

Archiwizacja i zakresy wykorzystania cyfrowych odwzorowań obiektów muzealnych

MARCIN WERLA

Długoterminowa archiwizacja

Dane wytworzone w procesie digitalizacji należy długoterminowo zabezpieczyć

- Dane wzorcowe
- Metadane
- Postacie użytkowe/prezentacyjne

Co jest celem takiego działania?

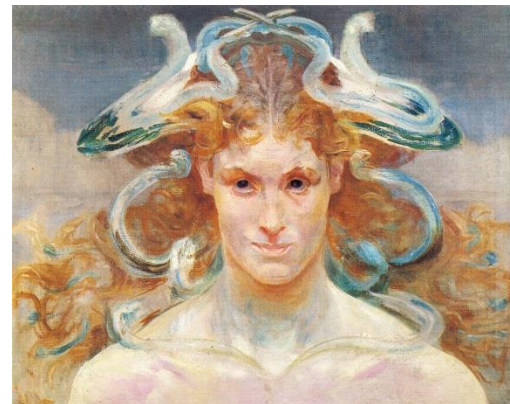
**Zapewnienie możliwości dostępu do informacji
zakodowanych w postaci danych cyfrowych**

Długoterminowa archiwizacja

DANE

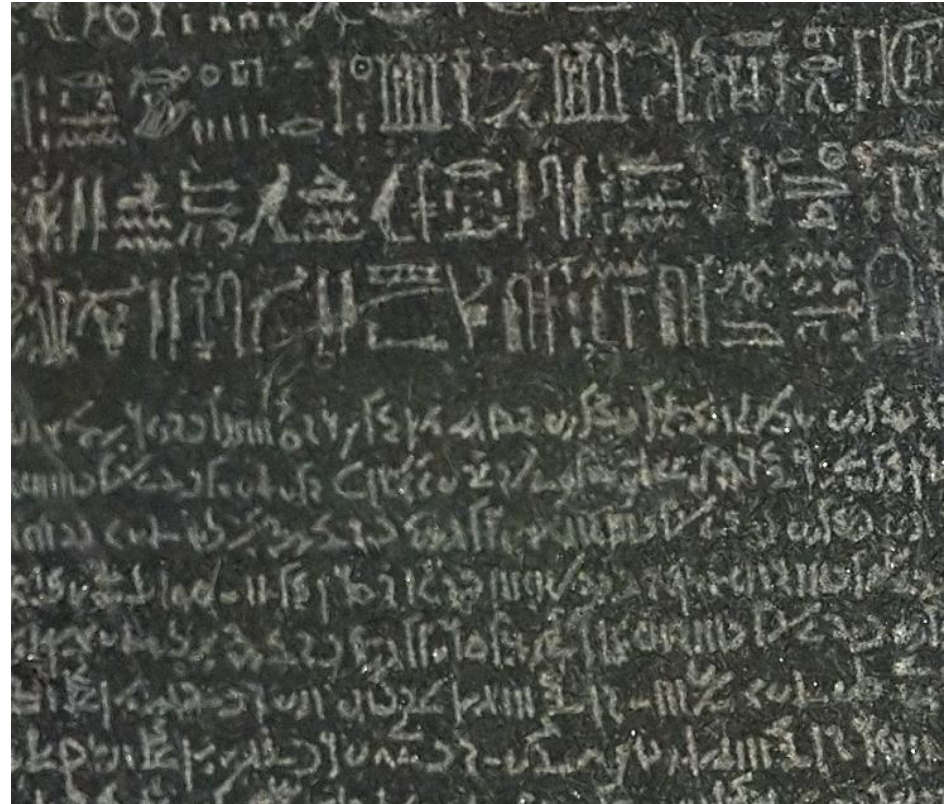
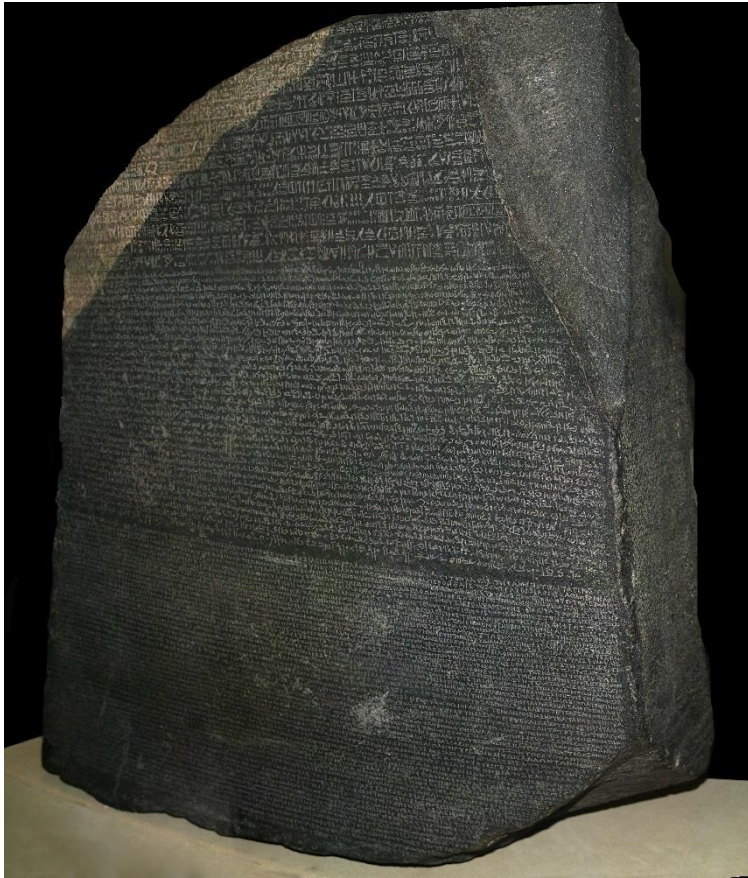
```
FF 08 FF E0 00 10 4A 46 | 49 46 00 01 01 01 02 58
02 58 00 00 FF DB 00 43 | 00 05 03 04 04 04 03 05
04 04 04 05 05 05 06 07 | 0C 08 07 07 07 07 0F 0B
0B 09 0C 11 0F 12 12 11 | 0F 11 11 13 16 1C 17 13
14 1A 15 11 11 18 21 18 | 1A 1D 1D 1F 1F 1F 13 17
22 24 22 1E 24 1C 1E 1F | 1E FF DB 00 43 01 05 05
05 07 06 07 0E 08 08 0E | 1E 14 11 14 1E 1E 1E 1E
1E 1E 1E 1E 1E 1E 1E 1E | 1E 1E 1E 1E 1E 1E 1E 1E
1E 1E 1E 1E 1E 1E 1E 1E | 1E 1E 1E 1E 1E 1E 1E 1E
1E 1E 1E 1E 1E 1E 1E 1E | 1E 1E 1E 1E 1E 1E 1E C2
00 11 08 04 E6 06 40 03 | 01 22 00 02 11 01 03 11
01 FF C4 00 1B 00 00 02 | 03 01 01 01 00 00 00 00
00 00 00 00 00 00 02 03 | 01 04 05 00 06 07 FF C4
00 1A 01 00 03 01 01 01 | 01 00 00 00 00 00 00 00
00 00 00 01 02 03 00 04 | 05 06 FF DA 00 0C 03 01
00 02 10 03 10 00 00 01 | D0 AF 2B 43 62 10 18 58
52 83 07 9D 3E C6 CB 2B | 16 CD AF 00 E1 E4 80 DA
CC 57 81 AE 25 47 8B 4E | B4 91 66 B8 D7 3A C3 29
9E 0D 94 49 D6 EB 82 C1 | 74 A0 08 B6 29 80 6C 8D
67 E3 CA E4 62 E6 D1 26 | 5B 75 C4 10 DC 5D 57 32
99 D5 3C D6 D6 B5 E1 6B | AB 02 B5 B8 41 32 36 BA
EC 62 D6 57 1C 18 2B 1C | CD 84 88 C7 0A 51 CF 24
```

