

ZARZĄDZANIE TRWAŁOŚCIĄ MATERIAŁÓW CYFROWYCH I METADANE W ARCHIWUM CYFROWYM

[DAWID ŻĄDŁOWSKI: Workflow Trends 2019]

KILKA SŁÓW O MNIE

WYSZUKIWANIE INFORMACJI

Archiwa cyfrowe

OCR

Konserwacja i opracowanie

TIFF

DIGITALIZACJA

Cyfrowa ochrona zbiorów

Archiwa historyczne

TIFF
PDF

PDF
TIFF
Archiwizacja

TIFF

METADANE

TIFF

UDOSTĘPNIANIE

Archiwa bieżące

Metadane

OCR Szkolenia archiwalne

PDF

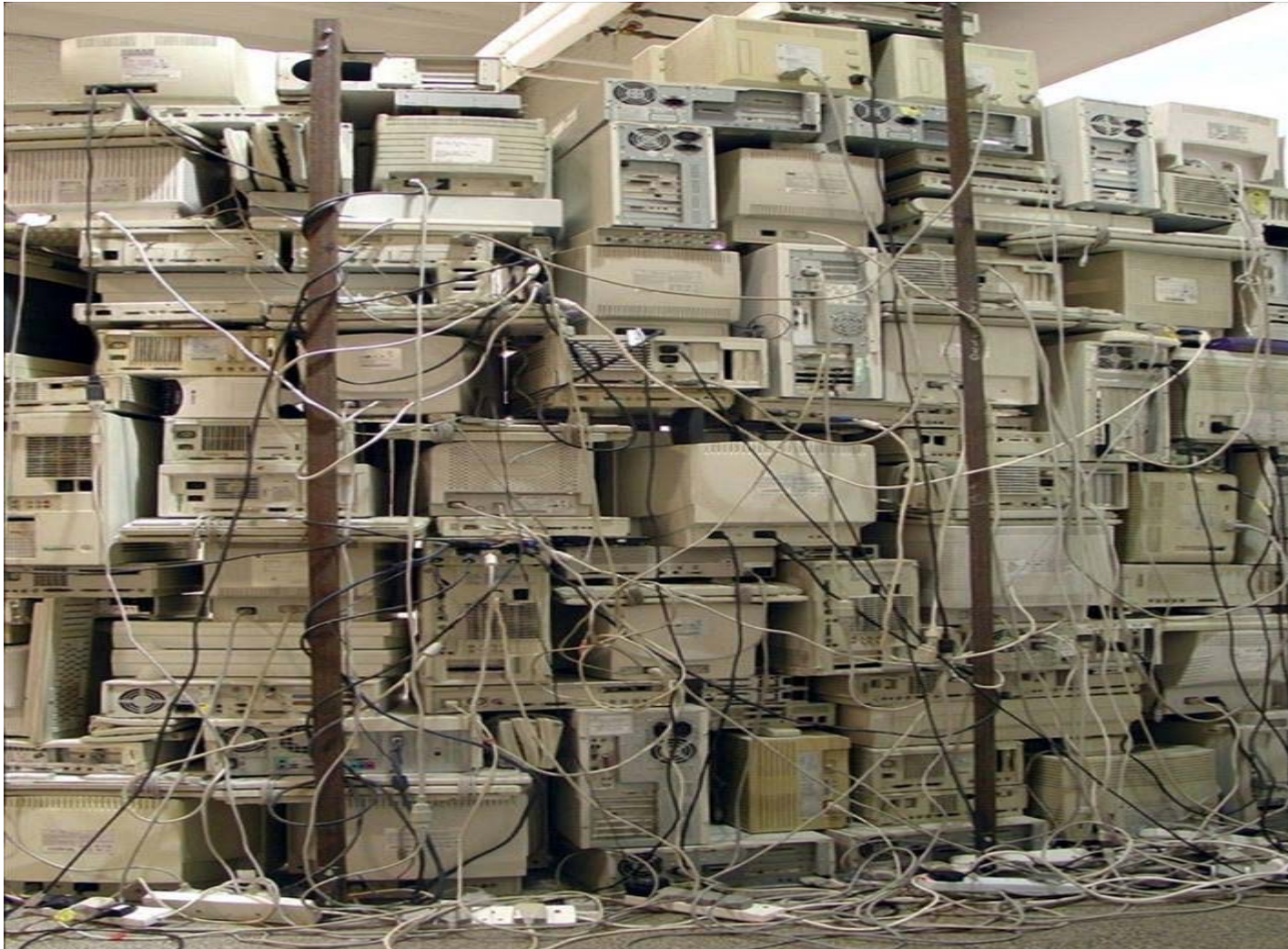
PDF
TIFF

AGENDA



- Archiwa cyfrowe
- Mechanizmy długotrwałej archiwizacji zasobów cyfrowych
- Metadane

Pierwsza wersja Archiwum cyfrowego



Archiwum cyfrowe - definicja



Pod pojęciem **archiwum cyfrowe (repozytorium)** należy rozumieć system złożony z osób oraz przyjętych rozwiązań organizacyjnych i technicznych, powołany w celu zgromadzenia, przechowania **oraz zapewnienia długoterminowego dostępu i użyteczności cyfrowego materiału.**

„Dobre” Archiwum cyfrowe

1. Umożliwia tworzenia i stosowania nowych formatów przedstawiania danych, informacji i wiedzy.
2. Brak ograniczeń co do czasu i miejsca.
3. Personalizuje informację np. reaguje na problem nadmiaru informacji.
4. Wspiera nowe formy pracy grupowej w zakresie tworzenia i wykorzystania informacji.
