

**Justyna Adamus**

### **Narzędzia klasyfikacji dokumentów elektronicznych: zarys problematyki**

Wraz z powstaniem i rozwojem nowych technologii informacyjnych i komunikacyjnych gospodarka zaczyna funkcjonować w tzw. rzeczywistości wirtualnej. Rzeczywistości, która jest każdemu z nas coraz bliższa, staje się codziennością, przez co warunkuje sposoby komunikacji publicznej i komfort życia obywateli. Administracja publiczna ulega tym tendencjom i zaczyna wdrażać nowe metody pracy i komunikacji z użytkownikiem.

Zachodzi konieczność archiwizacji oraz organizacji i przechowywania dokumentów elektronicznych, albowiem już dziś wiele istotnych informacji tworzonych w postaci elektronicznej ginie bezpowrotnie mimo, iż stanowią one ważny czynnik rozwoju życia społecznego, kulturalnego i gospodarczego narodu, stanowiąc przy tym tzw. dziedzictwo kulturalne. Ponadto przez wprowadzenia do szerokiego stosowania podpisu elektronicznego, co zgodnie z ustawą ma nastąpić w sierpniu 2006 roku, jak też rozwiązanie problemu wnoszenia opłat skarbowych przez Internet (tzw. cyfrowe znaczki skarbowe) możliwa będzie realizacja pełnych usług elektronicznych urzędów przez całą dobę, siedem dni w tygodniu, przez cały rok, zgodnie z formułą: „24/7/365”. Możliwość posługiwania się dokumentami w formie elektronicznej, sprawi, iż będzie potrzebne m.in. przyjęcie nowego systemu klasyfikowania e-dokumentów, który to system będzie miał na celu zapewnić odpowiednią identyfikację, porządkowanie i wyszukiwanie dokumentów oraz wsparcie dla obiegu dokumentów. System taki dodatkowo mógłby mieć zastosowanie w systemie informacji publicznych, gdyż w myśl ustawy z dnia 6 września 2001 r. o dostępie do informacji publicznej każdemu przysługuje prawo uzyskania informacji publicznej, wgląd do dokumentów urzędowych itd. W zakresie porządkowania i archiwizacji dokumentów wymagane jest zatem, przyjęcie nowej metodologii uwzględniającej potrzeby dokumentacji elektronicznej, nieco odmiennej od dokumentów papierowych oraz uwzględnienie wymagań w zakresie dostępu do informacji publicznych.

Postępujący proces informatyzacji sektora publicznego wymusza zmiany dotychczasowego modelu instrukcji kancelaryjnej oraz rzeczowych wykazów akt w celu tworzenia systemu archiwizacji e-dokumentów zapewniającego pełną informatyzację obiegu dokumentów w elektronicznej administracji w Polsce oraz systemu w pełni uwzględniającego potrzeby społeczeństwa informacyjnego. „Strategia Informatyzacji Rzeczypospolitej Polskiej – ePolska na lata 2004-2006”<sup>1</sup> przewiduje utworzenie zintegrowanej platformy usług administracji publicznej dla społeczeństwa informacyjnego pt. „Wrota Polski”. W jednym z działań (pkt 1.7.) planowane jest utworzenie rządowego systemu obiegu dokumentów dla wybranych procesów (II połowa 2005 r.), co ma się ściśle wiązać z modyfikacją instrukcji kancelaryjnych. W punkcie 3.6. zakłada się natomiast przygotowanie państwowego archiwum dokumentów elektronicznych (II połowa 2006 r.). Zapisy te wyznaczają nowe kierunki oraz zadania dla archiwistyki. Archiwiści powinni zatem, podjąć kroki w celu tworzenia zaleceń i wymagań dla zarządzania aktami, już od chwili ich wytworzenia lub otrzymania przez

