

METADANE, ZAGADNIENIA SŁOWNIKÓW KONTROLOWANYCH

*Zespół ekspertów powołany przez Narodowy Instytut
Muzealnictwa i Ochrony Zbiorów, grupa ds. metadanych*

**Lidia Karecka, Marcin Kłos, Karolina Tabak, Marcin Werla,
dr Tomasz Zaucha, Katarzyna Zielonka, dr hab. Łukasz Gawel**

*ze strony NIMOZ: **Anna Kuśmidrowicz-Król***



NARODOWY INSTYTUT MUZEALNICTWA I OCHRONY ZBIORÓW

Redakcja: Łukasz Gaweł

Skład i opracowanie techniczne:
Agnieszka Kalota, Tomasz Kalota

Warszawa 2012

Narodowy Instytut Muzealnictwa i Ochrony Zbiorów

ISBN 978-83-933790-6-4

Spis treści

Od redakcji	5
1. Udostępnianie informacji o zbiorach w kontekście ewidencji muzealiów	9
1.1. Udostępnienie informacji o zbiorach [Lidia Karecka]	10
1.2. Opis obiektu muzealnego a jego ewidencja [Lidia Karecka]	11
1.3. Trudności wynikające z braku konsekwencji w opisie obiektów [Lidia Karecka, Marcin Kłós, Katarzyna Zielonka]	13
2. Metadane	17
2.1. Schematy i standardy metadanych w muzeach [Tomasz Zaucha]	18
2.1.1. Dublin Core Metadata Element Set (DCMES) oraz DCMI Metadata Terms (DC Terms) [Marcin Werla]	19
2.1.2. Categories for Description of Works of Art (CDWA) [Tomasz Zaucha]	21
2.1.3. Spectrum [Marcin Kłós]	23
2.1.4. CIDOC CRM [Tomasz Zaucha]	25
2.1.5. Lightweight Information Describing Objects (LIDO) [Tomasz Zaucha]	33
2.2. Wybór optymalnego rozwiązania [Katarzyna Zielonka]	40
3. Kontrola słownictwa	43
3.1. Po co kontrolować słownictwo? [Katarzyna Zielonka]	44
3.2. Jak kontrolować słownictwo? [Karolina Tabak]	46
3.3. Wybór optymalnego rozwiązania	48
3.3.1. Gotowe rozwiązania [Karolina Tabak]	48
3.3.1.1. Art and Architecture Thesaurus (AAT) [Tomasz Zaucha]	48
3.3.1.2. The Union List of Artist Names (ULAN) [Karolina Tabak]	49
3.3.1.3. Getty Thesaurus of Geographic Names (TGN) [Tomasz Zaucha]	50
3.3.1.4. British Museum Materials Thesaurus i British Museum Object Names Thesaurus [Marcin Kłós]	50
3.3.1.5. Rijksbureau voor Kunsthistorische Documentatie (RKD) [Karolina Tabak]	54
3.3.1.6. Virtual International Authority File (VIAF) [Tomasz Zaucha]	55
3.3.1.7. ICONOCLASS [Tomasz Zaucha]	56
3.3.2. Budowa własnego słownika [Karolina Tabak]	58
4. Propozycje rozwiązań na poziomie krajowym	65
4.1. Krajowy agregator danych [Katarzyna Zielonka]	66

4.2. Standaryzacja metadanych [Katarzyna Zielonka].....	68
4.3. Baza tezaurów [Katarzyna Zielonka]	69
4.3.1. Scenariusz 1. Wyszukiwanie haseł poprzez przeglądarkę WWW [Marcin Werla]	70
4.3.2. Scenariusz 2. Skopiowanie całego tezaurusa do bazy danych własnego systemu muzealnego [Marcin Werla]	72
4.3.3. Scenariusz 3. Dynamiczna integracja systemu informatycznego muzeum i systemu tezaurusa [Marcin Werla]	73
4.4. Narodowy Uniwersalny Katalog Centralny (NUKAT) [Lidia Karecka]	75
4.5. Udostępnianie danych o zbiorach w Internecie [Marcin Werla, Tomasz Zaucha]	78
4.6. Sprawozdawczość [Katarzyna Zielonka]	83
Zakończenie	85

Od redakcji

Narodowy Instytut Muzealnictwa i Ochrony Zbiorów, działając w myśl założeń WPR Kultura+ jako centrum kompetencji do spraw digitalizacji muzealiów, rozpoczął w 2011 roku prace zmierzające do wypracowania standardów i dobrych praktyk, wspierających rozwijanie procesów digitalizacyjnych w muzeach.

W zeszłym roku, we współpracy z zespołem ekspertów złożonym z przedstawicieli wiodących muzeów w kraju oraz reprezentantów sektora nauk ścisłych, opracowaliśmy i udostępniiliśmy dokument *Zalecenia dotyczące planowania i realizacji projektów digitalizacyjnych w muzealnictwie*.

W tym pierwszym etapie prac standaryzacyjnych postawiliśmy sobie za cel identyfikację podstawowych problemów dotyczących planowania i wdrażania projektów digitalizacyjnych w muzeach. Zakres poruszonych zagadnień obejmował kwestie od pozyskiwania, gromadzenia i przechowywania danych cyfrowych dotyczących muzealiów, do ich udostępniania i roli w promocji zbiorów. Główną część zeszłorocznego raportu poświęciliśmy syntetycznemu przedstawieniu kluczowych etapów procesu cyfryzacji zbiorów muzealnych. Staraliśmy się zwrócić uwagę na newralgiczne punkty i decyzje, jakie należy podjąć w trakcie procesu.

Zgodnie z przyjętym na kolejne lata planem kontynuowania prac specjalistycznych zespołów nad różnymi aspektami procesu cyfryzacji muzealiów, w tym roku zorganizowaliśmy dwie grupy. W centrum uwagi pierwszej znalazły się zagadnienia dotyczące metadanych i słowników kontrolowanych, druga skupiła się na pozyskiwaniu cyfrowych odwzorowań obiektów muzealnych. Efektem niemal całorocznej pracy są dwa raporty, które niniejszym przedstawiamy.

Mamy nadzieję, że opracowania te nie tylko będą służyć praktyczną pomocą muzealnikom, mającym do czynienia z zadaniami związanymi z cyfryzacją zbiorów, ale będą także wskazaniem dla kadry zarządzającej, że wdrażanie digitalizacji wymaga równocześnie wprowadzenia pewnych wymogów organizacyjnych i jakościowych.

Chcielibyśmy, żeby dokumenty te przyczyniły się do promowania wysokich standardów dla procesów digitalizacji muzealiów w Polsce, ale także żeby wywołały w środowiskach muzealnych dyskusję i wymianę doświadczeń dotyczących digitalizacji, podnosząc jej rangę oraz wskazując konkretne rozwiązania systemowe i szczegółowe.

Wraz z zeszłoroczną, te trzy publikacje tworzą już pewien komplementarny zbiór zaleceń i wskazówek, które powinny wspierać sprawne i skuteczne zarządzanie procesem digitalizacji zbiorów muzealnych. Nasuwa się jednak jeszcze jedno, kluczowe pytanie: czego potrzeba, aby opracowania te nie pozostały jedynie teoretycznymi rozważaniami? Jak nadać im wymiar praktyczny? Opracowywanie teorii zdaje się być pozbawione sensu, jeśli nie znajdzie ona zastosowania w codziennej praktyce muzealnej i to nie tylko muzeów największych, ale wszystkich muzeów w Polsce. Weryfikacja nastąpi dopiero wówczas, gdy zalecenia teoretyczne sprawdzą

się w konfrontacji z realiami zarówno w dużych ośrodkach jak i w mniejszych. Sformułowane standardy powinno udać się wdrożyć zarówno w Muzeach Narodowych (dysponujących większymi zasobami ludzkimi, rozbudowaną strukturą organizacyjną, a przede wszystkim stosunkowo wysokimi funduszami), jak i w znacznie skromniej wyposażonych w te środki muzeach regionalnych. Nie sposób stwierdzić w toku rozważań teoretycznych czy poziom dostępu do nowych informacji oraz doświadczenie kadry muzealnej jest porównywalne we wszystkich placówkach w Polsce? Dlatego uważamy, że przedstawiane opracowania są pierwszym krokiem na drodze do profesjonalizacji procesu digitalizacji muzealiów, kolejnym powinna być strategia wprowadzenia tych założeń w życie. Jak już zdefiniowano w zeszłorocznym opracowaniu, określającym podstawowe problemy związane z planowaniem procesu digitalizacji, w skali jednostkowej niezbędne jest sformułowanie celów, które mają zostać osiągnięte, czasu ich realizacji, struktury zarządzającej związanej z tym działaniami i procesami, konieczny jest też wybór zespołu oraz przydział konkretnych zadań poszczególnym osobom.

Reasumując: konieczne jest prowadzenie profesjonalnego zarządzania procesem digitalizacji muzealiów w Polsce, a więc wypełnienie czterech podstawowych funkcji zarządzania: planowania, organizowania, kierowania i kontrolowania. Dla realizacji tego założenia niezbędna jest szczegółowa wiedza dotycząca poszczególnych instytucji muzealnych w zakresie: zarządzania strategicznego instytucją, stanu dokumentacji tradycyjnej, zaawansowania digitalizacji, świadomości celu tego procesu, planowanego wykorzystania metadanych, wykorzystywanych standardów itp. Z inicjatywy trzech podmiotów: Narodowego Instytutu Muzealnictwa i Ochrony Zbiorów, Instytutu Sztuki Polskiej Akademii Nauk oraz Instytutu Kultury Uniwersytetu Jagiellońskiego podjęto przeprowadzenie badania, które powinno dać odpowiedzi na wymienione pytania.

Dla zapewnienia jak najbardziej efektywnego postępu we wdrażaniu wielowymiarowego procesu digitalizacji, zarówno w skali pojedynczych podmiotów, jak i w ujęciu ogólnokrajowym, pomocne byłyby działania w kilku zakresach:

1. Identyfikacja potrzeb, i pomoc w wyposażaniu (doposażaniu) pracowni digitalizacyjnych oraz w rozbudowie w muzeach niezbędnej infrastruktury informatycznej.
2. Organizacja szkoleń dla osób pracujących w pracowniach digitalizacji.
3. Opracowanie systemowego harmonogramu działań digitalizacyjnych w każdej placówce muzealnej. W ramach tego harmonogramu uwzględnienie jasno określonych kryteriów ilościowych i jakościowych procesu cyfryzacji. Posiadanie zwartej strategii w tym zakresie powinno warunkować dotacje celowe ze środków publicznych.
4. Wypracowanie zasad wykorzystywania i udostępniania wizerunków cyfrowych muzealiów i dotyczących ich metadanych.
5. Opracowanie koncepcji centralnego repozytorium cyfrowego, lub współpracującego z siecią regionalnych.

6. Opracowanie schematu metadanych umożliwiającego wymianę informacji oraz ujednoczenie terminologii stosowanej w muzealnych bazach danych – wdrożenie tezaursów.

Do realizacji opisanych powyżej działań niewątpliwie konieczne są odpowiednie nakłady finansowe. Jednak poprzestanie na opracowaniach teoretycznych byłoby zatrzymaniem się w pół kroku, a efekty cyfryzacji dziedzictwa kulturowego pozostającego w zasobach muzealnych, w skali kraju niewymierne i niereprezentatywne.

1. Udostępnianie informacji o zbiorach w kontekście ewidencji muzealiów



1.1. Udostępnienie informacji o zbiorach

Muzea, gromadząc zbiory, są zobowiązane nie tylko do ich fizycznego udostępniania, ale także upowszechniania informacji o nich i okolicznościach, w jakich powstały. Obowiązek ten wynika zarówno z przepisów prawa¹, jak i oczekiwań społecznych. Zgodnie ze swoją misją muzea powinny, podążając za rozwojem wiedzy w dziedzinach dotyczących posiadanych zbiorów, opracowywać je naukowo i dzielić się informacjami o posiadanych muzealiach w możliwie pełny sposób. Tradycyjnie muzealia udostępniane są na wystawach stałych i czasowych. Na tego rodzaju ekspozycjach obiektom towarzyszy zazwyczaj krótka informacja, a całej prezentacji, występujący w różnych formach, opis kontekstowy tematu wystawy. Jednak w wielu muzeach na świecie nie udaje się pokazać w opisany sposób więcej niż 5% posiadanych zbiorów. Chcąc zaspokoić potrzeby poznawcze osób zainteresowanych zbiorami, muzealnicy tworzą książkowe katalogi, omawiające w sposób szczegółowy znajdujące się w muzeach eksponaty. Katalogi takie powinny towarzyszyć zarówno prezentacjom na wystawach stałych, jak i czasowych oraz docelowo obejmować cały zbiór muzealiów². Katalog zbiorów był do niedawna jedyną formą udostępniania informacji o całym zbiorze posiadanym przez muzeum³. Katalogi są postrzegane jako źródło wiedzy o kulturze wizualnej minionych czasów⁴ i rozstrzygają o jakości instytucji muzealnej⁵. Obecnie możliwe jest przekazywanie informacji o muzealiach przy użyciu nowych mediów: takich jak baza danych i Internet. Nowoczesne udostępnianie odbywa się za pomocą katalogów online. Są one generowane z systemów komputerowych przeznaczonych do ewidencjonowania i zarządzania zbiorami muzealnymi w poszczególnych instytucjach. Katalogi budowane dzięki systemom komputerowym umożliwiają wszechstronny, atrakcyjny przekaz informacji o kolekcjach muzealnych, ułatwiając wprowadzenie ich do światowego obiegu wiedzy.

¹ Ustawa o zmianie ustawy o muzeach z dnia 29 czerwca 2007 r., Dz. U. z dnia 31.07.2007, Nr 136 poz. 956

² Niestety obowiązek ten nie jest realizowany w kraju w sposób zadawalający. Temat ten został poruszony w artykule dr Bożeny Steinborn, *Katalogi, których nam brak*, *Muzealnictwo*, t. 49, 2008, s. 389-396.

³ Zazwyczaj noty katalogowe zawierają podstawowy opis muzealiów, bibliografię, i dociekania autora na tematy związane z przedmiotem. Dla różnych dziedzin są to różne rozważania. Najbardziej cenione są katalogi rozumowane, najmniej, choć i takich brakuje, będące spisem posiadanych artefaktów, rozbudowanym o szczegółowe dane.

⁴ B. Steinborn, op.cit. s. 389.

⁵ G.U. Grossmann, we wstępie do: A. Tacke: *Die Gemälde des 17. Jahrhunderts*. 1995, za B. Steinborn op.cit. przyp.22, s. 396.

1.2. Opis obiektu muzealnego a jego ewidencja

Na opis obiektu muzealnego składa się zespół jego cech fizycznych ułatwiających rozpoznanie, czyli jego identyfikację oraz ewidencjonowanie. Dla celów ewidencyjnych muzealnicy obligatoryjnie powinni stosować zasady opisane w Rozporządzeniu Ministra Kultury z dnia 30 sierpnia 2004 r. w sprawie zakresu, form i sposobu ewidencjonowania zabytków w muzeach⁶. Rozporządzenie wymienia obowiązkowe elementy opisu muzealiów w celu zastosowania ich w „dokumentach ewidencyjnych: karcie ewidencyjnej, inwentarzu, księdze depozytów, dokumentacji badań archeologicznych i innych badań terenowych”⁷. Są to „możliwe do ustalenia, następujące dane identyfikacyjne zabytku: określenie autorstwa lub wytwórcy, pochodzenie, wartość w dniu nabycia, czas i miejsca powstania, materiał, techniki wykonania, wymiary, ewentualnie jego waga oraz określenie cech charakterystycznych”⁸. Te minimalne informacje wymagane przy prowadzeniu ewidencji są jednakowo obowiązujące dla wszystkich rodzajów obiektów gromadzonych w muzeum: dla przedmiotów z zakresu sztuk pięknych, archeologii, etnografii, tworców przyrody, obiektów techniki; niemniej każda z tych dziedzin wymaga zaprojektowania indywidualnego opisu, a przy tym jednolicie skodyfikowanego. Opisany w Rozporządzeniu zakres informacji pozwala jedynie na identyfikację obiektu muzealnego. Opis powinien dostarczać jak najwięcej informacji o przedmiocie, przybliżając muzealium potencjalnemu badaczowi, niezależnie od tego czy będzie naukowcem, czy uczniem realizującym szkolny projekt. Wnikliwi obserwatorzy, profesjonalści, hobbisci, kolekcjonerzy oczekują szczegółowej wiedzy o przedmiocie i o kontekstach z nim związanych. Muzealnicy zatem, pracując nad powierzonym zbiorem, starają się zgromadzić jak najwięcej, często unikatowych wiadomości o kolekcji. Programy do ewidencji i zarządzania zbiorami nie tylko zapewniają możliwość gromadzenia i aktualizowania informacji⁹ o obiekcie, kolekcji i całym zasobie muzeum w sposób uporządkowany, ale również – przy użyciu odpowiedniej infrastruktury – umożliwiają ich publikację w Internecie. Przewagą systemu komputerowego nad tradycyjną księgą jest nieograniczona możliwość wprowadzania szczegółowych danych oraz przeszukiwanie zbioru według skomplikowanych zapytań¹⁰. Programy pozwalają nie tylko na tworzenie elektronicznego inwentarza, prowadzenie ewidencji, ale także umożliwiają pełne zarządzanie zbiorami. Dane wprowadzane do systemów można podzielić na te, które opisują muzealia i je dookreślają, podając powiązania, analogie, histo-

⁶ Rozporządzenie Ministra Kultury z dnia 30 sierpnia 2004 r. w sprawie zakresu, form i sposobu ewidencjonowania zabytków w muzeach, Dz.U. 2004, z dnia 16.09.2004, nr 202, poz. 2073.

⁷ Ibidem, §1p.2.

⁸ Ibidem, §3p.1.

⁹ W zależności od potrzeb instytucji i konfiguracji systemu bazodanowego.

¹⁰ L.M. Karecka, K. Tabak, *Jak to się robi w Muzeum Narodowym w Warszawie*, „Muzealnictwo”, t. 53, 2011, s.115-125.

rię oraz dane związane z ich zarządzaniem¹¹. Pierwsze wykorzystywane są do sporządzania różnorodnej dokumentacji, niezbędnej przy pracy w muzeach, w tym również not katalogowych oraz udostępniania za pomocą katalogów internetowych. Informacje odnoszące się do zarządzania dotyczą między innymi akcesji, odnotowywania przemieszczeń muzealiów, dokonywania skontrów oraz generowania okresowych statystyk i raportów. Każdy błąd w zapisie zmniejsza dostępność do informacji w wyniku nieuniknionej w tej sytuacji niezadawalającej trafności wyszukiwania przez zainteresowanych. Pamiętając o tym, że na muzealnikach ciąży odpowiedzialność za jakość przekazywanej wiedzy, do zagadnienia umieszczania informacji w bazie danych i następnie prezentowania ich w Internecie należy podejść w sposób drobiazgowo precyzyjny.

¹¹ O sposobie organizacji informacji w najpopularniejszych systemach bazodanowych można przeczytać m.in.: H. Ślówka, *Opis zabytku na podstawie systemu muzealnego MONA*, [w:] A. Seidel-Grześnińska i K. Stanicka-Brzezicka (red.), *Nowoczesne metody gromadzenia i udostępniania wiedzy o zabytkach. Cyfrowe spotkania z zabytkami 1*, Wrocław 2008, s. 94-101; O. Sieradzka-Malec, *Musnet – moduły systemu wpływające na zmianę jakości pracy muzealnika*, [w:] A. Seidel-Grześnińska i K. Stanicka-Brzezicka (red.), *Nowoczesne metody gromadzenia i udostępniania wiedzy o zabytkach. Cyfrowe spotkania z zabytkami 1*, Wrocław 2008, s. 86-93.

1.3. Trudności wynikające z braku konsekwencji w opisie obiektów

Przemyślana i konsekwentnie realizowana ewidencja zbiorów ma ogromne znaczenie w całościowym zarządzaniu kolekcjami muzealnymi, dlatego podstawowym zadaniem muzealników jest dbałość o jej poprawne prowadzenie. Mimo że muzea przykładają dużą wagę do uzupełniania wpisów w księgach inwentarzowych i kartach ewidencyjnych, to tradycyjną dokumentację ewidencyjną w polskich instytucjach cechuje duża dowolność w zakresie stosowanej terminologii i brak jednoznacznych zasad jej prowadzenia.

W tradycyjnej dokumentacji znajdujemy różnego rodzaju błędy oraz niekonsekwencje w opisach, będące m.in. skutkiem dawniejszego, niedostatecznego rozpoznania obiektu czy opracowywania zbiorów tego samego typu przez różne osoby bez uzgodnienia zasad i sposobu opisu muzealiów (może to powodować, że dwa podobne do siebie obiekty będą opisane w różny sposób). Zestawienie pojawiających się błędów bądź niejednoznacznych zapisów daje obraz skutków braku konsekwencji i wpływa na deprecjację samej czynności ewidencjonowania muzealiów. Z pozoru najłatwiejszym elementem opisu jest numer inwentarza – dla muzealium to identyfikator, numer unikatowy, który może być źródłem wielu pomyłek. W dokumentacji papierowej brak literowego identyfikatora muzeum (np. MNK) przy numerze inwentarzowym zasadniczo nie wpłynie na odczyt jej treści, ale przez lata powielany taki błąd może spowodować zatarcie proveniencji obiektu¹².

Trudności w ewidencjonowaniu zbiorów muzealnych nastręczają te zabytki, które są obiektami złożonymi z kilku elementów integralnie ze sobą powiązanych, jak np. dzbanek z przykrywką, stojący na podstawce lub narzędzia/przybory posiadające wspólną rękojeść; podobnie niełatwo jest zinwentaryzować kilkuselementowy serwis obiadowy¹³. Przykładowy zapis numeru ewidencyjnego wymienionego serwisu np. na 24 osoby może wyglądać następująco: CER 1023/1-150¹⁴, gdzie waza do zupy z tego serwisu może mieć numer CER 1023/73/a-b lub CER 1023/73/1-2, w zależności od przyjętego sposobu zapisu. Łatwo tu o pomyłkę np. przez powielenie numeru na dwóch wazach, albo pomylenie zapisów na kartach ewidencyjnych i do czasu szczegółowej kontroli, błąd może pojawić się również w publikacjach.

¹² O utracie informacji o pochodzeniu muzealiów w wyniku niekonsekwencji lub zaniechania ewidencjonowania pisał R. Olkowski, *O badaniu proveniencji muzealiów*, „Muzealnictwo”, t. 53, 2012, s. 54-85.

¹³ W tym przypadku mamy do czynienia z grupą obiektów – jednoznacznie wyznaczonym zbiorem o skończonej liczbie elementów, w którym poszczególne elementy składają się z elementów składowych, tworzących podzbiory, jak np. wspomniany dzbanek z pokrywką i podstawką.

¹⁴ Zawilości przy wprowadzeniu numerów inwentarzowych do inwentarza elektronicznego zostały opisane w: L. M. Karecka, *Zarządzanie wiedzą o zabytkach za pomocą systemu muzealnego MONA. Realizacja przepisów o ewidencjonowaniu* [w:] A. Seidel-Grzesińska i K. Stanicka-Brzezicka (red.), *Informatyka w historii sztuki. Stan i perspektywy rozwoju współczesnych metodologii Cyfrowe spotkania z zabytkami 2*, Wrocław 2009, s. 12-23.

Wydawałoby się również, że zapis autora dzieła (o ile jest znany) nie powinien sprawiać kłopotów. Tymczasem Adam Kowalski, w dokumentacji nieznormalizowanej może widnieć jako: Kowalski Adam, Kowalski A., A. Kowalski, Kowalski, oraz z inicjałem drugiego imienia jako: A.X.Kowalski, Adam X.Kowalski, Kowalski A.X. oraz z różnymi spacjami pomiędzy inicjałami (A. X. Kowalski, Adam X. Kowalski, Kowalski A. X.). Jeśli do tej listy dodamy literówki, to zapis popularnego nazwiska może być wprowadzony na kilkadziesiąt sposobów. Trzeba także zwrócić uwagę na nazwiska dwuczłonowe. I tak, posługując się podanym przykładem Kowalskiego, różnie można zapisać nazwisko artysty Alfreda Wierusza-Kowalskiego: Kowalski-Wierusz A., Wierusz-Kowalski A. itp.

Podobnych niekonsekwencji i błędów pojawia się wiele w zapisie tytułów. Znane są tzw. arkusze spisów z natury, w których wiele obrazów określono lakonicznym zapisem tytułu „Pejzaż”, „Krajobraz” „Scena rodzajowa” itp., co utrudnia identyfikację obrazu (zwłaszcza jeśli wymiary kilku *Pejzaży* w muzeum są podobne). Zdarzają się także błędy interpretacyjne w określaniu tytułów, jak np. w określaniu postaci świętych poprzez widoczne ich atrybuty na obrazie np. *Męczeństwo św. Jana Nepomucena* czy *Męczeństwo św. Liwiniusza*, patrona Gandawy. Jeśli do tak nieprecyzyjnych określeń zostanie dodany potencjalny, wcale nierzadki, „czeski błąd” w zapisie cyfr numeru, to w prosty sposób pojawia się „nowe” dzieło, a w efekcie wykazy zbiorów i listy braków stają się dokumentem o wątpliwej wartości.

Zdarza się, fatalne w skutkach, niestaranne przepisywanie z dawniejszych zapisów na różne dokumenty ewidencyjne tytułu dzieła bez oględzin tegoż przez osobę dokonującą nowego zapisu (a w razie opisywania dzieła zaginionego – bez zbadania jego dawnej dokumentacji). Częstym błędem są – wynikające także z braku rzetelnych oględzin – omyłki w określeniu podobrazia obrazów w technice olejnej: drewno myłone z blachą, płótno z papierem. Powoduje to albo „mnożenie bytów” albo niemożność odnalezienia istniejącego w zbiorach dzieła.

Wymienione przykłady niestarannego bądź niekonsekwentnego opisu obiektu w dokumentacji tradycyjnej najczęściej skutkują wytworzeniem nieprawdziwych lub nieweryfikowalnych informacji o posiadanym zbiorze (np. jego kompletności), co przekłada się na wadliwe raporty i statystyki, a w dalszej kolejności na powielanie wytworzonych błędów poprzez podawanie ich do wiadomości publicznej w różnego rodzaju publikacjach. Skutki braku konsekwencji w zakresie opisu muzealiów widoczne są w pełni w momencie zamiany tradycyjnych form ewidencjonowania na formę elektroniczną.

Zazwyczaj w mniejszym stopniu jest to zauważalne w ramach dokumentacji jednej kolekcji. Przy zestawieniu opisów z różnych kolekcji nawet w obrębie jednego Muzeum uwidaczniają się znaczące różnice w sposobie zapisu analogicznych informacji niemalże we wszystkich jednostkach informacji, począwszy od autora, a kończąc na wymiarach obiektów. Duża dowolność zapisu, charakterystyczna dla naturalnego języka opisu, która nie ma tak wielkiego znaczenia w przypadku tradycyjnych, papierowych form ewidencji, jest bardzo niekorzystna dla wyszukiwania informacji w bazodanowych systemach komputerowych, gdzie stosowanie pełno tek-

stowych opisów w ustrukturyzowanych schematach danych powoduje, że w praktyce stają się one w znacznym zakresie nieużyteczne.

Systemy do ewidencji i zarządzania zbiorami w swej podstawowej funkcji służą do opisu muzealiów i obsługi ich przemieszczeń, przeprowadzania kwerend oraz tworzenia katalogów zbiorów wydawanych drukiem i publikowanych online.

Nowoczesne programy, dzięki możliwości pełno tekstowego przeszukiwania, skutecznie filtrują zgromadzone dane i dostarczają wyniki obejmujące konkretnie poszukiwane informacje na zadany temat. W praktyce jednak, jakość uzyskiwanych w ten sposób informacji jest niezadowalająca. Poprawność wyników uzyskiwanych poprzez przeszukiwanie baz danych nie jest zależna jedynie od zastosowanych algorytmów programu, ale przede wszystkim warunkowana jest jakością materiału wyjściowego, inaczej ujmując: poziomem uporządkowania danych wprowadzanych do bazy. Pojawiające się błędy w wynikach wyszukiwania powstają w wyniku niekonsekwencji w opisie obiektów.