5. Umożliwia dostęp do całości lub części zarchiwizowanych Umożliwia tworzenia i stosowania nowych materiałów.
6. Otwarta architektura, interoperacyjność.
7. **Wiarygodność Cyfrowego Archiwum (ISO 16363:2012)**

„Dobry” Zasób cyfrowy

1. Dobry zasób cyfrowy jest tworzony na podstawie jasno określonej polityki.
2. Zasoby powinny być opisane w sposób pozwalający użytkownikowi uzyskać charakterystyki zasobów np. zakres zasobu, ograniczenia dostępu.
3. Zasobem cyfrowym trzeba dobrze zarządzać w trakcie całego cyklu.
4. Dobry zasób cyfrowy respektuje wszystkie prawa autorskie.

Model referencyjny OAIS (1)

Model referencyjny repozytoriów cyfrowych **Open Archival Information System** został stworzony na potrzeby archiwizacji i wymiany danych elektronicznych, zawierających informacje z badań przestrzeni kosmicznej. W 2003 model OAIS dostał normę postępowania w zakresie **długoterminowej archiwizacji danych cyfrowych (ISO 14721:2003)**. Jest on uznawany za uniwersalny model organizowania i funkcjonowania repozytoriów cyfrowych i stosowany do gromadzenia, przechowywania i udostępniania różnych typów dokumentów elektronicznych.

