

Marek Kuczyński

**Dokumentacja elektroniczna
i problemy z nią związane**



Dokumentacja elektroniczna to całkowicie nowy typ dokumentu. Cechy różniące ten rodzaj dokumentacji od dokumentacji tradycyjnej (papierowej) to m.in. oddzielenie informacji od nośnika (dokumenty wirtualne), nieograniczona możliwość reprodukcji, a co za tym idzie zanikanie pojęcia takiego jak „oryginał” czy „kopia” oraz wielowarstwowość informacji (pliki multimedialne). Czym jest więc dokument elektroniczny? Oto jedna z jego definicji stworzona przez największą rzeszę odbiorców dokumentacji elektronicznej – „internautów”:

Dokument elektroniczny (inaczej **dokument cyfrowy**, **dokument binarny**) - *dokument w postaci pliku tekstowego, graficznego, muzycznego, filmowego lub mieszanego będącego wynikiem pracy z danym programem komputerowym, dający się zapisać, a następnie odczytać. Posiada cechy dokumentu, czyli potwierdza prawdziwość zaistnienia danego wydarzenia, okoliczności, zjawiska oraz posiada cechy pliku.*¹

Aby uzmysłwić problemy związane z dokumentacją elektroniczną pozwolę sobie na przytoczenie jeszcze jednej definicji.

Dokumentacja aktowa patrz **Akta** – dokumentacja twórcy zespołu powstała w wyniku jego działalności, utrwalona za pomocą pisma niezależnie od techniki wykonania i formy zewnętrznej.²

Teraz należy odpowiedzieć sobie na pytanie: czy dokument elektroniczny wchodzi w zakres definicji Akta?

Moim zdaniem tak. Różni się on od dokumentacji papierowej tylko techniką wykonania czyli komputer a nie maszyna do pisania (czy długopis, pióro) oraz formą zewnętrzną, czyli nie papier (poszyt, księga itp.), a HDD (czy CD-ROM lub DVD).

Oprócz tego dokumentacja filmowa oraz nagrania (obecnie tworzone wyłącznie jako dokumentacja cyfrowa) automatycznie wchodzi w zakres dokumentacji elektronicznej ze względu na zdobycze technologiczne umożliwiające w sposób bezpośredni przenoszenie zapisu cyfrowego np. z taśm DV do zapisu „w pełni” cyfrowego (komputerowego) na płyty CD-ROM czy DVD.

W związku z powyższymi różnicami można wymienić trzy najważniejsze problemy związane z tą dokumentacją:

¹ http://pl.wikipedia.org/wiki/Dokument_elektroniczny

² Polski Słownik Archiwalny, Naczelna Dyrekcja Archiwów Państwowych, Warszawa 1974

Problem 1: wiarygodność (autentyczności) informacji

Problem ten należy rozbić na dwie części.

1. Zabezpieczenie autentyczności informacji podczas przekazywania dokumentacji elektronicznej pomiędzy jednostką organizacyjną a archiwum państwowym.

Rozwiązanie najprostsze - to odpowiednio opracowane standardy metadanych np. *Dublin Core Metadata Element Set (DCMES)* wraz z odpowiednimi procedurami przekazywania „e – dokumentów”, w skład których powinien wejść podpis elektroniczny. Dlaczego tak istotną sprawą są standardy?

- po pierwsze (i najważniejsze dla „e - urzędów”) – dla ujednoczenia wymiany informacji;
- po drugie (i najważniejsze dla „e - archiwisty”) – dla ewidencji dokumentacji elektronicznej;
- oraz po trzecie (i najważniejsze dla „e - użytkownika”) – dla sprawnego i szybkiego wyszukiwania informacji.

Samo przekazywanie „fizyczne” natomiast nie powinno być inne niż np. dokumentacji aktowej czyli spisy zdawczo-odbiorcze itp. (w końcu to podobna dokumentacja różniącą się jedynie techniką wykonania oraz formą zewnętrzną).

Podpis elektroniczny - co to jest i jak to działa?

Polska norma PN-I-02000:2002 definiuje podpis cyfrowy następująco: *Przekształcenie kryptograficzne danych umożliwiające odbiorcy danych sprawdzenie autentyczności i integralności danych oraz zapewniające nadawcy ochronę przed sfalszowaniem danych przez odbiorcę.*³ Podpis elektroniczny, choć różni się od tradycyjnego musi jednak spełniać te same funkcje, co podpis odręczny.