---

<sup>1</sup> [www.informatyzacja.gov.pl](http://www.informatyzacja.gov.pl)

instytucję, aż do momentu przekazania do archiwum państwowego, gdyż dzięki temu unikną wielu prac przy opracowania materiałów archiwalnych po przekazaniu ich do właściwego „archiwum historycznego”. Przy czym archiwiści powinni pozostać jedynie w roli konsultantów wobec zarządzania aktami tak, aby działania powodujące ujawnienie archiwalnych dokumentów były zakończone, zanim dokumenty te podlegać będą interpretacji<sup>2</sup>. Warunek ten wynika z konieczności zachowania pełnej neutralności archiwisty wobec zawartości intelektualnej przekazu (treści dokumentu). Z drugiej strony podkreśla się potrzebę rozwagi w wytwarzaniu, utrzymywaniu i zarządzaniu cyfrowymi zasobami wiedzy<sup>3</sup>. Niemniej jednak powinno się dążyć do współpracy archiwistów z twórcami dokumentów w celu wpłynięcia na stosowane przez tych ostatnich standardy i praktyki oraz uświadomienia im potrzeby zachowania elektronicznych materiałów archiwalnych stanowiących dziedzictwo kulturowe, jako ważną inwestycję w przyszłość.

Opracowanie przepisów kancelaryjno-archiwalnych będzie również wymagać uwzględnienia międzynarodowych norm i standardów w zakresie zarządzania dokumentami i aktami elektronicznymi dla zapewnienia jednolitości w organizacji czynności kancelaryjnych w skali światowej.

Wydarzenia polityczne, jakie miały miejsce w związku z wykryciem przez media usunięcia słów „lub czasopisma” z art., 36 ust. 3 pkt. 1 rządowego projektu ustawy o radiofonii i telewizji (tzw. „Afera Rywina”) pokazały jak ważne jest ustanowienie standardów techniczno-prawnych dla zarządzania e-dokumentami w trakcie ich tworzenia i bieżącego stosowania tak, aby nie było możliwe wprowadzanie zmian do dokumentów, których pochodzenie jest trudne lub niemożliwe do ustalenia. Unia Europejska wypracowała w tym zakresie model wymagań dla zarządzania elektronicznymi dokumentami<sup>4</sup>, natomiast Międzynarodowa Organizacja Normalizacji wydała normę ISO 15489. Powstanie podobnych unormowań w warunkach polskich jest niezbędne dla zagwarantowania odpowiedniej przejrzystości procesów w instytucjach rządowych i zaufania publicznego oraz zapewnienia autentyczności i odpowiedzialności ze strony twórców akt elektronicznych. Zasady zarządzania aktami nie tylko zagwarantują ich niezmiennosc, ale zapewnią także pełną kontrolę nad cyklem ich życia, dzięki odpowiednim elementom metadanych przypisanych do dokumentów, które to metadane będą również stanowić ważny potencjał informacyjny w procesie ich archiwizacji<sup>5</sup>. Sama filozofia zarządzania e-aktami i e-dokumentami, jakkolwiek bardzo ważna i interesująca, wykracza jednak poza ramy niniejszego artykułu. W dalszej części skupiono się natomiast na jednym z ważnych elementów metadanych jakim jest klasyfikacja akt w celu ich identyfikacji i porządkowania.

---

<sup>2</sup> Menne-Haritz Angelika: Dostęp do archiwów, czyli przeformułowanie archiwalnego paradygmatu // Archeion.- 2002, T. 104, s. 72

<sup>3</sup> Ochrona dziedzictwa cyfrowego: zalecenia / Oprac. National Library of Australia.- Warszawa, 2003, s. 24

<sup>4</sup> [www.cornwell.co.uk/moreq.html](http://www.cornwell.co.uk/moreq.html) - Model Requirements for the Management of Electronic Records (MoReq)

Przypomnijmy, iż organizację i porządkowanie dokumentów tradycyjnych (papierowych) w celu ich archiwizowania, zgodnie z Rozporządzeniem Ministra Kultury z dnia 16 września 2002 r. (Dz. U. nr 167, poz. 1375), określa instrukcja kancelaryjna ustalona i stosowana w instytucjach i urzędach, państwowych i niepaństwowych jednostkach organizacyjnych oraz podmiotach realizujących zadania publiczne i gospodarcze. Podstawą porządkowania akt są jednolite rzeczowe wykazy akt.

