

Andrzej Nowosad

## **METODY I TECHNIKI POZYSKIWANIA I PRZETWARZANIA INFORMACJI MEDIALNEJ NA POTRZEBY BIAŁEGO WYWIADU**

Analiza informacji medialnej jest, jeśli nie podstawowym, to jednym z najważniejszych motorów pozyskiwania danych o krajach i regionach. Wspecjalizowane jednostki, które przeglądają obcą prasę, znajdują się praktycznie przy każdym rządzie i placówce dyplomatycznej. To, w jaki sposób lokalna prasa ocenia działalność swej władzy, ma ogromny wpływ nie tylko na to, co o rządzie danego kraju myśli jego społeczeństwo, ale też na decyzje obcych państw wobec tych krajów. Innymi słowy – to, co się mówi i pisze o Polsce w Polsce jest powszechnie analizowane i przetwarzane w wyspecjalizowanych placówkach obcych krajów.

Aż 84 proc. ogółu mieszkańców Polski powyżej 15 roku życia negatywnie ocenia funkcjonowanie systemu politycznego w kraju. Tylko co dziesiąty Polak (10 proc.) nie ma zastrzeżeń co do funkcjonowania polskiego systemu politycznego. Także niewiele, bo 6 proc. osób nie ma wyrobionego zdania w tej sprawie. W Polsce, podobnie jak w innych krajach „młodej demokracji”, większość badanych niezmiennie wyraża niezadowolenie również ze sposobu, w jaki funkcjonuje demokracja

– wynika z badań IPOS i CBOS ([www.IPSOS.pl](http://www.IPSOS.pl), [www.CBOS.com.pl](http://www.CBOS.com.pl)). I nie należy mieć złudzeń. Taką samą opinię o Polsce usłyszymy od bazujących na źródłach jawnych analitykach obcych państw. 95 proc. informacji wywiady obcych państw czerpią bowiem ze źródeł ogólnodostępnych, 4 proc. z półjawnych, a tylko 1 proc. z tajnych (Doronin 2003, 28). Dziś, ze względu na koszty pozyskiwania informacji, to prasa powszechnie dostępna w Internecie stała się podstawowym źródłem informacji dla obcych rządów, a tym samym i ocen wystawianych za granicą polityce gospodarczej krajów, starających się o uzyskanie wiarygodności politycznej, kredytowej czy inwestycyjnej.

Nie jest tajemnicą, że w każdym kraju jedne media napędzają inne w spirali informacji, a dla „świadomości informacyjnej” społeczeństwa czy obcych państw nawet kaczka dziennikarska powtórzona kilka razy przez różne wiarygodne

źródła staje się „prawdą”. Tego typu „prawda” zwykle trafia do sprawozdań prasowych z ambasad obcych krajów za granicę i na jej podstawie w poszczególnych stolicach podejmuje się odpowiednie decyzje polityczne, choć w rzeczywistości może to być wielokrotnie powtarzana półprawda bądź kłamstwo. W marcu 2005 roku np. dziennikarze polscy cytowali wiadomości podawane na łamach innych mediów 6227 razy. Najwięcej nawiązań dotyczyło „Gazety Wyborczej”, na którą powoływano się 1086 razy, a na Rzeczpospolitą 1066 razy. „Rzeczpospolita” i „Gazeta Wyborcza” – liderzy w rankingu mediów najczęściej cytowanych poza przeglądami prasy (Witkowski, *Media cytowane*) – nie są jednak źródłami w pełni obiektywnymi a i zdarzają im się dość często kaczki dziennikarskie czy wręcz świadome przekłamania. Media te bowiem uczestniczą w Polsce w aktywnej grze politycznej, reprezentują interesy swych właścicieli oraz skupionych wokół nich grup lobbingu politycznego.

W pracy analityka informacji na potrzeby rządu najważniejsze jest ustalenie zarówno prawdziwości jak też i pierwoźródła informacji. Za Vriensem można uznać, że informacja to pojedyncze, odebrane sygnały i ich interpretacja, które w sposób jednostkowy mogą rozwijać się do danych i wprowadzać coś nowego dla analityka. Przy czym trzy sygnały można traktować już jako wiarygodne dane (Vriens 2004: 7). Warunkiem byłoby jednak, aby te trzy sygnały nie pochodziły z jednego źródła, bądź źródła opartego na pierwoźródle i – przede wszystkim – by informacja była wiarygodna.