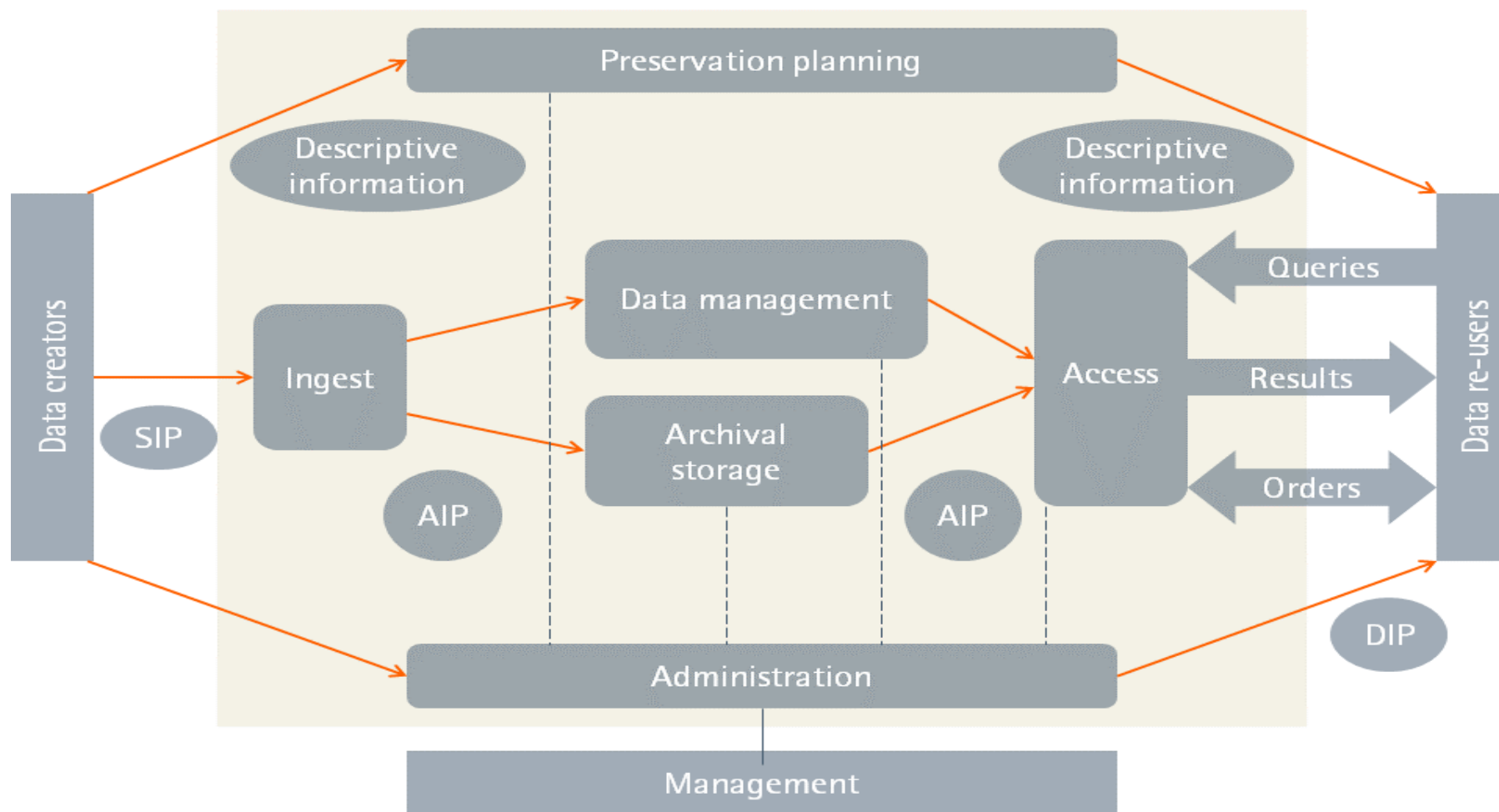
Model referencyjny OAIS (2)

W OAIS wyróżniamy tzw. **pakiet informacyjny (Information Package)**. Składa się on z dwóch komponentów, tj. **kontenera informacyjnego (Content Information)** oraz **informacji dotyczącej przechowywania jego zawartości (Preservation Description Information)**. Są to wszelkie informacje konieczne do odpowiedniego przechowania informacji treściowej (**kontenera informacyjnego**). Zalicza się tu 4 typy informacji, określane jako: **Provenance (historia)**, **Context (powiązania)**, **Reference ((identyfikatory)**, **Fixity (mechanizmy ochrony)**.

Model referencyjny OAIS (3)

W archiwach zgodnych z modelem OAIS istotne jest rozróżnienie pomiędzy **danymi cyfrowymi (Data Object)** a **obiektami informacyjnymi (Information Object)**. Niezbędnym w archiwum elektronicznym są **metadane** przechowywanych obiektów. W modelu referencyjnym OAIS określane są one terminem **Descriptive Information**. Metadane dostarczają informacji o **zawartości pakietu informacyjnego** oraz umożliwiają jego odnalezienie w archiwum.