INFORMACJE



**Cel: Zapewnienie możliwości dostępu do informacji
zakodowanych w postaci danych cyfrowych**

Długoterminowa archiwizacja



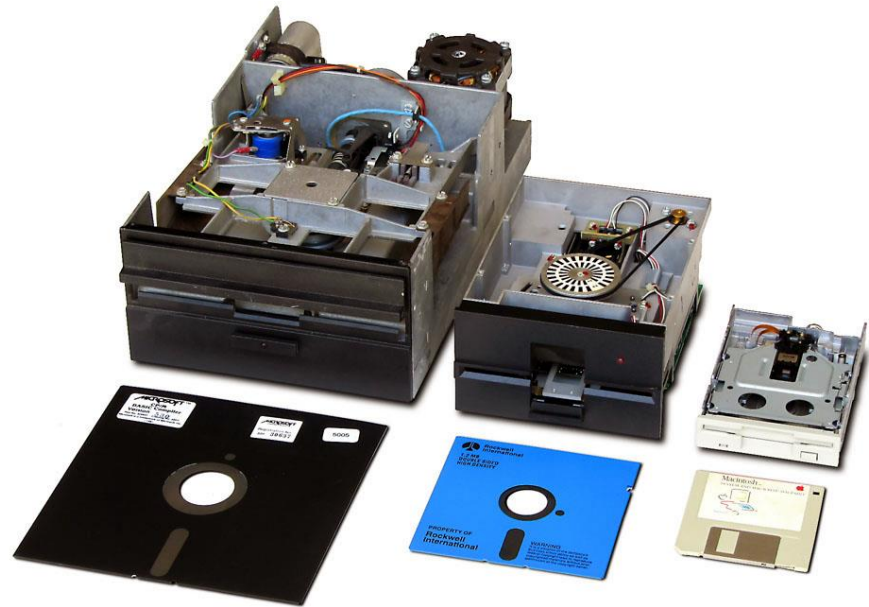
http://en.wikipedia.org/wiki/File:Rosetta_Stone.JPG