Powodu tej sytuacji bardzo często upatruje się w niekonsekwentnym posługiwaniu się terminologią przez osoby opracowujące zbiory. Jest to jednak zbyt uproszczenie. Nie wszystkie nieprawidłowości wynikają z niestaranności czy przyzwyczajień. Wiele problemów warunkowanych jest specyfiką języka polskiego, część uwarunkowaniami historycznymi, ale przede wszystkim brakiem standardów opisu. Poniżej zamieszczono wybrane przykłady, w celu zasygnalizowania ostatniego ze wskazanych problemów.

Brak usystematyzowanego nazewnictwa obiektów powoduje, że poszczególni opiekunowie zbiorów (w ramach jednej instytucji) nazywają obiekty tego samego typu bliskoznacznymi określeniami (np. wieniec dożynkowy, krutka, pępek, równianka, plon). Stosowanie różnych nazw powoduje, że potencjalne wyszukiwanie nie obejmie w pełni zbiorów znajdujących się w muzeum, pomimo że wprowadzono o nich informacje do bazy danych.

Brak uporządkowanego nazewnictwa materiału i techniki jako osobnych kategorii powoduje problemy już nie tylko z wyszukiwaniem, ale wątpliwości co do samej formy zapisu. W przypadku kategorii „materiał” częstą praktyką jest podawanie formy/nazwy przedmiotu w miejscu, gdzie powinna znaleźć się informacja o tworzywie, np.: „deska”, „klepka”, „belka”, „boazeria” – zamiast „drewno”. Kolejnym wyzwaniem jest unifikacja języka opisu materiału i techniki, stosowanego w przypadku sztuki, z językiem innych dziedzin, w tym np. techniki. Jeśli wyszukiwanie informacji o zbiorach ma się odbywać na zasobach całego muzeum, ta nawet najbardziej podstawowa unifikacja jest niezbędna. W innym przypadku wyniki wyszukiwania będą obarczone znaczącym błędem (np. „olej” – „farba olejna”; „gęsta ciecz pochodzenia roślinnego lub zwierzęcego” – „gęsta ciecz będąca produktem destylacji”).

Brak przyjętej formy zapisu wartości księgowych (wartość, wartość w dniu nabycia, wartość po przeszacowaniu, koszt zakupu etc.) skutkuje błędnymi sumami wartości.

Należy podkreślić, że jeśli w instytucji opis wspierany jest uporządkowaną terminologią (słownikami, indeksami, tezaurusami), możliwa jest migracja danych do innych baz, niezależnie od stosowanego standardu opisu.

Podsumowując, brak konsekwencji w zakresie prowadzenia ewidencji zbiorów, w pierwszej kolejności niesie ze sobą nieścisłości w zakresie opisu muzealiów, prowadząc następnie do wielu problemów z codziennym zarządzaniem kolekcją, ale przede wszystkim utrudnia późniejsze wykorzystanie wytworzonych zasobów, na przykład za pośrednictwem Internetu. Opracowywanie dokumentacji bez szczegółowo i jednoznacznie sprecyzowanych zasad uniemożliwia uzyskanie prawidłowych wyników przeszukiwana, co w konsekwencji wyklucza publikowanie informacji o obiektach bez wcześniejszych żmudnych korekt. W ten sposób proces dokumentacji zbiorów jest ograniczony jedynie do konieczności formalnego wywiązania się instytucji z przepisów dotyczących ewidencji, a powstały materiał nie posłuży realizowaniu jednego z podstawowych celów funkcjonowania muzeów – upowszechniania zbiorów. Z wymienionych powyżej powodów stosowanie wytycznych, tak w zakresie samej struktury zapisu metadanych, jak i sposobu zapisywania treści, ma znaczenie fundamentalne. W celu zapewnienia konsekwencji zapisu informacji niezbędne jest opracowanie instrukcji wprowadzania danych, w której zdefiniowane zostaną podstawowe parametry poszczególnych jednostek informacji (metadanych) oraz pełna kontrola słownictwa wsparta stosowaniem tezaurusów, słowników, indeksów oraz rozwijanych list kontrolnych. Pracownicy opracowujący zbiory powinni być zobligowani do stosowania zasad wskazanych w instrukcji.

2. Metadane

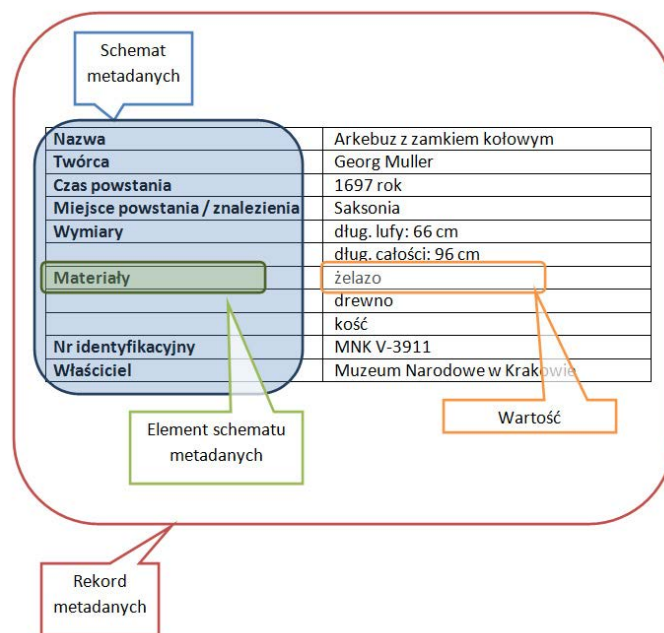


2.1. Schematy i standardy metadanych w muzeach

Przetworzenie opisów muzealiów na postać cyfrową, czyli ich digitalizacja, polegające na przeniesieniu treści dotychczas używanych w muzeach dokumentów analogowych (inwentarzy, kart katalogu naukowego, kart ewidencyjnych) do komputerowych baz danych, wymagało stworzenia odpowiednich do tego celu schematów (formatek, szablonów). Ich wspólnym rdzeniem były z reguły dotychczasowe normatywy, takie jak *Instrukcja do Rozporządzenia Ministra Kultury i Sztuki z dnia 18 kwietnia 1964 r. w sprawie prowadzenia inwentarza muzealiów*¹⁵. Potrzeba poszerzenia podstawowego opisu o dodatkowe informacje spowodowała, że schematy te zaczęły się rozrastać, czasem w dość spontaniczny sposób. Tego rodzaju indywidualne schematy sprawdzały się bardzo dobrze w ramach pojedynczych instytucji. Powstanie internetowych katalogów zbiorów i otwierająca się w ślad za tym możliwość zbierania (agregowania) danych z różnych muzeów, w celu wspólnej prezentacji online swoich kolekcji, rodzą konieczność unifikacji zarówno schematów metadanych, jak i zasad opisywania muzealiów. Co więcej, owa unifikacja nie może zostać przeprowadzona wyłącznie na poziomie krajowym, lecz wymaga dostosowania się do międzynarodowych schematów i standardów opisu (lub wprost do ich przyjęcia). Tylko wówczas zbiory polskich muzeów będą mogły z powodzeniem być prezentowane na wspólnych platformach internetowych wraz ze zbiorami muzeów zagranicznych, co pozwoli znacząco poszerzyć krąg odbiorców polskiego dziedzictwa narodowego.

Opis obiektów oparty jest na schemacie metadanych, który stanowi zestaw elementów definiujący zakres danych. Podstawową składową schematu jest element schematu metadanych, w środowisku informatycznym często zwany polem. Do konkretnego elementu schematu metadanych przypisana jest jego wartość, czyli treść wprowadzana do pola. Wypełnione dla danego obiektu pola stanowią rekord metadanych – zbiór elementów metadanych wraz z przypisanymi im wartościami w określonym kontekście. Znormalizowane schematy metadanych określane są mianem standardu.

¹⁵ *Instrukcja do Rozporządzenia Ministra Kultury i Sztuki z dnia 18 kwietnia 1964 r. w sprawie prowadzenia inwentarza muzealiów*, Dz. U. z dnia 16.09.2004, Nr 202, poz. 2073



Ilustracja 1. Przykładowa struktura schematu metadanych.
Opracowanie: M. Werla

Istnieje zaledwie kilka najważniejszych międzynarodowych schematów metadanych, stosowanych do opisu muzealiów, które warto uwzględnić w procesie unifikacji.

2.1.1. Dublin Core Metadata Element Set (DCMES) oraz DCMI Metadata Terms (DC Terms)

Dublin Core Metadata Element Set (DCMES)¹⁶ to piętnastoelementowy schemat metadanych opracowany z myślą o opisywaniu zasobów cyfrowych pod kątem udostępniania ich online. Jak nazwa wskazuje, schemat ten jest podstawowym zestawem elementów, wybranym z szerszego zbioru pól utrzymywanego przez Dublin Core Metadata Initiative (DCMI), znanego pod nazwą DCMI Metadata Terms¹⁷. W skład DCMES wchodzi bardzo ogólnie określone pola, takie jak: „Tytuł”, „Twórca”, „Opis”, „Język” czy „Format”. Definicja każdego z pól jest krótka i pozostawia wiele miejsca na interpretację oraz doprecyzowanie na potrzeby konkretnego zastosowania. Dla przykładu pole „Data” zdefiniowane jest jako „Moment lub okres w czasie, powiązany z wydarzeniem w cyklu życia zasobu”. Nie jest również wymagane obligatoryjne uzupełnienie wszystkich pól.

DCMI Metadata Terms jest schematem bardziej rozbudowanym, zawierającym wiele pól będących doprecyzowaniem czy też uściśleniem pól DCMES. Dobrym przykładem może być tutaj

¹⁶ *Dublin Core Metadata Element Set, Version 1.1.* Dok. elektroniczny: <http://dublincore.org/documents/dces/> [Dostęp: 31 X 2012].

¹⁷ *DCMI Metadata Terms.* Dok. elektroniczny: <http://dublincore.org/documents/dcmi-terms/> [Dostęp: 31 X 2012].

8 bardziej precyzyjnych znaczeniowo pól dotyczących daty, przeznaczonych na informacje takie jak data utworzenia, udostępnienia, modyfikacji czy obowiązywania praw autorskich.

Ogólne definicje elementów schematów DCMES i DC Terms oraz stosunkowo duża ich popularność są przyczyną częstego wyboru tych schematów jako podstawy funkcjonowania systemów gromadzących metadane z wielu różnych źródeł. Interoperacyjność pomiędzy różnymi bazami danych, osiągnięta poprzez zastosowanie wspólnego ogólnego schematu metadanych, może jednak zaowocować również niepożądanymi efektami. Dzieje się tak wtedy, gdy interpretacja pól wspólnego schematu jest odmienna dla poszczególnych baz danych (np. ogólnie zdefiniowane pole „Data” jest wypełniane datami różnego rodzaju). Jest to bowiem dopiero pierwszy z czterech poziomów interoperacyjności zdefiniowanych w kontekście schematu DCMES¹⁸.

Dla uniknięcia tego typu problemów w przypadku stosowania DCMES czy DC Terms, zaleca się opracowanie precyzyjnego schematu zwanego profilem aplikacji¹⁹. Schemat taki powinien zawierać listę elementów DC Terms, które będą stosowane w danej bazie, ale może zawierać też informacje o specyficznych polach dodatkowych, nie pochodzących z DC Terms. Dla każdego z pól schematu powinna być podana jednoznaczna definicja. Warto również określić czy dany element jest obligatoryjny, ile może mieć wartości oraz podać wymagania i wskazówki dotyczące wartości tego atrybutu (np. sposób zapisu wartości w polu „Data”, czy też zamkniętą listę dozwolonych wartości dla pola „Język”).

Przykładem tego typu profilu aplikacyjnego może być schemat Europeana Semantic Elements (ESE) opracowany na potrzeby portalu Europeana²⁰, agregującego dane o dostępnych online zbiorach europejskich instytucji kultury. Schemat ESE składa się z wszystkich elementów DCMES, kilkunastu dodatkowych elementów DC Terms oraz kolejnych kilku elementów specyficznych dla portalu Europeana, niezbędnych do zrealizowania poszczególnych funkcji, które portal ten oferuje swoim użytkownikom.

Dokumentacja schematu ESE składa się z:

- 1) specyfikacji poszczególnych pól schematu²¹;
- 2) wytycznych mapowania i normalizacji danych z lokalnych baz do ESE²²;
- 3) specyfikacji zapisu danych zgodnych ze schematem ESE w języku XML²³.

¹⁸ *Interoperability Levels for Dublin Core Metadata*. Dok. elektroniczny: <http://dublincore.org/documents/interoperability-levels/> [Dostęp: 31 X 2012].

¹⁹ *Guidelines for Dublin Core Application Profiles*. Dok. elektroniczny: <http://dublincore.org/documents/profile-guidelines/> [Dostęp: 31 X 2012].

²⁰ *Europeana*. Dok. elektroniczny: <http://europeana.eu> [Dostęp: 31 X 2012].

²¹ *Europeana Semantic Elements Specification, Version 3.4.1, 2012-24-02*. Dok. elektroniczny: <http://pro.europeana.eu/documents/900548/dc80802e-6efb-4127-a98e-c27c95396d57> [Dostęp: 31 X 2012].

²² *Metadata Mapping & Normalisation Guidelines for the Europeana Semantic Elements, Version 2.1.1, 2012-02-24*. Dok. elektroniczny: <http://pro.europeana.eu/documents/900548/683de455-27a7-4dd6-81c7-928120957dfa> [Dostęp: 31 X 2012].

²³ *Europeana Semantic Elements v 3.4 schematu XML*. Dok. elektroniczny: <http://www.europeana.eu/schemas/ese/ESE-V3.4.xsd> [Dostęp: 31 X 2012].

Pod koniec 2012 r. schemat ESE został w portalu Europeana zastąpiony semantycznym schematem Europeana Data Model (EDM)²⁴, który zawiera wszystkie elementy ESE. Osoby zainteresowane dokładniejszymi informacjami na temat schematów metadanych stosowanych przez Europeana, mogą zapoznać się z bezpłatnym kursem e-learningowym „Współpraca z Europeaną”, dostępnym w ramach Federacji Bibliotek Cyfrowych²⁵.

UWAGA: Schemat DCMES jest uznany jako norma ISO 15836:2009²⁶. Ostatnia oficjalna polska wersja tej normy to PN-ISO 15836:2006, jest ona jednak przestarzała i zawiera błędy rzeczowe, wynikające z niepoprawnego tłumaczenia oryginalnego dokumentu. Prace nad nową wersją polskiej normy są w toku, jednak do czasu ich ukończenia lepiej korzystać z tłumaczenia specyfikacji DC Terms opracowanego w 2009 r. przez Bibliotekę Narodową²⁷.

2.1.2. Categories for Description of Works of Art (CDWA)

CDWA²⁸ jest standardem opisu dzieł sztuki, który od lat 90. XX w. powstawał pod egidą Getty Research Institute w USA. Jest on zasadniczo zaliczany do tego typu schematów metadanych, które nazywa się strukturalnymi, a więc takimi, które wyliczają zestawy elementów, kategorii bądź pól opisu, mogących tworzyć rekord lub innego rodzaju obiekt informacyjny, określają nazwy, długości, powtarzalność i inne cechy pól oraz ich wzajemne relacje. Zestaw, który oferuje CDWA jest bardzo obszerny, gdyż liczy ponad 500 różnych elementów (kategorii danych), jakimi da się opisywać dzieła sztuki. Naturalnie, w muzealnych systemach bazodanych nie używa się w praktyce rekordów składających się z aż tylu danych, lecz dokonywany jest wybór tych tylko, które są niezbędne do prawidłowego i pełnego zinventaryzowania bądź opisanie tych dzieł, które planuje się umieścić w bazie danych. Twórcy standardu CDWA uznali jednakże, że istnieje 36 podstawowych elementów schematu (oznaczonych określeniem *core*, tj. rdzeniowe), bez których opis jakiegokolwiek dzieła z zakresu sztuk pięknych obyć się nie może. Wszystkie kategorie schematu CDWA ułożono w formie hierarchicznego drzewa, co znacząco ułatwia poruszanie się po tak obszernym dokumencie.

CDWA, prócz tego że jest schematem strukturalnym, może być zarazem rozpatrywany jako standard treściowy metadanych, to jest taki, który zawiera reguły i wytyczne służące do wypełnienia zadanego schematu treścią ukształtowaną poprawnie i wyczerpująco. Znajdujemy tutaj szeroką definicję danego elementu schematu, omówienie, punkty dyskusyjne, rekomendacje

²⁴ *Europeana Data Model (EDM) Documentation*. Dok. elektroniczny: <http://pro.europeana.eu/web/guest/edm-documentation> [Dostęp: 31 X 2012].

²⁵ *Federacja Bibliotek Cyfrowych*. Dok. elektroniczny: <http://fbc.pionier.net.pl/elearning/pl/> [Dostęp: 31 X 2012].

²⁶ *ISO/R 843:1968. International system for the transliteration of Greek characters into Latin characters*. Dok. elektroniczny: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=5214 [Dostęp: 31 X 2012].

²⁷ *Terminy Metadanych DCMI*. Dok. elektroniczny: <http://www.bn.org.pl/download/document/1261049421.pdf> [Dostęp: 31 X 2012].

²⁸ *Categories for the Description of Works of Art*. Dok. elektroniczny: http://www.getty.edu/research/publications/electronic_publications/cdwa/ [Dostęp: 31 X 2012].

i wskazówki, a wreszcie przykłady prawidłowego opisu. Taki sposób zredagowania standardu pozwala na daleko idące ujednoczenie wszystkich danych, które zostaną wytworzone z jego użyciem. Nie ma tu miejsca na dowolne interpretowanie przez osoby katalogujące zakresu danych, które mogłyby lub powinny zostać wpisane do odpowiedniego pola. Dzięki temu dane wytworzone w rozmaitych lokalnych systemach, lecz z użyciem schematu CDWA powinny być spójne, a tym samym łatwe do ewentualnej przyszłej agregacji i do łącznego prezentowania na wspólnych portalach internetowych²⁹.

```
<?xml version="1.0"?>
<obraz>
  <autor>Boznańska, Olga</autor>
  <datowanie>1894</datowanie>
</obraz>
```

Ilustracja 2. Przykład prostego dokumentu XML; informacja (pogrubioną czcionką) ujęta jest pomiędzy znacznikami (w nawiasach trójkątnych). Opracowanie: T. Zaucha

Na podstawie CDWA został przygotowany schemat XML, nazwany CDWA Lite³⁰ (ang. *lite* – uproszczony). XML (Extensible Markup Language), który użyto do wytworzenia CDWA Lite jest uniwersalnym językiem pozwalającym na zapis i przesył informacji. Dane są ujmowane w tzw. znaczniki (tzw. tagi, rozpoznać je można dzięki trójkątnym nawiasom) i można je układać w hierarchiczne struktury.

Zadaniem CDWA Lite było umożliwienie eksportu danych wytworzonych w oparciu o systematykę CDWA, w sposób ustandaryzowany, z użyciem powszechnie stosowanego protokołu OAI-PMH (Open Archive Initiative – Protocol for Metadata Harvesting)³¹. Ten powszechnie używany protokół komunikacyjny, czyli zespół reguł i zasad umożliwiających połączenie pomiędzy komputerami, służy wymianie metadanych wyrażonych w dowolnym schemacie zapisanym w języku XML, w tym również CDWA Lite. Co więcej, protokół ten wspiera wymianę metadanych w kilku schematach równocześnie, przy czym za niezbędne minimum uznany jest schemat Dublin Core Metadata Element Set. CDWA Lite użyty został następnie do stworzenia niemieckiego schematu Museumdat³², lecz ostatecznie rozwój obydwu schematów został zarzucony wskutek przyjęcia schematu LIDO (Lightweight Information Describing Object)³³.

²⁹ Przykład adaptacji standardu CDWA zob. T. Zaucha, *Wzorcowa postać metadanych obiektu muzealnego na przykładzie Categories for Description of Works of Art (CDWA)* [w:] E. Bunsch et al., *Zalecenia dotyczące planowania i realizacji projektów digitalizacyjnych w muzealnictwie*; red. D. Galas, dokumenty uzupełniające: s. 1–15, Dok. elektroniczny: <http://www.nimoz.pl/pobierz/305.html> [Dostęp: 31 X 2012].

³⁰ *CDWA Lite*: Dok. elektroniczny: http://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.html [Dostęp: 31 X 2012].

³¹ *Open Archives Initiative Protocol for Metadata Harvesting*. Dok. elektroniczny: <http://www.openarchives.org/pmh/> [Dostęp: 31 X 2012].

³² *XML Schema zur Bereitstellung von Kerndaten in museumsübergreifenden Beständen*. Dok. elektroniczny: <http://www.museumdat.org> [Dostęp: 31 X 2012].

³³ *Lightweight Information Describing Object*. Dok. elektroniczny: <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/> [Dostęp: 31 X 2012].

2.1.3. Spectrum

Standard Spectrum (The UK Museum Collections Management Standard)³⁴ służy do dokumentacji i zarządzania informacjami o kolekcjach. Jest to schemat otwarty, co oznacza, że jego użytkownicy mogą kształtować go w dowolny sposób, rozwijać i skracać na własne potrzeby. W celu zapewnienia zgodności struktury danych gromadzonych przez brytyjskie muzea, a co za tym idzie normalizacji obejmującej pewien obowiązkowy schemat metadanych, pierwotnie Rada Muzeów, Bibliotek i Archiwów (Museum, Libraries and Archives Council – MLA) stanowiąca obecnie część Arts Council England³⁵, określiła system akredytacji (Museum Accreditation Scheme)³⁶, obejmujący 8, z 21 procedur schematu jako obowiązkowe minimum.

Ten schemat metadanych został opublikowany po raz pierwszy w 1994 r. Zarządzany przez powołane w 1977 r. Stowarzyszenie ds. Dokumentacji Muzealnej (Museum Documentation Association – MDA), przeorganizowane w 2008 r. w organizację charytatywną o nazwie Powiernictwo Kolekcji (Collections Trust), która jest wspierana przez Arts Council England oraz organizację Muzea i Galerie Szkocji (Museums Galleries Scotland). Schemat stale poddawany jest aktualizacjom i poprawkom, co czyni go „standardem żywym”; ma to swoje wady i zalety. Aktualna wersja to Spectrum 4.0³⁷. Według deklaracji Collection Trust ze standardu korzysta ponad 7000 muzeów w 40 krajach³⁸. Warto zwrócić uwagę, że część druga SPECTRUM 3.2 nie uległa zmianie w zaktualizowanej wersji 4.0. Ponadto wersja 4.0 otrzymała bardziej przejrzysty układ treści i została wzbogacona o diagramy procedur stosowanych w muzealnictwie, obejmujące pełny cykl życia obiektu od chwili przyjęcia do muzeum.

Pierwszą znaczącą różnicą w stosunku do innych schematów jest to, że elementy metadanych (pola nie są ponumerowane, jak np. w omawianym już CDWA. Schemat składa się z dwóch zasadniczych elementów: procedur i grup informacji. Ponadto podręcznik schematu zawiera podstawowe informacje na temat tego, jak używać standardu oraz jak pozyskiwać dane i je zabezpieczać. Procedury umożliwiają dokumentację procesów związanych z zarządzaniem obiektami (począwszy od przyjęcia do muzeum, przez włączenie do zbiorów, wypożyczenia, kontrole, po informację o wykreśleniu ze zbiorów czy zwrocie). Każda procedura wykorzystuje pewien określony zestaw elementów schematu. Grupy informacji stanowią zestawy związa-

³⁴ *SPECTRUM, The UK Museum Documentation Standard, obecnie dostępna wersja 4.0.* Dok. elektroniczny: <http://www.collectionslink.org.uk/spectrum-standard?phpMyAdmin=OYNyINPdn3sQmoXugKH1gcCLSW0> [Dostęp: 31 X 2012].

³⁵ *Arts Council England.* Dok. elektroniczny: <http://www.artscouncil.org.uk/> [Dostęp: 31 X 2012].

³⁶ Opublikowany w październiku 2011 roku schemat akredytacji dla muzeów i galerii w Wielkiej Brytanii - *Accreditation Scheme for Museums and Galleries in the United Kingdom: Accreditation Standard, 2011.* Dok. elektroniczny: http://www.artscouncil.org.uk/media/uploads/pdf/accreditation_standard_english_web.pdf [Dostęp: 31 X 2012]. Starsza wersja dokumentu Museums, Libraries and Archives Council, Accreditation Standard, The Accreditation scheme for Museums in the United Kingdom, version 01, 2004.

³⁷ *Możliwość zakupu wersji publikowanej drukiem za 43.99 funty brytyjskie. SPECTRUM 4.0: The UK Museum Collections Management Standard.* Dok. elektroniczny: <http://www.collectionslink.org.uk/ct-retail/product/view/2/12> [Dostęp: 31 X 2012].

³⁸ Ibidem.

nych ze sobą elementów schematu, ułatwiające strukturyzację danych i nawigację po schemacie. Elementy schematu są podstawowym ogniwem tego standardu i pozwalają dokumentować różnego rodzaju informacje o obiektach i kolekcjach. Określenie elementu schematu obejmuje zarówno nazwę jak i dane z nią związane. Każdy element schematu jest opisany przy pomocy: definicji, sposobu zapisu, przykładów, zastosowania i grupy informacji, do której jest przyporządkowany. W schemacie istnieje 45 grup informacji, z tego jedna służy do rozbudowywania 19 innych, odnoszących się do procedur oraz może być wykorzystana do budowania nowych procedur, nie ujętych w schemacie. 17 grup informacji obejmuje dokumentację danych o obiekcie, 3 – o rekordzie danych, 1 – o odwołaniach (obejmujących np. bibliografię) oraz 4 grupy informacji dotyczące osób, miejsc, społeczności i lokalizacji obiektu. Wspomniane ostatnie 4 grupy znajdują zastosowanie jako rozwinięcie elementów schematu w kilkunastu innych grupach informacji (np. w przypadku elementu schematu „Autor”, powiązana z nią grupa informacji „Osoba”, dostarcza szereg informacji o autorze – daty życia, imiona i przydomki, notę biograficzną, powiązania z innymi obiektami w muzeum, powiązania z miejscami, innymi osobami etc.).

Procedury wskazane w schemacie mają ograniczone zastosowanie w polskich muzeach i należy uważnie przemyśleć ewentualne dostosowanie wewnętrznej organizacji pracy przyjętej w instytucji do rozwiązań brytyjskich. Z pewnością wdrożenie procedur może zapewnić sprawne zarządzanie informacją w muzeum, jednakże musi być realizowane konsekwentnie. Należy przy tym zaznaczyć, że zmiana wypracowanych przez lata, sprawnie działających instrukcji kancelaryjnych i zasad realizacji poszczególnych działań w muzeach na uniwersalne procedury Spectrum, nie musi być wcale rozwiązaniem optymalnym. Wszystko zależy od specyfiki instytucji. Niezależnie od ostatecznej decyzji, o jednym warto pamiętać – wprowadzenie procedur, które obowiązują jedynie na papierze, mija się z celem i od razu powinno zostać odrzucone jako kierunek działania.

Opisywany schemat pozwala dokumentować dane o zbiorach archeologicznych i przyrodniczych jedynie w ograniczonym stopniu. Dotyczy to również procedur obowiązujących przy okazji konserwacji. Dlatego należy liczyć się z koniecznością poszerzenia schematu o niezbędne grupy i jednostki informacji. Niewątpliwą zaletą schematu jest możliwość tworzenia rejestrów osób, podmiotów, społeczności i miejsc, które mogą być powiązane relacjami między sobą (miejsce i podmiot związane z osobą) oraz z obiektami i wydarzeniami. Dzięki temu możliwe jest tworzenie rejestrów (np. osób), współdzielonych przez wszystkich użytkowników bazy danych w instytucji (raz wprowadzony do bazy danych rekord osoby, jest dostępny dla wszystkich użytkowników), co pozwala przypisać jednej osobie wiele ról (w związku z różnymi obiektami i wydarzeniami), takich jak np. producent, projektant, organizator, sprzedawca, darczyńca, oferent etc. Należy zaznaczyć, w kontekście wymiany danych, że tylko część elementów schematu o zbiorach standardu Spectrum została zmapowana na użytek schematu wymiany danych znanego jako Lekka Informacja Opisująca Obiekt (Lightweight Information Describing Object – „LIDO”).