Aby otrzymać pierwoźródło informacji można zastosować metodę Bradforda, który w 1934 r., będąc bibliotekarzem, opracował ogólną zasadę pozwalającą na porządkowanie dokumentacji w sposób racjonalny i ekonomiczny. T. Lafouge (1993: 26), a za nim Martinett (1999: 80-81) tak opisują prawo Bradforda:

Jeśliby czasopisma naukowe uporządkować rosnąco pod względem produktywności tematycznych materiałów, to istnieje liczba dodatnie  $k$  i jądro  $r_o$  periodyków publikujących  $R(r_o)$  artykułów na dany temat w taki sposób, że jeśli chce się otrzymać  $nR(r_o)$  (przy czym  $n$  oznacza jakąś całość) artykułów na ten sam temat, należy wziąć pod uwagę liczbę  $P$  periodyków, którą można obliczyć za pomocą wzoru:

$$P = r_o + K r_o + K^2 r_o + \dots + K^{(n-1)} r_o$$

Co oznacza, że ogół środków masowego przekazu można podzielić na grupy podstawowych przeglądów dotyczących  $r_o$  tzn.

$$K r_o, K^2 r_o, \dots, K^{(n-1)} r_o,$$

z których każdy dostarcza  $R r_o$  artykułów na dany temat (Martinett 1999: 80 – 81, Lafouge T., 1993).

Metoda Bradforda może być szczególnie użyteczna w pracy z hipermediami w hipertekście oraz w wykrywaniu gazet, które bazują na nakręcaniu spirali informacyjnej „kaczek” prasowych.

Następnie źródła informacji można podzielić pod względem wiarygodności na:

- 1) wiarygodne – informacje z całą pewnością można uznać za prawdziwe,
- 2) wiarygodne, w przypadku których istnieje ryzyko pomyłki lub subiektywizmu – informacje są zwykle prawdziwe, ale mogą być subiektywnie interpretowane,

- 3) mało pewne – informacje są zawsze niepewne i winny być sprawdzone,
- 4) podejrzane albo subiektywne – informacje niepewne i wymagają ostrożności (Kwieciński 1999: 70, Wersig G., 2002; 2004).

Dla dyplomacji wiarygodność informacji zależy przede wszystkim od wiarygodności samego źródła informacji. Dla państwowych analityków Internet jest bardzo ważny, ale hipertekst wszelkie segmenty informacji i ich nośniki pod względem wiarygodności traktuje równorzędnie (Caching, K. Ch, Yahiko K. 2002: 276-290; Cooper M.D., 1983: 9-26; Cook, M. & Cook, C. 2000), zaś stosowanie metody Bradforda do setek tysięcy internetowych źródeł jest żmudne, czasochłonne i tym samym kosztowne. Pod tym względem zatem pozyskanie wiarygodnej informacji z taniego Internetu staje się więc nie tylko bardzo drogie, ale i niebezpieczne. W dyplomacji wysoki poziom entropii w opracowywanych informacjach jest o wiele bardziej niebezpieczny niż informacyjny głód (Doronin 2003, 185; Cronin B., Overfelt K., Fouchereaux K., Manzvanzvikie T., Cha M., & Sona E., 1994).

Dlatego też w przypadku źródeł internetowych, zanim zastosuje się metodę Bradforda, najlepiej skorzystać wstępnie z innej kategorii klasyfikacji zawartości przekazów:

- 1) odnoszących się do treści (co się mówi),
- 2) odnoszących się formy (jak się mówi).

Inaczej mówiąc, informacje należy posegregować według: podziału materialnego – co odpowiada formie i podziału formalnego – co odpowiada treści. W klasie kategorii treściowych wyróżnia się kryterium tematyki. Taka kategoria znajduje się w każdym kluczu klasyfikacyjnym stosowanym w analizie zawartości informacji. Liczba kategorii tematycznych w takich kluczach waha się od kilku do kilkudziesięciu (por. Pamula S. 1996; Pisarek W. 1983).