Długoterminowa archiwizacja

Cel: Zapewnienie możliwości dostępu do informacji zakodowanych w postaci danych cyfrowych

Trzy poziomy działania:

- Zapewnienie trwałości nośnika i trwałości zapisu danych na tym nośniku
- Zapewnienie możliwości odczytu danych z nośnika
- Zapewnienie możliwości odkodowania informacji zawartych w danych



<http://oldcomputers.net/>

Długoterminowa archiwizacja

Zapewnienie trwałości nośnika i trwałości zapisu danych na tym nośniku – niezbędne są:

- Świadomość czasu życia nośnika danych i odpowiednio częste przegrywanie danych ze starych nośników na nowe
 - Przy dużej ilości danych i/lub dużej liczbie nośników takie działania stają się poważnym wyzwaniem
 - Niezbędne może być też odświeżanie zapisu na danym nośniku, czyli odczyt i ponowny zapis danych
- Weryfikowanie poprawności zapisu jeszcze w trakcie czasu życia pojedynczego egzemplarza nośnika
 - Nie wszystkie nośniki żyją tak długo jak by chciał tego producent
 - Poprawność zapisu weryfikowana jest przy pomocy sum kontrolnych (np. SHA-512)

91 1240 6175 1111 0000 4574 9087

- Ale co gdy wykryjemy błąd zapisu cyfrowego na nośniku?

Długoterminowa archiwizacja

Zapewnienie trwałości nośnika i trwałości zapisu danych na tym nośniku – niezbędne są:

- Przynajmniej dwie niezależnie składowane kopie danych - niezależnie czyli
 - Na innym nośniku / w innym urządzeniu
 - W innym pomieszczeniu / w innym budynku / w innym mieście

Zapewnienie możliwości odczytu danych z nośnika - musimy posiadać urządzenie do odczytu używanych przez nas nośników

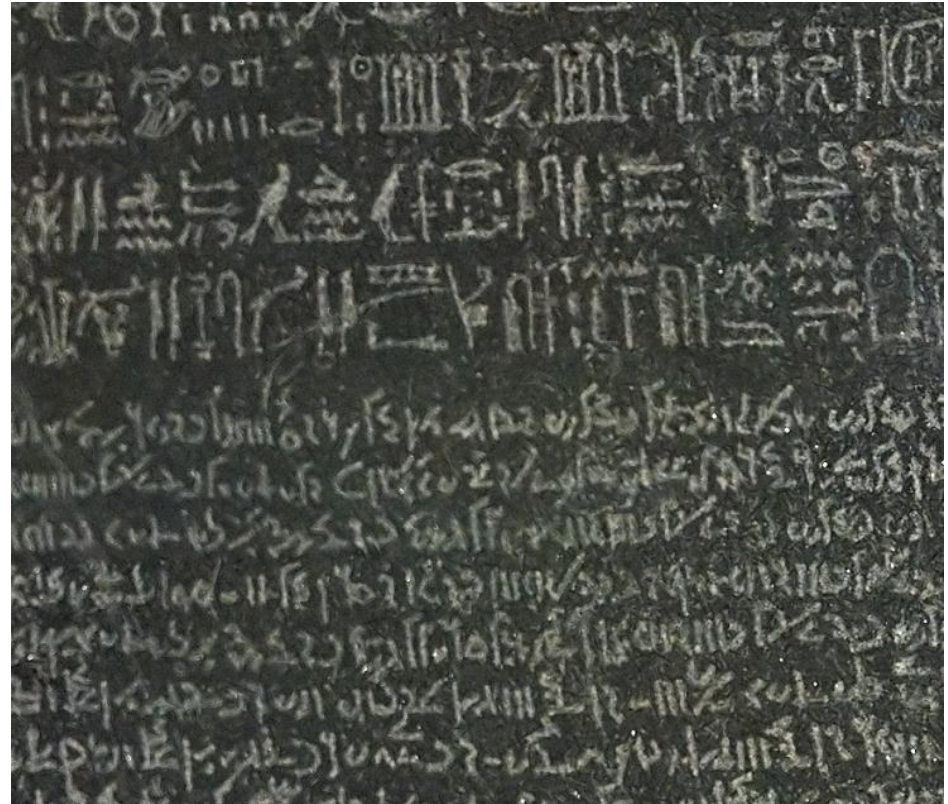
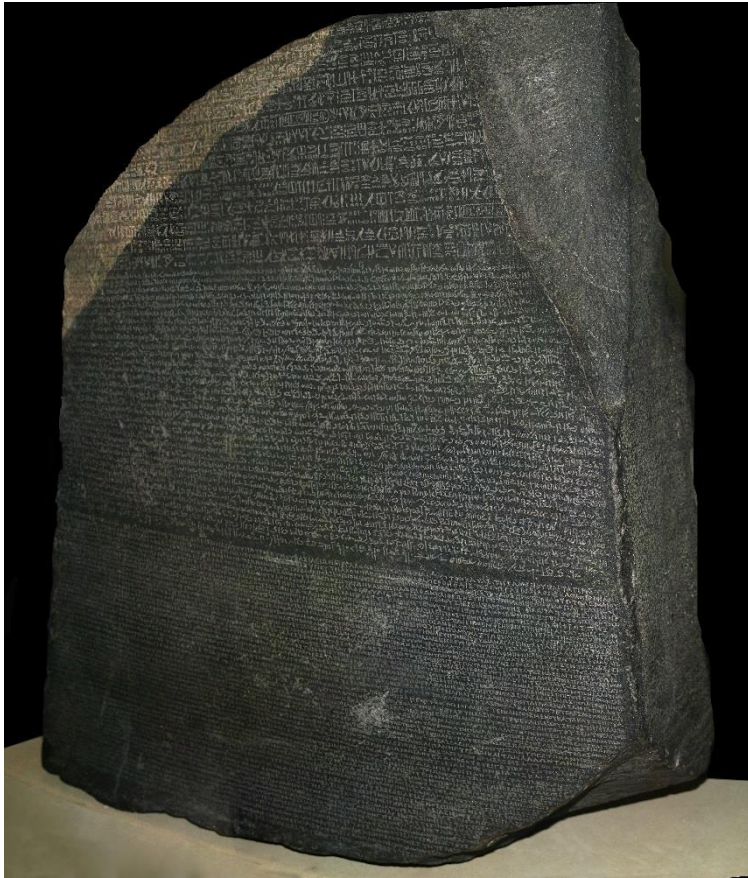
- Sprawne i zgodne ze współczesnymi technologiami IT
- Musimy uważać, czy technologia której używamy do zapisu danych nie wychodzi z użycia
 - Jeżeli tak, to należy dokonać przeniesienia danych na nośniki nowszej generacji