2.1.4. CIDOC CRM

CIDOC CRM³⁹ jest sformalizowanym schematem pojęciowym: zbiorem pojęć z zakresu dóbr kultury i zbiorem relacji charakterystycznych dla tej dziedziny wiedzy, które mogą zachodzić pomiędzy owymi pojęciami. Schemat taki nazywany jest w teorii informacji ontologią.

Został on przygotowany przez International Committee for Documentation (CIDOC; skrót utworzono od francuskiej nazwy komitetu: Comité international pour la documentation), będący jednym z komitetów Międzynarodowej Rady Muzeów (ICOM). Schemat ten tworzony był od roku 1996, początkowo przez jedną z grup roboczych komitetu, zajmującą się standardami dokumentacji⁴⁰, a później przez wyodrębnioną grupę zajmującą się jedynie tym standardem. Ontologia zaczerpnęła swą nazwę od skrótu komitetu, natomiast CRM jest skrótem od Conceptual Reference Model. Pojęcia tej ontologii mogą być przedstawione jako zestaw twierdzeń, tworzących model rzeczywistości, dający się interpretować przez maszyny liczące. Aktualna wersja CIDOC CRM (5.0.4) została wydana w listopadzie 2011 r. W 2006 ontologia ta została formalnie zaakceptowana przez International Organization for Standardization jako norma międzynarodowa ISO 21127:2006⁴¹.

Różnica jaka istnieje pomiędzy CIDOC CRM a dotychczas przedstawionymi schematami, takimi jak np. CDWA czy Spectrum, jest zmianą zasadniczą, która uwzględnia wymogi sieci semantycznej, sformułowane przez World Wide Web Consortium (W3C)⁴². Podstawowy cel tych zmian to poszerzenie możliwości wyszukiwania informacji w sieci, natomiast środkiem jest sformułowanie i ujęcie w językowe rygory własności pojęć lub relacji łączących pojęcia w taki sposób, by mogły być one interpretowane przez komputery. Maszynowa interpretacja owych własności pozwoli na kojarzenie takich informacji, które dotąd nie funkcjonowały w jakimkolwiek powiązaniu i formułowaniu nowego rodzaju zapytań. Podstawowym elementem przedstawiającym twierdzenia o rzeczach jest – w tym schemacie – trójka składająca się z podmiotu, przedmiotu oraz łączącego je predykatu, za pomocą którego podmiot orzeka o przedmiocie: predykat przedstawia pewną charakterystyczną właściwość podmiotu, natomiast przedmiot wyraża wartość owej właściwości podmiotu.

³⁹ *CIDOC CRM*. Dok. elektroniczny: <http://www.cidoc-crm.org/> [Dostęp: 31 X 2012].

⁴⁰ Documentation Standards Working Group (DSWG).

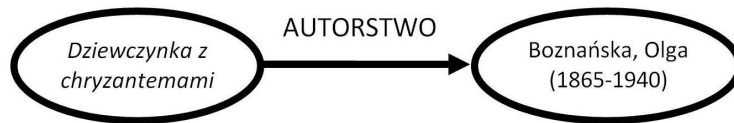
⁴¹ *ISO 21127:2006. Information and documentation - A reference ontology for the interchange of cultural heritage information*. Dok. elektroniczny: http://www.iso.org/iso/catalogue_detail?csnumber=34424 [Dostęp: 31 X 2012].

⁴² *Semantic Web*. Dok. elektroniczny: <http://www.w3.org/standards/semanticweb/> [Dostęp: 26 XI 2012].



Ilustracja 3. Trójka złożona z podmiotu, przedmiotu i predykatu.
Opracowanie: T. Zaucha.

Z użyciem tego modelu można np. wyrazić twierdzenie, że autorem obrazu jest Olga Boznańska:

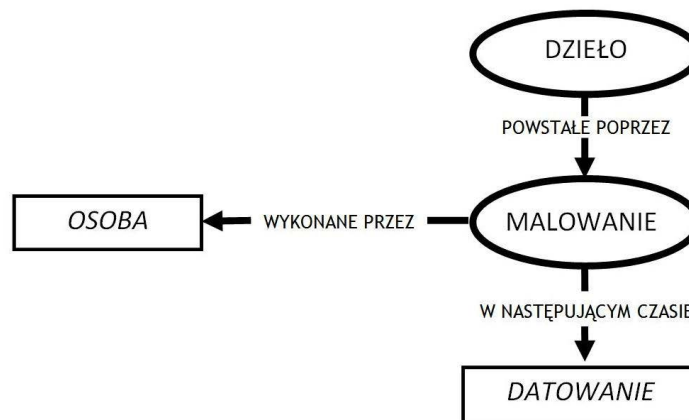


Ilustracja 4. Trójka przedstawiająca informację o atrybucji dzieła.
Opracowanie: T. Zaucha.

Taki sposób przedstawiania informacji rozbija sztywną, tabelaryczną strukturę informacji i pozwala na zastąpienie jej triadami, które mogą w dalszej kolejności łączyć się w długie łańcuchy zależności i skojarzeń, tworząc skomplikowane sieci znaczeń.

	OBRAZ
DATA POWSTANIA	DATOWANIE
ARTYSTA	OSOBA

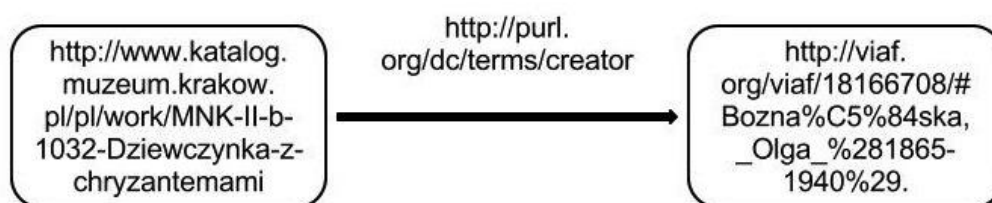
Ilustracja 5. Przykład struktury tabelarycznej pozwalającej na zapisywanie datowań i atrybucji.
Opracowanie: T. Zaucha.



Ilustracja 6. Przykład schematu graficznego służącego do zapisywania datowań i atrybucji.
Opracowanie: T. Zaucha.

Jednym z założeń tej zmiany jest rozszerzenie koncepcji dotychczas stosowanych w sieci adresów URL (Uniform Resource Locator) i wprowadzenie nowego typu unikatowych i stałych identyfikatorów URI (Uniform Resource Identifier). W tym ujęciu adresy URL są podtypem identyfikatorów URI. Zarówno URL, jak i URI mogą być używane jako jednoznaczne identyfikatory, jednak adresy URL powinny pozwalać na zlokalizowanie obiektów w sieci, podczas gdy identyfikatory URI nie muszą dawać takiej możliwości. Identyfikatorem URI można oznaczyć dowolny byt, a nie tylko to, co da się zlokalizować w Internecie. Ponadto identyfikatory URI mogą występować w roli podmiotu, predykatu i przedmiotu, jako odwołania do rzeczy lub pojęć. Te założenia są bardzo istotne, ponieważ pozwalają tworzyć przy użyciu rozmaitych, wspólnie używanych zasobów, które są jednoznaczne, a co więcej: mogą być jednoznacznie interpretowane także przez aplikacje.

Poniższy przykład wyraża w postaci schematu graficznego twierdzenie: „*Dziewczynka z chryzantemami* została namalowana przez Olę Boznańską”. W roli podmiotu występuje tu odwołanie do tego dzieła [w celach poglądowych użyto adresu URL], w roli predykatu identyfikator URI elementu Creator w Dublin Core (DCMI Metadata Terms), a w roli przedmiotu identyfikator rekordu w VIAF.



Ilustracja 7. Trójka wyrażająca informację o atrybucji dzieła z użyciem odwołań do zasobów sieciowych.
Opracowanie: T. Zaucha

Stosując zasady sieci semantycznej, można budować szczególnego rodzaju ontologie, które opisują określone zakresy życia. Przykładem tego rodzaju ontologii może być np. *curate*, której zadaniem jest dostarczenie narzędzi do tworzenia narracji wykraczających poza opis i interpretację pojedynczego obiektu muzealnego, lecz pozwalających na poszerzoną interpretację grup eksponatów, pewnych zjawisk ogólnych, tak jak to czyni się przy budowaniu wystaw muzealnych⁴³.

CIDOC CRM jest właśnie tego rodzaju ontologią dziedzinową, której zadaniem jest opis szeroko pojętych dóbr kultury i zjawisk w zakresie dziedzictwa kulturowego. CIDOC CRM składa się zasadniczo z 86 klas oraz 137 właściwości⁴⁴. Klasa jest tutaj rozumiana jako kategoria przedmiotów, które posiadają jedną lub więcej wspólnych cech. Zakres klasy jest sformułowa-

⁴³ *Curate: An ontology for describing museum narratives*. Dok. elektroniczny: <http://decipher.open.ac.uk/curate/introduction> [Dostęp: 31 X 2012].

⁴⁴ Pełna dokumentacja ontologii: *The CIDOC Conceptual Reference Model*. Dok. elektroniczny: <http://www.cidoc-crm.org/> [Dostęp: 31 X 2012].

ny w postaci noty, która po prostu opisuje, jakie pojęcia lub rzeczy kwalifikują się do danej klasy, odwołując się do powszechnego rozumienia terminu będącego nazwą klasy. W jednolitym systemie oznakowania elementów CIDOC CRM symbole klas mają postać liczb poprzedzonych literą E (od *Entity* – ang. jednostka, byt); terminy oznaczające klasy mają z reguły postać rzeczownikową, gdyż klasy pełnią w tym systemie semantycznym rolę podobną do gramatycznych podmiotów. Można je też definiować w sposób samodzielny, gdyż klasy są niezależnymi elementami tej ontologii.

CIDOC CRM Class Hierarchy

<u>E1</u>	CRM Entity
<u>E2</u>	- Temporal Entity
<u>E3</u>	- - Condition State
<u>E4</u>	- - Period
<u>E5</u>	- - - Event
<u>E7</u>	- - - - Activity
<u>E8</u>	- - - - - Acquisition Event
<u>E9</u>	- - - - - Move
<u>E10</u>	- - - - - Transfer of Custody
<u>E11</u>	- - - - - Modification
<u>E12</u>	- - - - - - Production
<u>E79</u>	- - - - - - Part Addition
<u>E80</u>	- - - - - - Part Removal
<u>E13</u>	- - - - - - Attribute Assignment
<u>E14</u>	- - - - - - Condition Assessment
<u>E15</u>	- - - - - - Identifier Assignment
<u>E16</u>	- - - - - - Measurement
<u>E17</u>	- - - - - - Type Assignment
<u>E65</u>	- - - - - - Creation
<u>E83</u>	- - - - - - - Type Creation
<u>E66</u>	- - - - - - Formation
<u>E85</u>	- - - - - - Joining
<u>E86</u>	- - - - - - Leaving
<u>E87</u>	- - - - - - Curation Activity
<u>E63</u>	- - - - - - Beginning of Existence
<u>E67</u>	- - - - - - Birth
<u>E81</u>	- - - - - - Transformation
<u>E12</u>	- - - - - - <i>Production</i>
<u>E65</u>	- - - - - - <i>Creation</i>
<u>E83</u>	- - - - - - - <i>Type Creation</i>

Ilustracja 8. Fragment hierarchii klas w ontologii CIDOC CRM (kursywą odnotowano te klasy, które pojawiają się powtórnie, jako rozwinięcie pewnych klas wyższego rzędu). Źródło: Definition of the CIDOC Conceptual Reference Models, Version 5.0.4, ed. N. Crofts, M. Doerr, T. Gill, S. Stead, M. Stiff, s.l., 2011, s. xxii; dok. electron. http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf; [odczyt: 31 X 2012].

Druga grupa elementów – właściwości, służy do określania relacji, jakie zachodzą pomiędzy klasami. Właściwości formułowane są z zasady w formie zdań bezpodmiotowych, gdyż ich rola jest w pewnym sensie analogiczna do roli orzeczenia w składni zdania. W nomenklaturze CIDOC CRM właściwości oznacza się liczbą poprzedzoną literą P (*Property* – ang. własność, właściwość). O ile klasy są niezależnymi pojęciami samymi w sobie, o tyle użycie własności jest ograniczone poprzez formalnie określone zakresy wchodzenia w związki z odpowiednimi klasami. Znaczenie relacji, podobnie jak w przypadku klas, opiera się na pospolitym ich rozu-

mieniu, a opisują je noty tekstowe (tzw. *scope notes*, tj. noty zawierające definicję i/lub określające zakres pojęcia).

P4 has time-span (is time-span of)

Domain: [E2 Temporal Entity](#)
 Range: [E52 Time-Span](#)
 Quantification: many to one, necessary, dependent (1,1:1,n)

Scope note: This property describes the temporal confinement of an instance of an E2 Temporal Entity.

The related E52 Time-Span is understood as the real Time-Span during which the phenomena were active, which make up the temporal entity instance. It does not convey any other meaning than a positioning on the “time-line” of chronology. The Time-Span in turn is approximated by a set of dates (E61 Time Primitive). A temporal entity can have in reality only one Time-Span, but there may exist alternative opinions about it, which we would express by assigning multiple Time-Spans. Related temporal entities may share a Time-Span. Time-Spans may have completely unknown dates but other descriptions by which we can infer knowledge.

Examples:

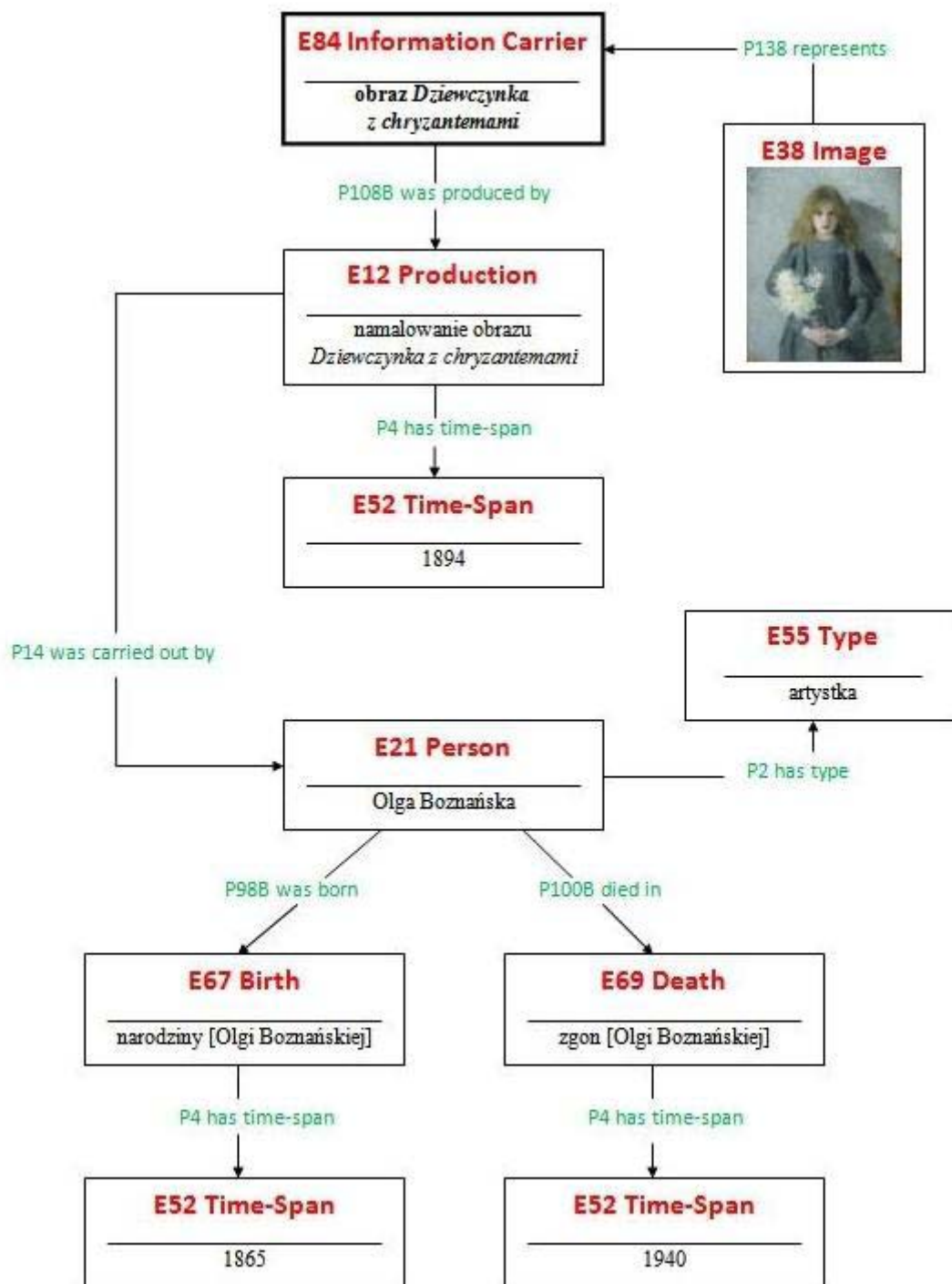
- the Yalta Conference (E7) *has time-span* Yalta Conference time-span (E52)

Ilustracja 9. Przykład właściwości P4 has time-span (is time-span of), która służy do wyrażania pojęć dotyczących rzeczy rozciągających się w czasie i tworzy relacje z pojęciem E52 Time-Span (zakres czasu).

Źródło: Definition of the CIDOC Conceptual Reference Models, Version 5.0.4, ed. N. Crofts, M. Doerr, T. Gill, S. Stead, M. Stiff, s.l., 2011, s. 38; dok. electron. http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf; [odczyt 31 X 2012].

Z klas i właściwości – przy zastosowaniu innych jeszcze reguł językowych szczegółowo opisanych w specyfikacji ontologii CIDOC CRM – można tworzyć wielką liczbę twierdzeń, którymi da się opisać obszerną dziedzinę dóbr kultury, i które będzie można wymieniać pomiędzy aplikacjami i interpretować (np. przez wyszukiwarki) z zachowaniem zapisanych znaczeń i związków logicznych. Naturalnie, nie ma konieczności użycia wszystkich pojęć i związków, jakie daje do dyspozycji CIDOC CRM. Można zastosować uproszczenia, ale ontologia ta pozwala też na dodawanie rozszerzeń, które byłyby konieczne dla poprawnego opisanie wszelkich, jakże różnorodnych aspektów dziedzictwa kulturowego.

Poniższy prosty przykład ilustruje użycie ontologii CIDOC-CRM do zapisania twierdzenia, które można wyrazić zdaniem: „*Dziewczynka z chryzantemami*, dzieło, którego reprodukcję oglądamy, zostało namalowane w roku 1894 przez Olę Boznańską, artystkę, urodzoną w roku 1865, a zmarłą w roku 1940”, bądź suchym zapisem inwentarzowym: *Dziewczynka z chryzantemami* (obraz), wyk. Olga Boznańska, art. mal., (*1865 +1940).



Ilustracja 10. Schemat ilustrujący użycie ontologii CIDOC-CRM do opisu obrazu O. Boznańskiej, *Dziewczynka z chryzantemami*. Opracowanie: T. Zaucha.

Ontologie tego rodzaju jak CIDOC-CRM wymagają języka (opartego na składni XML), który pozwalałby w sposób formalny wyrazić użyte w dokumentach terminy i pojęcia. Tego rodzaju językami są RDF (Resource Description Framework), czy stanowiący jego rozwinięcie OWL (Web Ontology Language; OWL jest anagramem skrótu i oznacza w języku angielskim sowę). Ten ostatni język posiada znacznie poszerzony słownik opisu pojęć i klas, więcej relacji typów

i charakterystyk opisujących własności. Tego rodzaju język nie jest już prostym schematem XML, pozwalającym na sformatowanie informacji, lecz przedstawieniem wiedzy. Dostarcza narzędzi pozwalających na budowę aplikacji, które będą mogły analizować ową wiedzę i tworzyć w logiczny sposób nowe twierdzenia, wywodzone z pierwotnie podanych.

Ciekawym i wartym odnotowania przykładem implementacji ontologii CIDOC CRM w języku OWL DL (to jedna z trzech odmian tego języka) jest Erlangen CRM, który wziął nazwę od Friedrich-Alexander-Universität Erlangen-Nürnberg, gdzie został sformułowany⁴⁵. Ta implementacja ontologii CIDOC CRM ma już kilka wdrożeń. Jednym z nich jest CLAROS – zbiorcza platforma pozwalająca na prezentację zabytków związanych ze starożytnościami greckimi i rzymskimi⁴⁶.



claros:coordinates-current	41.89999961853027, 12.483333110809326
claros:coordinates-find	41.89999961853027, 12.483333110809326
crm:P102_has_title	Per l'esequie del Ser.mo S. Duca Alessa(n)dro Farnese
crm:P108i_was_produced_by	http://id.clarosnet.org/BSR/0015816/production
crm:P138_represents	Rome
crm:P138i_has_representation	Image of Per l'esequie del Ser.mo S. Duca Alessa(n)dro Farnese
crm:P2_has_type	graphic Enchings 1600-1620 http://id.clarosnet.org/type/object/graphic Etching
crm:P3_has_note	21st plate of Ornamenti di fabbriche antichi et moderni dell'alma citta di Roma, published in Rome, as 2nd part in the 2nd edition of the Vestigi dell'antichità di Roma by Etienne Du Pérac, by Andrea and Michelangelo Vaccaro, after 1607, probably after 1614. This edition also includes Antiquarum Statuarum Urbis Romae

Ilustracja 11. Opis obiektu na platformie CLAROS; do etykietowania poszczególnych danych użyto nazw elementów CIDOC CRM. Źródło: dok. elektroniczny: <http://id.clarosnet.org/BSR/0015816>; [odczyt: 31 X 2012].

Erlangen CRM jest także używany przez platformę wiedzy przygotowywaną przez Poznańskie Centrum Superkomputerowo-Sieciowe w ramach projektu SYNAT⁴⁷. Użyto go również do tworzenia katalogu internetowego zbiorów Germanisches Nationalmuseum w Norymberdze⁴⁸.

⁴⁵ Erlangen CRM / OWL – An OWL DL 1.0 implementation of the CIDOC Conceptual Reference Model (CIDOC CRM). Dok. elektroniczny: <http://erlangen-crm.org>. [Dostęp: 31 X 2012].

⁴⁶ CLAROS: The world of art on the semantic web. Dok. elektroniczny: <http://clarosnet.org>. [Dostęp: 31 X 2012]. Warto odwiedzić także stronę zawierającą wymagania techniczne odnośnie publikacji danych w tym agregatorze, gdzie podano cztery podstawowe szablony zapisu informacji dot. obiektów, miejsc, okresów i osób ClarosWeb MediaWiki. Dok. elektroniczny: <http://explore.clarosnet.org/wiki/>. [Dostęp: 31 X 2012].

⁴⁷ SYNAT. Zespół Bibliotek Cyfrowych PCSS. Dok. elektroniczny: <http://dl.psync.pl/activities/projekty/synat/>. [Dostęp: 31 X 2012]; Mazurek C., Sielski, K., Stroński, M., Walkowska, J., Werla, M., & Węglarz, J., *Transforming a Flat Metadata Schema to a Semantic Web Ontology: The Polish Digital Libraries Federation and CIDOC CRM Case Study* [w:] *Intelligent Tools for Building a Scientific Information Platform (Studies in Computational Intelligence)* ISBN 978-3642248085, 2012.

⁴⁸ Objektkatalog der Sammlungen des Germanischen Nationalmuseums Nürnberg. Dok. elektroniczny: <http://objektkatalog.gnm.de/>. [Dostęp: 31 X 2012].

Poniższe przykłady przedstawiają hasło katalogowe i odpowiadający mu schemat XML sformułowany z użyciem Erlangen CMR.

OBJEKT-KATALOG DER SAMMLUNGEN DES GERMANISCHEN NATIONALMUSEUMS

[Highlights](#)
[Recherche](#)
[Informationen](#)
[Impressum](#)





Größere Ansicht

Madonna vom Wohnhaus des Veit Stoß (Skulptur aus Lindenholz)

Inventarnummer: Pl.O.217
Hersteller: Veit Stoß
Datierung: um 1520
Ort: Nürnberg
Material/Technik: Lindenholz - unbemalt; Rückseite gehöhlt; freiplastisch
Maße: H. 141,0 cm
Sammlung: Skulptur bis 1800

Beschreibung:
 Maria, in ein weites Gewand mit Mantel gekleidet, hält vor sich das Christuskind, das sie mit der rechten Hand um den Leib faßt, während die Linke, die gleichzeitig den Mantel emporrafft, es von unten unterstützt. Das Haupt der Maria ist leicht nach rechts geneigt, lange Haarwellen fließen über die Schultern herab. Sie trägt eine hohe Blätterkrone. Das Christkind hält in der Rechten eine Birne, während die Linke seinen rechten Fuß faßt. Die Beruhigung in der Bewegung der Falten, die zu einem geschlossenen Umriß mit langgezogenen, durchlaufenden Binnenformen führt, wie auch der pralle Kinderkörper und das rundliche Gesicht Mariä sind als Einfluß der neuen Sehweise Dürers zu werten.

- IHRE RECHERCHE

Freitext:
wunderburggasse

- OBJEKTANSICHT

Objekt / Inventarnr.:
Pl.O.217

- AKTIONEN

Navigation:
Zurück zur Ergebnisliste

- LINKED DATA

Objektidentifikatoren:
<http://id.gnm.de/obj/Pl.O.217> ...
 urn:moi:d95c67c0-ba75-5 ...