Na podstawie art. 3 ust. 2 ustawy o podpisie elektronicznym z 18 września 2001 r. (Dz. U. nr 130, poz.1450) bezpieczny podpis to taki który:

- a) jest przyporządkowany wyłącznie do osoby składającej ten podpis,
- b) jest sporządzany za pomocą podlegających wyłącznej kontroli osoby składającej podpis elektroniczny bezpiecznych urządzeń służących do składania podpisu elektronicznego i danych służących do składania podpisu elektronicznego,

³ PN-I-02000:2002

c) jest powiązany z danymi, do których został dołączony, w taki sposób, że jakakolwiek późniejsza zmiana tych danych jest rozpoznawalna,⁴

Podpis elektroniczny opiera się na tzw. funkcji skrótu i szyfrze asymetrycznym. Tworzy ona unikalną dla danej wiadomości sumę kontrolną o długości niezależnej od wielkości wiadomości. Najważniejszą cechą tej funkcji jest to, że każda zmiana w wiadomości (nawet jednego bitu) spowoduje nieprzewidywalną zmianę skrótu, a co za tym idzie można łatwo zweryfikować czy w dokumencie elektronicznym nastąpiły jakiegokolwiek zmiany.

Należy również pamiętać, że „e-dokument” bez ważnej sygnatury elektronicznej (podpisu elektronicznego) jest w Polsce nieważny, nie może więc być środkiem dowodowym w postępowaniu sądowym.

Na podstawie tego ostatniego stwierdzenia archiwa nie powinny przejmować dokumentacji elektronicznej (a przynajmniej pewnej części takiej dokumentacji) bez podpisu elektronicznego ze względu na charakter i zakres działalności, ponieważ (w świetle prawa) nie mogły by wydawać uwierzytelnionych „kopii” dokumentów.

2. Zabezpieczenie autentyczności informacji będącej już w posiadaniu archiwum czyli ni mniej, ni więcej zabezpieczenie danych przed próbą ich zmiany przez ... pracowników archiwum. Dlaczego tylko przed pracownikami? Ponieważ możliwości technologiczne umożliwiają takie zabezpieczenie danych, że „e- użytkownik” nie będzie miał fizycznego kontaktu z „e – dokumentem”.

Tutaj należy sobie zadać pytanie: Czy możemy w tak zaistniałej sytuacji (dokumentacja elektroniczna jest już w archiwum) obyć się bez podpisu elektronicznego? Odpowiedź brzmi TAK. Należy wówczas stworzyć odpowiednie procedury wypożyczania takiej dokumentacji. Procedury te powinny być bardzo rygorystycznie przestrzegane z powodu jednej z cech dokumentacji elektronicznej, a mianowicie łatwości jej kopiowania oraz modyfikowania. Oto jedna z możliwości zabezpieczenia „e – dokumentów”:

- wydzielony magazyn z ograniczonym dostępem;
- wszystkie nośniki dokumentacji elektronicznej (np. CD-ROM, DVD) umieścić w odpowiednich ochronnych pudełkach wraz z paskiem magnetycznym oraz bramką elektromagnetyczną w drzwiach takiego magazynu (tak jak ma to miejsce np. w bibliotekach multimedialnych);

⁴ Dz.U. z dnia 18 września 2001r. nr 130 poz. 1450 art. 3 ust.2

- otwarcie takiego pudełka powinna odbywać się wyłącznie w jednym pomieszczeniu posiadającym komputer umożliwiający odczyt danych ale uniemożliwiający ich skopiowania (brak nagrywarek CD-ROM i DVD oraz wyłączone porty USB).
- użytkownik nie ma możliwości fizycznego kontaktu z nośnikiem (taką możliwość dają rozwiązania sieciowe).

Problem 2: przechowywania danych

Ponieważ dokumentacja elektroniczna nie jest trwale przypisana do nośnika, tak jak ma to miejsce przy dokumentacji tradycyjnej (papierowej), problem może dotyczyć wyłącznie nośnika, na którym został umieszczony „e – dokument”. Polska norma PN-ISO 11799:2006 z czerwca br. w sprawie wymagań dotyczących warunków przechowywania materiałów archiwalnych i bibliotecznych reguluje (w zakresie wilgotności oraz temperatury) przechowywanie materiałów archiwalnych w tym taśm magnetycznych, dysków magnetycznych oraz dysków optycznych.