Zarówno w pozyskiwaniu, jak i przetwarzaniu oraz przekazywaniu informacji ważne jest, by robić to w sposób usystematyzowany, a nie chaotyczny, z wykorzystaniem zasady słów kluczowych, korzystając wyłącznie z profesjonalnie przesortowanych danych według komunikacyjnej klasyfikacji wagi informacji. Za Schiwelym można wyróżnić trzy sposoby przeprowadzenia pomiaru informacji:

- 1) pomiar nominalny – przyporządkowanie badanych jednostek do wyodrębnionych kategorii. Nie ma jednak tutaj mowy o tym, jakie relacje zachodzą między tymi kategoriami,
- 2) pomiar porządkowy – uporządkowanie kategorii ze względu na to, „ile” zmiennej w sobie zawierają. Dzięki temu osiągamy przybliżoną wizję skali, która reprezentuje naszą zmienną, a punktu przypisane jednostkom obrazują na jakim poziomie skali są one zlokalizowane,
- 3) pomiar interwałowy – jeżeli istnieje jednolita jednostka pomiaru dla skali, za pomocą której da się zmierzyć różnice pomiędzy wynikami (Schiwely 1997, wyd. polskie 2001: 89-90).

Adepci sztuki zbierania i segregacji informacji oraz analizy zjawisk bardzo często niesłusznie rezygnują z danych na wyższym poziomie pomiaru, przechodzą na poziom niższy, aby ułatwić sobie opracowywanie raportów. Pomiaru należy dokonywać tak precyzyjnie, jak to tylko możliwe, aby nie tracić informacji. Do

tego celu w analizie informacji może posłużyć roboczy konspekt, znany między innymi z metod analizy zawartości prasy. Konspekt powinien brać pod uwagę cztery wskaźniki:

- 1) wskaźnik syntaktyczności i semantyczności ogranicza się do przedstawienia samej treści, ale nie posuwa się do bezpośredniego opisywania jej ukrytych intencji,
- 2) wskaźnik obiektywności – kategorie mają być tak precyzyjne, aby różni analitycy uzyskali takie same wyniki,
- 3) wskaźnik systematyczności ma zapewnić eliminację analizy cząstkowej i stroniczej,
- 4) wskaźnik ilościowy – najbardziej charakterystyczny dla analizy zawartości (por. Pisarek 1983: 29-30).

Pozostaje tu do spełnienia jeszcze jeden istotny warunek – źródła informacji i ich pochodne nie będą dublowane. Ogólnie rzecz ujmując, analiza zawartości ma być przydatną techniką służącą wnioskowaniu na podstawie systematycznego i obiektywnego identyfikowania precyzyjnie określonych informacji (por. Pamula 1996: 141). Pięć podstawowych zasad identyfikowania informacji i zjawisk to:

- 1) dać słowom (zjawiskom) właściwe znaczenie,
- 2) sprowadzić je do ich realnych proporcji,
- 3) stworzyć granice rozumowania aby nie zgubić się w zwodniczej przypadkowości,
- 4) określić wzajemne stosunki między informacjami (zjawiskami), aby ich koordynacja logiczna nie przekraczała analitycznych możliwości,
- 5) obserwować każdą informację (zjawisko) i wszystkie okołoinformacje (około-zjawiska) z całą ich relatywnością; unikać tworzenia kategorycznych wniosków i stwierdzeń, bo jest to tylko nieświadome pragnienie systematyzacji, czyli *de facto* dążenia do niepoznawalności badanych zjawisk. Konspekt bazy danych powinien zatem zawierać też Gerbnerowskie cztery podstawowe formy prezentacji informacji: istnienia, priorytetów, wartości i relacji zjawisk.

Tab. 1. Płaszczyzny analizy zawartości według G. Gerbnera

Płaszczyzny analizy	Istnienie	Priorytety	Wartości	Relacje
Założenia do zakresu analizy	Co jest?	Co jest ważne?	Co jest słuszne lub niesłuszne?	Co jest z czym i jak powiązane?
Pytania	Co jest dostępne dla publiczności, jak często i jak wiele?	W jakim kontekście lub porządku ważności?	W jakim świetle, z jakiej pozycji, z jakimi ocenami?	W jakiej całościowej strukturze logicznej, przyczynowej itp.?
Terminy i miary:	Uwaga (obecność, złożoność, frekwencja)	Nacisk (uporządkowanie, regenerowanie, skalowanie, intensywność)	Tendencja (krytyczna i różnicująca cechy i jakości)	Struktura (korelowanie, grupowanie, struktura akcji)

Źródło: G. Gerbner, L. Gross, W. Melody, *Communication Technology and Social Policy*, New York 1973, s. 565; [za:] T. Goban-Klas, *Media i komunikowanie masowe. Teorie i analizy prasy, radia, telewizji i Internetu*, Warszawa – Kraków 2002, s. 194.