Rohdaten:
 OWL/XML, RDF/XML,
 RDF/JSON, N-Triple, Turtle

Ilustracja 12. Nota katalogowa dot lipowej rzeźby Madonna z domu Wita Stwosza w zbiorach Germanisches Nationalmuseum. Źródło: dok. elektron. <http://objektkatalog.gnm.de/objekt/Pl.O.217>, [odczyt: 31 X 2012].

```

- <rdf:RDF>
- <owl:Ontology rdf:about="">
  <owl:imports rdf:resource="http://erlangen-crm.org/120111/">
  <owl:Ontology>
- <ecrm:E84_Information_Carrier rdf:about="http://id.gnm.de/obj/Pl.O.217">
  <owl:sameAs rdf:resource="urn:moi:d95c67c0-ba75-5564-82e4-5f753bc5ded1"/>
- <ecrm:P48_has_preferred_identifier>
  - <ecrm:E42_Identifier rdf:nodeID="bn0">
    <ecrm:P3_has_note rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Pl.O.217</ecrm:P3_has_note>
  </ecrm:E42_Identifier>
  </ecrm:P48_has_preferred_identifier>
- <ecrm:P102_has_title>
  - <ecrm:E35_Title rdf:nodeID="bn1">
    - <ecrm:P3_has_note rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
      Madonna vom Wohnhaus des Veit Stoß (Skulptur aus Lindenholz)
    </ecrm:P3_has_note>
  </ecrm:E35_Title>
  </ecrm:P102_has_title>
+ <ecrm:P65_shows_visual_item></ecrm:P65_shows_visual_item>
+ <ecrm:P43_has_dimension></ecrm:P43_has_dimension>
+ <ecrm:P70i_is_documented_in></ecrm:P70i_is_documented_in>
+ <ecrm:P70i_is_documented_in></ecrm:P70i_is_documented_in>
+ <ecrm:P104_is_subject_to></ecrm:P104_is_subject_to>
- <ecrm:P50_has_current_keeper>
  - <ecrm:E40_Legal_Body rdf:nodeID="bn11">
    - <ecrm:P131_is_identified_by>
      - <ecrm:E82_Actor_Appellation rdf:nodeID="bn12">
        <ecrm:P3_has_note rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Germanisches Nationalmuseum</ecrm:P3_has_note>
      </ecrm:E82_Actor_Appellation>
    </ecrm:P131_is_identified_by>
  </ecrm:E40_Legal_Body>
  </ecrm:P50_has_current_keeper>

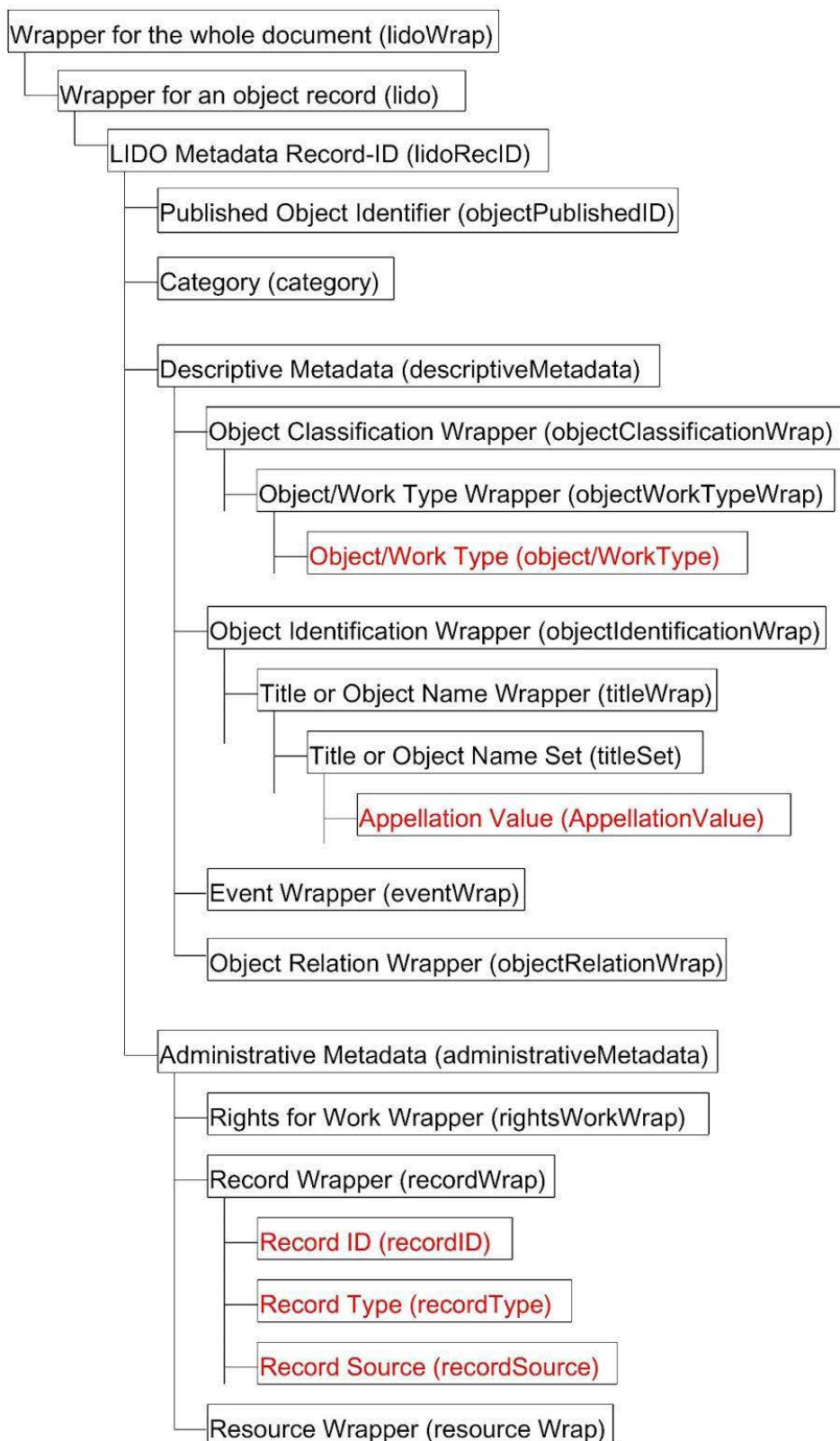
```

Ilustracja 13. „Surowe” dane (Rohdaten) rzeźby Madonna z domu Wita Stwosza w zbiorach Germanisches Nationalmuseum zapisane w Erlangen CMR. Źródło: dok. elektron. <http://objektkatalog.gnm.de/objekt/Pl.O.217/rdf>, [odczyt: 31 X 2012].

2.1.5. Lightweight Information Describing Objects (LIDO)

Rozmaitość zbiorów zgromadzonych w muzeach i nie dająca się przewyciężyć różnorodność schematów metadanych bądź systemów informatycznych używanych przez różne instytucje stoi u źródeł powstania standardu Lightweight Information Describing Objects (LIDO), który został przygotowany przez grupę roboczą Data Harvesting and Interchange, działającą w obrębie komitetu CIDOC przy Międzynarodowej Radzie Muzeów (ICOM). Jest to schemat służący przekazywaniu posiadanych informacji o muzealiach w ujednolicony sposób, tak, by nadawały się one do wyszukiwania i kwerend (w praktyce oznaczać to będzie przeważnie przekazywanie danych do portali internetowych). Schemat LIDO obejmuje te informacje, które zwykle znajdujemy w katalogach zbiorów lub wystaw, natomiast poza jego zakresem pozostaje cała grupa danych służących w elektronicznych systemach ewidencji zbiorów do procedur administracyjnych (typu: akcesja, wypożyczenia, inwentaryzacja kontrolna – skontrum i in.). Zakres danych udostępnianych przy użyciu schematu LIDO jest pozostawiony do decyzji instytucji udostępniającej; obowiązkowych jest zaledwie kilka zupełnie nieodzownych pól: z danych opisowych: typ obiektu (obraz, rzeźba itp.), nazwa/tytuł obiektu; z danych administracyjnych: identyfikator rekordu, typ rekordu i źródło rekordu.

LIDO ma postać schematu XML, tak jak np. CDWALite, czy wypracowany przez Fachgruppe Dokumentation des Deutschen Museumsbundes schemat Museumdat. Ogłoszony w listopadzie 2010 r. LIDO wywodzi się zresztą z obydwu wspomnianych schematów i docelowo ma je zastąpić. Ponadto zestawiono go w taki sposób, by utrzymać zgodność z danymi ujętymi zgodnie z brytyjskim schematem Spectrum. Tak jak ma to miejsce w schematach tego rodzaju, LIDO posiada strukturę hierarchiczną. Poniższy rysunek przedstawia rozwinięcie głównych „węzłów” tego schematu, ukazując, w których miejscach ulokowane są elementy obowiązkowe (wyróżnione kolorem czerwonym).



Ilustracja 14. Podstawowe elementy hierarchii schematu LIDO. Opracowanie: T. Zaucha na podstawie: E. Coburn, R. Light, G. McKenna, R. Stein, A. Vitzthum, LIDO – Lightweight Information Describing Objects Version 1.0, s.l., 2010, s. 11-26.

Lakoniczne określenia – w nawiasach podano odpowiadające im znaczniki (tagi) XML – znajdują wyjaśnienie w pełnej specyfikacji schematu LIDO⁴⁹. Jest on rekomendowany jest jako kompatybilny z CIDOC CRM, jakkolwiek w jego specyfikacji brakuje jeszcze odniesień do elementów tej ontologii.

W zilustrowanej wyżej strukturze można wyróżnić pięć głównych grup informacji, zestawionych pod następującymi „węzłami”:

- 1) dane typologiczne – w obrębie pakietu (ang. *wrapper* – dosł. obwoluty) danych opisowych, poniżej „węzła” Object Classification, takie jak typ obiektu, kategoria;
- 2) dane identyfikacyjne – w obrębie pakietu danych opisowych, pod „węzłem” Object Identification, czyli m.in. nazwa/tytuł, napisy, miejsce przechowywania, wymiary;
- 3) zdarzenia – w obrębie pakietu danych opisowych, pod „węzłem” Events, gdzie w rozbudowanej strukturze może zostać zawarta informacja o wszelkiego rodzaju zdarzeniach dotyczących obiektu;
- 4) związki – w obrębie pakietu danych opisowych, pod „węzłem” Relations, głównie dotyczące szeroko pojętej tematyki dzieła oraz obiektów pokrewnych;
- 5) metadane administracyjne – pakiet danych administracyjnych (rozwiniecie „węzła” Administrative Metadata).

Specyfikacja LIDO, poza informacją o ogólnych zasadach budowy i podstawowej strukturze zawiera także szczegółowy opis każdego z elementów używanych do budowy schematu opisu obiektu. Poniżej przedstawiono zawarty w specyfikacji (wersja pdf) przykład takiego opisu.

⁴⁹ LIDO. Dok. elektroniczny: <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/lido-technical/specification/> [Dostęp:31 X 2012].

Element objectDescriptionSet	
Description	
<i>Title</i>	Description/Descriptive Note Set
<i>General</i>	Wrapper for a description of the object, including description identifier, descriptive note and sources.
<i>How to record</i>	Includes usually a relatively brief essay-like text that describes the content and context of the object / work, including comments and an interpretation that may supplement, qualify, or explain the physical characteristics, subject, circumstances of creation or discovery, or other information about it. If there is more than one descriptive note, repeat this element.
<i>Notes</i>	A reference to a text resource holding the description may be given in description identifier.
Technical	
<i>Tag-name</i>	objectDescriptionSet
<i>Attributes</i>	type; sortorder
<i>Required?</i>	No
<i>Repeatable?</i>	Yes
<i>Super-elements</i>	objectDescriptionWrap
<i>Sub-elements</i>	descriptiveNoteID; descriptiveNoteValue; sourceDescriptiveNote
<i>Equivalentents</i>	CDWA Lite descriptiveNoteSet
	museumdat descriptiveNoteSet
	SPECTRUM [none]

Wskazówki, co do rodzaju zawartości danego pola - tutaj: jak powinna wyglądać nota do obiektu

Symbol znacznika używanego w schemacie XML

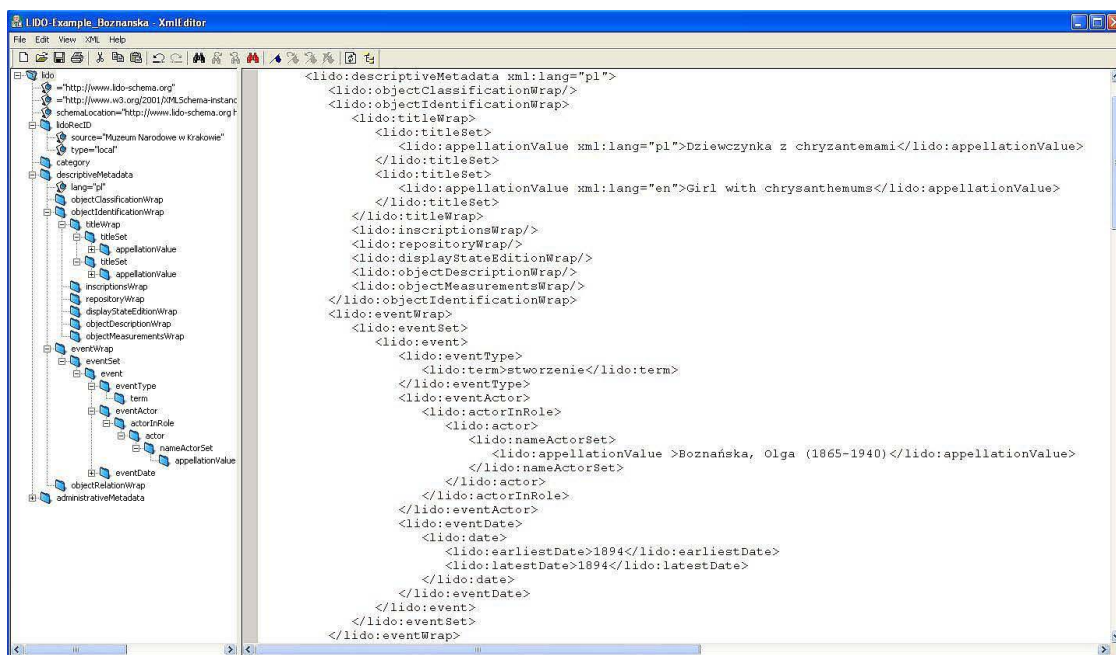
Elementy, których rozwinięciem może być ten dany element oraz elementy, których można użyć na niższych szczeblach hierarchii

Mapowanie, czyli odpowiedniki w innych schematach

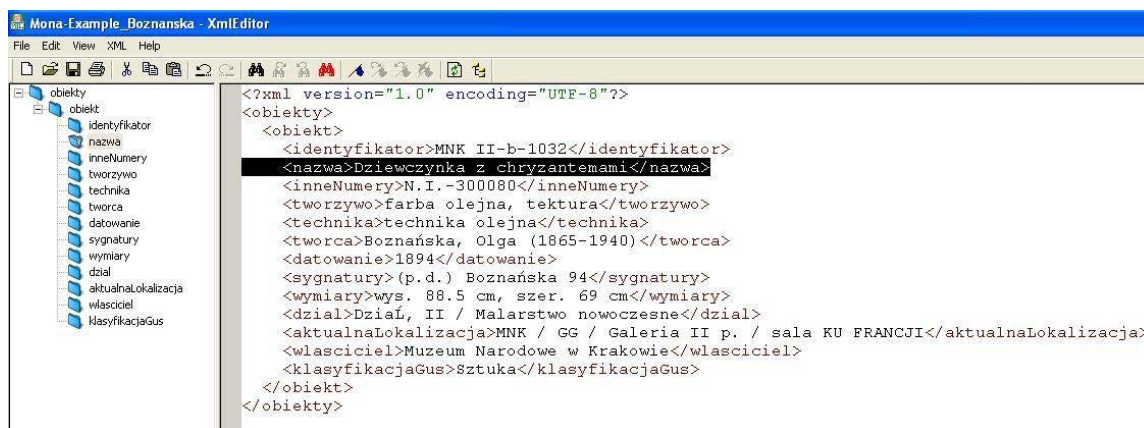
Ilustracja 15. Opis pojedynczego elementu (*objectDescriptionSet*) w schemacie LIDO. Opracowanie: T. Zaucha na podstawie: E. Coburn, R. Light, G. McKenna, R. Stein, A. Vitzthum, *LIDO – Lightweight Information Describing Objects Version 1.0, s.l., 2010, s. 71.*

Łatwo zauważyć, że w głównych grupach danych, tworzących strukturę LIDO, nie znajdujemy tradycyjnych elementów opisu obiektu muzealnego, takich jak: technika i materiał, osoba twórcy, miejsce i czas powstania. Te ważne dane są jednak zapisywane, a używa się do tego elementów z pakietu zdarzeń. To jest bardzo istotna cecha schematu LIDO, który został zbudowany w taki sposób, by spełniać założenia ontologii CIDOC-CRM. Na podanym powyżej – przy omówieniu zasad CIDOC CRM – przykładzie obrazu Olgi Boznańskiej, zilustrowano w jaki sposób można zapisać informację na temat autorstwa dzieła; to i cały szereg innych zdarzeń, jakie dotyczą obiektu, jego autorów (czy ogólniej: czynników sprawczych, zwanych aktorami, od łac. *ago* – działam), miejsc i dat z nim związanych. Co więcej, te same elementy można zastosować nie tylko do samego dzieła, jego powstania i późniejszych losów. Można je także użyć w części dotyczącej relacji, czyli tematyki i innego rodzaju związków. Jest to zwrot w kierunku sieci semantycznej, znacznie poszerzający możliwości interpretacyjne, gdyż jedno i to samo wydarzenie (osoba, miejsce) może występować i łączyć się w różny sposób z różnymi obiektami: odsiecz wiedeńska (zdarzenie historyczne) wiąże ze sobą zarówno mahometańską chorągiew zdobytą przez Sobieskiego i oddaną do Loreto (obecnie Zamek Królewski na Wawelu), jak i obraz Matejki, przedstawiający triumf Sobieskiego (w Muzeum Watykańskim).

Struktura LIDO, wyrażona w języku XML, wygląda tak, jak w podanym niżej przykładzie. Informacja w nim zawarta mówi o tym, że obraz o tytule polskim *Dziewczynka z chryzantemami* (w wersji angielskiej: *Girl with chrysanthemums*) został namalowany przez Olgę Boznańską w roku 1894. Lewy panel przedstawia znaną nam z Ilustracji. 14. hierarchię, zaś prawy – odpowiadające elementom hierarchii linie kodu. Dla porównania, kolejna ilustracja przedstawia prostą strukturę pliku XML wygenerowanego z systemu Mona 3W, z rzeczywistego rekordu opisującego obraz.



Ilustracja 16. Przykładowy schemat pliku XML w formacie LIDO (obraz *Dziewczynka z chryzantemami*).
Opracowanie: T. Zaucha.



Ilustracja 17. Plik XML wyeksportowany z systemu Mona 3W (obraz *Dziewczynka z chryzantemami*).
Opracowanie: T. Zaucha.

W szeregu przypadków zestaw informacji może być powtarzalny, co oznacza, że po wielokroć można zapisać różne informacje tego samego rodzaju (w powyższym przykładzie zapisano dwakroć tytuł, w dwu wersjach językowych). Jest to szczególnie ważne wówczas, kiedy dla jednego obiektu trzeba odnotować kilka różnego rodzaju wydarzeń (np. zaprojektowanie, wy-

konanie, przeróbka), z których każde może mieć inny zestaw danych (kto, gdzie, kiedy, rodzaj wydarzenia). Ponadto informacje można zapisywać przy pomocy powtarzalnych elementów, takich chociażby jak: *Appellation Value*, gdzie odnotowuje się nazwę własną (np. tytuł dzieła), bądź *Term/Label*, w którym umieszcza się po prostu określenie danego pojęcia, a który bywa używany – jako element podrzędny – w dwudziestu rozmaitych kategoriach (co oznacza, że LIDO jest strukturą polihierarchiczną). LIDO pozwala także na stosowanie atrybutów opisujących pewne dane; np. tytuł można opatrzyć atrybutem *pref*, który może przybrać dwojaką wartość: *preferred* dla głównego tytułu albo *alternate* dla tytułów obocznych. Można również używać pewnych ogólnych atrybutów XML, np. wskazujący na użyty język atrybut `xml:lang`. Możliwość obsłużenia wpisów w wielu językach jest zresztą jedną z istotnych cech LIDO. LIDO pozwala także na przechowywanie i prezentację w portalach informacji „wzbogaconej”, tzn. odnośników do zasobów zewnętrznych, które pozwalają na uzupełnienie danych zawartych w źródłowej bazie danych. Zwykle są to odnośniki do tezaurusów zawierających zstandaryzowaną terminologię lub innych źródeł, pozwalających na sprecyzowanie bądź ujednoczenie informacji (np. do AAT, VIAF, ICONCLASS i in.⁵⁰). Poniższy przykład pokazuje jak w opartym na LIDO wirtualnym muzeum instrumentów muzycznych (MIMO – Musical Instruments Museum Online⁵¹) termin geograficzny „Korea” został uzupełniony linkiem do portalu internetowego Geonames.org, dzięki czemu można stwierdzić, że opisywany zabytek (flet ze zbiorów Ethnologisches Museum w Berlinie-Dahlem) pochodzi nie z Półwyspu Koreańskiego, lecz z indonezyjskiej wyspy Kura (inne nazwy: Korea, Koera).

Record Identifier `spk:obj-1634455`

Data Provider

Acronym SPK
Provider [Stiftung Preußischer Kulturbesitz - Ethnological Museum](#)

Identification Classification **Events** Digital resources Original record

Event 1

Event type Production
Event (display) Herstellung am Ort
Event name Herstellung

Place (display) Korea

Place 1

Event Place Name [Korea](#)
Location in Geonames <http://www.geonames.org/1639066>

Ilustracja 18. Rekord obiektu w MIMO – Musical Instruments Museum Online; flet ze zbiorów Ethnologisches Museum, Staatliche Museen zu Berlin, nr inv. VII c 775. Źródło. dok. elektron. http://www.mimo-db.eu/mimo/infodoc/Ged/View.aspx?eid=SPK_BERLIN_DE_EM_OBJID_1634455&searchId=32f01ecc-b1fa-4c51-a90d-7bdef34b8350, [odczyt: 31 X 2012].

⁵⁰ *Art & Architecture Thesaurus*. Dok. elektroniczny: <http://www.getty.edu/research/tools/vocabularies/aat/index.html> [Dostęp: 31 X 2012].; *Virtual International Authority File*. Dok. elektroniczny: <http://viaf.org> [Dostęp: 31 X 2012]; *Iconclass*. Dok. elektroniczny: <http://www.iconclass.org> [Dostęp: 31 X 2012].

⁵¹ *Musical Instruments Museum Online*. Dok. elektroniczny: <http://www.mimo-international.com/> [Dostęp: 31 X 2012].

Należy zaznaczyć, że pierwszym i podstawowym sposobem użycia ontologii takich jak CIDOC CRM, nie jest budowa systemów zarządzania zbiorami. Środowiska, w których gromadzona jest informacja o obiektach muzealnych (i nie tylko o nich) dzielą się zasadniczo na dwa typy:

- 1) lokalne systemy informatyczne, zarządzane z reguły przez pojedynczą instytucję i służące do wprowadzania danych i bieżącego zarządzania zbiorem
- 2) systemy dostępu, które powinny dostarczać informację w ujednolicony sposób.

Ta informacja spoczywa zasadniczo w systemach lokalnych i jest udostępniana poprzez fizyczny eksport/import zestandaryzowanych danych, lub też poprzez wysyłanie zapytań wprost do systemu lokalnego (o ile działają odpowiednie mechanizmy bieżącej translacji po stronie systemu lokalnego bądź lub po stronie systemu dostępu).

Kompatybilny z CIDOC CRM schemat LIDO, kwalifikuje się przede wszystkim – jako płaszczyzna współpracy znaczeniowej (*semantic interoperability*) – do obsługi drugiego typu systemów. Umożliwia on łączne zestawianie danych (agregację), które mogą pochodzić z rozmaitych, często bardzo odmiennych, systemów lokalnych. Zatem rekomendowaną cechą tych systemów powinna być możliwość eksportu danych do struktur kompatybilnych z CIDOC CRM, to znaczy takich, które będą mogły w sposób automatyczny wymieniać dane z innymi strukturami opartymi na tej ontologii (bez utraty znaczeń). Dla zachowania rzeczywistej zgodności systemu informatycznego z CIDOC CRM przewidziano w specyfikacji tej ontologii procedurę weryfikacyjną, polegającą m.in. na konieczności przedstawienia próbnych danych w sposób możliwy do zweryfikowania kompatybilności przez dowolną stronę trzecią lub wymogu uzyskania certyfikatu.

LIDO wydaje się być istotnym schematem nie tylko dlatego, że jest krokiem ku unifikacji schematów wymiany danych: zastąpi amerykański CDWA Lite i niemiecki Museumdat, które nie będą dalej rozwijane, może też być łatwo przełożony na brytyjskie Spectrum. Jest ważny także dlatego, iż dzięki odpowiedniemu sformułowaniu znaczników i definicji elementów, ów XML-owy schemat stosunkowo łatwo da się odnieść do standardów treściowych (*data content standards*), czyli mówiąc wprost: do zestawu informacji, jakim przyjęto opisywać muzealia w danej instytucji. Poprzez swą genezę LIDO odwołuje też użytkownika wprost do elementów CDWA, który jest uznanym, międzynarodowym standardem treściowym (mówi z jakich elementów ma – a raczej: może – składać się opis zabytku). Ułatwia to znacznie dialog pomiędzy muzealnikiem a informatykiem, który jest niezbędny do pełnego i poprawnego powiązania schematu źródłowego, który został zastosowany w danym systemie lokalnym oraz schematu docelowego (w tym przypadku: LIDO).

2.2. Wybór optymalnego rozwiązania

W kontekście przedstawionych powyżej światowych schematów i standardów metadanych oraz dobrych praktyk, tworzenie od podstaw nowych rozwiązań, dostosowanych do potrzeb różnych zbiorów zgromadzonych w polskich muzeach, wydaje się zadaniem nieuzasadnionym. Należy mieć jednak na uwadze, że elementy schematów definiujące procedury zarządzania kolekcją nie mogą zostać bezpośrednio użyte na gruncie polskim między innymi z uwagi na odmienną przepisów w tym zakresie. W tej sytuacji optymalnym rozwiązaniem byłoby posługiwanie się opracowanym standardem opisu i zarządzania muzealiami opartym na istniejących międzynarodowych schematach metadanych, uwzględniających specyfikę procedur i historyczne uwarunkowania polskich instytucji muzealnych. Podejmowane obecnie przez NIMOZ działania mogą w przyszłości zaowocować ogólnopolskimi rozwiązaniami. Brak narodowych standardów opisu muzealiów nie ogranicza całkowicie instytucji w tym zakresie. Możemy już dzisiaj wskazać na kilka potencjalnych rozwiązań, które możliwe są do wdrożenia.

Zanim muzeum dokona wyboru najkorzystniejszego sposobu opisu zbiorów, konieczne jest przeprowadzenie analizy potrzeb instytucji, z uwzględnieniem planów rozwoju. W pierwszej kolejności należy określić jaki jest cel ewidencji zbiorów. Czy zadanie to ma ograniczać się jedynie do inwentaryzacji, czy ma także wspomagać całościowe zarządzanie kolekcją, czy w końcu ma służyć promocji zbiorów. Wybrane rozwiązanie powinno być uzależnione od rodzaju gromadzonych w muzeum obiektów i wymagań w zakresie ich opisu. Drugi kluczowy element analizy powinien dotyczyć bieżącego potencjału instytucji w zakresie finansów i posiadanego zaplecza sprzętowego, a przede wszystkim kadr. Zastosowanie rozbudowanego schematu metadanych może okazać się niemożliwe do wdrożenia z uwagi na ograniczenia istniejącego oprogramowania, czy brak stosownych umiejętności u opracowujących zbiory. Rzetelna weryfikacja powyższych elementów pozwoli na dobór optymalnych rozwiązań.

Muzeum posiadające program do ewidencji zbiorów lub zainteresowane jego zakupem, powinno w pierwszej kolejności rozważyć wdrożenie jednego z międzynarodowych standardów opisu muzealiów i zarządzania kolekcją, jak wyżej wymienione Spectrum czy CDWA. Jest to rozwiązanie najbardziej korzystne, mimo że wymaga znacznego nakładu finansowego, wyspecjalizowanych kadr, wprowadzenia znaczących zmian w dotychczasowym funkcjonowaniu instytucji, a przede wszystkim stosunkowo długiego okresu wdrożenia.

Pierwszym elementem w ramach tego rozwiązania, na który trzeba zwrócić szczególną uwagę, to sam schemat metadanych. Nie jest możliwe, jak to już było wcześniej podkreślane, bezpośrednie przełożenie schematu bez wcześniejszej jego modyfikacji. Dotyczy to zarówno konieczności wprowadzenia innych elementów opisu (wynikających z obowiązujących w muzealnictwie polskim procedur), ale także konieczności dostosowania schematu dla każdej instytucji. Przy tego typu rozwiązaniach należy liczyć się z koniecznością znacznego przekształcenia dotychczasowo użytkowanej bazy danych lub wręcz jej wymiany na inną. Zakup

nowego oprogramowania obsługującego rozbudowane standardy opisu i zarządzania kolekcją, z uwagi na wysoką cenę, będzie wymagał przeprowadzenia procedur zamówienia publicznego. Należy pamiętać, iż samo wskazanie schematu metadanych nie spowoduje, że firma będzie w stanie automatycznie dostosować oprogramowanie do wymagań danego muzeum. W tej sytuacji konieczne będzie szczegółowe opracowanie przedmiotu zamówienia określające bardzo dokładnie zdefiniowanie poszczególnych elementów schematu, uwzględniające specyfikę funkcjonowania instytucji. Implementacja nowego oprogramowania wiąże się także z migracją dotychczasowych danych, które – z uwagi na modyfikację struktury opisu – będą wymagały szczegółowych korekt.

Poniesiony przez muzeum znaczny koszt, zarówno w wymiarze finansowym jak i nakładu pracy, w perspektywie czasu może przynieść instytucji wymierne korzyści na wszystkich polach związanych z zarządzaniem zbiorami. Efektem udanego wdrożenia jest możliwość wymiany zgromadzonych danych i publikacji informacji o zbiorach zarówno za pośrednictwem własnej strony internetowej, jak i wspólnych platform, takich jak Europeana.