Dotychczasowe sposoby organizacji akt oraz jednolite rzeczowe wykazy akt w obecnym kształcie nie będą w stanie skutecznie sprostać wymogom zarządzania aktami elektronicznymi. Grupowanie elektronicznych dokumentów w obrębie akt spraw, a następnie teczek odpowiadających podziałowi rzeczowemu, powinno odbywać się automatycznie, co będzie można osiągnąć dzięki zastosowaniu odpowiednich narzędzi klasyfikowania elektronicznych dokumentów w procesie aktotwórczym. W tym celu możliwe jest przyjęcie nowych ujednoczonych metod i narzędzi klasyfikacji w obrębie całego systemu gospodarczego w państwie.

Klasyfikacja polega na systematycznej identyfikacji i porządkowaniu aktywności gospodarczych i/lub akt wewnątrz kategorii zgodnych z przyjętą strukturą, metodami oraz zasadami proceduralnymi reprezentowanymi w schemacie klasyfikacji<sup>6</sup>. Klasyfikacja dokumentów elektronicznych jest o tyle istotna, iż umożliwi ona grupowanie dokumentów w akta spraw a następnie formowanie elektronicznych teczek aktowych, ułatwiających zarządzane aktami, podobnie jak ma to miejsce w przypadku porządkowania dokumentów tradycyjnych (papierowych), albowiem istotą archiwizacji jest taka organizacja dokumentów, która umożliwia szybkie odnajdywanie konkretnych spraw. Tworzenie elektronicznych teczek będzie opierać się na numerze akt lub zakresie numerów odpowiadających symbolom klasyfikacyjnym i/lub przy uwzględnieniu kryteriów czasowych. Przy czym oznaczanie dokumentów kodem zgodnym z systemem klasyfikacji następować będzie w procesie „przechwytywania” dokumentów przez system zarządzania dokumentami elektronicznymi, czyli w sposób automatyczny. W środowisku elektronicznym fizyczne miejsce pojedynczych dokumentów w takim wymiarze jak w świecie realnym nie istnieje (przechowywanie akt w odpowiedniej tezcze). Elektroniczne sprawy i elektroniczne tecki są wirtualne, w tym sensie, że realnie niczego nie zawierają, składają się natomiast z atrybutów (wartości) metadanych przypisanych do dokumentów<sup>7</sup>, dlatego tak ważne jest przygotowanie odpowiednich narzędzi klasyfikacji e-dokumentów. Dzięki temu, oprogramowanie odpowiadające za zarządzanie elektronicznymi aktami pozwoli użytkownikom przeglądać tecki i akta oraz sprawnie nimi zarządzać.

### **Klasyfikacja dokumentów w systemie informacji gospodarczej w Australii**

W Australii zaprojektowano kompleksowy system zarządzania informacją gospodarczą (Managing Business Information). Jednym z narzędzi tworzenia takiego systemu jest **schemat**

---

<sup>5</sup> Prace przygotowawcze do archiwizacji dokumentów elektronicznych w oparciu o doświadczenia australijskie i brytyjskie / Kazimierz Schmidt // Archeion.- T. 106; Metadane: sposób na uporządkowanie Internetu / Marek Nahotko.- Kraków, 2004, s. 50-55

<sup>6</sup> [www.cornwell.co.uk/moreq.html](http://www.cornwell.co.uk/moreq.html) - Model Requirements for the Management of Electronic Records (MoReq)

<sup>7</sup> [www.cornwell.co.uk/moreq.html](http://www.cornwell.co.uk/moreq.html) - Model Requirements for the Management of Electronic Records (MoReq)

**klasyfikacji gospodarki** (ang. business classification scheme). Schemat ten jako element systemu **informacji gospodarczej** ma być narzędziem służącym do opisu dokumentów dla agencji rządowych. Tworzony jest na podstawie analizy procesów, działań i funkcji gospodarczych. Klasyfikacja ta ma służyć organizacji materiałów ze względu na funkcje organizacji, gdyż **układ strukturalny, uznano za niefunkcjonalny**. W polskiej praktyce natomiast najczęściej spotykanym kryterium układu akt jest układ kancelaryjny zgodny ze strukturą organizacyjną instytucji.