Tomasz Goban-Klas podaje, że płaszczyzna pierwsza – istnienia – dotyczy tego „co jest?” przedstawiane. Płaszczyzna druga – priorytety – dotyczy ważności tematów, wątków, postaci w systemie ogółu przekazów. Mierzy się w niej nacisk, jaki został położony na prezentowane treści. Trzecia płaszczyzna – wartości – dotyczy systemu wartości i tendencji. Mierzy się zatem ukierunkowanie, tendencyjność oraz jakość cech przypisywanych różnym wątkom i zagadnieniom. Płaszczyzna czwarta – relacje – koncentruje się na najbardziej złożonych asocjacjach w obrębie analizowanego materiału, jak również i na syntezie rezultatów pozostałych wymiarów (Goban-Klas 2002: 194). Metodologia opracowania w ten sposób posegregowanych już materiałów powinna zawierać następujące kategorie w bazie danych:

- 1) informacja bazowa do dalszego opracowania analitycznego,
- 2) bieżąca informacja o faktach,
- 3) kategorie subiektywno-oceniające, zawierające ocenę i przestrogi, przy czym przy analizie artykułu należy zwrócić szczególną uwagę nie tylko na fakty, ale również na osobisty stosunek autora do opisywanego problemu.

Wobec tak skategoryzowanych danych można już zastosować podstawowe grupy sortowania informacji medialnej na potrzeby analizy pod kątem:

- 1) analizy przyczynowo-skutkowej – jeśli z a wynika b, a b wynika z c, to obserwując c, można określić a. Tego typu analiza przyczynowo-skutkowa prowadzi do tego, że obserwując zewnętrzne oznaki, można dojść do zrozumienia istotnych przyczyn obserwowanych zjawisk, których nie widzi się, ani nie odbiera w sposób bezpośredni.

- 2) analizy eksperckiej – szczególną grupę źródeł stanowią wypowiedzi prasowe ekspertów z jakiejś dziedziny. Praca z tą kategorią źródeł jest zawsze trudna, a to ze względu na przeciwstawne rezultaty. Tam, gdzie spotyka się trzech specjalistów, są co najmniej cztery zdania na temat omawianego zjawiska. Sąd eksperta pozwala jednak na nowo spojrzeć na istniejący problem, otrzymać materiały bazowe i dotrzeć do nieznanych do tej pory źródeł informacji. Ogólna ocena otrzymanych przy tym danych, przy ogólnej trudności, zwykle jest bardzo wysoka.

- 3) analizy realnych powiązań i możliwości – jest to określenie sieci powiązań między źródłami informacji (ze sfer politycznych czy gospodarczych) konieczne do przeprowadzenia dokładnej analizy informacyjnych związków pomiędzy źródłami. Należy w tej analizie wy badać związki personalno-towarzyskie i biznesowe między np. właścicielami środków masowych informacji czy pracujących dla nich dziennikarzy z organami władzy, kręgami bankowymi, grupami lobbingsowymi, współpracownikami przedsiębiorstwa czy jego szefostwa itd. (Doronin 2003, 18-31, Gerbner, Gross, Melody 1973: 553-573, Hasanali, Leavitt 2004).