Archiwum Cyfrowe w modelu OAIS

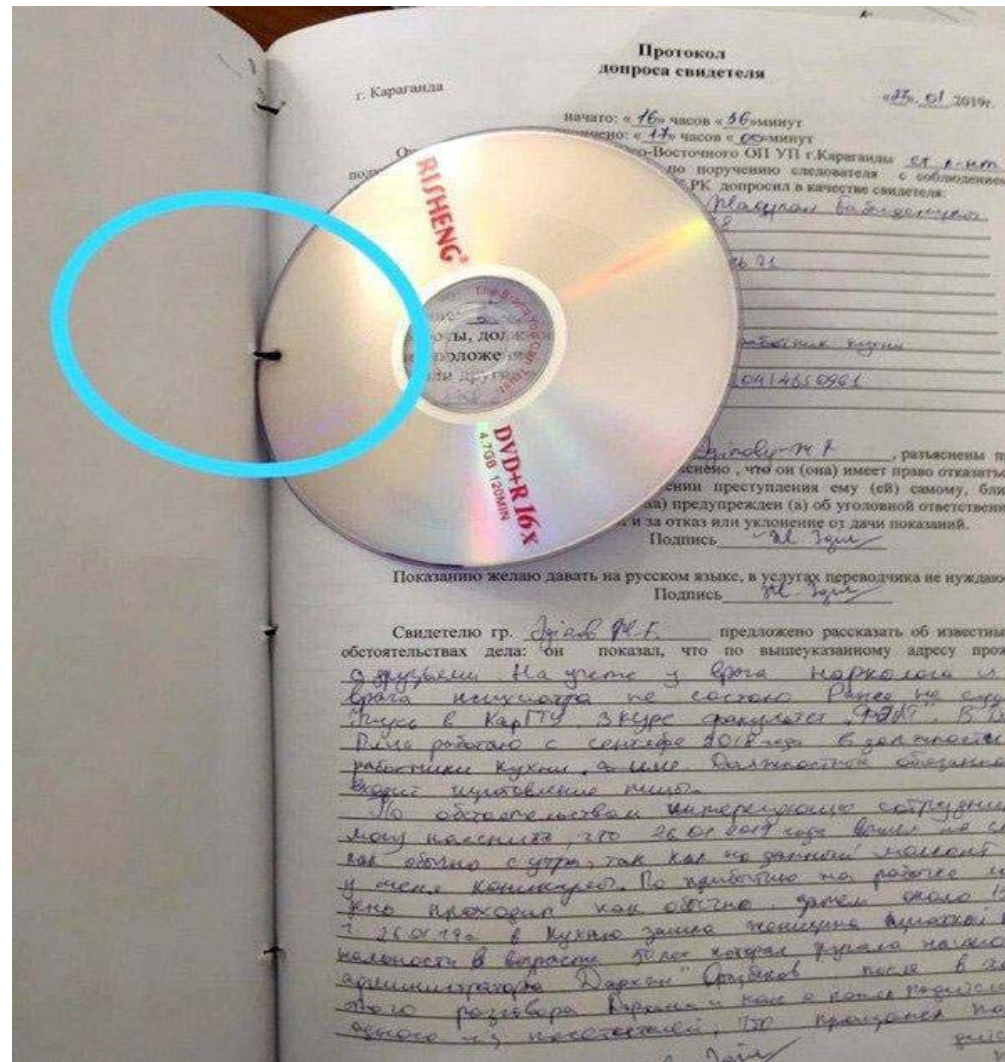


Mechanizmy długotrwałej archiwizacji zasobów cyfrowych

Długoterminowa archiwizacja zasobów cyfrowych to **proces ochrony autentyczności, integralności oraz poufności obiektów cyfrowych.**

Natomiast trwała ochrona informacji cyfrowych definiujemy jako **zbiór rozwiązań służących do zapewnienia ciągłości dostępu do materiałów w okresie, w którym taka ciągłość jest pożądana.**

Archiwizacja danych cyfrowych razem z metadanymi w postaci analogowej



Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – kopia zapasowa

- **Reguła 3-2-1**
- **Kopie cykliczne** - cyklicznie wykonywanie kopii wszystkich zgromadzonych danych w systemie.
- **Kopie przyrostowe** – robimy kopie danych zmodyfikowanych lub tych dodanych między backupami.

Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – Problemy

Problem Nr. 1 - Niska trwałość nośników dokumentów cyfrowych.