Schemat klasyfikacji gospodarki porządkuje elektroniczne akta oraz zapewnia organizację informacji dla nowego systemu przechowywania dokumentów, ale odpowiada też szerokim zadaniom, jakie stawia się przed systemem informacji gospodarczej. Metodologia klasyfikacji jest zgodna z australijskim standardem AS ISO 15489 i ma sprzyjać podnoszeniu poziomu praktycznych narzędzi zapewniających odpowiednią organizację akt<sup>8</sup>.

System klasyfikacji gospodarki stosowany w Australii pozwala na tytułowanie, klasyfikację oraz zachowanie akt i innych informacji gospodarczych, a także zapewnia kontrolę używanego słownictwa i konsekwentne generowanie opisu dokumentów produkowanych podczas działalności gospodarczej, co zapewni także właściwe przechowywanie, porządkowanie i zarządzanie tą informacją w przyszłości (np., gdy trafią do archiwów).

Przyjęto, iż organizacja akt poprzez funkcje i zadania jest bardziej stabilna dla klasyfikacji niż struktura organizacyjna instytucji, która często ulega zmianie, w przeciwieństwie do funkcji. W takim przypadku akta będą klasyfikowane ze względu na funkcje, dla jakich powstały (kontekst powstania akt), co powinno być także ważniejsze niż przedmiot zawarty w dokumencie (to, o czym są akta). W klasyfikacji niezbędna jest analiza działań i procesów gospodarczych, której rezultatem jest ujawnienie zależności pomiędzy różnymi sferami gospodarki, a także wytwarzaną przez nie dokumentacją.

W schemacie klasyfikacji hasła klasyfikacyjne są ze sobą powiązane relacjami odzwierciedlającymi struktury hierarchiczne. Schemat klasyfikacji zawiera trzy składniki, czy też poziomy:

- poziom nadrzędny, obejmujący kategorie, zwany funkcją,
- poziom podrzędny w stosunku do funkcji, zwany działaniem,
- poziom podrzędny i bardziej specyficzny w stosunku do działania, zwany transakcją lub poziom deskryptora przedmiotowego, czy też poziom tematu.

Klasyfikacja akt przebiega czteropoziomowo (klasyfikacja od ogółu do szczegółu) i przedstawia się następująco:

#### **FUNKCJA – DZIAŁANIE – TEMAT – PODTEMAT**

W oryginale: „function – activity – topic – subtopic”. Klasyfikacja akt na wszystkich poziomach pozwala na zorganizowanie akt w sposób hierarchiczny i powiązanie dokumentów tak, aby umożliwić wyszukiwanie przez różne poziomy klasyfikacji.

---

<sup>8</sup> [www.naa.gov.au/recordkeeping/control/tools.pdf](http://www.naa.gov.au/recordkeeping/control/tools.pdf) - Overview of Classification Tools – July 2003 / National Archives of Australia

## **Opracowanie (klasyfikowanie zwane tu także tytułowaniem) akt na kolejnych poziomach.**

### **Poziom I – FUNKCJA**

Pierwszy poziom klasyfikacji stanowi **funkcja komórki**, czy agendy w organizacji. Klasyfikowanie na tym poziomie polega na tym, iż twórca dokumentu (autor) lub system komputerowy będzie pobierać hasło typu „funkcja” z dostępnych dla danej organizacji haseł wzorcowych, czy autorytatywnych i przypisywać je do akt.