## **Korzystanie z informacji w Internecie**

Wywiad analityczny za pomocą Internetu jest jednym z najbardziej perspektywicznych kierunków pracy służb bezpieczeństwa. Wielu nazywa już te działania wywiadem komputerowym (Hulnick 2004: 165-173). Wielkie zainteresowanie meto-

dami wywiadu analitycznego przejawiają zarówno państwowe, jak i niepaństwowe służby. Jest to związane z tym, że w Internecie znajdują się ogromne bazy informacji, które mogą być w orbicie zainteresowań zarówno pierwszych, jak i drugich. Według niektórych źródeł zasoby Internetu to dziś 550 mld dokumentów, z których 40 proc. jest dostępne bezpłatnie. Nawigację w tej sieci informacyjnych zasobów ułatwia ponad milion systemów wyszukiwawczych, katalogów i baz danych. Dlatego też w wielu krajach świata, a szczególnie w Rosji, USA, Niemczech i także i w Polsce utworzono specjalne wydziały analitycznego wywiadu w Internecie. Mają one za zadanie wyłapanie ważnych informacji w „cyfrowym oceanie”.

W Europie Zachodniej i USA pozyskiwanie informacji za pomocą Internetu staje się już bardzo dochodowym biznesem. Według danych prasowych tylko we Francji działa kilkadziesiąt firm, które zajmują się wyszukiwaniem dokumentów, w tym również tablic i rysunków istniejących w przestrzeni Internetu. Francuzi na przykład pracują w systemie inżynierii lingwistycznej MAAG, zorientowanej na informacyjno-analityczne zabezpieczenie takich kluczowych gałęzi francuskiej gospodarki, jak przemysł aerokosmiczny, transport i energetyka (Doronin 2003, 170).

### Programy do globalnych przeszukiwań

Do globalnego wyszukiwania w Internecie wykorzystuje się specjalne procesory poboru danych. W tym kontekście termin „procesor” nie ma nic wspólnego z mikroprocesorem, jest to część programu, która określa, w jaki sposób sam program segreguje, zarządza i analizuje dane wg w/w kategorii. Procesor poboru danych wykorzystuje zabezpieczenie programowe, które otrzymało nazwę „robot”. „Robot” wyprowadza niezbędną informację, wykorzystując cały arsenał środków lingwistycznej, semantycznej i statystycznej analizy. Działając w sposób autonomiczny, procesory danych przechwytyują każdą informację, wobec której powstało zapytanie, gdy ta tylko pojawi się w Internecie (Doronin 2003, 171).

Dziś istnieją dziesiątki takich systemów na świecie, między innymi: LEXI-MAPPE, SAMPLER, CANDIDE (The Hume Machine), HENOCH, IRIT, NEURODOC, NEUROWEB, PÉRICLES, SEMIOMAP itd. (Ich opis patrz: *Outils d'analyse textuelle*).

Pierwszy tego typu odtajniony procesor to francuski TAIGA (*Traitement automatique d'information geopolitique d'actualité*) – automatyczne opracowanie aktualnej informacji geopolitycznej. Ten programowy kompleks opracowano pierwotnie dla potrzeb francuskiego wywiadu, służył tam 11 lat – do 1987 – po czym został przekazany do wykorzystania komercyjnego. Zadania, które stawiają mu dziś cywilni specjaliści, zostały takie same: „Buszować po Internecie w celu przechwycenia cennej informacji z baz danych o patentach, notatkach agencji informacyjnych i publikacjach konferencji naukowych” (*TAIGA, Traitement Automatique*).

Kolejny francuski program do analizy jakościowej przygotowany przez firmę Acetic i naukowców z Uniwersytetu w Paryżu znany jest jako program TROPES – *Une logique d'Intelligence Artificielle*.

Wybór poszukiwanej informacji w TROPES jest zsynchronizowany z kluczowymi słowami i pojęciami łączącymi się jeśli chodzi o sens ze sobą. Na przykład *mirage* łączy się ze słowami „myśliwiec”, „samolot”, a słowo premier ze słowami „minister”, „polityk”, „rząd”. TROPES, *L'analyse de textes haut de gamme à la portée de tous*, pozwala w tym samym czasie jednocześnie analizować na wielu płaszczyznach teksty informacyjne wielkości setek tysięcy stron znormalizowanego maszynopisu, daje możliwość tworzenia „scenariuszy” pracy informacyjnej, na podstawie których automatycznie dokonuje się nie tylko poszukiwanie, ale również ukierunkowane grupowanie wyszukanych danych na wielu płaszczyznach. (*Tropes, Analyse de textes*).