Z uwagi na brak odpowiednio przeszkolonej kadry, wiele instytucji nie będzie gotowe wdrożyć tak dużego rozwiązania, co prowadzi do oczywistego pytania: czy tak rozbudowany schemat metadanych jest niezbędny każdemu muzeum. W małych placówkach tego typu koncepcja na pewno nie znajdzie zastosowania. Rozwiązaniem pośrednim jest więc dostosowanie użytkownika lub nowo wdrażanego oprogramowania do jednego ze schematów wymiany danych. W przeciwieństwie do wyżej opisanego, drugie rozwiązanie wiąże się z koniecznością dostosowania znacznie mniejszej liczby wymaganych elementów schematu, co znacząco ułatwia ich wdrożenie, a także późniejszą obsługę. Należy pamiętać, że schematy wymiany danych ze względu na swój charakter nie oddają wszystkich niuansów opisu każdego typu zbiorów i nie definiują wszystkich elementów związanych z zarządzaniem kolekcją. Dzięki zastosowaniu tych samych elementów schematu wdrożonych przez inne muzea, instytucja będzie mogła bez konieczności każdorazowego przygotowania materiałów eksportować dane i umożliwiać ich publikację we wspólnym środowisku sieciowym.

Muzeum, które nie może wdrożyć jednego ze standardów opisu lub wymiany danych, nie powinno rezygnować z prób strukturyzowania informacji. Niezależnie od sposobu prowadzenia ewidencji, każda z instytucji powinna wdrożyć wewnętrzną instrukcję opracowywania i opisu zbiorów. Nie oznacza to, że od razu należy opracować rozbudowaną dokumentację szczegółowo dzielącą najdrobniejsze możliwe elementy charakteryzujące obiekt. Ważne jest by wewnętrzny schemat metadanych jasno precyzował minimalny zakres opisu i rodzaje elementów schematu oraz sposób ich uzupełniania, a przede wszystkim, aby był zrozumiały dla wszystkich pracowników.

Instrukcje opisu obiektów niezależnie od stopnia ich szczegółowości, by właściwie funkcjonować, powinny spełniać następujące kryteria:

- 1) określać elementy schematu służące opisowi wszystkich typów zbiorów;
- 2) definiować poszczególne elementy schematu;
- 3) określać sposoby wprowadzania danych;
- 4) określać pola eksploatacji powstałego opisu;
- 5) określać sposób wprowadzania zmian w instrukcji;
- 6) określać sposób weryfikacji wprowadzonych danych.

Należy pamiętać, że rozporządzenie MKiDN dotyczące ewidencji zbiorów determinuje jedynie minimalny zakres informacji, które w momencie opisu muzealium należy spełnić. Wymagane rozporządzeniem jednostki opisu nie spełniają kryteriów wszystkich typów obiektów, ponieważ zostały zdefiniowane przede wszystkim dla zbiorów artystyczno-historycznych. Zasady opisu zbiorów muszą być zróżnicowane i uwzględniać specyfikę całości zbiorów gromadzonych w danej instytucji, jednakże przy opracowaniu instrukcji należy pamiętać o konieczności określenia ich wspólnych pól. Kolejnym krokiem jest zdefiniowanie zawartości poszczególnych pól, tak aby każdy z opisujących nie miał wątpliwości jaki rodzaj informacji w nich zawrzeć. Z kolei określenie sposobu wprowadzania danych dotyczyć będzie wskazania czy dane pole wspomagane jest słownikiem, listami kontrolowanymi czy też jest polem uzupełnianym dowolnym tekstem. Wewnętrzne regulacje powinny jasno definiować jakim celom mogą służyć dane: czy powstają tylko na potrzeby ewidencji, czy też posłużą późniejszemu udostępnianiu. Niezwykle ważnym elementem każdej instrukcji jest wcześniejsze określenie zasad wprowadzania w niej zmian, a także sposobu weryfikacji wprowadzonych informacji. W ten sposób opracowana instrukcja pozwoli w stosunkowo szybki sposób przygotować materiał do udostępnienia.

Dzięki zastosowaniu odpowiednich algorytmów możliwe będzie wyeksportowanie informacji, a konieczność ręcznego wprowadzania zmian będzie w znacznym stopniu zminimalizowana. Stosowanie jasnych zasad zapisu informacji w dokumentacji ewidencyjnej powinno przynieść muzeum korzyści, ułatwiając znacznie bieżącą pracę, między innymi przeprowadzanie kwerend.

Niezależnie od tego, które z rozwiązań zostanie zastosowane, instytucje powinny dążyć do konsekwentnego zapisu danych. Nie wszystkie z rekomendowanych rozwiązań wymagają dużych nakładów finansowych. Żadne nie będzie możliwe bez zaangażowania pracowników. Poniesiony trud może przynieść wymierne korzyści, co już dziś dostrzega część polskich instytucji, wdrażających nowe rozwiązania.

3. Kontrola słownictwa



3.1. Po co kontrolować słownictwo?

Standardy definiują poszczególne kategorie informacji oraz ich zawartość, nie precyzują jednak stosowanej terminologii. Powszechnie stosowane swobodne i zwyczajowe wolno-tekstowe opisy są obciążone błędami i niekonsekwentnym zastosowaniem terminologii⁵².

Dzieje się tak przede wszystkim ze względu na swoistość języka polskiego, zwłaszcza takie jego elementy jak różnorodność form gramatycznych, fleksja, gęstość synonimów, homonimów, antonimów i pojęć bliskoznacznych. W tych samych instytucjach w opisie stosowane są równoległe takie sformułowania jak „lawowanie”/„lawowany”, „glazura”/„szkliwo”, „ściągać”/„trzępień”/„pręt”. Trudno uniknąć także wieloznaczności terminów, spotykając się np. z określeniem „flet” nie wiemy, czy autor miał na myśli instrument muzyczny, typ wazonu czy też może typ kielicha. Komplikacje pojawiają się również w zakresie zapisu pojęć złożonych: „stopa kielicha mszalnego”, czy „kielich mszalny – stopa”; „gwintowe łączenia”, czy „łączenie – gwintowanie”.

Duże znaczenie dla stosowanego aparatu pojęciowego mają uwarunkowania historyczne, a zwłaszcza stosowanie języka branżowego. W języku stosowanym przez historyków sztuki występują tradycyjnie stosowane terminy, które wydają się nierozdzielnie związane z opisem dzieł sztuki. Zdarza się, że kustosze niewłaściwie stosują niektóre terminy; dla przykładu: przy opisie konstrukcji mebla, forniru czy intarsji, podają gatunki drzew („jesion”, „brzoza”, „czarny dąb”), a nie poprawne terminy określające poszczególne gatunki drewna („drewno jesionowe”, „drewno brzozowe”, „drewno dębowe”). Odnosząc się do obrazów wykonanych w technice olejnej na podłożu drewnianym, zawsze spotkamy się z określeniem „olej na desce”. Jednakże odwołując się do *Słownika Języka Polskiego* termin ten nie jest do końca jednoznaczny. Rzeczownik „olej” oznacza: „gęstą ciecz pochodzenia roślinnego lub zwierzęcego; gęstą ciecz będącą produktem destylacji; farbę olejną lub obraz namalowany takimi farbami”⁵³; natomiast „deska” to: „obustronnie płaski kawałek pnia drzewa przepiłowanego wzdłuż, lub narta”⁵⁴.

Zbiory muzeów nie ograniczają się jedynie do zabytków sztuki. Gromadzone w nich obiekty reprezentują różne dziedziny dziedzictwa materialnego, przy opracowywaniu których muzealnicy posługują się właściwym im aparatem pojęciowym. Etnograf, ze względu na odmienne podejście do danego obiektu, przy opisywaniu poszczególnych przedmiotów zawsze będzie odnotowywał informację dotyczącą rodzaju wyrobu („wyrób ręczny”, „wyrób fabryczny”, „wyrób domowy”), co nie wystąpi nigdy w przypadku not opracowanych przez specjalistów innych dyscyplin. Powracając do pojęcia „olej” warto zwrócić uwagę, że dla historyka sztuki

⁵² M. Kłos, K. Zielonka, *Strukturyzacja języka opisu. Doświadczenia Międzymuzealnej Grupy ds. Digitalizacji DigiMuz.* [w:] E. Herden, A. Seidel-Grzezińska i K. Stanicka-Brzezicka (red.), *Dobra kultury w Sieci. Cyfrowe spotkania z zabytkami 3*, Wrocław 2012, s. 111-120.

⁵³ *Słownik Języka Polskiego*. Dok. elektroniczny: <http://sjp.pwn.pl/> [Dostęp: 31 X 2012].

⁵⁴ Ibidem.

będzie on jednoznaczny z pojęciem „technika olejna”, natomiast przedstawiciele innych dziedzin będą raczej identyfikować ten termin z „olejem syntetycznym”, „naturalnym”, „mineralnym”, „roślinnym” itd. – wymienią „olej” nie jako technikę, ale materiał.

Niespójna terminologia wynika także z różnic regionalnych. Często dla tego samego przedmiotu (zjawiska) w poszczególnych ośrodkach (regionach) stosuje się odmienne nazwy. Obok występowania dwóch wymiennych, równoznacznych terminów, zdarzają się także przypadki stosowania sformułowania, które nie jest akceptowane przez inne środowisko, ze względu na błąd merytoryczny. Nazwą, która budzi wiele wątpliwości jest „fajansowa płytką ścienna” lub „kafel”, które w środowisku pomorskim określane są mianem „flizy”, natomiast termin „kafel” oznacza element pieca. W przypadku zbiorów etnograficznych odmienności regionalne w stosowaniu innych nazw, wynikają z braku opracowania jednolitej naukowej terminologii i posługiwania się głównie określeniami potocznymi lub gwarowymi.

Duży problem stanowi często wybór nazwy związanej z daną lokalizacją, ponieważ najczęściej poprzez przemiany historyczne i ustrojowe, z tymi samymi miejscowościami związane były różne nazwy. Doskonałym przykładem jest Kaliningrad dawniej funkcjonujący pod nazwą polską Królewiec i niemiecką Königsberg. Przy opisywaniu obiektu, który powstał w XVII wieku w Królewcu (Prusach Książęcych), miejsce powstania zgodnie z wytycznymi *Komisji Standaryzacji Nazw Geograficznych poza Granicami Rzeczypospolitej Polskiej* należałoby określić Kaliningrad (Rosja), co wydaje się od razu błędem logicznym⁵⁵.

Dbanie o jakość opisu z powyższych względów nie może opierać się tylko i wyłącznie na stosowaniu schematów i koniecznie musi zostać wsparte narzędziami weryfikującymi zastosowaną terminologię.

⁵⁵ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 6, Białoruś, Rosja, Ukraina*. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_06.pdf [Dostęp: 31 X 2012].

3.2. Jak kontrolować słownictwo?

Muzea są instytucjami gromadzącymi i przetwarzającymi wiedzę. Najistotniejszą częścią tego dynamicznego zasobu są informacje o zabytkach. Zapisane są one w tradycyjnych księgach inwentarza, a także w inwentarzu elektronicznym. W przeciwieństwie do ksiąg papierowych, dane zapisane w formie cyfrowej są danymi aktywnymi nie tylko podlegają ciągłej aktualizacji (są poprawiane i uzupełniane), ale są przeszukiwane w celu uzyskania różnych informacji. Sposób organizacji tych informacji determinuje efekt każdej kwerendy. W systemach muzealnych służących do ewidencjonowania zabytków, charakter wprowadzanych informacji może być dwojaki, w zależności od definicji konkretnego pola. Z jednej strony są to pola wolno-tekstowe, w które użytkownik może wpisać dowolną informację. W polach takich jak „tytuł”, „opis fizyczny”, „opis ikonograficzny” czy „glossa” nie jest możliwe zastosowanie zamkniętego zasobu słownictwa, choć wymagane jest, aby w każdym muzeum określono standardy zapisywania tychże. Z drugiej strony, informacja zapisywana jest w polach słownikowych, są to miejsca gdzie z uwagi na powtarzalność użycia oraz nierozbudowany charakter poszczególnych jednostek leksykalnych daje się zastosować zamknięty zasób słownictwa. Postać słownikową mają informacje merytoryczne, takie np. jak: twórca, tworzywo (materiał), technika, rodzaj (klasyfikacja przedmiotowa), miejsca (powstania, znalezienia, użytkowania), słowa kluczowe, a także informacje administracyjne, odnoszące się do statusu danego obiektu (w tym np.: lokalizacja, kontrahenci, działy muzeum).

Sposób organizacji danego słownika może przyjąć różnoraki charakter. Każdy ze słowników, w zależności od ilości haseł, przewidywanego przyrostu i możliwości kontroli może przyjąć inną formę haseł kontrolowanych lub tezaursów.

Hasła kontrolowane są inaczej nazywane listami autoryzowanymi lub indeksami. Głównym założeniem takiej listy jest alfabetyczne uszeregowanie i stosowanie określonych zasad przy tworzeniu każdego hasła. Indeksy często przybierają formę hierarchiczną, co pozwala na zobrazowanie prostych zależności. Przykładem zastosowania prostej hierarchii są słowniki nazw pomieszczeń, tutaj możliwe staje się odzwierciedlenie struktury danego budynku i danego magazynu, możliwe jest wyszukanie wszystkich obiektów z danego zbioru/ kolekcji, aktualnie znajdujących się w konkretnym pomieszczeniu, co może być szczególnie pomocne przy przeprowadzaniu skontrum.

Przyrost danych odbywa się w sposób kontrolowany. Zaleca się, aby w danej instytucji była wyznaczona osoba, która weryfikuje każde nowo wprowadzone hasło, bądź sama je wprowadzała. Forma indeksu często stosowana jest w przypadku haseł osobowych i korporatywnych, pozwala na wyszukanie z jednoczesnym częściowym podglądem listy, co sprawia, że rozwiązanie to jest szczególnie popularne w internetowych katalogach zbiorów. Użytkownik internetowy nie zna słownictwa danego zbioru, możliwość podglądu poszczególnych stron indeksu ułatwia wyszukiwanie/przeglądanie. W wewnętrznym muzealnym systemie bazodanowym

większość zapytań do bazy ma charakter nieco bardziej ukierunkowany, stąd forma słownika – pola wyszukiwawczego – jest bardziej uzasadnione.

Tezaury to słowniki, których hierarchiczna struktura obrazuje wzajemne zależności semantyczne poszczególnych pojęć. Umożliwiają one określanie preferowanych i niepreferowanych pojęć poprzez zastosowanie odsyłaczy zarówno pomiędzy zależnościami równoległymi jak i dalszymi⁵⁶. W tego typu słownikach powinno się – przewidzieć osobne pola na wpisywanie oboczności, co jest szczególnie ważne z przypadku haseł osobowych i stosowanych różnych form pisowni. Zawartość pola oboczności powinna być ujęta w głównym wyszukiwaniu, lecz z zaznaczeniem, że jest to askryptor. W przypadku istnienia złożonych zależności relacyjnych pomiędzy hasłami, stosowane są słowniki polihierarchiczne, szczególnie przydatne w rozbudowanych i szybko przyrastających słownikach, jak np./ słowniki słów kluczowych, a także w słownikach nazw geograficznych, czy nazw przedmiotów.

Słowniki o charakterze administracyjnym powinny być słownikami zamkniętymi i zarządzanymi przez jedną osobą, która jako jedyna może dokonywać w nich modyfikacji. W przypadku bardziej rozbudowanych słowników – jak słownik słów kluczowych – stosowanie zależności hierarchicznych (hasła nadrzędne, podrzędne) może bardzo ułatwić uporządkowanie całego zasobu. Stworzenie szczegółowego szkieletu hierarchii jest procesem czasochłonnym i wymagającym zaangażowania grupy osób, w tym przedstawicieli poszczególnych specjalizacji. Korzystnym rozwiązaniem jest podzielenie pracy na etapy – poziomy hierarchii, poczynając od bardzo ogólnych podziałów (przykładowo – nawiązując do Iconclass⁵⁷ – „historia”, „biblia”, „natura”, „literatura”, „mitologia” itp.), a następnie stopniowe opracowywanie konkretnych gałęzi danego działu.

Praca nad kontrolą każdego słownika powinna być poprzedzona ustaleniem podziału zadań i wyznaczeniem osób kontrolujących słownik, w miarę potrzeby stosowanie kontroli uprawnień przy pracy nad słownikami. Niekorzystnym rozwiązaniem jest, aby każdy pracownik miał swobodny dostęp zarówno do rejestracji nowych, jak i do procesu usuwania istniejących haseł. Jeśli nie jest możliwe zaangażowanie jednej konkretnej osoby do administracji treści, należy wyznaczyć reprezentantów w poszczególnych działach. Przy pracy nad porządkowaniem istniejącego zasobu należy wyznaczyć zadania oraz zasady pracy – jednym z głównych zadań jest eliminacja homonimii i synonimii. System bazodanowy powinien być wyposażony w narzędzia ułatwiające ujednolicanie poszczególnych haseł. Następnym zadaniem powinno być zre-

⁵⁶ Szczegółowy sposób budowania tezaurów opisano w dokumentach uzupełniających do *Zaleceń Dotyczących Planowania i Realizacji Projektów Digitalizacyjnych w Muzealnictwie*.

M. Kłos, *Dobre praktyki tworzenia tezaurów* [w:] *Zalecenia Dotyczące Planowania i Realizacji Projektów Digitalizacyjnych w Muzealnictwie*, NIMOZ, Warszawa 2011, s. 115-158. Dok. elektroniczny: <http://www.nimoz.pl/pobierz/305.html> [Dostęp: 31 X 2012]. Przy budowie słownika pomocne mogą być wybrane publikacje i normy: L.A.Bielicka, *Metodyka i organizacja opracowania tezaury*, Warszawa 1988; E. Ścibor, J. Tomasiak-Beck, *Metodyka budowy tezaurów*, Warszawa 1995; J. Woźniak-Kasperek, *Podstawy budowy tezaury: poradnik*, Warszawa 2005; *Information and documentation - Thesauri and interoperability with other vocabularies. cz. 1: Thesauri for information retrieval*, Geneve 2011 (ISO 25964-1:2011).

⁵⁷ *Outline of the Iconclass system*. Dok. elektroniczny: <http://www.iconclass.org/help/outline> [Dostęp: 31 X 2012].

dagowanie wszystkich haseł zgodnie z przyjętym standardem. Tak zredagowany słownik powinien być administrowany przez wybraną osobę, aby nie był możliwy jakikolwiek przyrost haseł w niepoprawnej formie.

3.3. Wybór optymalnego rozwiązania

3.3.1. Gotowe rozwiązania

W procesie unifikacji stosowanej terminologii warto uwzględnić międzynarodowe tezaury opisyjące różnorodne aspekty dziedzictwa kulturowego. Rodzaje stosowanych słowników terminologicznych w europejskich muzeach zostały opisane i scharakteryzowane w raporcie w ramach projektu ATHENA⁵⁸. Poniżej przedstawiono jedynie wybrane i najpopularniejsze rozwiązania.

3.3.1.1. Art and Architecture Thesaurus (AAT)

Słownik AAT⁵⁹ został przygotowany przez The Getty Research Institute w Los Angeles (*Vocabulary Program*) i zawiera pojęcia z zakresu sztuk pięknych, architektury, rzemiosła artystycznego, kultury materialnej oraz materiałów archiwalnych, w liczbie ponad 34 tysięcy. Z zakresu AAT wyłączone są pojęcia inne niż ogólne; nie znajdziemy tam np. nazw własnych. Jest to słownik wielojęzyczny, tzn. językiem podstawowym jest angielski (US), lecz terminy słownikowe posiadają odpowiedniki także w innych językach (hiszpański, niderlandzki). Podstawowymi jednostkami słownika są rekordy zawierające dane na temat poszczególnych pojęć, takie jak terminy określające pojęcie (ze wskazaniem terminów deskryptorowych), definicja pojęcia, miejsce w hierarchii słownika, źródła, pojęcia pokrewne, unikatowy identyfikator liczbowy pojęcia i inne. Rekordy łączą się w hierarchiczną strukturę, od najbardziej ogólnych do coraz bardziej szczegółowych. Co więcej, konstrukcja słownika pozwala na to, by każde pojęcie mogło być usytuowane w obrębie rozmaitych szerszych kontekstów. Jest to zatem słownik polihierarchiczny: do jednego i tego samego rekordu (pojęcia) można dojść różnymi ścieżkami. Zakres AAT został ujęty w siedem głównych gałęzi (faset):

- 1) „pojęcia pokrewne” (np. „piękno”, „symetria”);
- 2) „cechy fizyczne” (np. „kształt”, „rozmiar”, „faktura”, „kolor”);
- 3) „style i okresy” (np. „barokowy”, „impresjonistyczny”);
- 4) „czynniki” (np. „drukarze”, „manufaktury”, „bakterie”);
- 5) „działania” (np. „tkanie”, „rysowanie”, „wystawa”, „korozja”);
- 6) „tworzywa” (np. „drewno”, „emulsja”, „metal”);
- 7) „obiekty” (np. „autoportrety”, „rzeźba”, „krzesła”).

⁵⁸ M.-V. Leroi, J. Holland, *Identification of existing terminology resources in museums, ECP-2007-DILI-517005, ATHENA, 31 July 2009*. Dok. elektroniczny: <http://www.athenaeurope.org/index.php?en/149/athena-deliverables-and-documents> [Dostęp: 31 X 2012].

⁵⁹ *Art & Architecture Thesaurus*. Dok. elektroniczny: <http://www.getty.edu/research/tools/vocabularies/aat/> [Dostęp: 31 X 2012].

Słownik AAT jest nadal rozbudowywany i aktualizowany. Można go przeglądać online z użyciem wyszukiwania tekstowego lub poprzez nawigację po strukturze albo używać w wersji licencjonowanej, np. w celu wbudowania we własny system elektronicznej ewidencji zbiorów.

3.3.1.2. The Union List of Artist Names (ULAN)

ULAN⁶⁰ to kolejny słownik opracowany przez The Getty Research Institute w Los Angeles w ramach *Vocabulary Program*. Zawiera on zarówno hasła osobowe, dotyczące artystów, jak i hasła korporatywne. Prace nad stworzeniem słownika rozpoczęły się już w 1984 r., głównie z myślą o stworzeniu ustandaryzowanego i kontrolowanego słownictwa dla potrzeb projektów prowadzonych przez J. Paul Getty Trust. W ciągu lat 90. XX w. struktura słownika była zmieniana tak, aby odpowiadała światowym standardom dotyczącym budowy tezaurusów. W 1994 r. opublikowano papierową wersję pt. *Union List of Artist Names*⁶¹. Obecnie słownik liczy ponad 293 000 haseł. Każde hasło (a w zasadzie każdy rekord) artysty otrzymuje unikalny numer (ID), co jest bardzo pomocne przy ewentualnej identyfikacji haseł.

Słownik umożliwia przeszukiwanie poprzez nazwisko twórcy – wpisanym precyzyjnie lub z wykorzystaniem standardowego maskowania znaków (*, ?), przeszukiwanie po numerze ID, wg kryterium roli oraz narodowości. Dwie ostatnie wartości mogą być wpisane ręcznie lub wybrane z podpiętych indeksów. Słownika ULAN nie można uznać za tezaurus polihierarchiczny, choć jego twórcy przewidzieli możliwość wyszukiwania wg dość prostej hierarchii o następujących szczeblach tego samego poziomu: hasła korporatywne, nie artyści, artyści, nieznani twórcy.

Każdy rekord zawiera zapis preferowany np. „Chełmoński, Józef (Polish painter and illustrator, 1850-1914)”, zapis zgodny z porządkiem naturalnym (*display*), inne zapisy nazwiska o charakterze askryptorów a także dodatkowe informacje o artyście m.in.: „rola”, „narodowość”, „płeć”, „osoby powiązane z artystą”, „odniesienia do innych słowników osobowych uwzględniających danego artystę”.

Słownik ULAN, ze względu na bardzo dużą ilość haseł, stanowi jeden z najpopularniejszych źródeł dotyczących zapisu haseł osobowych i korporatywnych. Choć z pewnością słownik nie jest kompletny, może (a wręcz powinien) stanowić jedno z pierwszych źródeł poszukiwań zapisu danego nazwiska artysty.

⁶⁰ *Union List of Artist Names*, Dok. elektroniczny: <http://www.getty.edu/research/tools/vocabularies/ulan/> [Dostęp: 31 X 2012].

⁶¹ J.M. Bower (project manager), M. Baca (senior editor), *Union List of Artist Names*, G.K. Hall, New York 1994.

3.3.1.3. Getty Thesaurus of Geographic Names (TGN)

Tezaurus nazw geograficznych to jeszcze jeden ważny słownik opracowany przez Getty Research Institute⁶². Jest on przeznaczony do zapisu wszelkiego rodzaju miejsc, jakie mogą być związane z dobrami kultury w jakikolwiek sposób i występujących w jakiegokolwiek roli. Polskie przepisy dotyczące ewidencjonowania muzealiów wymagają podania miejsca powstania dzieła, lecz słowniki geograficzne, takie jak TGN, mogą służyć również do odnotowywania miejsca znalezienia (istotne np. w przypadku zabytków archeologicznych), miejsca pozyskania (zabytki etnograficzne) i innych. Zakres TGN jest bardzo szeroki: ujęto tam podmioty polityczne lub administracyjne, takie jak: „kraje”, „regiony” i „miejscowości”, „cesarstwa”, „królestwa”, „księstwa” i inne domeny, także religijne lub plemienne. Słownik mieści w sobie także nazwy z zakresu geografii fizycznej: „kontynenty”, „wyspy”, „góry i niziny”, „morza”, „jeziora i rzeki”. Ujęto w nim także miejsca dziś opuszczone, państwa, które przestały istnieć, a nawet kilka miejsc pozaziemskich (zapewne służących do opisu kolekcji mineralogicznych). TGN zawiera blisko 900 000 nazw miejsc. Każde z nich opisano w odrębnym rekordzie, opatrzonym unikatowym identyfikatorem numerycznym, przy użyciu informacji takich jak nazwa lub nazwy, jeżeli jest ich więcej – w tym nazwy w wielu językach, nazwy historyczne itd. Ponadto rekord zawiera „współrzędne geograficzne”, „noty objaśniające”, „typ miejsca” i inne jeszcze dane dodatkowe. Struktura TGN jest hierarchiczna, ułożona zasadniczo wg aktualnego porządku świata, jednak zastosowanie polihierarchii pozwoliło uwzględnić także sytuacje historyczne, kiedy to dane miejsce wchodziło w skład większej całości (kraju, regionu), innej niż dzisiejsza.

3.3.1.4. British Museum Materials Thesaurus i British Museum Object Names Thesaurus

Cyfrowa dokumentacja ewidencyjna zbiorów British Museum rozpoczęła się w 1979 r. od działu etnografii, obejmując do 1988 r. stopniowo wszystkie działy. Od 1993 r. pozyskiwane przez muzeum obiekty ewidencjonowane są elektronicznie. W 1999 r. struktura danych została dostosowana do standardu Spectrum. W Internecie dostępne są dwa opracowane przez muzealników z Londynu tezaury: British Museum Materials Thesaurus⁶³ i British Museum Object Names Thesaurus⁶⁴.

British Museum Materials Thesaurus został opublikowany w 1997 roku⁶⁵ i w tej wersji obejmuje 2413 terminów (2129 deskryptorów i 284 askryptory oraz posiada 513 relacji między deskryptorami). Powstał na potrzeby elektronicznej dokumentacji ewidencyjnej zbiorów British

⁶² *Getty Thesaurus of Geographic Names*. Dok. elektroniczny: <http://www.getty.edu/research/tools/vocabularies/tn/index.html> [Dostęp: 31 X 2012].

⁶³ *British Museum Materials Thesaurus*. Dok. elektroniczny: http://www.collectionslink.org.uk/assets/thesaurus_bmm/matintro.htm?phpMyAdmin=OYNyINPdn3sQmoXugKH1gcCLSW0#Editors [Dostęp: 31 X 2012].

⁶⁴ *British Museum Object Names Thesaurus*. Dok. elektroniczny: http://www.collectionslink.org.uk/assets/thesaurus_bmon/Objintro.htm [Dostęp: 31 X 2012].

⁶⁵ T. Szrajber, *The British Museum Materials Thesaurus*, MDA, ISBN-10 1900642026, ISBN-13 978-1900642026.

Museum, głównie w celu przeprowadzania kwerend w oparciu o tworzywo, z którego powstał obiekt.