### **Poziom II – DZIAŁANIE**

Osoba indeksująca lub system wybiera na tym poziomie hasło odzwierciedlające czynność, czyli działanie, któremu towarzyszy dokument (nazwa czynności lub działalności). Poziom ten w języku klasyfikacyjnym obejmuje symbole (słownictwo), które charakteryzują partykularne działania instytucji.

### **Poziom III – TEMAT**

Na trzecim poziomie opisu wybierane będzie właściwe hasło dla dokumentu, które opisz temat (przedmiot), którego dotyczy dokument. Polega to na opisanu **transakcji**, jaka jest zawarta w dokumencie. Hasło na tym poziomie opisuje to, co niezbędne dla wyszukiwania dokumentów przez użytkowników (wchodzimy w szczegóły treści dokumentu).

### **Poziom IV – PODTEMAT**

Poziom ten również odpowiada za charakterystykę **treści pisma**, jest jednak opisem bardziej szczegółowym w stosunku do poziomu trzeciego (często na poziomie tym stosuje się opis poprzez swobodne słowa lub frazy w języku naturalnym) lub tezaurs przedmiotowy.

Wszystkie poziomy muszą być ze sobą spójne, a zastosowanie właściwych haseł na wszystkich poziomach powinna zapewniać odpowiednia **gramatyka języka** klasyfikacji.

Projektowanie zasobu leksykalnego (terminów klasyfikacyjnych) dla pierwszego, drugiego i trzeciego poziomu klasyfikacji musi poprzedzać analiza funkcji i działalności organizacji, czyli analiza tego, za co dana jednostka odpowiada w systemie gospodarczym. Natomiast dla czwartego potrzebny jest zasób słownictwa obejmującego najróżniejsze tematy, dlatego właściwe jest tutaj zaadaptowania jakiegoś istniejącego języka informacyjno-wyszukiwawczego, np. **tezaurusu przedmiotowego**<sup>99</sup>.

Do terminów w schemacie klasyfikacji gospodarki może być też załączony opis, który wyjaśnia użycie terminu i ma wspomagać użytkowników we właściwym klasyfikowaniu akt lub podczas wyszukiwania.

---

<sup>99</sup> Tezaurs jest słownikiem kontrolowanym języka deskryptorowego. Tezaursy od wielu lat stosowane są w bibliotekarstwie do opisu rzeczowego dokumentów. Poleca się także zastosowanie tezaursów w sieci WWW dla poprawienia nawigacji i wyszukiwania [zob. Architektura informacji w serwisach internetowych / Louis Rosenfeld, Peter Morville.- Beijing [i in.], 2003, s. 217-233

Użytkownicy mogą przeszukiwać system informacji gospodarczej przechodząc od terminów poziomu pierwszego (funkcja) do bardziej szczegółowych (działalność). Wyszukiwanie odbywa się w języku naturalnym. Wiele schematów klasyfikacji przewiduje na trzecim i czwartym poziomie opis poprzez tekst w języku naturalnym – swobodne słowa lub frazy (język swobodnych słów kluczowych). W takim przypadku hasła na tym poziomie nie są prezentowane explicite w klasyfikacji i tym samym nie ma kontroli nad redagowaniem tych haseł (brak słownika, czy kartoteki wzorcowej). Dzięki budowie hierarchicznej klasyfikacji wyszukiwania akt można prowadzić poprzez kilka dróg wyszukiwawczych.