NOEMIC, *The search engines*, który zamienił system TAIGA, nie tylko skanuje, ale również automatycznie łączy źródła, opracowując otrzymaną informację z prędkością jednego miliarda znaków na sekundę, bez względu na to, czy jest ona podana w formie gotowej bazy danych, czy też na przykład przekazywana elektronicznie przez agencję informacyjną w jakimkolwiek języku (Guisnel 2004). Jeśli postawi się mu zadanie wychwycenia na przykład wszystkich związków pomiędzy firmami amerykańskimi i firmami w Hongkongu, funkcjonującymi w sferze mikroelektroniki w ciągu ostatnich 15 lat, to do wypełnienia zadania wystarczy mu kilka godzin – zwraca uwagę Doronin (Doronin 2003: 172).

Analogiczny amerykański system TOPIC również początkowo był przygotowany dla potrzeb wywiadu wojskowego. Pojawił się on w rezultacie długich prac badawczych pod kontrolą CIA. Dziś także ten system został sprzedany do wykorzystania komercyjnego, a wszystkie prawa z nim związane należą do kalifornijskiej firmy „Verity”, światowego lidera w zbieraniu dokumentów i danych (Guisnel 2004, Doronin 2003: 203).

Większość tych programów działa według kilku podstawowych modeli ustalania relacji między informacjami i ich źródłami. Podstawowe modele to:

- 1) model hierarchiczny,
- 2) model danych sieciowych,
- 3) model relacji,
- 4) model obiektów zorientowanych.

Model hierarchiczny funkcjonuje od początku lat 60. Przedstawia wszystkie elementy według klasyfikacji czasowej i klasyfikacji ważności w tzw. grafach i drzewach grafów. Opiera się on na zasadzie hierarchii typów obiektów, jeden typ obiektu jest główny, a pozostałe znajdują się niżej w hierarchii i są mu podporządkowane. Między obiektem głównym i obiektami podporządkowanym obiektami powstaje więź wzajemna „jeden w stosunku do wielu” (Gatnar 1998, Kraaij, Wesel 2004, *Outils d'analyse textuelle*, Schwab 2005).

Model danych sieciowych jest związany z działaniem sieci komputerowych i powstał w latach 60. W modelu tym pojęcia obiektu głównego i podporządkowanych są rozszerzone. Każdy obiekt może być zarówno główny, jak i podporządkowany. Tym samym każdy obiekt może tworzyć klasyfikację danych, jak też być członkiem zbioru. Sieciowy model bazy danych jest podobny do hierarchicznego, ale charakter związków między obiektami jest zasadniczo inny.

W sieciowym modelu przyjęto wolny związek pomiędzy elementami różnych poziomów, to znaczy dopuszcza się związki „wiele w stosunku do wielu”. Systemem takim jest np. Cronos Plus.

W modelu relacji (*relational*) każdy zapis w bazie danych zawiera informację odnoszącą się (*relatet*) tylko do konkretnego obiektu. Ponadto z danymi dwóch typów można pracować jak z jedną całością, zbudowaną na znaczeniach związanych między sobą danych. Niezależność fizyczna modelu relacyjnego polega na tym, że model danych nie posiada żadnych opisów fizycznych. W rzeczywistości fizyczne przedstawienie stosunku i dróg dostępu opisywane jest niezależnie od opisu logicznego schematu relacji.

Model obiektów zorientowanych, w odróżnieniu od wyżej opisanych, gdzie informacje i procedury chroniono oddzielnie (dane i związki między nimi w bazie danych, a procedury w programie), pozwala chronić procedury obróbki, wraz z danymi. Taka wspólna ochrona uważana jest za krok naprzód w metodologii zarządzania danymi. Model obiektowo-zorientowanych utrzymuje związki typu „wiele w stosunku do wielu” (Doronin 2003, 206-210).

Warto zaznaczyć, że w celu uzyskania informacji o zjawiskach w ruchu wykorzystuje się działania na zbiorach rozmytych, gdzie zachodzą:

- 1) dychotomia, gdzie  $a = b$  i  $a \neq b$ ,
- 2) kolejność, gdzie  $a = b$ ,  $a > b$  i  $a < b$ ,
- 3) przechodniość, gdzie  $a > b \wedge b > c \Rightarrow a > c$ ,
- 4) asymetria, gdzie  $a > b \vee a < b$ .