Tezaurus obejmuje specjalistyczną terminologię, archaizmy i lokalne określenia dla niektórych materiałów, co wynika z powszechności zastosowania różnego rodzaju słownictwa. Podstawą opracowania były scalone indeksy danych użytych do opisanie materiału w ewidencji elektronicznej muzeum. Stworzony w ten sposób tezaurus nie jest naukową systematyką, a odzwierciedleniem języka i terminologii stosowanej przez kuratorów British Museum. Tezaurus podzielono na trzy główne grupy, do których należą poszczególne materiały (uporządkowane w węższych podrzędnych grupach): „organiczny” (*Organic*), „nieorganiczny” (*Inorganic*) i „materiał przetworzony” (*Procseed Material*). Przypisując materiały trudne do zaklasyfikowania tylko do jednej z grup, kierowano się zasadą użyteczności przyjętego rozwiązania dla użytkowników. W ramach grupy „materiałów organicznych” główne kategorie to „materiały zwierzęce” i „materiały roślinne”. Kolejne kategorie to „węglowodory” (obejmujące m.in.: „asfalt”, „bitumin”, „gagat”, „węgiel”, „dziegieć”, „smołę”) oraz „klejnoty organiczne” (obejmujące m.in.: „bursztyn”, „kopal”, „koral”, „macica perłowa”, „perła, muszla”, „roślinna kość słoniowa” (tagua). W grupie „materiałów nieorganicznych”, dwie podstawowe kategorie to „minerały” i „kamienie”, jednakże – jak wskazano powyżej – kamienie ozdobne pochodzenia organicznego włączono do grupy „materiały organiczne”. „Materiały przetworzone” obejmują materiały poddane obróbce, przy czym nie określono zakresu obróbki, który stanowi kryterium do włączenia materiału do omawianej grupy. Do „materiałów przetworzonych” włączono „metale”, przyjmując, że w większości przypadków są one poddane obróbce przed użyciem do produkcji. Ważną grupą „materiałów przetworzonych” są „tekstyli”, obejmujące wszelkie tkaniny niezależnie od sposobu wykonania. Ponadto, do „materiałów przetworzonych” należą „materiały syntetyczne”, obejmujące „tworzywa sztuczne” (np. „wełnę syntetyczną”). W „materiałach nieorganicznych” i „materiałach przetworzonych” występuje ponadto grupa „lime”, odnosząca się do różnych mineralnych i przemysłowych form tlenku wapnia, których rozróżnienie wymaga przeprowadzenia specjalistycznych analiz.

Podsumowując, należy pamiętać, że British Museum Materials Thesaurus to jeden z wielu możliwych układów hierarchiczno-relacyjnych pojęć z zakresu materiał/tworzywo i należy podejść do niego krytycznie, uwzględniając specyfikę gromadzonych zbiorów i potrzeby rodzimnej instytucji w tym zakresie. Decyzja o wdrożeniu tezaury bez edycji wymaga szczegółowej analizy jego struktury i treści, zwłaszcza pod kątem sprawdzenia, które typy materiałów potraktowano jako główne, a które jako marginalne. Układ tezaury umożliwia dodawanie nowych pojęć, jednakże ewentualna zmiana już istniejącego układu relacji (np. przez zmianę powiązań hierarchicznych) będzie wymagać dużego nakładu pracy.

W latach 80. XX wieku powołano w British Museum grupę roboczą, której zadaniem była analiza terminologii używanej w nazewnictwie obiektów gromadzonych w muzeum, w celu stworzenia tezaury, systematyzującego je w układzie hierarchicznym. Bazowy zestaw terminów

pozyskano z dokumentacji źródłowej, takiej jak: słowniki, karty katalogowe i katalogi. W trakcie prac korzystano z konsultacji kuratorów oraz szeregu publikacji, a architektura samego tezaurusa została oparta na normie ISO 2788:1986 *Documentation – Guidelines for the establishment and development of monolingual thesauri* (zastąpiona w 2011 r. przez ISO 25964-1:2011). Utworzony tezaurus został udostępniony online w 1999 r. dla celów niekomercyjnych.

Tezaurus ma strukturę hierarchiczną i składa się 5611 terminów (deskryptorów i askryptorów) podzielonych na 33 grupy (deskryptory główne), przy czym twórcy tezaurusa uwzględnili zasady ewidencji obowiązujące w British Museum, przyjmując, że jeden obiekt może być opisany kilkoma terminami np. „rytualny sztylet”, zostanie opisany przy pomocy dwóch deskryptorów: „sztylet”, „wyposażenie religijne/rytualne”. Jest to istotne choćby ze względu na opis obiektów, które użytkowane były w różnych celach. Wszystkie pojęcia w tezaurusie muszą być unikatowe, dlatego homonimy (w języku polskim np. zamek – budowla, urządzenie, zamek krawiecki) są rozróżniane przez użycie określników (w języku angielskim np. *smoking-pipe* – fajka do palenia, *musical pipe* – piszczałka).

Jako zasadę przyjęto stosowanie nazw w liczbie pojedynczej, chyba, że powszechna jest liczba mnoga (w języku polskim np. drzwi, nożyczki, spodnie). Znak łącznika (-) służy do obrazowania relacji między elementami nazwy bądź zwraca uwagę na dwufunkcyjność przedmiotu. Natomiast zastosowanie apostrofu (‘) wiąże się z regułami pisowni angielskiej bądź wskazania relacji, która nie może być zapisana inaczej (np. koło garncarskie – *potter’s wheel*). W przypadku nazw obcych (bez użycia obcych znaków/liter), nazwa taka jest obowiązującym terminem, dla którego angielski ekwiwalent jest pojęciem szerszym (deskryptor szerszy) bądź pojęciem kojarzeniowym (deskryptor kojarzeniowy). Litery i znaki spoza alfabetu obowiązującego w pisowni języka angielskiego zostały zachowane tam, gdzie było to możliwe, a pisownia angielska została podana jako niepreferowany termin (askryptor), w celu poprawienia efektywności wyszukiwania (niepreferowany termin posiada odniesienie do terminu preferowanego).

W tezaurusie jako zasadę przyjęto nie uwzględnianie nazw elementów przedmiotu („uszek filiżanki”) i wprowadzanie ich pod nazwą całego przedmiotu („filiżanka”). Takie rozwiązanie zapewnia konsekwentny i zgodny zapis nazw we wszystkich rekordach odnoszących się do jednego obiektu, niezależnie od tego, czy poszczególne elementy zostały oddzielone od całości. O kompletności obiektu informuje natomiast opis fizyczny przedmiotu (lub jego elementu), zawarty w polu „opis”. Od tej zasady istnieje kilka wyjątków dla następujących przypadków: gdy element obiektu jest ważniejszy w wyszukiwaniu niż nazwa całego obiektu (np. dachówka – w języku angielskim *roof-tile*); gdy nieznana jest nazwa całego przedmiotu („uchwyt”, „ostrze”, „wieko”); gdy nazwa składnika jest tak odległa od całego przedmiotu, że użycie nazwy całego przedmiotu może być mylące (np.: nit ze statku jest wpisywany jako „nit”, a nie „statek”).

W celu odzwierciedlenia szczególnego przeznaczenia, bądź zastosowania przedmiotów wprowadzono w tezaurusie znaczną liczbę węższych pojęć (węższych deskryptorów) jak np. „łyżeczka do kawy” (ang. *coffee-spoon*). Zasada ta ogranicza się jedynie do oryginalnie obowiązują-

jących terminów, a nie specjalnie tworzonych pojęć na potrzeby tezaurusa. Termin „wyposażenie” używany jest w tezaursie jako dopełnienie poszczególnych aktywności lub koncepcji. Dla przykładu, „instrument muzyczny” jest węższym pojęciem „wyposażenia muzycznego” (deskryptor szerszy).

Dla pozostałości lub części roślin stosowany jest termin „szczątki roślinne” (ang. *vegetal remains*); analogicznie występuje termin „szczątki zwierzęce” (ang. *animal Remains*) oraz „szczątki ludzkie” (ang. *animal remains*). Dla produktów spożywczych, w tym przedmiotów z nich wykonanych (wówczas w pole nazwy wprowadzane jest więcej terminów z tezaurusa) stosuje się termin „produkt spożywczy” (ang. *food product*). Dla wyróżnienia próbek różnych materiałów stosowany jest deskryptor „próbka” (ang. *sample*). Surowce przeznaczone do produkcji powinny być wprowadzane pod jedną z wyżej wymienionych grup (*animal remains, vegetal remains, sample*), a konkretne określenie surowca powinno zostać wprowadzone do pola materiał. Dla niezidentyfikowanych obiektów archeologicznych stosowany jest termin „artefakt” (ang. *artefact*), posiadający wąską grupę węższych deskryptorów.

W tezaursie nie są stosowane rzeczowniki zbiorowe z wyjątkiem rzeczownika „zestaw” (ang. *set*, np. *chess-set* – zestaw do gry w szachy, *tea-set* – serwis do herbaty).

W tezaursie nie utworzono złożonych terminów dla zabawek i modeli (przy czym jest kilka wyjątków od tej zasady). Należy je dodać do nazwy obiektu jaki reprezentują, np. model statku należy zapisać przy użyciu dwóch deskryptorów: „model”, „statek”. Ponieważ terminy „model” i „figura” (ang. *figure*) są bliskoznaczne, stosowana jest dla nich relacja kojarzeniowa (asocjacja, deskryptory kojarzeniowe), przy czym przyjęto zasadę, że termin „figura” odnosi się do podmiotów świata ożywionego, a „model” – przedmiotów świata nieożywionego.

W przypadku obiektów stanowiących: „replikę”, „falszerstwo”, „odlew”, należy uwzględnić taki termin obok nazwy samego przedmiotu, np.: „replika, miecz”. Określenia tego typu zgromadzone są hierarchicznie pod deskryptor główny „produkcja/powielanie” (ang. *production/replication*).

Nazwy, których nie udało się zaklasyfikować do innych grup, przypisano do deskryptora głównego „rozmaite wyposażenie” (ang. *Miscellaneous equipment*). Pojęcia „tekstylija/tkanina” (ang. *textile/cloth*) – odnoszące się do tekstylii, nie mają relacji z innymi deskryptorami w tezaursie, do ich przeszukiwania służy grupa „tekstylija” w tezaursie „materiał”.

Podsumowując, bogactwo zbiorów gromadzonych przez British Museum przekłada się na wielkość tezaurusa nazw przedmiotów. Obecnie w Internecie dostępna jest kolekcja British Museum online z ponad 2 000 000 obiektów, w opisach których zastosowano British Museum Object Names Thesaurus i British Museum Materials Thesaurus.

Przeszukując prezentowane zbiory można samodzielnie ocenić efektywność obu tezaursów.

Do omawianego tezaurusa mają zastosowanie uwagi zawarte w podsumowaniu opisu British Museum Materials Thesaurus.

3.3.1.5. Rijksbureau voor Kunsthistorische Documentatie (RKD)

RKD⁶⁶ stanowi jedną z najważniejszych światowych instytucji gromadzących dokumentację dotyczącą sztuki europejskiej, poczynając od średniowiecza do współczesności. Załącznikiem kolekcji RKD była dokumentacja dotycząca sztuki holenderskiej. Duża część zbiorów jest udostępniona w postaci internetowych baz danych również w postaci słowników RKD artists⁶⁷.

Słownik zawiera hasła osobowe dotyczące artystów, a także marszandów, kolekcjonerów, historyków sztuki. Obecnie obejmuje około 250 tys. haseł. Nowe rekordy rejestrowane są każdego dnia. Słownik jest dwujęzyczny (angielsko-holenderski), słownik słów kluczowych, nazwy geograficzne, role artystów i rodzaje służące jako słowniki powiązane z głównym słownikiem osobowym prowadzone są wyłącznie w języku holenderskim. Najobszerniej w słowniku opracowane są hasła osobowe artystów holenderskich i flamandzkich.

Twórcy słownika umożliwiają użytkownikowi zarówno przeszukiwanie proste wg nazwiska twórcy, ale także przeszukiwanie w kontekście dotyczącym szeroko pojętej działalności artysty, z możliwością samodzielnego określenia precyzji wyszukiwania z użyciem takich kryteriów jak:

„Fragment nazwy” – pole, które pozwala na przeszukiwanie wg frazy lub kombinacji znaków występujących w nazwisku; pole to jest również przydatne, kiedy użytkownik nie jest pewny pisowni danego nazwiska; „monogram”; „szkoła/narodowość”; „miejsce urodzenia”; „miejsce śmierci”; „miejsca działalności”; „data urodzenia”; „data śmierci” – z możliwością podania przybliżonego okresu; „okres działalności”; „rola artysty”; „temat (wyszukiwanie wg słów kluczowych, typów ikonograficznych); „tworzywo”; „rodzaj”; „nazwa projektu badawczego”.

Lista wyników obejmuje deskryptor wraz z listą askryptorów, standard zapisu odpowiada hasłom preferowanym w słowniku ULAN Getty. Wygodną formą korzystania ze słownika jest możliwość wydruku lub ściągnięcia listy wyników do pliku tekstowego, co ułatwia aktywne korzystanie z bazy.

Precyzyjnie opracowane hasła osobowe zawierają biogram artysty z bardzo rozbudowaną sekcją dotyczącą *oeuvre* i miejsc działalności, bibliografią, dokumentacją źródeł i – co bardzo istotne – z listą artystów powiązanych („nauczyciele”, „uczniowie”, „naśladowcy”).

Słownik jest szczególnie przydatny w przypadku artystów holenderskich i flamandzkich, choć wiele haseł artystów europejskich również odczekało się szczegółowego opracowania. Przy budowie własnego słownika osobowego stanowi świetne komplementarne narzędzie dla słownika ULAN Getty. Jego przewagą jest naukowe opracowanie każdego hasła, co pozwala na jego podstawie budować słownik z pełną dokumentacją dotyczącą danego artysty.

⁶⁶ *Rijksbureau voor Kunsthistorische Documentati*. Dok. elektroniczny: <http://website.rkd.nl/home> [Dostęp: 31 X 2012].

⁶⁷ *RKDartists&*. Dok. elektroniczny: <http://www.rkd.nl/rkddb/%28S%281i0tn0qng0j1q0vex2ajcwg1%29%29/default.aspx> [Dostęp: 31 X 2012].

3.3.1.6. Virtual International Authority File (VIAF)

VIAF⁶⁸ powstał na potrzeby sieci semantycznej, a celem jego działania jest ujednoczenie dostępu do słowników nazw osobowych w zasobach sieciowych różnych instytucji, przede wszystkim w bibliotekach. Serwis VIAF zapoczątkowały Biblioteka Kongresu (Library of Congress), Niemiecka Biblioteka Narodowa (Deutsche Nationalbibliothek), Biblioteka Narodowa Francji (Bibliothèque Nationale de France) oraz Online Computer Library Center (OCLC). VIAF pozwala przezwyciężyć nieuniknione różnice, jakie powstały przy tworzeniu wzorcowych list haseł (ang. *authority files*) w różnych instytucjach centralnych, działających, jak dotąd, zwykle na szczeblu krajowym. Po powstaniu w ostatnich latach globalnych agregatorów (czyli katalogów zbiorowych (takich jak np. WorldCat prowadzony przez OCLC⁶⁹), owe różnice powodowały oczywiste trudności w uzyskiwaniu jednolitych wyników wyszukiwania. VIAF przezwycięża tę trudność nie poprzez próby ujednoczenia zapisu nazw osobowych we wszystkich instytucjach, ale poprzez platformę, która zbiera wszystkie warianty krajowe i prezentuje je wspólnie, pod jednym, unikatowym identyfikatorem numerycznym.

VIAF ID: 102895066 (Personal)
Permalink: <http://viaf.org/viaf/102895066>

Preferred Forms

- 100 0 _ †a Marcus Aurelius, †c Emperor of Rome, †d 121-180
- 100 0 _ †a Marc Aurèle, †c empereur de Rome, †d 121-180
- 100 0 _ †a Marc Aurèle, †c empereur romain, †d 0121-0180
- 100 1 _ †a Aurelius Antoninus, Marcus, †c imperatore romano, †d 121-180
- 100 0 0 †a Marco Aurelio, †c Emperador de Roma
- 100 1 _ †a Marcus Aurelius, Emperor of Rome †g Roman emperor and patron, 121-180 CE
- 100 1 _ †a Aurelius Antoninus, Marcus
- 100 0 _ †a Aurelius Antoninus, Marcus, †c imperatore romano, †d 121-180
- 100 1 _ †a Marcus Aurelius, Antoninus, †c římský císař, †d 121-180
- 100 0 _ †a Марк Аврелий, †c Римский Император, †d 121-180
- 200 _ | †a Marc Aurèle †c empereur romain †f 0121-0180
- 100 0 _ †a Mark Aurel †c Römisches Reich, Kaiser †d 121-180
- 100 0 _ †a Marcus Aurelius Antoninus, †c empereur romain
- 100 0 _ †a מרקוס אורליוס, †c קיסר רומן, †d 121-180
- 100 0 _ †a Marek Aureliusz †c (cesarz rzymski ; †d 121-180).

NUKAT Center (Poland)
NUKAT|n 96103790

Ilustracja 19. Przykład rekordu w VIAF, ilustrujący różne sposoby zapisu nazwiska cesarza Marka Aureliusza, przez rozmaite biblioteki, w różnych wersjach, językach i alfabetych. Wersja polska, oznaczona biało-czerwoną flagą, pochodzi z katalogu NUKAT. Źródło: dok. elektron. <http://viaf.org/viaf/102895066> [odczyt: 31 X 2012].

Dzięki takiemu rozwiązaniu dotychczasowe listy haseł wzorcowych mogą współistnieć. Wystarczy jedynie dla każdej nazwy osobowej, niezależnie od tego jak wygląda w danej instytucji jej lokalnie ustandaryzowany zapis, powołać się dodatkowo na identyfikator, aby pozbyć się

⁶⁸ VIAF: *The Virtual International Authority File*. Dok. elektroniczny: <http://viaf.org> [Dostęp: 31 X 2012].

⁶⁹ WorldCat.org: *The World's Largest Library Catalog*. Dok. elektroniczny: www.worldcat.org [Dostęp: 31 X 2012].

potencjalnych niejednoznaczności. Prócz identyfikatora numerycznego, w serwisie VIAF znajduje też zastosowanie stały odnośnik (*permalink*), pozwalający na bezpośrednie odwołanie do danego rekordu.

Warto zauważyć, że VIAF czerpie dane nie tylko z zasobów bibliotek, gdzie nazwiska artystów trafiają się o tyle, o ile artyści owi występują w roli autorów materiałów bibliotecznych (książek, ilustracji książkowych, grafik) lub są im poświęcone jakieś pozycje. Jedną z instytucji współpracujących z tym serwisem jest Getty Center, który udostępnia dane z opisanego wyżej Union List of Artist Names (ULAN), dzięki czemu jest to narzędzie istotne również do celów ewidencjonowania dzieł sztuki.

Wizerunki	
	
Nazwisko / Nazwa	Boznańska, Olga (1865-1940)
Opis	1886-1889 studia w Akademii Monachijskiej; 1898 osiadła w Paryżu; malowała głównie portrety, martwe natury; zwracała szczególną uwagę na barwę, wykazując predylekcję dla tonów chłodnych i srebrzysto-szarych, unikała wyraźnych linii i konturów
Rola domyślna	malarz
Kontrahent	
Bibliografia	<i>Virtual International Authority File</i> . Dublin (Ohio, USA): OCLC Online Computer Library Center, Inc.. 2012-06-15. Dostępny w internecie: http://viaf.org/ . http://viaf.org/viaf/18166708 .

Ilustracja 20. Przykład hasła w słowniku nazw osobowych, w którym umieszczono odwołanie do VIAF
Opracowanie: T. Zaucha.

3.3.1.7. ICONOCLASS

Nazwa tego słownika jest akronimem od ICONographic CLASsification System – System Klasyfikacji Ikonograficznej⁷⁰. Jest to dzieło prof. Henri van de Waala z uniwersytetu w Lejdzie, który pracował nad nim od lat 50. XX w. Iconclass został wydany drukiem w latach 1972-1985. Książkowe wydanie, które składa się z siedmiu tomów zawierających cały system, trzech tomów indeksu alfabetycznego do systemu oraz obszernej bibliografii ikonograficznej, która uporządkowana jest wg klasyfikacji Iconclass. System ten nie jest *de facto* słownikiem, który za-

⁷⁰ *Iconclass: a multilingual classification system for cultural content*. Dok. elektroniczny: www.iconclass.nl [Dostęp: 31 X 2012].

wierałby właściwą terminologię, lecz zbiorem tematów i motywów, jakie możemy odnaleźć w szeroko pojętej sztuce zachodniej. Wszystkie one zestawione są w obszerną taksonomię, liczącą około 28 tysięcy jednostek, która obejmuje dziewięć podstawowych grup:

1. Religia i magia.
2. Natura
3. Istota ludzka, Człowiek w ogólności.
4. Społeczeństwo, Cywilizacja, Kultura.
5. Abstrakcyjne idee i pojęcia.
6. Historia.
7. Biblia.
8. Literatura.
9. Mitologia Klasyczna i Historia starożytna.

Każda z powyższych faset rozwija się na jednostki niższego rzędu, w których do pierwszej cyfry, dodana jest kolejna i kolejna (trzeci znak w tej notacji, czyli trzeci poziom hierarchii oznaczony jest literą). Tak np. symbol 73 B 141 oznacza temat: „*Pasterze podążają do Betlejem*”, a jest on rezultatem następującego rozwinięcia hierarchii:

```

7      Biblia
73     └─ Nowy Testament
73 B 1     └─ Historia narodzenia Chrystusa
73 B 14    └─ Zwiastowanie narodzenia Chrystusa pasterzom
73 B 141   └─ Pasterze podążają do Betlejem
    
```

Ilustracja 21. Fragment hierarchicznej struktury słownika Iconclass. Opracowanie: T. Zaucha.

Każde hasło jest ponadto dopełnione słowami kluczowymi, które ułatwiają wyszukanie potrzebnego przedstawienia i zawiera odnośniki do podobnych tematów.

Iconclass jest najobszerniejszą taksonomią ikonograficzną, co więcej, jest to narzędzie wielojęzyczne. Podstawowe, książkowe wydanie tej klasyfikacji, przygotowane zostało w języku angielskim, jednak dotąd przetłumaczono ją już na języki: niemiecki, francuski, włoski oraz fiński. W tym momencie, poza wydaniem książkowym, istnieje także elektroniczna wersja Iconclass, którą utrzymuje Rijksbureau voor Kunsthistorische Documentatie w Hadze. Dzięki wyczerpującemu opisowi tematów ikonograficznych Iconclass jest bardzo dobrym narzędziem do opisu zbiorów sztuki przedstawiającej w zbiorach muzealnych i nie tylko. Jednym z największych zasobów, które używają tej klasyfikacji jest Deutsches Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg, jedna z największych światowych fototek. W Polsce Iconclass został zaimplementowany do celów klasyfikacji zbiorów w Muzeum Narodowym w Krakowie (we fragmentarycznym tłumaczeniu na jęz. polski)⁷¹.

⁷¹ *Katalog zbiorów [Muzeum Narodowego w Krakowie]*. Dok. elektroniczny: www.katalog.muzeum.krakow.pl [Dostęp: 31 X 2012].

3.3.2. Budowa własnego słownika

Słownictwo używane w muzealnym systemie do ewidencjonowania zbiorów dotyczy zarówno podstawowego opisu danego muzealium, jego naukowego opracowania, a także administracji i zarządzania kolekcją. Ilość i zakres stosowanych słowników w danym muzeum będzie warunkowany przyjętym schematem opisu zbiorów, użytym systemem bazodanowym. Powinny one przede wszystkim odpowiadać wewnętrznym potrzebom danej instytucji: słowniki merytoryczne powinny być dostosowane do charakteru gromadzonych zbiorów, natomiast słowniki administracyjne powinny odpowiadać wewnętrznej strukturze organizacyjnej. Pierwsza grupa powinna być opracowywana wspólnie dla muzeów o podobnej charakterystyce, druga – z racji specyfiki – będzie tworzona indywidualnie zgodnie z potrzebami danej instytucji. Poniżej przedstawiono przykładowe słowniki używane w Muzeum Narodowym w Warszawie:

Słowniki Merytoryczne

Twórcy

Nazwy geograficzne

Tworzywa

GUS

Techniki

Rodzaje

Słowa Kluczowe

Rola (twórcy/miejsca)

Wystawy

Bibliografia

Słowniki Administracyjne

Działy muzeum

Pomieszczenia

Klasyfikacja

Cel wydania/przyjęcia

Kontrahenci

W zależności od rodzaju słownika, a przede wszystkim stopnia jego złożoności, poszczególne słowniki przybiorą formę prostych list autoryzowanych, indeksów, czy rozbudowanych tezaurusów.

Idealnym rozwiązaniem jest opracowanie standardów budowy słownictwa jeszcze przed rozpoczęciem budowy inwentarza elektronicznego oraz przed podjęciem decyzji o zakupie oprogramowania do ewidencjonowania zbiorów. Wskazane jest, aby oprzeć się na istniejących rozwiązaniach, a nie tworzyć nowe zasady, które mogą okazać się niespójne z największymi dostępnymi zasobami słownikowymi/ bazami terminologicznymi. Pracę nad każdym słownikiem należy rozpocząć od kwerendy dotyczącej wiodących słowników i kartotek haseł wzorcowych w danej dziedzinie. Duże ogólnie dostępne tezaury, przykładowo dotyczące terminologii sztuk pięknych (Getty AAT – Art&Architecture Thesaurus⁷²), mogą nie zawierać specjalistycznego słownictwa, które określa zabytki w zbiorze danego muzeum. Wówczas nieocenionym źródłem wiedzy są opiekunowie kolekcji, na co dzień korzystający z elektronicz-

⁷² *Art & Architecture Thesaurus*. Dok. elektroniczny: <http://www.getty.edu/research/tools/vocabularies/aat/index.html> [Dostęp: 31 X 2012].

nych lub papierowych słowników odpowiednich dla swojej dziedziny, które warto włączyć do listy preferowanych źródeł. Tworząc słownik warto uwzględnić zarówno polskie jak i zagraniczne rozwiązania.

Poniżej przedstawiono przykładową listę źródeł pomocnych przy tworzeniu haseł geograficznych⁷³:

1. Nazwy geograficzne miejscowości Polski⁷⁴.
2. Nazewnictwo geograficzne świata:
 - Zeszyt 1. Ameryka, Australia i Oceania⁷⁵.
 - Zeszyt 2. Bliski Wschód⁷⁶.
 - Zeszyt 3. Afryka⁷⁷.
 - Zeszyt 4. Azja Południowa⁷⁸.
 - Zeszyt 5. Azja Środkowa i Zakaukazie⁷⁹.
 - Zeszyt 6. Białoruś, Rosja, Ukraina⁸⁰.
 - Zeszyt 7. Azja Południowo-Wschodnia⁸¹.
- Źródła pomocnicze:
 - Zeszyt 8. Antarktyka⁸².
 - Zeszyt 9. Azja Wschodnia⁸³.

⁷³ Przykłady z zestawienia *Dobrych praktyk w zakresie opisu muzealiów w Inwentarzu elektronicznym*, przygotowanego przez pracowników Muzeum Narodowego w Warszawie.

⁷⁴ Nazwy geograficzne miejscowości Polski. Dok. elektron. <http://ksng.gugik.gov.pl> [Dostęp: 31 X 2012].

⁷⁵ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 1, Ameryka Australia i Oceania*, Główny Urząd Geodezji i Kartografii, Warszawa 2004. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_01.pdf [Dostęp: 31 X 2012].

⁷⁶ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 2, Bliski Wschód*, Główny Urząd Geodezji i Kartografii, Warszawa 2004. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_02.pdf [Dostęp: 31 X 2012].

⁷⁷ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 3, Afryka*, Główny Urząd Geodezji i Kartografii, Warszawa 2004. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_03.pdf [Dostęp: 31 X 2012].

⁷⁸ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 4, Azja Południowa*, Główny Urząd Geodezji i Kartografii, Warszawa 2005. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_04.pdf [Dostęp: 31 X 2012].

⁷⁹ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 5, Azja Środkowa i Zakaukazie*, Główny Urząd Geodezji i Kartografii, Warszawa 2005. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_05.pdf [Dostęp: 31 X 2012].