Długoterminowa archiwizacja

Zapewnienie możliwości odcodowania informacji zawartych w danych

- Zapisując dane w celu długoterminowego przechowywania należy wybierać bardzo dobrze znane, popularne formaty zapisu danych – najlepiej jeżeli są to tzw. otwarte standardy/formaty (np. TIFF, JPEG-2000)
 - Zwiększa to znacznie szansę na odczytanie tych danych w przyszłości
 - <http://opendatahandbook.org/en/appendices/file-formats.html>
- Należy monitorować, czy formaty danych z których korzystamy nie wychodzą z użycia i w razie konieczności migrować dane do nowych formatów
 - Przy dużej ilości danych takie działanie staje się poważnym wyzwaniem – potrzeba nie tylko odczytać i zapisać dane, ale również dokonać ich konwersji – niezbędna jest więc moc obliczeniowa
 - Przy zmianie formatu zapisu danych należy uważać aby nie stracić i nie zmienić zawartych w nich informacji

Długoterminowa archiwizacja



http://en.wikipedia.org/wiki/File:Rosetta_Stone.JPG

Długoterminowa archiwizacja

Cel: Zapewnienie możliwości dostępu do informacji zakodowanych w postaci danych cyfrowych

Trzy poziomy działania:

- Zapewnienie trwałości nośnika i trwałości zapisu danych na tym nośniku
- Zapewnienie możliwości odczytu danych z nośnika
- Zapewnienie możliwości odkodowania informacji zawartych w danych

Straszne? Trudne? Nudne?

Jedną z opcji działania jest outsourcing

Długoterminowa archiwizacja

Outsourcing archiwizowania danych – na co warto zwrócić uwagę?

- Sposób dostarczania danych do archiwum i dostępu do tego archiwum
 - Np. parametry łącza internetowego pomiędzy instytucją archiwizującą dane a firmą wykonującą archiwizację
- Gwarancje niezawodności dostępu i bezpieczeństwa składowania danych, czyli to w jaki sposób firma radzi sobie z omówionymi wcześniej wyzwaniami
- Maksymalna oferowaną pojemność archiwum
- Stabilność firmy i jej pozycja na rynku