Metoda zagęszczeń w podzbiorach rozmytych pozwala na badanie zjawisk, należących do różnych skupień i na dodatek na sprawdzanie między innymi miar: odległości, asocjacji i korelacji według zasad taksonomii numerycznej.

Dla wyznaczenia sposobu działania układu wystarczy znajomość samej macierzy wyrażającej metody sposobu działania jego elementów (macierzy transformacji –  $t$ ), oraz macierzy struktury układu, czyli sieci sprzężeń między elementami ( $s$ ), od czego zależy działanie badanego układu jako całości.

W analizie informacji na potrzeby rządów należy pamiętać, że stałe zmiany stanów układu to proces ich rozwoju, a kształtująca się formacja polityczna czy społeczno-gospodarcza może być rozpatrywana i jako stan, i jako proces. Matematycznie przedstawione czasowe prawo ruchu układu wyraża się w postaci wektorowego równania, wyznaczającego relacje między stanami wejść i wyjść w jednej chwili ( $t$ ) a stanem tych wejść i wyjść w chwilach późniejszych. W przypadku na przykład przebiegu działania mającego charakter jednorazowy w czasie, równanie algorytmiczne przybiera charakter:

$$X_{t+\theta} = TS(X_t)$$

$$Y_{t+\theta} = ST(Y_t)$$

(gdzie  $t$  oznacza chwilę początkową, a  $\theta$  czas reakcji)” (por. Nowakowski, Sobczak 1971: 45-72).

Aby uniknąć błędów analitycznych można przyjąć sprawdzoną już metodę klasyfikacji danych Roberta T. Dattoli (1969) i założyć, że zbiór badanych obiektów jest



podzielony na  $K$  tzw. klas początkowych:  $K_{1,1}, K_{1,2}, \dots, K_{1,k}$ . Dla każdej z tych klas zostaje wyznaczony jej profil zgodnie z zasadą klasyfikacji dla klasy spójnej (por. eksperymenty Cambridge Language Research Unit z lat 60. Needhama i Parkera-Rhodos; Dąbrowski, Laus-Mączyńska 1978: 62-69).

Metodologia pracy ze źródłami ze środków masowej informacji powinna się opierać na:

- 1) określeniu celu zbierania danych,
- 2) określeniu źródeł informacji,
- 3) określeniu wiarygodności informacji
- 4) zbiorze informacji bazowej według kluczy kategoryzacyjnych i map semantycznych.

Cykl analityczny zaś powinien się rozpocząć od:

- 1) sformułowania tematu badawczego,
- 2) zebrania i przesortowania odpowiedniej informacji,
- 3) przetworzenia informacji według kluczy kategoryzacyjnych,
- 4) syntezy informacji kluczowych,
- 5) analizy informacji kluczowych,
- 6) zredagowaniu pod względem treści i kontekstu.

Jakie narzędzia w analizie informacji medialnej na potrzeby państwa wykorzystamy – czy tradycyjne „papier i nożyczki”, czy zaawansowane programy komputerowe, które przesortują materiał badany za analityka – nie ma to już tak istotnego znaczenia. Ważne, by analiza była rzeczowa i kompetentna, a analityk nie uległ informacyjnej entropii i faktów medialnych nie uznawał w całości za fakty rzeczywiste.

## Bibliografia

- Caching K. Ch, Yahiko K. (2002), *A Semantic Model for Hypertext Data Caching*, [w:] *Lecture Notes In Computer Science*; vol. 2503, Proceedings of the 21st International Conference on Conceptual Modeling, London, s. 276-290.
- Cook M. & Cook C. (2000), *Competitive Intelligence*, London.
- Cooper M. D. (1983), *The structure and future of the information economy*, *Information Processing & Management*, vol. 19, s. 9-26.
- Cronin B., Overfelt K., Fouchereaux K., Manzvanzvikie T., Cha M., & Sona E. (1994), *The Internet and competitive intelligence: a survey of current practice*, „International Journal of Information Management”, vol. 14 (3), s. 204-222.
- Dąbrowski M., Laus-Mączyńska K. (1978), *Metody wyszukiwania i klasyfikacji informacji*, Warszawa.
- Доронин А. (2001), *Аналитическая обработка материалов открытого доступа*, Стратегический менеджмент, [www.smanagement.ru](http://www.smanagement.ru).
- Доронин А. (2003), *Бизнес – разведка*, 2 издание, „Ось-89”, Москва.
- Gatnar E. (1998), *Symboliczne metody klasyfikacji danych*, Warszawa.