Rozwiązanie tego problemu (stosowanym w innych krajach np. Szwecji) jest okresowe przenoszenie dokumentacji elektronicznej z nośnika na nośnik. Nie należy tutaj sugerować się badaniami nad żywotnością np. płyt CD (ok. 20 lat dla nośników renomowanych firm), ponieważ postęp technologiczny w tej dziedzinie jest tak ogromny, że 20 lat to prawie wieczność.

Oto dwa przykłady:

Przykład 1:

Pierwsze płyty CD-ROM weszły na rynek w 1982 roku, czyli ponad 20 lat temu. Czy jest możliwy odczyt tych danych? Jak najbardziej TAK oczywiście pod warunkiem odpowiedniego ich przechowywania. Czy obecne nośniki „z górnej półki” (cena jednostkowa ok. 2,5zł) wytwarzane są w lepszej technologii niż te z przed 20 lat? Oczywiście TAK.

Przykład 2:

Gdybyśmy zaczęli przejmować dokumentację elektroniczną na płytach CD-ROM (pojemność: od 650MB do 870MB) do w chwili pojawienia się (w archiwach) pierwszych komputerów z tym napędem czyli około 12 – 13 lat temu, to po upływie 8 lat całość dokumentacji zostałaby przeniesiona na płyty DVD (od 4,7GB do nawet 17GB) posiadające większą pojemność, jednak posiadające ten sam mankament co płyty CD-ROM, a mianowicie

podatność na uszkodzenia mechaniczne. To z kolei w najbliższej przyszłości cała dokumentacja przenoszona będzie najprawdopodobniej na napędy pamięci flash (pojemność od 256MB do nawet 8GB), które w porównaniu z płytami DVD mają następujące zalety: nie są podatne na bodźce mechaniczne, zajmują mniej miejsca, mają możliwość dodatkowego zabezpieczenia danych za pomocą hasła. Obecnie ich wadą jest zbyt wysoka cena.

A jakie nośniki pamięci masowej zostaną wprowadzone do użytku za 20 lat? Może będzie to „Matrix” czyli macierze dyskowe, które posiadają ogromne pojemności co jest już zauważalne obecnie. Jeszcze do niedawna przecież poczta elektroniczna miała zagwarantowaną przestrzeń dyskową około 5MB, w chwili obecnej 1GB z możliwościami wykupienia opcji bez limitu pojemności w skład której wchodzi dodatkowo miejsce na stronę www oraz pliki.

Problem 3: udostępnianie dokumentacji elektronicznej

Gdy już archiwa uporają się z dwoma pierwszymi problemami czeka je następne wyzwanie.

W jakiej formie udostępniać dokumentację elektroniczną? O ile przy dokumentacji papierowej nie ma takiego problemu, o tyle przy tego rodzaju dokumentacji problem jest dość poważny i jest związany z jej różnorodnością (tekst, dźwięk, wizja, bazy danych).

Należy tutaj określić pewne standardy przekazywania dokumentacji elektronicznej z podziałem na jej rodzaj. Standardy te powinny być jak najprostsze, aby uniezależnić zapis dokumentacji elektronicznej od platformy systemów operacyjnych (Windows, Linux, Unix) oraz programistycznej (Pascal, C++, Visual Basic, SQL itp.). Takie standardy powinien wypracować wraz z instytucjami zainteresowanymi Polski Komitet Normalizacji. Dlaczego właśnie ta jednostka organizacyjna? Pozwolę sobie na kolejną definicję a zarazem odpowiedź na pytanie. Co jest celem normy?