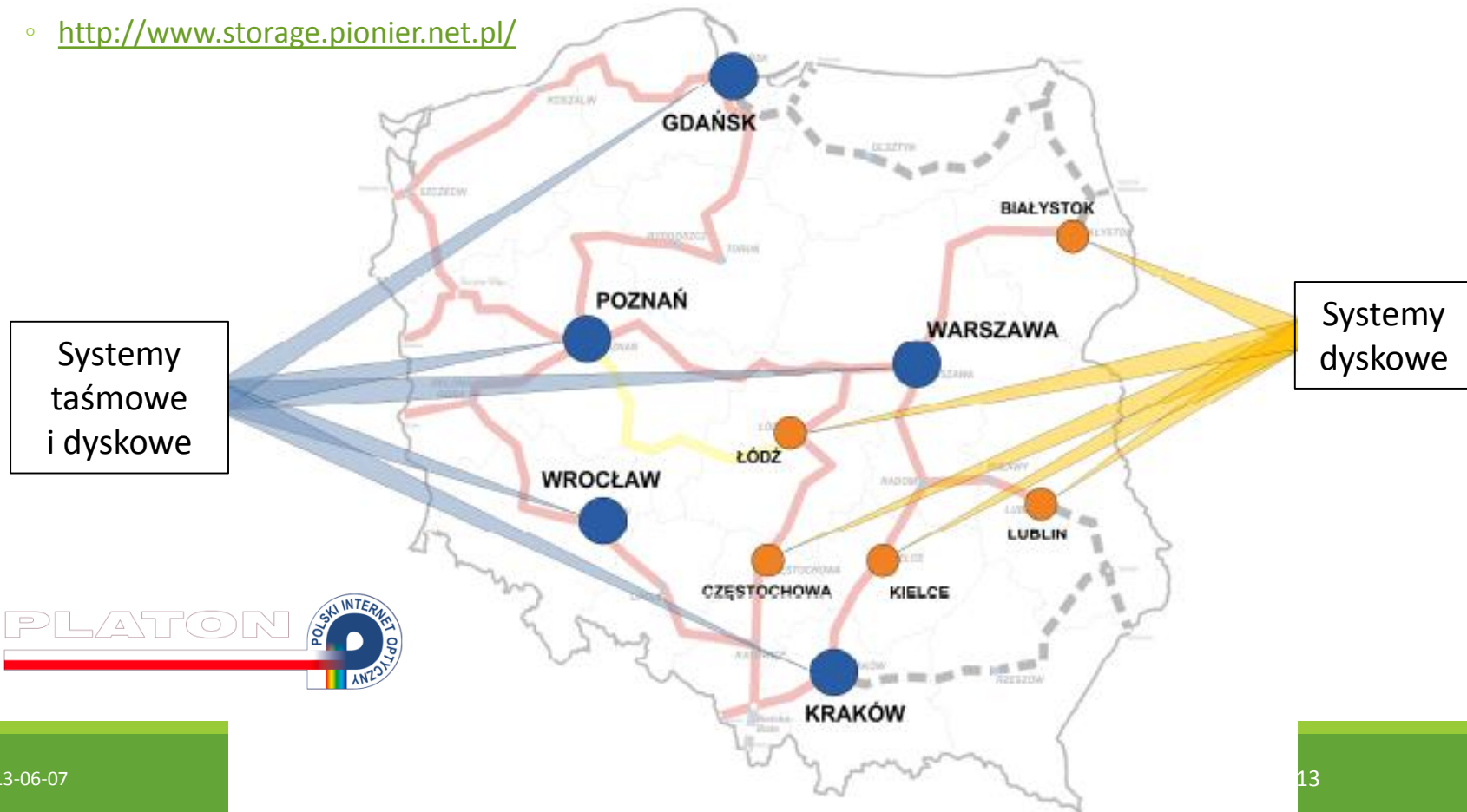
Długoterminowa archiwizacja

Outsourcing archiwizowania danych – na co warto zwrócić uwagę?

- Koszty, w tym:
 - koszt składowania danych (wyrażony np. w cenie za składowanie 100 GB przez jeden rok)
 - koszt transferu danych (może być kosztem dodatkowym poza samym kosztem składowania)
 - zmianę kosztów w przypadku wzrostu ilości składowanych danych (nie musi być to zależność liniowa)
 - Przykładowe ceny za rok składowania (pierwszy TB / kolejne TB, źródło: <http://duracloud.org/pricing>, w NBP średni kurs USD na dziś 1 \$ = 3,1392 PLN):
 - Jedna kopia: 1 500 \$ / 1 000 \$
 - Dwie kopie: 1 635 \$ / 1 135 \$ lub 2 500 \$ / 1 700 \$
 - Jedna kopia + zaawansowane uprawnienia dostępu: 5 900 \$ / 1 000 \$
 - Dwie kopie + zaawansowane uprawnienia dostępu: 6 035 \$ / 1 135 \$ lub 6 900 \$ / 1 700 \$

Długoterminowa archiwizacja

- Rozwiązanie krajowe: Usługa Powszechnej Archiwizacji funkcjonująca w ramach Platformy Obsługi Nauki PLATON i sieci PIONIER
 - <http://www.storage.pionier.net.pl/>



Zakresy wykorzystania cyfrowych odwzorowań obiektów muzealnych

Podstawowy podział

- Wykorzystanie wewnątrz instytucji macierzystej
- Wykorzystanie poza instytucją – w szczególności **udostępnianie on-line**