⁸⁰ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 6, Białoruś, Rosja, Ukraina*, Główny Urząd Geodezji i Kartografii, Warszawa 2005. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_06.pdf [Dostęp: 31 X 2012].

⁸¹ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 7, Azja Południowo-Wschodnia*, Główny Urząd Geodezji i Kartografii, Warszawa 2006. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_07.pdf [Dostęp: 31 X 2012].

⁸² Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 8, Antarktyka*, Główny Urząd Geodezji i Kartografii, Warszawa 2006. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_08.pdf [Dostęp: 31 X 2012].

⁸³ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 9, Azja Wschodnia*, Główny Urząd Geodezji i Kartografii, Warszawa 2006. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_09.pdf [Dostęp: 31 X 2012].

Zeszyt 10. Morza i oceany⁸⁴.

Zeszyt 11. Europa, Część I⁸⁵.

Zeszyt 12. Europa, Część II⁸⁶.

Nazwy państw świata, ich stolic i mieszkańców (wyd. 3, 2007)⁸⁷.

Getty Thesaurus of Geographic Names⁸⁸.

W oparciu o wzorcowe rozwiązania należy również zaprojektować lub odpowiednio zmodyfikować funkcjonalności używanego oprogramowania. Należy uwzględnić możliwość rozbudowy istniejących słowników w formie indeksów do postaci tezaursów, co umożliwi wpisywanie dodatkowych informacji, a także do wprowadzania w powiązaniu z hasłem preferowanym listy askryptorów. Co ważne, oboczności uznane za askryptor, jeśli mają formę osobnego hasła słownikowego, na zasadzie odsyłacza, powinny kierować do hasła głównego lub mieć jedynie formę hasła podrzędnego w stosunku do deskryptora.

Poniższy przykład pokazuje hasło główne, występujące w słowniku „Arndt, Wilhelm (1750- 1813)”; oboczności, które nie występują na liście słownika, ale dają możliwość kwerendowania, pole przeznaczone na źródła, gdzie podawany jest odpowiedni kod słownika zewnętrznego lub przypis bibliograficzny do słownika papierowego.

⁸⁴ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 10, Morza i oceany*, Główny Urząd Geodezji i Kartografii, Warszawa 2008. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_10.pdf [Dostęp: 31 X 2012].

⁸⁵ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 11, Europa Część I*, Główny Urząd Geodezji i Kartografii, Warszawa 2009. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_11.pdf [Dostęp: 31 X 2012].

⁸⁶ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazewnictwo Geograficzne Świata, Zeszyt 12, Europa, Część II*, Główny Urząd Geodezji i Kartografii, Warszawa 2010. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/zeszyty/zeszyt_12.pdf [Dostęp: 31 X 2012].

⁸⁷ Komisja Standaryzacji Nazw Geograficznych Poza Granicami Polski przy Głównym Geodecie Kraju, *Nazwy Państw Świata, ich stolic i mieszkańców*, Główny Urząd Geodezji i Kartografii, Warszawa 2007. Dok. elektroniczny: http://ksng.gugik.gov.pl/pliki/nazwy_panstw_swiate.pdf [Dostęp: 31 X 2012].

⁸⁸ *Getty Vocabularies*. Dok. elektroniczny: http://www.getty.edu/research/conducting_research/vocabularies/tgn/ [Dostęp: 31 X 2012].

Słownik twórców _ □ X

Wyszukiwanie

Podaj nazwę (nazwisko, imię):

	Nazwisko i imię lub nazwa	rola
	Arndt, Leo (1857-post 1909)	grafik
<input checked="" type="checkbox"/>	Arndt, Wilhelm (1750-1813)	grafik
	Arnold, Carl Johann (1829-post 1858)	grafik

Dane szczegółowe

Nazwisko i imię lub nazwa:

Oboczności:

Źródła:

Rola główna:

Biogram:

Uwagi dodatkowe:

Ilustracja 22. Hasło główne słownika twórcy. Opracowanie: K. Tabak.

RKDartists&:52683
ULAN:500115789
TGN:7006952
Engen, Dictionary of Victorian Engravers (1979)
VIAF ID:96063195

Ilustracja 23. Przykładowe wartości pola źródła. Opracowanie: K. Tabak.

ULAN	Union List of Artist Names® Online
TGN	Getty Thesaurus of Geographic Names® Online
RKDartists&	RKDartists&
CERL	CERL Thesaurus File
VIAF	The Virtual International Authority File
NUKAT	Katalog NUKAT
PND	Personennamendatei (Deutsche Nationalbibliothek)

Ilustracja 24. Kody zewnętrznych słowników. Opracowanie: K. Tabak.

Proces budowy słownika przebiega równolegle z procesem wprowadzania i opracowywania kolejnych rekordów muzealiów. Na początkowym etapie uniemożliwienie użytkownikom wprowadzania nowych haseł może okazać się kłopotliwe. Dobrym rozwiązaniem jest zatrudnienie, w zależności od ilości osób pracujących w systemie, merytorycznego administratora, który zajmowałby się wprowadzaniem nowych haseł wg zgłoszeń pracowników, zgodnie z ustalonymi wcześniej wytycznymi. Jeśli jednak to rozwiązanie nie jest możliwe, należy przeprowadzić szkolenie pracowników tak, aby słownik był wprowadzony w sposób jednolity, a także należy przeprowadzać okresowe kontrole jakości danych.

W praktyce wiele muzeów już od kilku, a niekiedy od kilkudziesięciu lat, prowadzi elektroniczną ewidencję swoich zbiorów. Stąd ważniejszym pytaniem od tego: „Jak zacząć budowę słownika?”, jest „Jak uporządkować słownik istniejący?”. Analogicznie, jak w przypadku nowego słownika, należy określić preferowane źródła. Drugą, równie ważną czynnością, jest skrupulatna analiza słownika i identyfikacja najczęściej występujących błędów oraz określenie ich skali.

Najczęściej występujące błędy w słownikach:

- występowanie synonimów;
- błędne hasła, niewykorzystane w żadnym rekordzie;
- niepotrzebne dublowanie haseł powiązanych z tym samym artystą np.

Angel, Philips (1616-post 1683);

Angel, Philips (1616-post 1683) – kopia;

Angel, Philips (1616-post 1683) – wg;

- stosowanie zdrobnień (blacha, blaszka);
- brak hierarchii, wymienne stosowanie haseł ogólnych i szczegółowych, np. analogiczne obiekty opisywane są różnymi atrybutami: „tworzywo: drewno”; „tworzywo: drewno różane”.

W przypadku rozbudowanych słowników tworzonych w sposób całkowicie dowolny, pomocą mogą się okazać rozwiązania systemowe, które pozwalają na częściową automatyzację procesu porządkowania.

Narzędzia systemowe pomocne przy porządkowaniu słownika:

- możliwość ujednolicania haseł słownikowych;
- generowanie list haseł niewykorzystanych;
- możliwość hurtowego usuwania wytypowanych haseł (funkcja powinna być dostępna wyłącznie dla administratora);
- możliwość blokowania haseł poprawnych;
- generowanie statystyk dotyczących pracy w słowniku tj. listy przyrostów, listy haseł usuniętych, zmodyfikowanych;
- możliwość wprowadzania trójstopniowej kontroli uprawnień do każdego słownika (możliwość rejestracji, modyfikacji, usuwania);
- możliwość określania haseł poufnych np. związanych z proveniencją, które nie będą widoczne dla osób o ograniczonych uprawnieniach;
- możliwość ustalania zależności hierarchicznych między hasłami w tym również możliwość przenoszenia łącznie całych gałęzi.

Oprogramowanie może okazać się bardzo pomocne, jednakże porządkowanie i ujednolicanie każdego słownika wymaga dużego nakładu pracy. Stąd istotnym elementem procesu jest częste szkolenie użytkowników i prowadzenie okresowych kontroli.

4. Propozycje rozwiązań na poziomie krajowym



4.1. Krajowy agregator danych

Na przestrzeni ostatnich dwóch lat obserwujemy znaczący postęp w zakresie digitalizacji zbiorów muzealnych i coraz częstsze udostępnianie ich za pomocą internetowych katalogów. Zmiany te należy wiązać z wprowadzeniem Wieloletniego Programu Rządowego Kultura+, lecz samo dofinansowanie sektora muzealnego nie przyniosłoby dzisiejszych rezultatów, gdyby nie wzrastająca świadomość muzealników. Liczne inicjatywy wzbudzają szerokie zainteresowanie w środowisku, przyczyniają się także do narastania dyskusji w obliczu pytań i wątpliwości dotyczących całego procesu cyfryzacji. Złożoność zagadnień digitalizacji obiektów muzealnych, brak jednolitych standardów oraz rozproszenie inicjatyw powodują, że coraz częściej podkreśla się konieczność wprowadzenia jednolitych ogólnopolskich wytycznych w tym zakresie.

Konieczność wprowadzenia standaryzacji procesów nie wynika tylko z oczekiwania środowiska muzealnego, ale przede wszystkim warunkowana jest dokumentami unijnymi, a zwłaszcza: Konkluzjami Komisji Europejskiej z 27 października 2011 w sprawie digitalizacji i udostępniania w Internecie dorobku kulturowego oraz w sprawie ochrony zasobów (2011/711/UE)⁸⁹ oraz Konkluzjami Rady Unii Europejskiej z dnia 10 maja 2012 r. w sprawie digitalizacji dorobku kulturowego i udostępniania go w Internecie oraz w sprawie ochrony zasobów cyfrowych⁹⁰.

W wymienionych dokumentach podkreśla się między innymi konieczność tworzenia i rozwijania krajowych agregatorów. Wnioskuje się o stosowanie wspólnych standardów digitalizacji, a zwłaszcza opieranie procesu na dostatecznych opisowych i technicznych metadanych, powiązanych z trwałymi identyfikatorami, które stanowią integralny element zapewniający długoterminową ochronę zasobów cyfrowych. Zwraca się również uwagę na konieczność konsolidacji krajowych strategii i celów digitalizacji poprzez wprowadzenie jasnych kryteriów doboru obiektów do digitalizacji, sposobów ich prezentacji w Internecie, a także nadzorowanie postępów w digitalizacji.

Utworzenie krajowego agregatora umożliwiającego migrację danych dla obiektów muzealnych będzie wymagać wprowadzenia ramowych rozwiązań i wyznaczenia instytucji koordynującej cały proces, która przyjąłaby odpowiedzialność między innymi za:

- utworzenie platformy (portalu) do publikowania zbiorów wraz z jej administrowaniem;
- opracowanie i opublikowanie procedur procesu digitalizacji;
- opracowanie instrukcji i schematu metadanych, odpowiadających polskiemu prawodawstwu (Rozporządzenie o ewidencji z 2004 r. i ewentualne nowsze), w celu umożliwienia

⁸⁹ *Zalecenie Komisji z dnia 27 października 2011 r. w sprawie digitalizacji i udostępniania w Internecie dorobku kulturowego oraz w sprawie ochrony zasobów cyfrowych*, 2011/711/UE. Dok. elektroniczny: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:283:0039:0045:PL:PDF> [Dostęp: 31 X 2012].

⁹⁰ *Konkluzje Rady z dnia 10 maja 2012 r. w sprawie digitalizacji dorobku kulturowego i udostępniania go w Internecie oraz w sprawie ochrony zasobów cyfrowych*, 2012/C 169/02. Dok. elektroniczny: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2012:169:0005:0008:PL:PDF> [Dostęp: 31 X 2012].

wymiany danych o zbiorach między muzeami na poziomie krajowym i międzynarodowym;

- opracowanie słowników terminologicznych wspierających opis obiektów;
- stworzenie zunifikowanych zasad kreowania identyfikatora URI;
- prowadzenie monitoringu postępu procesu digitalizacji, a zwłaszcza systemu kontroli jakości wytworzonych zasobów cyfrowych;
- archiwizacja danych portalu;
- prowadzenie cyklicznych szkoleń dla muzeów, na szczeblu krajowym i regionalnym, z zakresu organizacji procesu digitalizacji i jego realizacji.

4.2. Standaryzacja metadanych

Udostępnienie za pośrednictwem Internetu dziedzictwa kulturowego gromadzonego w polskich muzeach musi zostać poprzedzone wypracowaniem wspólnego schematu metadanych. Jak to wielokrotnie podkreślano w niniejszym dokumencie, jedynie opracowanie jednorodnego zestawu elementów schematu opisu i jasne zdefiniowanie poszczególnych pól może stanowić podstawę do agregacji i budowy jednorodnego portalu.

Możliwe jest zastosowanie co najmniej kilku rozwiązań, począwszy od opracowania polskiego schematu metadanych, przyjęcia jednego ze stosowanych standardów w muzeach światowych lub ograniczenia się do schematu wymiany danych. Każde z zastosowanych rozwiązań ma swoje wady i zalety, a jego wybór będzie zależał przede wszystkim od realnych celów jakie wyznaczy instytucja koordynująca cały proces digitalizacji w polskich muzeach. Tworzenie systemu standaryzacji opisu muzealiów powinno nie tylko bazować na wieloletnich światowych doświadczeniach, ale opracowane narzędzie musi uwzględniać wymagania stawiane europejskim muzeom w zakresie agregacji danych.

Kompleksowe opracowanie rozbudowanego schematu metadanych umożliwiającego nie tylko dokumentację, ale też całościowe zarządzanie kolekcją będzie wymagać szczegółowego określenia rodzaju stosowanych kategorii opisu i ustalenia jednolitych procedur odnoszących się do bieżących zadań muzeum (akcesja, wypożyczenia, itp.). Pełna standaryzacja powinna zapewnić sprawne zarządzanie informacją w muzeum, może być jednak trudna do przeprowadzenia ze względu na długi czas realizacji i niechęć części środowiska przyzwyczajonego do wypracowanych modeli działań. Rozwiązanie takie nie musi zostać wprowadzone obowiązkowo dla wszystkich instytucji. Schematy metadanych stosowane w muzeach muszą wskazywać jedynie wybrane elementy obowiązkowe. W ten sposób poszczególne muzea mogą modyfikować i kształtować wewnętrzne standardy zgodnie z wewnętrznymi potrzebami odnoszącymi się zwłaszcza do rodzaju gromadzonych zbiorów. Zachowanie wspólnych minimalnych elementów schematu pozwala na standaryzację podstawowego opisu muzealiów, co umożliwi migrację informacji o wszystkich obiektach. Zgodność z promowanymi w Europie schematami wymiany powinna umożliwić agregację danych do międzynarodowych portali, takich jak Europeana.

4.3. Baza tezaurusów

Efektywne przeszukiwanie zasobu udostępnionego na portalu prezentującym obiekty dziedzictwa kulturowego przechowywane w polskich muzeach, nie może być oparte jedynie na wspólnym schemacie metadanych. Dopelnienie stanowić muszą kontrolowane słowniki terminologiczne. Rodzaj i ilość wymaganych słowników będzie warunkowana przez dobór konkretnego schematu metadanych. Przedstawione w dokumencie międzynarodowe rozwiązania skłaniają do postawienia tezy, że tworzenie od podstaw przynajmniej części słowników wydaje się nieuzasadnione. W zakresie stosowania terminologii odnoszącej się do nazw geograficznych, autorów, materiałów, nazw obiektów, czy tematów ikonograficznych należałoby przetłumaczyć popularnie stosowane tezaury. Za tłumaczeniem funkcjonujących już międzynarodowych tezaurusów przemawia idąca za tym możliwość łatwiejszej agregacji i wymiany danych, zarówno w celach udostępniania jak i prowadzenia badań naukowych.

Budowanie od podstaw ogólnokrajowych słowników jest zadaniem niezwykle pracochłonnym, a przy tym wymagającym dużego nakładu środków i zaangażowania wielu specjalistów. Nie tylko należałoby rozpoznać jakie słownictwo jest stosowane w poszczególnych instytucjach, uwzględniając wszystkie dziedziny, uporządkować poszczególne terminy i stworzyć listy, a przede wszystkim opracować zasady budowania poszczególnych tezaurusów. Odnosząc się do doświadczeń np. British Museum, które tworzyło słowniki kilkanaście lat oraz mając na uwadze konieczność jak najszybszego wprowadzenia ogólnokrajowych rozwiązań, takie zadanie nie tylko może okazać się trudne, ale wręcz niemożliwe do realizacji w wyznaczonym terminie.

Tworzenie polskojęzycznych słowników nie może opierać się jedynie na tłumaczeniu gotowych rozwiązań. Powszechnie stosowane międzynarodowe tezaury muszą zostać uzupełnione o terminy stosowane w polskich instytucjach. Uzupełnienia wymagać będą zwłaszcza nazwy osobowe, czy geograficzne, jednakże w odniesieniu do zasad budowania haseł stosowanych w przyjętym rozwiązaniu.

Z powyższych względów nie rekomenduje się pojedynczym instytucjom budowy rozbudowanych tezaurusów. Zaleca się porządkowanie stosowanej terminologii poprzez eliminację synonimów, powtórzeń, błędnych zapisów itp., a dopiero w dalszej kolejności wdrażanie wspólnie wypracowanych ogólnopolskich rozwiązań.

Przed instytucją koordynującą digitalizację w polskich muzeach stoi nie tylko konieczność doboru odpowiednich międzynarodowych rozwiązań i dostosowania ich do rodzimych potrzeb, ale także opracowanie i wdrożenie narzędzi umożliwiających ich implementację.

Aby tezaurus mógł być wykorzystywany w codziennej pracy, musi być on oczywiście łatwo dostępny dla osób opracowujących zbiory, a najlepiej, gdy jest w intuicyjny sposób zintegrowany z systemem informatycznym wykorzystywanym w danym muzeum. Praktycznych sposo-

bów korzystania z tezaury może być wiele, a ostateczny sposób pracy determinowany jest zarówno przez możliwości techniczne środowiska informatycznego, w którym ma miejsce opracowywanie zabytków, jak i przez preferencje oraz umiejętności osoby opracowującej.

Poniżej przedstawiono trzy najpopularniejsze scenariusze korzystania z tezaury wraz ze wskazaniem wad i zalet poszczególnych podejść. W każdym z tych scenariuszy przyjęto założenie, że istnieje pewien tezaurus dość powszechnie wykorzystywany przez więcej niż jedno muzeum i pracownik opracowujący zbiory chce z niego korzystać. Efektem użycia tezaury w każdym ze scenariuszy ma być zawarcie w stosownym polu opisu obiektu hasła z tezaury oraz jednoznacznego identyfikatora tego hasła. Użycie jednoznacznych identyfikatorów znacznie ułatwia integrowanie danych pochodzących z różnych baz korzystających ze wspólnego tezaury i jest właściwie niezbędne (np. w sytuacjach, gdy słowo będące hasłem z tezaury jest wieloznaczne). Wówczas identyfikator jednoznacznie określa, o które znaczenie chodziło osobie opracowującej obiekt.

W poniższych przykładach wykorzystano opisane wcześniej tezaury VIAF⁹¹ oraz Iconclass⁹², które dostępne są na otwartych licencjach (odpowiednio: Open Data Commons Attribution License v1.0⁹³ oraz Open Database License v1.0^{94,95}).

4.3.1. Scenariusz 1. Wyszukiwanie haseł poprzez przeglądarkę WWW

W tym scenariuszu tezaurus nie jest w żaden sposób zintegrowany z muzealnym systemem informatycznym. Opracowujący obiekt wchodzi na stronę WWW tezaury, na której wyszukuje interesujące go hasło, a następnie kopiuje je wraz z identyfikatorem w stosowne pole opisu zabytku w lokalnym systemie. Dla przykładu:

Strona WWW umożliwiająca przeszukiwanie:

VIAF: <http://viaf.org/>

Iconclass: <http://www.iconclass.org/help/outline>

Strona WWW z wynikami zapytania:

VIAF (dla zapytania 'Lempicka, Tamara): <http://viaf.org/viaf/search?query=local.names+all+%22Lempicka,%20Tamara%22&stylesheet=/viaf/xsl/results.xsl&sortKeys=holdingscount&maximumRecords=100>

⁹¹ *Virtual International Authority File*. Dok. elektroniczny: <http://viaf.org> [Dostęp: 31 X 2012].

⁹² *Iconclass*. Dok. elektroniczny: <http://www.iconclass.org> [Dostęp: 31 X 2012].

⁹³ *Open Data Commons Attribution License (ODC-By) v1.0*. Dok. elektroniczny: <http://opendatacommons.org/licenses/by/1.0/> [Dostęp: 31 X 2012].

⁹⁴ *Open Database License (ODbL) v1.0*. Dok. elektroniczny: <http://opendatacommons.org/licenses/odbl/1.0/> [Dostęp: 31 X 2012].

⁹⁵ Uwaga: Przy korzystaniu z tezaury pochodzącego spoza macierzystej instytucji należy mieć pewność, że licencja na jakiej ten tezaurus jest dostępny pozwala na jego użycie.

Iconclass (dla zapytania „zrozumienie” – ang. *understanding*):

http://www.iconclass.org/rkd/0/?q=understanding&q_s=1

Przykładowy wynik:

VIAF:

Hasło: Lempicka, Tamara de, 1898-1980

Identyfikator: 66489937

URL: <http://viaf.org/viaf/66489937>

Iconclass:

Hasło: Understanding, Comprehension

Identyfikator: 52 C 4

URL: <http://www.iconclass.org/rkd/52C4/>

Warto tu zwrócić uwagę, że w przypadku VIAF, który jest tak naprawdę połączeniem wielu baz danych osób i instytucji (głównie bazy biblioteczne, ale również np. Getty ULAN), w ramach jednego wyniku prezentowane jest kilka równoważnych pisowni danego hasła. Co więcej, może być tak, że żaden ze sposobów zapisu nie odpowiada konwencji przyjętej w danym muzeum i hasło skopiowane z takiego tezaurusa musi być odpowiednio dostosowane. To jeszcze bardziej podkreśla kluczową rolę jednoznacznych identyfikatorów poszczególnych haseł tezaurusa i konieczność przechowywania ich również we własnym systemie muzealnym.

Zalety omawianego rozwiązania:

- brak konieczności integracji muzealnego systemu informatycznego z systemami obsługującymi tezaursy (zarówno początkowej integracji, jak i utrzymywania zgodności systemów w przypadku modyfikacji interfejsu programistycznego systemu danego tezaurusa);
- możliwość korzystania z wielu tezaursów równocześnie;
- łatwość wprowadzania nowych tezaursów;
- korzystanie z najnowszej dostępnej wersji tezaurusa.

Wady omawianego rozwiązania:

- konieczność równoczesnego korzystania z dwóch systemów i przełączania się między nimi;
- uzależnienie od dostępności tezaurusa online – problemy w przypadku przerw w dostępności strony WWW tezaurusa, powolnego funkcjonowania tezaurusa lub problemów z łączem internetowym muzeum.

4.3.2. Scenariusz 2. Skopiowanie całego tezaurusa do bazy danych własnego systemu muzealnego

Jeżeli tezaurus wykorzystywany przy opracowywaniu zbiorów muzeum jest dostępny nie tylko w postaci wyszukiwarki na stronie WWW, ale również w postaci pliku (plików) z danymi do pobrania, to możliwe powinno być jednorazowe skopiowanie zawartości takich plików do bazy danych słowników/tezaurów lokalnego systemu muzealnego. Dla przykładu, wspomniane w poprzednim scenariuszu tezaury można pobrać w całości z następujących stron:

VIAF: <http://viaf.org/viaf/data/>

Iconclass: <http://www.iconclass.org/help/lod>

Format w jakim dostępne są dane może być różny dla poszczególnych tezaurów, jednak powszechnie stosowaną praktyką jest udostępnianie tezaurów, niezależnie od formatów specyficznych, również w sposób zgodny z modelem RDF⁹⁶ z wykorzystaniem ontologii SKOS⁹⁷. Obydwa podane powyżej tezaury są właśnie w taki sposób w całości dostępne.

W zależności od stosowanego systemu muzealnego oraz od umiejętności personelu IT muzeum, taka operacja importu tezaurusa może być wykonana samodzielnie w ramach muzeum lub też może być przedmiotem zlecenia dla producenta czy dostawcy systemu muzealnego lub też innej zewnętrznej firmy. W przypadku gdy muzeum zamawia nowy system informatyczny, warto rozważyć wpisanie w wymagania dotyczące tego systemu oczekiwania, iż system będzie od razu posiadał w swojej bazie danych stosowne tezaury. W tym scenariuszu praca osoby opracowującej zbiory zamyka się w całości w systemie muzealnym i nie ma konieczności ręcznego kopiowania danych pomiędzy stroną WWW tezaurusa, a lokalnym systemem. Dzięki temu można oczekiwać przyspieszenia tempa pracy oraz spadku liczby omyłek związanych z kopiowaniem danych.

Zalety omawianego rozwiązania:

- praca w jednym systemie informatycznym, bez konieczności przełączania się i ręcznego kopiowania danych;
- pełne uniezależnienie od zmian w systemie informatycznym tezaurusa oraz od jego dostępności online;

Wady omawianego rozwiązania:

- konieczność wykonania operacji kopiowania tezaurusa do bazy danych lokalnego systemu informatycznego;
- posiadanie własnej kopii tezaurusa, która nie jest automatycznie aktualizowana, gdy aktualizowany jest tezaurus dostępny w Internecie. W przypadku istotnych zmian i długo-

⁹⁶ *Resource Description Framework (RDF)*. Dok. elektroniczny: <http://www.w3.org/RDF/> [Dostęp: 31 X 2012].

⁹⁷ *SKOS Simple Knowledge Organization System*. Dok. elektroniczny: <http://www.w3.org/2004/02/skos/> [Dostęp: 31 X 2012].

trwałego stosowania tezaury, konieczne może być okresowe uwspólnianie tezaury lokalnego z tezaurem dostępnym online.

4.3.3. Scenariusz 3. Dynamiczna integracja systemu informatycznego muzeum i systemu tezaury

Opisane powyżej scenariusze wykorzystywały kolejno dwie cechy udostępnianych w Internecie tezaurów: możliwość wyszukiwania online poprzez stronę WWW tezaury oraz możliwość pobrania całego tezaury i zaimportowania go do wykorzystywania w systemie muzealnym, bez konieczności komunikacji online.

Trzeci scenariusz jest rozwiązaniem pośrednim, które przy odpowiedniej realizacji może cechować się częścią zalet obydwóch wcześniejszych propozycji. Zastosowanie tego rozwiązania wymaga jednak odpowiedniego przygotowania systemu informatycznego wykorzystywanego w muzeum, co nie zawsze jest możliwe.

Scenariusz ten opiera się na wzbogaceniu systemu muzealnego o możliwość wyszukiwania i pobierania online haseł ze zdalnego zewnętrznego tezaury (lub nawet wielu tezaurów, w zależności od elementu opisu, który jest wypełniany). Wzbogacenie takie jest możliwe dzięki wykorzystaniu interfejsu programistycznego (API) serwisów udostępniających tezaury, jednak wymaga najczęściej zaangażowania dostawcy oprogramowania muzealnego. Przykładowa dokumentacja API dla tezaury VIAF dostępna jest pod adresem <http://www.oclc.org/developer/documentation/virtual-international-authority-file-viaf/using-api>, jednak szersze jej omówienie wykracza znacznie poza zakres niniejszego dokumentu.

Zalety omawianego rozwiązania:

- korzystanie z najnowszej dostępnej wersji tezaury;
- praca w jednym systemie informatycznym, bez konieczności przełączania się i ręcznego kopiowania danych.

Wady omawianego rozwiązania:

- uzależnienie od dostępności tezaury online – problemy w przypadku przerw w dostępności strony WWW tezaury, powolnego funkcjonowania tezaury lub problemów z łączem internetowym muzeum. Ta wada może być jednak zminimalizowana poprzez odpowiednie zaprojektowanie modułu integrującego system informatyczny tezaury i system muzealny;
- rozwiązanie najbardziej złożone pod względem informatycznym, może wymagać istotnego nakładu pracy przy wdrażaniu i wiązać się z dodatkowymi pracami w sytuacji zmian API systemu tezaury.

Przedstawione powyżej trzy scenariusze różnią się od siebie złożonością ich wdrożenia i utrzymania oraz wygodą użytkownika. Scenariusz pierwszy jest zdecydowanie najprostszy do wprowadzenia w życie, gdyż nie wymaga integracji odrębnych systemów informatycznych – muze-

alnego i tezaury. Niestety integracja ta zastąpiona jest dodatkowymi zadaniami dla personelu, co może owocować wolniejszym tempem pracy i zwiększoną liczbą błędów w stosunku do wyników osiągniętych w scenariuszach 2. i 3.