Celem normy jest ułatwienie realizacji zadań wynikających z potrzeb społecznych i gospodarczych poprzez tworzenie wzorcowych rozwiązań w zakresie:

- *bezpieczeństwa ludzi, środowiska i mienia,*
- *likwidowania barier w handlu,*
- *utrwalania osiągnięć techniki,*
- *upowszechniania postępu technicznego,*
- *zwiększenia efektywności w gospodarce,*
- *tworzenia podstawy do rozstrzygania sporów między dostawcą a odbiorcą,*

- *tworzenia podstawy odniesienia przy zawieraniu umów cywilno-prawnych.*⁵

Pytanie rodzi pytanie. Dlaczego tak istotną sprawą jest normalizacja przekazywanej dokumentacji elektronicznej pod względem jej rodzaju? Oto kilka przykładów:

Przykład 1:

Wojewódzki Ośrodek Dokumentacji Geodezyjno – Kartograficznej w Olsztynie posiada w swoim zasobie dokumentację kartograficzną w formie tradycyjnej oraz cyfrowej. Ogólna ilość arkuszy map to 160 000. Zdecydowana większość map istnieje w zapisie cyfrowym w programie EWMAPA.

EWMAPA jest relacyjno – obiektowo - warstwowym programem grafiki komputerowej umożliwiającą prowadzenie graficznej bazy danych, edycję map i opracowań graficznych.

Dane programu są przechowywane w pięciu podstawowych strukturach: działki, warstwy, obiekty, szrafury, rastry.

Czy archiwa państwowe mają zamiar przejmować tę dokumentację wraz z oprogramowaniem do jego tworzenia? Kto ten program kupi i będzie go aktualizował? Prostszy rozwiązaniem (przed przekazaniem map do archiwum) byłoby przecież przeniesienie ich do formatu tif lub jpg (program posiada takie możliwości). Pomijamy wówczas całą „otoczkę programistyczną”.

Przykład 2:

Regionalna Rozgłośnia w Olsztynie „Radio Olsztyn” S.A. w swoim zasobie posiada przede wszystkim zbiór dokumentacji dźwiękowej obejmujący obecnie około 3 500 jednostek archiwalnych, które zostały w większości przeniesiono z zapisu tradycyjnego na format zapisu cyfrowego. Dane dźwiękowe zapisywane są w formacie mp2 oraz wav (starsze nagrania). Obecny stan zasobu archiwum elektronicznego to 0,6 tera bajta danych. Rocznie przybywa ok. 100GB.

Do ewidencji nagrań służy program NOA umożliwiające m.in. cyfryzację i archiwizację materiałów dźwiękowych. Jest również narzędziem do obróbki dźwięku podczas procesu przetwarzania z formy analogowej do w pełni cyfrowej. Materiały pozyskane w ten sposób są opisywane i katalogowane w bazie danych wraz z odnośnikami do składnic danych gdzie się fizycznie znajdują.

Tutaj problem jest już rozwiązany. Wystarczy podpisać odpowiednią umowę na przekazanie zasobu np. na nośnikach DVD (spełniających odpowiednie kryteria jakościowe) w formacie

⁵ <http://www.pkn.pl/index.php?pid=pytania>

mp2 lub wav i archiwa mogą taką dokumentację udostępniać na dowolnym sprzęcie komputerowym posiadającym kartę dźwiękową bez specjalistycznego oprogramowania.

Przykład 3:

Telewizja Polska S.A. Oddział Terenowy w Olsztynie posiada w swoim zasobie zbiór 435 taśm zawierających 10 433 jednostek ewidencyjnych (materiałów programowych). Całość zasobu jest w postaci cyfrowej znajdujący się na taśmach cyfrowych DV. Zapis DV jest zapisem cyfrowym na taśmie magnetycznej. Maksymalny czas zapisu na taśmie wynosi 270 minut.

Czy archiwa państwowe chcą przejmować zasób w ten sposób? Przy pomocy jakich urządzeń będzie on odtwarzany? Prostszy rozwiązaniem byłoby przeniesienie nagrań telewizyjnych do formy „w pełni cyfrowej” (komputerowej) np. w formacie mpge-2 lub mocno ostatnio reklamowanym HDV (format zapisu cyfrowego wysokiej rozdzielczości). I znowu wystarczyłby dowolny komputer z kartą dźwiękową bez specjalnego dodatkowego sprzętu oraz oprogramowania.

Przykład 4:

Urzędy miast, gmin, instytucje wymiaru sprawiedliwości

Zasób (urzędów miast i gmin) to przede wszystkim bazy danych takie jak: ewidencja szkół, ewidencja ksiąg dotyczących urodzeń, małżeństw i zgonów, ewidencja stowarzyszeń, ewidencja budowli ochronnych, ewidencja gruntów i budynków.