Udostępnianie zasobów cyfrowych

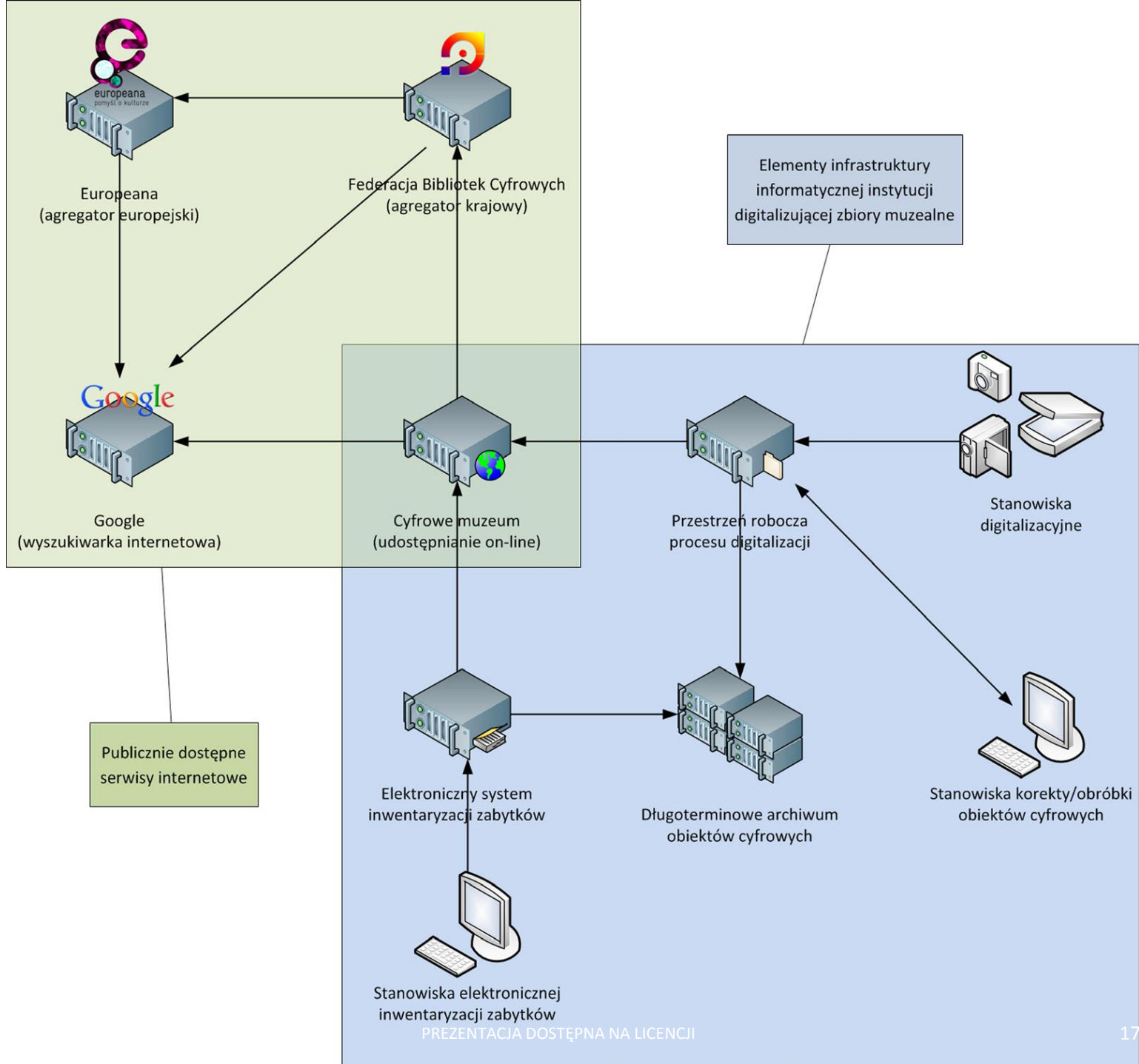
Pierwszy krok: określenie docelowej grupy użytkowników, odbiorców udostępnianych treści – ma to wpływ na:

- Postać użytkową obiektów cyfrowych
 - Właściwości, jakie postać ta powinna mieć
- Schemat metadanych i stopień szczegółowości jego wypełnienia
 - Zarówno dla metadanych opisowych, jak i technicznych
- Dodatkowe funkcje, jakie powinien oferować system udostępniający wizerunki cyfrowe wraz z metadanymi
- Niezbędny poziom zabezpieczenia udostępnianych danych
- Pożądany lub wymagany poziom upowszechnienia i udostępniania danych
 - Otwarte standardy wymiany danych

Udostępnianie zasobów cyfrowych

Co wchodzi w skład technicznej infrastruktury związanej z udostępnianiem zasobów cyfrowych on-line