- Gerbner G., Gross L., Melody W. (1973), *Communication Technology and Social Policy*, New York, s. 553-573.
- Goban-Klas T. (1997), *Public relations czyli promocja reputacji. Praktyka działania*, Warszawa.
- Goban-Klas T. (1999), *Media i komunikowanie masowe, teorie i analizy prasy, radia, telewizji i Internetu*, Warszawa – Kraków.
- Guisnel J. (2004), *Espionage on the Internet*, „Primedius Views”, [www.prime-dius.com](http://www.prime-dius.com).
- Hasanali F., Leavitt P. (2004), *Competitive Intelligence: A Guide for Your Journey to Best-practice Processes*, APQC, Houston.
- Hulnick A. S. (2004) *Espionage: Does It Have a Future in the 21st Century?*, „Brown Journal of World Affairs”, Winter/Spring, vol. XI, issue 1, s. 165-173.
- Kahaner L. (1997), *Competitive Intelligence*, New York.
- Kraaij, Wessel (2004), *Variations on Language Modeling for Information Retrieval*, Taaluitgeverij Neslia Paniculata/CTIT Ph.D. thesis series No. 04-62, Rotterdam.
- Kwieciński M. (1999), *Wywiad gospodarczy w zarządzaniu przedsiębiorstwem*, Warszawa-Kraków.
- Lafouge T. (1993), *Bibliométrie – Esquisse d'un panorama*, „Science dla la société”, № 29, mai.
- Martinet B., Marti Y. M. (1999), *Wywiad gospodarczy. Pozyskiwanie i ochrona informacji*, Warszawa (*L'Intelligence économique, Les yeux et les oreilles de l'entreprise*, Paris 1995).
- Nowakowski J., Sobczak W. (1971), *Teoria informacji*, Warszawa.
- Ocena systemu politycznego w Polsce*, [www.ipsos.pl](http://www.ipsos.pl).
- Outils d'analyse textuelle – Outils "scientométriques, infométriques, bibliométriques"* *Searching, Mapping, Veille Stratégique – "Competitive Intelligence"*, <http://www.uhb.fr/>.
- Pamula S. (1996), *Metoda analizy zawartości prasy*, Częstochowa.
- Pisarek W. (1983), *Analiza zawartości prasy*, Kraków.
- Schiwely W. P. (1997), *The craft of political research*, wyd. polskie 2001: *Sztuka prowadzenia badań politycznych*, Poznań.
- Schwab D. P. (2005), *Research Methods for organizational Studies*, Mahwah, New Jersey, London.
- Sobczak W., Malina W. (1978), *Metody selekcji informacji*, Warszawa.
- TAIGA, Traitement Automatique de l'Information Géopolitique d'Actualité*, [www.teaser.fr](http://www.teaser.fr).
- Tan P. & Kumar V. (2002), *Discovery of web robot sessions based on their navigational patterns*, “Data Mining and Knowledge Discovery”, vol. 6, s. 9-35.
- Tropes, Analyse de textes & Sémantique*, <http://www.acetic.fr>.
- Vriens D. (2005), *The Role of Information and Communication Technology in Competitive Intelligence*, [w:] Vriens, Dirk *Information and Communication Technology for Competitive Intelligence*, Amsterdam-Nijmegen, s. 1-32.

- Wersig G. (2002), *Problemstellungen für die Organisation des Information-  
saustausches*, [w:] *Spannungsfeld: Internationaler Informationstransfer – Chancen  
und Risiken globaler Vernetzung*, Berlin, [www2.kommwiss.fu-berlin.de](http://www2.kommwiss.fu-berlin.de).
- Wersig G. (2004), *Überleben in der Informationsgesellschaft*, Vortrag auf der 21. Mes-  
se-Fachtagung „Ohne Fleiß kein Preis“, 16.-17. Juni 2004, Wiesbaden.
- Witkowski P. (2005), *Media cytowne w marcu*, [www.medialink.pl](http://www.medialink.pl).