Zasób (instytucji wymiaru sprawiedliwości) to m.in. innymi Centralny Rejestr Ksiąg Wieczystych, który na podstawie Dz. U z dnia 17 września 2001r. w sprawie prowadzenia ksiąg wieczystych i zbiorów dokumentów może być prowadzony **dodatkowo** w postaci elektronicznej. Jak również Centralny Rejestr Skazanych, Centralny Rejestr Sądowy, Centralny Rejestr Firm.

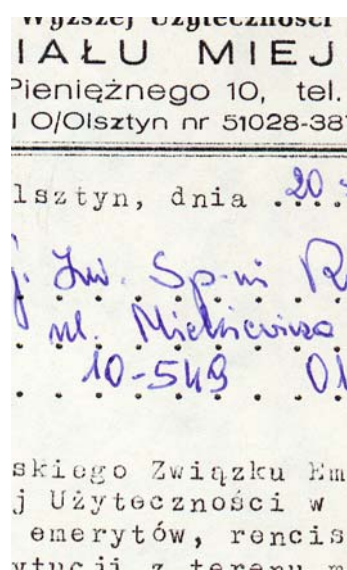
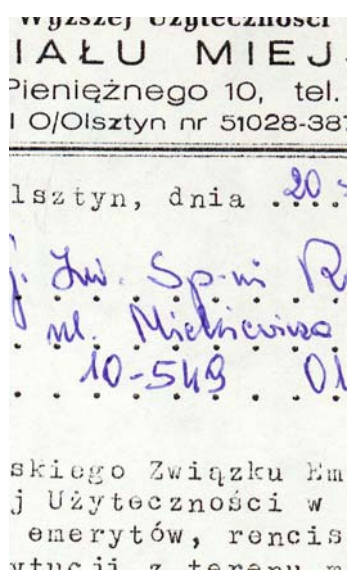
Czy archiwa państwowe chcą przejmować bazy danych wraz z aplikacjami do ich odczytywania? Kto to oprogramowanie kupi? Na terenie działania Archiwum Państwowego w Olsztynie znajduje się 112 samorządowych jednostek organizacyjnych. Każda z tych jednostek może kupować dowolne oprogramowanie dla swoich potrzeb. Średnio takich baz jest 5 (informacje tam zawarte mają kategorię archiwalną „A”). W rezultacie archiwum byłoby zmuszone do przejścia około 500 baz danych wraz z oprogramowaniem (tylko z jednostek samorządowych). Część z tych baz wykona jest w architekturze klient-serwer. Czy

archiwa państwowe stać na zakup odpowiedniego sprzętu (zdarzają się sytuacje, że na jednym serwerze nie można postawić jednocześnie kilku systemów zarządzania bazami danych) wraz z oprogramowaniem? A zatem odpowiedź brzmi – NIE. Trzeba znowu zastosować rozwiązania najprostsze. Czyli przejmowanie baz danych w formatach bez aplikacji np. dbf (taka wymiana informacji odbywa się bezproblemowo pomiędzy GUS a NDAP – baza JOPIS) lub wręcz bezpośrednia konwersja baz danych do plików tekstowych.

Przykład 5: (ostatni)

Uświadamiający jak ważny jest nie tylko format zapisu danych ale również jego parametry.

Oto dwa obrazy wykonane w formacie „jpg” oraz „tif” o tej samej rozdzielczości



Różnica między nimi jest kolosalna jeżeli chodzi o zajmowaną przestrzeń dyskową a co za tym idzie szybkości ich otwierania w aplikacjach do tego przeznaczonych. Czy jednak (dla zwykłego śmiertelnika) są widoczne różnice, jeśli chodzi o jakość obrazu? Zapytanie jest ważne i przede wszystkim dotyczy tych materiałów archiwalnych, gdzie najważniejsza jest **treść** a nie **obraz**.

Wszystkie ww. przykłady nasuwają ostateczne pytanie: Czy jest sens przejmowania dokumentacji, której w żaden sposób nie będziemy mogli udostępnić?

mgr Marek Kuczyński

Archiwum Państwowe w Olsztynie