- System informatyczny, w którym gromadzone i przetwarzane są udostępniane dane
- Serwery, na których system ten działa
- Sieć komputerowa, którą system ten jest połączony z potencjalnymi użytkownikami
 - Zarówno łączy po stronie instytucji jak i po stronie użytkowników
- Urządzenia końcowe wykorzystywane przez użytkowników



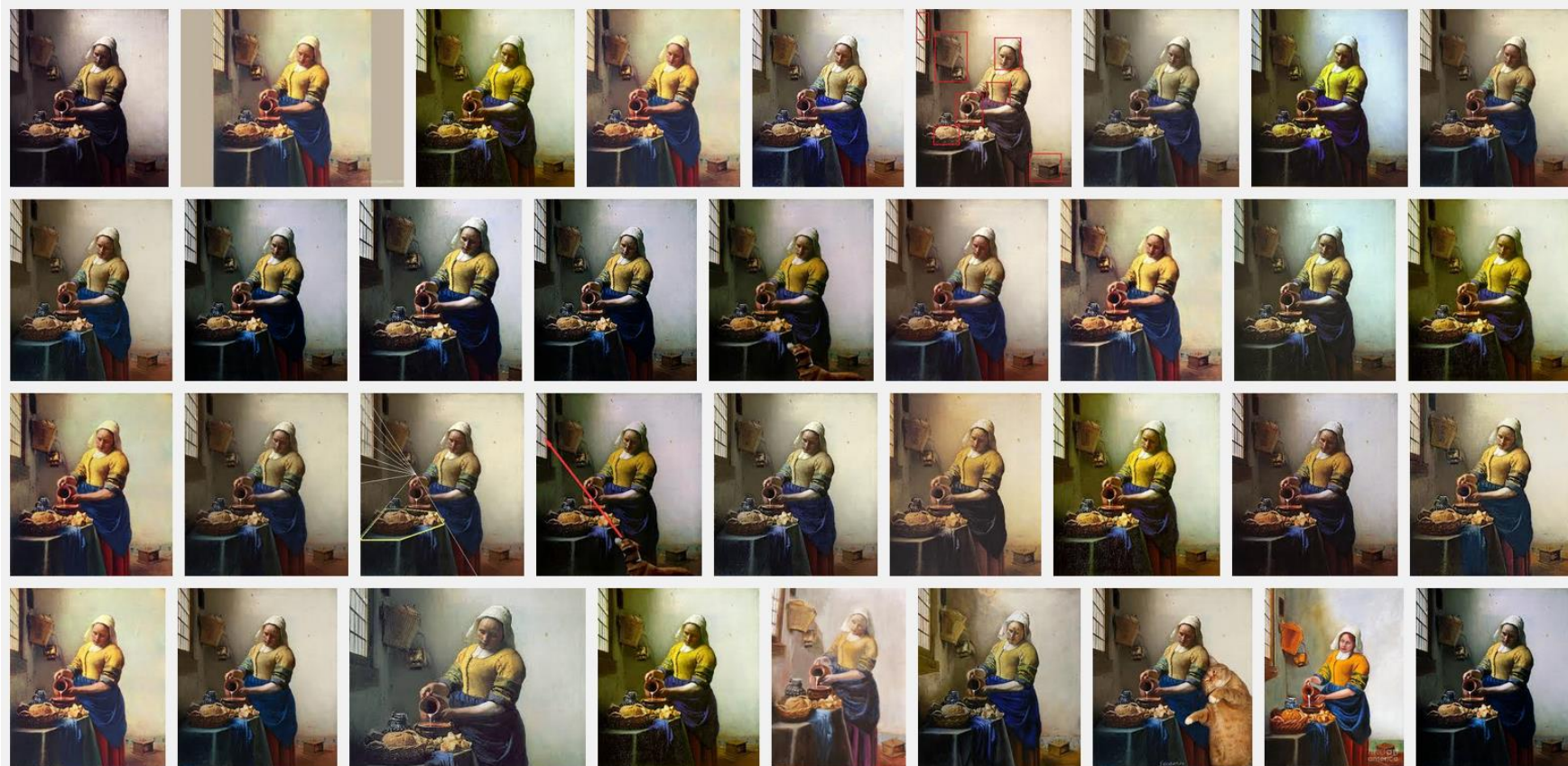
Udostępnianie zasobów cyfrowych



Problem żółtej „Mleczarki”