### **Tezaurus funkcji.**

Dla usprawnienia narzędzi klasyfikacji proponuje się organizację zasobu leksykalnego w postaci **tezaurusa**. Australijski tezaurus funkcji stworzono do prezentacji terminów klasyfikacyjnych tak, aby ukazać różnego rodzaju relacje pomiędzy terminami klasyfikacyjnymi. Terminy te są powiązane ze względu na zależności semantyczne, hierarchiczne i skojarzeniowe. System relacji jest narzędziem umożliwiającym zlokalizowanie haseł klasyfikacyjnych dla różnych akt. W teaurusie mogą występować wszystkie rodzaje haseł, aż do czwartego poziomu szczegółowości, co pozwala na wszechstronność klasyfikowania i wyszukiwania. Ważną cechą tezaurusa jest to, iż występują w nim także terminy, których użycie nie jest zalecane (nie mogą służyć do klasyfikowania) tzw. askryptory. Od takich haseł, uznanych za nie preferowane prowadzi się automatyczne odnośniki do haseł właściwych (zalecanych do stosowania). Tworzy się więc, system odsyłaczy od terminów nie przyjętych do stosowania do terminów preferowanych (zob., zob. też)<sup>10</sup>.

Taka podwójna organizacja terminów klasyfikacyjnych, właściwa językom deskryptorowym, tj. schemat klasyfikacji oraz tezaurus, zapewnia dobrą organizację leksyki języka i dzięki temu usprawnia proces opracowywania oraz wyszukiwania. Schemat klasyfikacji jest reprezentacją w układzie hierarchicznym natomiast funkcją tezaurusa jest prezentacja w układzie alfabetycznym aktywnych terminów powiązanych hierarchicznie. Każda organizacja, czy instytucja powinna otrzymać narzędzia klasyfikacyjne dostosowane do jej działalności gospodarczej, czyli stosownie do dokumentów będących rezultatem jej specyfiki działalności. Organizacje czy instytucje, dla których przeznaczona jest klasyfikacja, powinny mieć możliwość zaadaptowania różnych narzędzi klasyfikacyjnych w zależności od sfer działalności gospodarczej, w jakich funkcjonują. I tak w Australii przygotowana jest oficjalna wersja tezaurusa terminów ogólnych zawarta w **Keyword AAA Thesaurus of General Terms**<sup>11</sup>. Schemat klasyfikacji obejmuje pewien hierarchiczny szkielet klasyfikacji dla organizacji i może być dalej rozbudowywany o bardziej szczegółowe zagadnienia. Stosowanie takiego narzędzia umożliwia ustalanie licencji zawieranej z Archiwum Narodowym Australii.

---

<sup>10</sup> [www.naa.gov.au/recordkeeping/control/functions\\_thesaur/intro.html](http://www.naa.gov.au/recordkeeping/control/functions_thesaur/intro.html) - Developing a Functions Thesaurus: Guidelines for Commonwealth Agencies

<sup>11</sup> [www.naa.gov.au/recordkeeping/control/KeyAAA/summary.html](http://www.naa.gov.au/recordkeeping/control/KeyAAA/summary.html) - Keyword AAA: A Thesaurus of General Terms

Standard metadanych dla Agencji Rządowych w Australii<sup>12</sup> stworzony w celu ujednoczenia elementów metadanych wymaganych dla zarządzania aktami, wśród wielu elementów, przewiduje również element dla klasyfikacji i opisu rzeczowego akt. Narzędzia klasyfikacji opisane powyżej mają być wsparciem dla procesu zarządzania aktami poprzez metadane. Przy czym Keyword AAA w schemacie metadanych ma być stosowany w połączeniu z tezaurem terminów nawiązujących do specyficznych lub rdzennych funkcji gospodarczych organizacji, czyli specjalistycznego tezaurusa danej agencji lub komórki i występują w elemencie odpowiadającym za klasyfikację akt. W schemacie metadanych występuje też element przedmiot, dla którego zalecane jest zastosowanie słownika haseł przedmiotowych Biblioteki Kongresu tj. **Library of Congress Subject Headings (LCSH)**<sup>13</sup>.

---

<sup>12</sup> [www.archives.sa.gov.au/files/management\\_standard\\_metadata.pdf](http://www.archives.sa.gov.au/files/management_standard_metadata.pdf) - SARKMS Metadata Scheme Final AMW 290503.doc

<sup>13</sup> [www.carl.org/tlc/crs/shed0014.htm](http://www.carl.org/tlc/crs/shed0014.htm) - Library of Congress Subject Headings