Problem Nr. 2 - utrata dostępu do dokumentów w wyniku zdarzeń losowych tj. pożar, powódź, atak hakerów, wirusy.

Problem Nr. 3 – utrzymanie prawidłowego zapisu i opisu dokumentów umożliwiające identyfikację oraz odczyt.

Problem Nr. 4 - Utrata informacji kontekstowych grozi utratą czytelności i wiarygodności danych.

Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – Trwałość nośników (1)

- CD/DVD
- Dyski HDD/SSD/Macierze dyskowe
- Biblioteki taśmowe

Warunki klimatyczne

| Rodzaj dokumentacji | Właściwa temperatura powietrza (w stopniach Celsjusza) | | Dopuszczalne wahania dobowe temperatury powietrza (w stopniach Celsjusza) | Właściwa wilgotność względna powietrza (w % RH) | | Dopuszczalne wahania dobowe wilgotności względnej powietrza (w % RH) |
|---|--|------|---|---|------|--|
| | min. | maks | | min. | maks | |
| 1. Papier | 14 | 18 | 1 | 30 | 50 | 3 |
| 2. Dokumentacja audiowizualna: | | | | | | |
| 2.a. Fotografia czarno-biała (negatywy i pozytywy) | 3 | 18 | 2 | 20 | 50 | 5 |
| 2.b. Fotografia kolorowa (negatywy i pozytywy), taśma filmowa | 3 | 18 | 2 | 20 | 50 | 5 |
| 2.c. Taśmy magnetyczne do analogowego zapisu obrazu lub dźwięku | 8 | 18 | 2 | 20 | 50 | 5 |
| 3. Informatyczne nośniki danych | 12 | 18 | 2 | 30 | 40 | 5 |

Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – Odświeżanie nośnika

W przypadku stwierdzenia oznak degradacji nośnika konieczne staje się tzw. **odświeżenie nośnika**, a więc przekopiowanie danych na nowy nośnik tego samego typu. Robimy do momentu kiedy dany nośnik osiągnie swój maksymalny okres trwałości.

Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – Zmiana generacji nośnika

Jeśli określony rodzaj staje się przestarzały, wtedy przenosimy dane ze wszystkich nośników starego typu na **nowszą generację, niezależnie od stanu technicznego poszczególnych egzemplarzy.**

Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – Migracja/Emulacja

Migracja: konwersja dokumentu z oryginalnego formatu do nowszego ale tak aby nie było utraty danych treściowych i strukturalnych.

Emulacja: to nie jest migracja danych, tylko otoczenia, które jest potrzebne do ich odczytu.

Mechanizmy długotrwałej archiwizacji zasobów cyfrowych – Formaty zapisu dokumentów cyfrowych

- Pliki tekstowe: **RFT, SGML, HTML, XML**
- Pliki graficzne: **TIFF, JPG, PNG, PDF**
- Multimedia: **FLAC, MPEG-2, WAV, AIFF, MP3**
- Pliki interaktywne: **format Flash**

Metadane - „Dobre” Metadane

1. Dobre metadane powinny być dobrane do użytkowników.
2. Dobre metadane wspomagają interoperacyjność
3. Dobre metadane stosują narzędzia kontroli treści takie jak **słowniki kontrolowane** i **słowniki haseł wzorcowych**.
4. Dobre metadane zawierają jasne określenie warunków i zasad wykorzystywania obiektu cyfrowego.
5. Dobre metadane techniczne i administracyjne ułatwiają długoterminowe zarządzanie obiektami w zasobach.

„Dobre” obiekty cyfrowe

1. Dobry obiekt zapisywany jest w formacie, który zapewni mu powtórne wykorzystanie.
2. Dobry obiekt cyfrowy musi się dać zachować dla przyszłości. Musi być dostępny przez długi okres
3. Dobry obiekt musi być interoperacyjny.
4. Dobry obiekt powinien mieć stały, unikalny identyfikator.
5. Dobry obiekt powinien mieć potwierdzoną autentyczność.



Dziękuję za uwagę

Dawid Żądłowski

Archiwista24

sirsilis@gmail.com