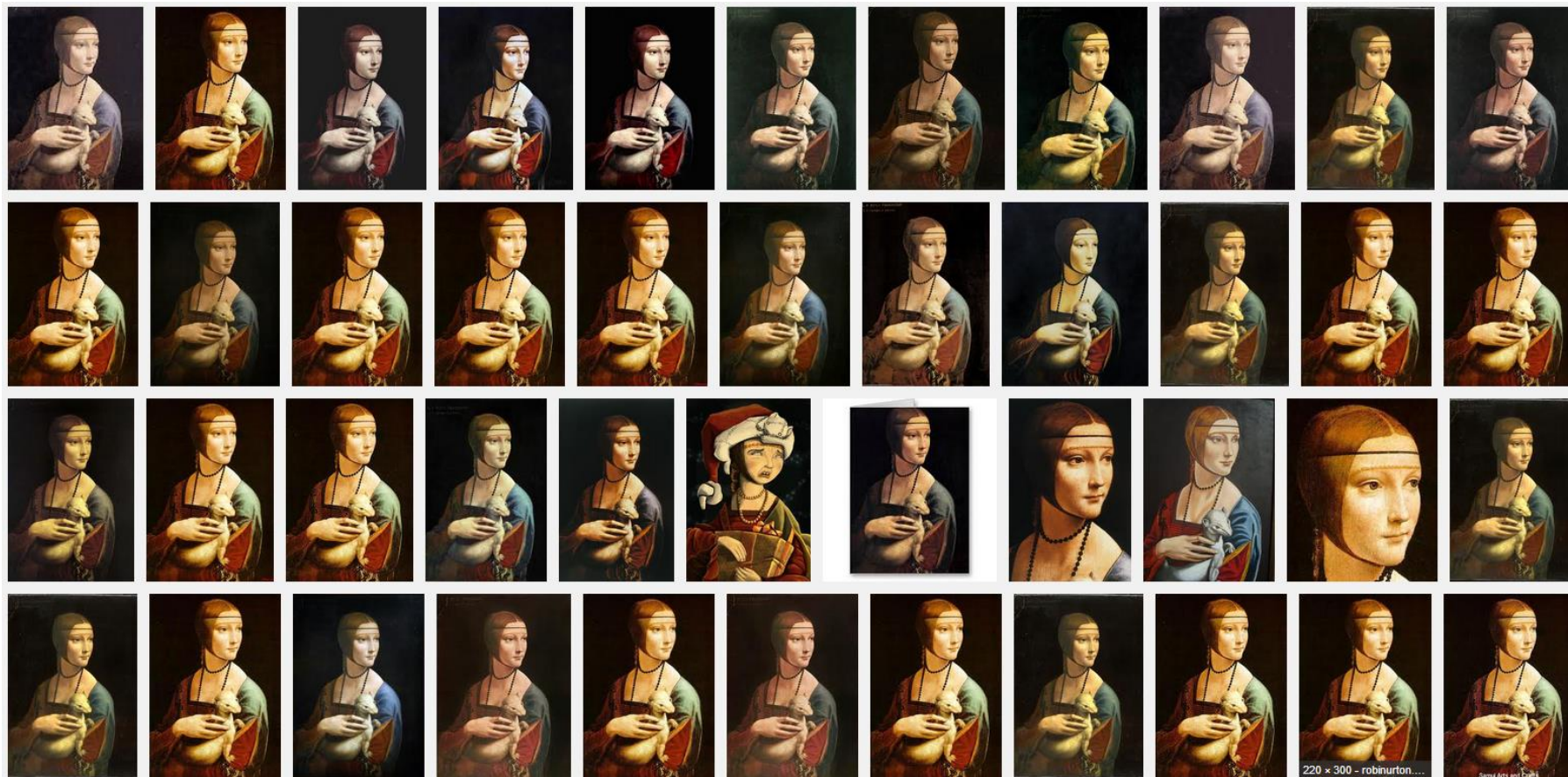
<http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5>

Udostępnianie zasobów cyfrowych



<https://www.google.pl/search?q=milkmaid+vermeer&tbm=isch>

Udostępnianie zasobów cyfrowych



<https://www.google.pl/search?q=lady+with+ermine&tbm=isch>

Udostępnianie zasobów cyfrowych

Zasady otwartych instytucji kultury (v.0.5)

- 1. Jeśli to tylko możliwe, cyfrowe informacje o utworach (metadane) powinny być udostępniane na zasadach oświadczenia Creative Commons Zero.**
- 2. Utrzymuj w domenie publicznej cyfrowe reprezentacje utworów należących z domeny publicznej, nie ograniczaj dostępu poprzez nowe ograniczenia prawne.**
- 3. Gdy publikujesz dane, oświadczyć jednoznacznie jakie są twoje życzenia i oczekiwania odnośnie ponownego wykorzystania opisów, całej kolekcji danych i jej podzbiorów.**
- 4. Gdy udostępniasz dane, używaj otwartych formatów, które są przygotowane do przetwarzania automatycznego**
- 5. Dąż do możliwości zaangażowania internetowej publiczności w nowatorski sposób**

Więcej informacji: <http://fbc.pionier.net.pl/pro/informacje-ogolne/zasady-otwartych-instytucji-kultury-openglam-principles/>

Dziękuję za uwagę

KONTAKT: MARCIN WERLA (MWERLA@MAN.POZNAN.PL)