IPN Szczecin (Pseudonim TW, sygnatura teczek):

Pseudonim	Sygnatura	Pseudonim	Sygnatura
Adam	IPN Sz 0010/1197	Ely	IPN Sz 0010/1998
Adam	IPN Sz 0010/1355	Energetyk	IPN Sz 0010/2994
Adamowicz	IPN Sz 0010/3000	Fachowiec	IPN Sz 0010/1571
Agata	IPN Sz 0010/3041	Feliks	IPN Sz 0010/3072
Alek	IPN Sz 0010/1942	Franek	IPN Sz 0010/1661
Anatol	IPN Sz 0010/3038	Franek	IPN Sz 0010/2826
Andrzej	IPN Sz 0010/1999	Grażyna	IPN Sz 0010/2374
Andrzej	IPN Sz 0010/3117	Grażyna	IPN Sz 0010/3113
Andrzej 1	IPN Sz 0010/944 t. 1–2	Grzejnik	IPN Sz 0010/2001
Ania	IPN Sz 0010/1517	Handlowiec	IPN Sz 0010/3033
Anna	IPN Sz 0010/1501	Hanka	IPN Sz 0010/1210
Bandera	IPN Sz 0010/1742	Harry Coss	IPN Sz 0010/2222
Barbara	IPN Sz 0010/1747	Henia	IPN Sz 0010/1424
Bartek	IPN Sz 0010/3429	Hodowca	IPN Sz 0010/3104
Bogdan	IPN Sz 0010/2334	Jacek	IPN Sz 0010/1932
Bolek	IPN Sz 0010/3177	Jadwiga	IPN Sz 0010/3039
Buk	IPN Sz 0010/2013	Jakub	IPN Sz 0010/2775
Bukowski	IPN Sz 0010/3076	Jan	IPN Sz 0010/3071
Byk	IPN Sz 0010/3067	Jaz	IPN Sz 0010/2384
Cepelia	IPN Sz 0010/3078	Jodłowski	IPN Sz 0010/1744
Chłodziarz	IPN Sz 0010/1743	Józef	IPN Sz 0010/1663
Czarny	IPN Sz 0010/2228	Justyna	IPN Sz 0010/1423
Dag	IPN Sz 0010/3077	Kaktus	IPN Sz 0010/3568
Danuta	IPN Sz 0010/1720	Kazimierz	IPN Sz 0010/2016
Dorota	IPN Sz 0010/2208	Klemens	IPN Sz 0010/2213
DS/SD	IPN Sz 0010/2995	Kola	IPN Sz 0010/2425
Edward/Leszek	IPN Sz 0010/2705	Konrad	IPN Sz 0010/2372
Ekspert	IPN Sz 0010/3040	Konrad	IPN Sz 0010/3105

Pseudonim	Sygnatura	Pseudonim	Sygnatura
Konsultant	IPN Sz 0010/3058	Robert	IPN Sz 0010/1564
Korek/Żuraw	IPN Sz 0010/2807	Roman	IPN Sz 0010/1660
Kostrzewa	IPN Sz 0010/983	Ropa	IPN Sz 0010/1119
Kot	IPN Sz 0010/3179	Róża	IPN Sz 0010/3070
Kowalewski	IPN Sz 0010/4055	Rubin	IPN Sz 0010/1673
Lustrator	IPN Sz 0010/2999	Rysio	IPN Sz 0010/2014
Marek	IPN Sz 0010/1118	Siwek	IPN Sz 0010/1738
Marek	IPN Sz 0010/1422	Sławek	IPN Sz 0010/2072
Marek	IPN Sz 0010/2373	Słowik	IPN Sz 0010/3116
Marek	IPN Sz 0010/3176	Specjalista	IPN Sz 0010/2728
Marian	IPN Sz 0010/3101	Stary	IPN Sz 0010/1988
Markiewicz	IPN Sz 0010/1569	Stefan	IPN Sz 0010/3075
Marta	IPN Sz 0010/1662	Stefania	IPN Sz 0010/2106
Mateusz	IPN Sz 0010/2017	Strop	IPN Sz 0010/2806
Maurycy	IPN Sz 0010/3035	Sylwester	IPN Sz 0010/1933
Mek	IPN Sz 0010/2452	TA/AT	IPN Sz 0010/2206
Mirka	IPN Sz 0010/1056	Tara	IPN Sz 0010/3069
Monika	IPN Sz 0010/2229	Teodor	IPN Sz 0010/2007
Motor	IPN Sz 0010/2209	Tomasz	IPN Sz 0010/2776
Nauczyciel	IPN Sz 0010/1596	Tor	IPN Sz 0010/1659
Nr 13	IPN Sz 0010/1664	Tulipan	IPN Sz 0010/2811
Odra	IPN Sz 0010/1973	Wera	IPN Sz 0010/2205
Ogrodnik	IPN Sz 0010/3081	Wiktor	IPN Sz 0010/1997
Ola	IPN Sz 0010/1987	WJ	IPN Sz 0010/3118
Oleg	IPN Sz 0010/1354	Wolski	IPN Sz 0010/4172
Olek	IPN Sz 0010/2377	Zefir	IPN Sz 0010/3068
Oneks	IPN Sz 0010/1594	Zenek	IPN Sz 0010/2051
Orzeł	IPN Sz 0010/2000	Zetor	IPN Sz 0010/2207
Paweł	IPN Sz 0010/3124	Zygmunt	IPN Sz 0010/3048
Piotr	IPN Sz 0010/1737	Zygmunt	IPN Sz 0010/3073
Prezes	IPN Sz 0010/2015	Zygmunt	IPN Sz 0010/3165
Proat	IPN Sz 0010/3102	Żak	IPN Sz 0010/2375
Rak	IPN Sz 0010/2727	Żuk	IPN Sz 0010/3074
Rak	IPN Sz 0010/3036		

Protokół z kontroli kompleksowej pracy Wydziału VI WUSW w Szczecinie, 8. Stan pracy z osobowymi źródłami informacji. AIPN, Sz 0012/375 t. 175