Scenariusz trzeci wydaje się być natomiast rozwiązaniem najbardziej efektywnym z punktu widzenia możliwości oferowanych osobie opracowującej obiekty muzealne. Niestety jego wdrożenie i utrzymanie jest najbardziej skomplikowane pod względem informatycznym, co oczywiście przekłada się również na koszty realizacji takiego scenariusza.

Wszystkie z opisanych powyżej scenariuszy zakładają jednokierunkową współpracę pomiędzy dwoma systemami, czyli pobieranie danych z tezaury do oprogramowania muzealnego. Nie jest to jednak jedyny możliwy model współpracy. Istnieją bardziej złożone rozwiązania, gdzie wspomniana komunikacja w niektórych sytuacjach jest dwukierunkowa. Przykładem funkcjonującego w ten sposób systemu jest katalog NUKAT⁹⁸.

⁹⁸ NUKAT. Dok. elektroniczny: <http://www.nukat.edu.pl> [Dostęp: 31 X 2012].

4.4. Narodowy Uniwersalny Katalog Centralny (NUKAT)

Katalog NUKAT oficjalnie działa od 21 października 2002 r. w Bibliotece Uniwersyteckiej w Warszawie. Oparty jest na koncepcji zbiorowego katalogowania w bazie centralnej i udostępniania bibliotekom gotowych danych do kopiowania. Jest przedsięwzięciem ogólnokrajowym, realizowanym przez biblioteki naukowe i akademickie stosujące systemy biblioteczne Virtua i Horizon oraz Bibliotekę Narodową. Początki tego Katalogu sięgają 1991 r., kiedy to Biblioteka Uniwersytecka w Warszawie rozpoczęła budowę Kartoteki Haseł Wzorcowych (khw), w tym języka haseł przedmiotowych (jhp) KABA (polskiego odpowiednika RAMEAU i LCSH w systemie kartkowym). W 1993 r. została zapoczątkowana Centralna Kartoteka Haseł Wzorcowych (CKHW), a w 1995 r. Centralny Katalog Czasopism (CKTCz). Obie te bazy, jako jedyne z utworzonych w latach 90., zostały później wykorzystane w katalogu centralnym NUKAT i są w nim nadal rozwijane.

W 1996 r. utworzono w BUW Centrum Formatów i Kartotek Haseł Wzorcowych, które przejęło opiekę i nadzór nad CKHW, zajmując się kontrolą wprowadzanych rekordów haseł, koordynacją współpracy, opracowywaniem materiałów metodycznych w zakresie formatów, kartotek haseł wzorcowych i jhp KABA oraz organizacją i prowadzeniem szkoleń. Koncepcja zakłada wykorzystanie CKHW jako podstawy katalogu centralnego. Opis przedmiotowy dokumentów katalogowanych w bazie NUKAT jest budowany w 3 językach haseł przedmiotowych (jhp):

- jhp KABA – wykorzystywanym głównie przez biblioteki akademickie;
- jhp BN (Biblioteki Narodowej) – wykorzystywanym przez BN oraz biblioteki publiczne;
- polskiej wersji MeSH (Medical Subject Headings) – wykorzystywanym przez biblioteki medyczne.

W 2001 r. nastąpiła zmiana nazwy na Centrum NUKAT. W czerwcu 2002 r. przeniesiono zawartość Centralnej Kartoteki Haseł Wzorcowych do podzbioru CKHW NUKAT, w lipcu uruchomiono pełne współkatalogowanie w katalogu NUKAT. Współpraca z NUKAT pod warunkiem pracy w VTLS, polega na tworzeniu centralnej bazy kartotek wzorcowych (CKHW) na serwerze Biblioteki Uniwersyteckiej w Warszawie, tworzeniu centralnej bazy czasopism (CKTCz) na serwerze Uniwersytetu Gdańskiego oraz katalogowaniu i wzajemnym przejmowaniu rekordów w lokalnych bazach VTLS wg zasad ujednoczonych przez Centrum Formatów i Kartotek Haseł Wzorcowych. Katalog centralny jest obsługiwany przez system biblioteczny Virtua i oprogramowanie typu portal WWW Vectors firmy VTLS. Katalogi lokalne bibliotek współpracujących są utrzymywane na lokalnych serwerach i obsługiwane przez systemy użytkowane przez te biblioteki.

Rekordy wszystkich nowo katalogowanych książek i czasopism są wprowadzane w NUKAT, tam walidowane i zatwierdzane, kopiowane do baz poszczególnych Bibliotek, zaopatrywane w siglum każdej czynnie współpracującej biblioteki, która rekord skopiowała. Na bieżąco aktualizowane kopie są udostępniane w kilku użytkowanych w polskich bibliotekach systemach,

w celu ułatwienia dostępu do rekordów kartoteki haseł wzorcowych bibliotekom nieuczestniczącym we współkatalogowaniu, które polega na tworzeniu nowych rekordów, bądź modyfikowaniu rekordów już istniejących. Współpraca obejmuje także opracowywanie rekordów haseł wzorcowych formalnych (khw). W systemach zautomatyzowanych baza khw pełni rolę odsyłaczy poprzez tworzenie powiązań pomiędzy hasłami w formie ujednocionej z formami odrzuconymi. Dzięki temu łatwiejsze staje się dla czytelników wyszukiwanie np. poprzez hasła autorskie.

Informacje o zbiorach 109 polskich bibliotek naukowych i akademickich uczestniczących we współkatalogowaniu w katalogu centralnym NUKAT, przekazywane są do WorldCat – światowej bazy danych o zbiorach bibliotecznych i ich właścicielach. Opisy publikacji z bazy NUKAT udostępniane są także do przeszukiwania w wyszukiwarce Google – w Google Scholar (<http://www.scholar.google.pl>) i w Google Books (<http://www.books.google.pl>), a rekordy KHW z bazy NUKAT przekazywane są do wirtualnej międzynarodowej kartoteki haseł wzorcowych VIAF (<http://www.viaf.org>) – dzięki VIAF pracownicy oddziałów katalogowania z całego świata mogą zapoznać się z wariantami nazw używanych w opisach dokumentów w różnych krajach i trafić do odpowiednich narodowych kartotek haseł wzorcowych, skąd mogą pobrać przydatne im dane. Biblioteki wprowadzają do NUKAT dane o swoich bieżących i dawnych nabytkach. Na końcu każdego opisu dokumentu wyszukanego w bazie NUKAT znajduje się lista nazw bibliotek udostępniających ten dokument, a kliknięcie na nazwę biblioteki pozwala sprawdzić dostępność publikacji w danej bibliotece. (NUKAT zaprosił również do współpracy biblioteki naukowe posiadające inne zintegrowane systemy biblioteczne, tj. Horizon, Aleph, Innopac oraz Prolib.)

Wprowadzenie systemu Virtua przez NUKAT i wszystkie biblioteki Konsorcjum Bibliotek VTLS oznacza permanentne prace nad ujednocnianiem zasad tworzenia katalogów lokalnych oraz katalogu centralnego. W 2003r. zastąpiono system VTLS jego nową wersją – Virtua/VTLS, co pozwoliło udostępnić katalog poprzez portal internetowy.

Zdaniem Marii Burchard, kierownika NUKAT, „zgromadzone w NUKAT ujednoczone nazwy osobowe, nazwy ciał zbiorowych nie powinny być kojarzone wyłącznie z katalogami bibliotecznymi. Mają one bowiem wymiar uniwersalny i powinny być wykorzystywane jako klucze dostępu do różnego rodzaju baz informacyjnych, w przyszłości także baz archiwalnych i muzealnych. Dodany do rekordu bibliograficznego temat dokumentu pojawia się we wszystkich kopiach tego rekordu, jeśli biblioteki sobie tego życzą. Można je oglądać z każdego zakątka świata i dla każdego powinny być zrozumiałe. Dlatego jedyne reguły, których należy restrykcyjnie przestrzegać, to międzynarodowe normy i standardy. Ich stosowanie gwarantuje międzynarodowe porozumienie⁹⁹.

⁹⁹M. Burchard, *Narodowy Uniwersalny Katalog Centralny NUKAT a przemiany w bibliotekach*, w: Biblioteki Naukowe w kulturze i Cywilizacji, s. 181-189. Dok. elektroniczny: http://library.put.poznan.pl/2005/pdf/4_2.pdf [Dostęp: 31 X 2012].

Centrum NUKAT, Biblioteka Narodowa, a także duże biblioteki uczelniane (BJ, BUW, UMK,) i biblioteki publiczne stosują Format MARC21, standard, który jest rozwijany przez Bibliotekę Kongresu.

Język Hasel Przedmiotowych KABA pozwala wyrazić treść dokumentu za pomocą słownictwa języka naturalnego i jest do tej pory jedynym w Polsce językiem informacyjno-wyszukiwawczym prezentowanym w formie kartoteki wzorcowej. Kartoteka ta jest elementem całego systemu kartotek wzorcowych służących do obsługi zautomatyzowanych katalogów bibliotecznych, niezależnie od stosowanego oprogramowania¹⁰⁰. Założono, że KABA będzie językiem kompatybilnym ze znanymi językami hasel przedmiotowych (jhp) tworzonymi przez największe biblioteki narodowe świata: Library of Congress Subject Heading (LCSH) Biblioteki Kongresu w Waszyngtonie, Repertoire de Vedettes-Matiere Biblioteki Uniwersytetu Laval'a w Quebec oraz Repertoire d'Autorite-Matiere Encyclopedique et Alphabetique Unifie (RAMEAU) Narodowej Biblioteki Francji¹⁰¹.

Źródłem bezpośrednim języka KABA jest język RAMEAU. Hasła przedmiotowe języka KABA wchodzi w analogiczne relacje, mają identyczne zasady stosowania i funkcje jak ich odpowiedniki w RAMEAU. Są dodatkowo powiązane relacjami ekwiwalencji z francuskojęzycznymi hasłami RAMEAU i angielskojęzycznymi hasłami LCSH. System VTLS umożliwi znalezienie odpowiedników polskojęzycznych hasel języka KABA w RAMEAU i LCSH. Dzięki temu możliwe jest wykorzystywanie danych ze światowych baz bibliograficznych stosujących języki RAMEAU i LCSH¹⁰². Dla wszystkich hasel przedmiotowych języka KABA tworzone są rekordy kartoteki hasel wzorcowych przedmiotowych (khwp) w formacie USMARC.

¹⁰⁰ Język KABA: aktualny stan badań / Danuta Domalewska, // W: Opracowanie przedmiotowe dokumentów z zakresu nauk ścisłych: matematyczno-przyrodniczych i technicznych. Język hasel przedmiotowych KABA: teoria, praktyka, przyszłość. Kazimierz Dolny, 20-22 września 2006 roku. - [Warszawa]: Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja „Elektronicznej Biblioteki”, 2006. - (EBIB Materiały konferencyjne nr 15).- ISBN 83-921757-6-X. Dok. elektroniczny: <http://www.ebib.info/publikacje/matkonf/kaba/domalewska.php> [Dostęp: 31 X 2012].

¹⁰¹ K. Kruszyńska *Oddział Opracowania Rzeczowego Zbiorów Biblioteki Głównej AGH*. Dok. elektroniczny: http://www.uci.agh.edu.pl/bip/65/15_65.htm [Dostęp: 31 X 2012].

¹⁰² Kompatybilność JIW nie jest cechą stopniowaną, chociaż pełna kompatybilność jest ideałem, podobnie jak pełna przekładalność języków naturalnych. Tę cechę JHP KABA osiągnięto dużym nakładem wysiłku intelektualnego i środków finansowych. Dzięki nim przestaje być językiem lokalnym, a włącza się w systemy leksykalne o międzynarodowym, a nawet światowym, zasięgu. w G.M., Wilczyńska, J. Woźniak-Kasperek, *Kompatybilność - zaleta czy obciążenie? Dwugłos o języku hasel przedmiotowych KABA*. *Bibliotekarz* 2005, nr 9, s. 14.

4.5. Udostępnianie danych o zbiorach w Internecie

Jednym z zadań na najbliższą przyszłość, zanim zostanie stworzony krajowy agregator danych pobieranych z systemów informatycznych poszczególnych muzeów, jest stworzenie zunifikowanych zasad kreowania identyfikatora URI. Coraz intensywniejsza automatyczna wymiana danych pomiędzy różnymi systemami informatycznymi i powstanie Sieci Semantycznej, której jednym z istotnych warunków jest stosowanie stałych identyfikatorów URI, rodzi poważny problem, który wyczerpujący został omówiony przez twórców ontologii CIDOC CRM:

„W powstającej, globalnej infrastrukturze informatycznej poświęconej dziedzictwu kulturalnemu, instytucje dostarczające danych muszą wytworzyć URI dla każdej rzeczy, jaką zamierzają opisać. Unikatowy identyfikator nie jest jednakże użyteczny bez wyjaśnienia jego znaczenia. Inicjatywa Linked Open Data (LoD) ma na celu dostarczenie autoryzowanego zestawu danych opisowych wspierających precyzyjną identyfikację obiektów. Ta metoda została prędko przyjęta przez wiele organizacji, w tym OCLC. Powstaje jednak nowy problem. LoD i Sieć Semantyczna nie posiadają określonej polityki odnośnie tego, kto będzie tworzył URI dla powszechnie znanych rzeczy, takich jak Szekspir czy Akropol. Każdy może to zrobić. O ile te szczególne przypadki są zewnętrzne w stosunku do muzeów, wiele muzeów dostarcza duże ilości danych do cyfrowych serwisów agregujących. Te z kolei tworzą URI dla każdej informacji, jaką otrzymują z muzeów i każdy z agregatorów tworzy własne URI, nawet jeżeli dotyczy on tej samej rzeczy. Tak więc miliardy rzeczy otrzymują po kilka identyfikatorów, każdy od różnej agencji i każdy niezależnie od innych identyfikatorów oznaczających tę samą rzecz. W rezultacie, aby połączyć wszelkie informacje o jednym obiekcie należy zorientować się, które jeszcze identyfikatory w morzu dziesiątków miliardów identyfikatorów odnoszą się do obiektu, którym jesteśmy zainteresowani. Większość agregatorów nie ma prerogatyw do weryfikowania bądź ustanawiania identyfikatorów, ponieważ nie ma dostępu do samych obiektów. Muzea jednakowoż, zachowują tożsamość obiektów w obrębie swoich inwentarzy, co jest częścią ich naukowej i prawnej misji, i jako takie mogą wytwarzać identyfikatory, które mają należyte ugruntowanie. Jeżeli zatem muzea nie zajmą się sposobem, w jaki agregatory Sieci Semantycznej stosują do identyfikacji obiektów, rozwijać się będzie konkurencja, autoryzowanych i nieautoryzowanych treści o muzealnych kolekcjach. Skala problemu będzie tak wielka, że nikt nie będzie potrafił rozwikłać tego bałaganu. Uważamy zatem za rzecz zupełnie konieczną, by obiekty muzealne były identyfikowane w taki sposób, aby weryfikacja identyfikatorów w ogniesieniu do obiektów realnego świata była procesem prostym i efektywnym. Ponadto powinno być rzeczą prostą dbanie o owe identyfikatory”¹⁰³.

Uznając wagę powyżej przytoczonych argumentów, w kontekście udostępniania danych na temat obiektów muzealnych w Internecie, wydaje się rzeczą pilną podjęcie prac nad ustanowie-

¹⁰³ *Call for Comments - Linked Open Data Recommendation for Museums*. Dok. elektroniczny: http://www.cidoc-crm.org/URIs_and_Linked_Open_Data.html, [Dostęp: 31 X 2012].

niem ogólnopolskiego mechanizmu generowania i przydzielania muzealiom identyfikatorów URI. Takie unikalne i certyfikowane sieciowe identyfikatory powinny być nieodłączną częścią opisu obiektów, zwłaszcza w przypadku udostępniania tych opisów w Internecie. Dzięki temu niezależnie od późniejszych losów wzorcowych danych opisowych, w szczególności niezależnie od ich wykorzystania w rozmaitych serwisach agregujących, wiadomo byłoby o jaki obiekt chodzi.

Naturalnymi uczestnikami procesu udostępniania danych muzealiów w Internecie są: muzeum, udostępniające dane poprzez swoje strony WWW oraz odwiedzający te strony – odbiorcy udostępnianych danych. W celu wdrożenia opisanej powyżej koncepcji globalnie unikatowych wiarygodnych sieciowych identyfikatorów muzealiów konieczne byłoby wprowadzenie do tego układu jeszcze jednego uczestnika – zaufanej trzeciej strony, która odpowiadałaby za utrzymanie tych identyfikatorów i koordynowała ich nadawanie. Strona ta powinna również umożliwiać uzyskanie podstawowych informacji o danym obiekcie na podstawie jego identyfikatora. W przypadku połączenia systemu identyfikatorów z krajowym agregatorem danych muzealnych, mogłyby to być informacje pobierane automatycznie z systemu informatycznego danego muzeum. Oczywiście, wprowadzenie wspomnianej trzeciej strony nie powinno zaburzyć standardowego sposobu korzystania ze stron WWW muzeum, a jedynie dać nowe możliwości.

To, jak taki proces mógłby wyglądać z punktu widzenia użytkownika końcowego, można prześledzić już dziś na przykładzie wymiany danych pomiędzy Cyfrowym Muzeum Narodowym w Warszawie¹⁰⁴ i Federacją Bibliotek Cyfrowych¹⁰⁵. Na stronach Cyfrowego MNW znaleźć można między innymi informacje o obrazie *Mistyczne zaślubiny św. Katarzyny* autorstwa Francesco Bissolo¹⁰⁶. W zbiorze danych podstawowych tego obrazu na stronach Cyfrowego MNW widnieje informacja, iż numer inwentarza dla tego obiektu to M.Ob.953. Typowy internauta, znający numer inwentarza obiektu, ale nie znający Cyfrowego MNW, mógłby zacząć jego wyszukiwanie od serwisu Google podając wspomniany numer jako zapytanie: <http://www.google.pl/search?q=M.Ob.953>. W tym przypadku pierwszy z wyników wyszukiwania prowadzi do strony w anglojęzycznej wersji Wikipedii (wyszukiwanie przeprowadzono w dniu 1 listopada 2012 r.), na której można zobaczyć cyfrowy wizerunek wspomnianego obrazu¹⁰⁷. Ponadto na stronie Wikipedii zawarto podstawowe metadane obrazu (m.in. tytuł, autora, datowanie, medium, rozmiar i numer inwentarzowy MNW) oraz informację, że wizerunek cyfrowy został pobrany ze stron Cyfrowego MNW. Porównanie cyfrowego wizerunku oraz metadanych na stronach Wikipedii oraz Cyfrowego MNW pokazują, że metadane na stronach muzeum są

¹⁰⁴ *Muzeum Narodowe w Warszawie*. Dok. elektroniczny: <http://www.cyfrowe.mnw.art.pl> [Dostęp: 31 X 2012].

¹⁰⁵ *Federacji Bibliotek Cyfrowych*. Dok. elektroniczny: <http://fbc.pionier.net.pl> [Dostęp: 31 X 2012].

¹⁰⁶ *Mistyczne zaślubiny św. Katarzyny autorstwa Francesco Bissolo* [w:] *Cyfrowe Muzeum Narodowe w Warszawie*. Dok. elektroniczny: <http://cyfrowe.mnw.art.pl/dmuseion/docmetadata?id=3626> [Dostęp: 31 X 2012].

¹⁰⁷ *Bad title*, [w:] *Wikipedia*. Dok. elektroniczny: http://en.wikipedia.org/wiki/File:Bissolo_Marriage_of_St._Catherine.jpg [Dostęp: 31 X 2012].

znacznie bogatsze, a sam wizerunek cyfrowy obrazu został przy wstawianiu do Wikipedii wykadrowany, aby usunąć wzornik kolorów, ramę i dolną uszkodzoną część obrazu. Nawet zmiana zapytania przesłanego do Google z samego numeru inwentarza na frazę „M.Ob.953 MNW” nie powoduje zmiany pierwszego miejsca na liście wyników wyszukiwania¹⁰⁸ – jest to efekt bardzo wysokiej pozycji Wikipedii w indeksach wyszukiwawczych Google. Choć w tym przypadku sytuacja jest o tyle lepsza, że na drugim miejscu po Wikipedii pojawia się link do właściwej strony Cyfrowego MNW. Jak wspomniano wcześniej, Cyfrowe MNW dostarcza dane do Federacji Bibliotek Cyfrowych. FBC z kolei udostępnia między innymi usługę lokalizowania wizerunków cyfrowych na podstawie identyfikatora¹⁰⁹. Użytkownik poszukujący danych o jakości wyższej niż ta, którą oferuje Wikipedia, znający serwis FBC i mający świadomość, iż serwis ten udostępnia dane pobierane wyłącznie bezpośrednio z instytucji kultury, mógłby zacząć poszukiwanie obiektu o numerze inwentarza M.Ob.953 od wykorzystania usługi lokalizowania wizerunków cyfrowych. Jest to możliwe m.in. poprzez dołączenie numeru inwentarza do podanego wcześniej adresu usługi lokalizacji, czyli wejście na stronę <http://fbc.pionier.net.pl/id/M.Ob.953>. Efektem takiego wywołania jest natychmiastowe przekierowanie na właściwą stronę serwisu Cyfrowego MNW¹¹⁰ prezentującą obszerne dane na temat obiektu oraz jego pełny wizerunek cyfrowy.

Przytoczony powyżej przykład pokazuje, że dane udostępniane na stronach WWW muzeum będą kopiowane do innych serwisów, również bez wiedzy pracowników muzeum, a w dodatku często w sposób niekompletny i obniżający jakość tych danych. Przy udostępnianiu online danych o muzealiach ważne jest zawieranie w nich trwałego odnośnika do tych danych (do strony na której są one prezentowane), gdyż w przypadku kopiowania danych z takim odnośnikiem do innych serwisów, odbiorca końcowy nadal będzie miał szansę na dotarcie do oryginalnej informacji.

Wprowadzenie dodatkowego serwisu, który będzie gromadził (kopiował) dane z baz muzealnych bez obniżania jakości tych danych i promowanie tego serwisu jako zaufanego źródła informacji o zbiorach polskich muzeów może zwiększyć szansę na to, że użytkownicy końcowi będą wybierać profesjonalne źródła informacji zamiast sięgać wyłącznie do Wikipedii. Dodatkowo serwis taki może oferować usługę pozwalającą w łatwy sposób przekształcić numer inwentarzowy czy inny identyfikator obiektu w adres WWW zgodny z koncepcją PURL¹¹¹, czyli adres stały i niezależny od zmian technologicznych zachodzących na stronach WWW danego muzeum, będący w praktyce przekierowaniem do aktualnego opisu obiektu prezentowanego na stronach WWW muzeum.

¹⁰⁸ Dok. elektroniczny: <https://www.google.pl/search?q=M.Ob.953%20MNW> [Dostęp: 31 X 2012].

¹⁰⁹ *Federacji Bibliotek Cyfrowych*. Dok. elektroniczny: <http://fbc.pionier.net.pl/id> [Dostęp: 31 X 2012].

¹¹⁰ *Cyfrowe Muzeum Narodowe w Warszawie*. Dok. elektroniczny: <http://cyfrowe.mnw.art.pl/dmuseion/docmeta-data?id=3626> [Dostęp: 31 X 2012].

¹¹¹ *PURL – ang. Persistent URL, pol. trwały URL*. Dok. elektroniczny: <http://www.purl.oclc.org/> [Dostęp: 31 X 2012].

Dodatkowej dyskusji wymaga możliwość wykorzystania numeru inwentarzowego muzealium jako unikatowego składnika adresu typu PURL (np. ostatni człon w adresie <http://fbc.pionier.net.pl/id/M.Ob.953>). W celu zapewnienia stu procentowej pewności, co do unikatowości identyfikatorów, wydaje się konieczne wprowadzenie dodatkowych elementów identyfikatora obiektu muzealnego. Powszechnie przyjętym w Internecie podejściem w tego typu sytuacjach jest prefiksowanie identyfikatora unikatowego dodatkowym identyfikatorem tego kontekstu. W tym przypadku mowa oczywiście o identyfikatorze dla muzeum. Wariantów wyboru takich identyfikatorów może być kilka:

- wykorzystanie już nadanego muzeum urzędowego identyfikatora, np. numeru NIP – prosty w zastosowaniu, ale mało przyjazny dla użytkowników i może prowadzić do częstych omyłek jeżeli wykorzystany zostanie długi i skomplikowany identyfikator (taki jak właśnie NIP);
- wykorzystanie adresu domenowego podstawowego serwisu WWW muzeum – prostszy w obsłudze i bardziej czytelny niż numer NIP, ale należy pamiętać że czasem (zwłaszcza w przypadku mniejszych instytucji) zdarzają się sytuacje konieczności zmiany adresu serwisu WWW;
- nadanie muzeum nowego identyfikatora przeznaczonego na potrzeby tworzenia globalnie unikalnych identyfikatorów muzealiów – rozwiązanie najbardziej uniwersalne i niezależne, ale wymaga prowadzenia rejestru takich identyfikatorów.

Ustalenia wymaga również składnia identyfikatora, ale wydaje się, że rozsądne będzie po prostu skorzystanie z przyjętego w Internecie schematu URN¹¹². Schemat ten przeznaczony jest do jednoznacznego identyfikowania zasobów w sieci, jednak nie mówi nic o lokalizacji tych zasobów. Składnia identyfikatora URN jest następująca:

urn:<identyfikator przestrzeni>:<identyfikator specyficzny dla przestrzeni>,

gdzie identyfikator przestrzeni to ciąg znaków precyzujący o jakiego rodzaju URN chodzi. Organizacja IANA oficjalnie przydziela takie identyfikatory przestrzeni¹¹³, jednak nie ma ich obecnie zbyt wielu i w praktyce często identyfikator taki ustala się samodzielnie (np. Polish Museum Object Identifier, w skrócie: plmoi). Zgodnie z wcześniejszymi rozważaniami element „identyfikator specyficzny dla przestrzeni” powinien składać się z dwóch części: identyfikatora muzeum oraz identyfikatora obiektu.

Podsumowując, przykładowy unikatowy sieciowy identyfikator wspomnianego wcześniej obrazu ze zbiorów Muzeum Narodowego w Warszawie mógłby wyglądać w sposób następujący:

urn:plmoi:mnwa:M.Ob.953

¹¹² URN – ang. *Uniform Resource Name*, pol. ujednolicona nazwa zasobu. Dok. elektroniczny: <http://www.tools.ietf.org/html/rfc2141> [Dostęp: 31 X 2012].

¹¹³ *Internet Assigned Numbers Authority (IANA)*. Dok. elektroniczny: <http://www.iana.org/assignments/urn-namespaces/urn-namespaces.xml> [Dostęp: 31 X 2012].

gdzie poszczególne składowe to:

urn – prefiks URN

plmoi – element oznaczający typ URN – polski identyfikator obiektów muzealnych

mnwa – element identyfikujący muzeum

M.Ob.953 – element identyfikujący obiekt w kontekście muzeum, czyli numer inwentarzowy.

Dodatkowo, w przypadku uruchomienia agregatora danych z systemów informatycznych polskich muzeów, oferującego usługę lokalizacji obiektów analogiczną do wspomnianej wcześniej usługi Federacji Bibliotek Cyfrowych, znajomość takiego sieciowego identyfikatora obiektu pozwalałaby na wygenerowanie adresu typu PURL dla tego obiektu. Dla przykładu, gdyby usługa lokalizacji agregatora muzealnego była dostępna pod adresem (obecnie fikcyjnym) <http://data.nimoz.pl/id/> to adres typu PURL dla przykładowego obiektu miałby postać <http://data.nimoz.pl/id/urn:plmoi:mnwa:M.Ob.953.>, a jego wywołanie prowadziłoby do przekierowania na odpowiednią stronę WWW serwisu Muzeum Narodowego w Warszawie. Tak skonstruowany adres mógłby być oficjalnym adresem WWW obiektu muzealnego przekazywanym do zewnętrznych serwisów razem z kopią danych dotyczących obiektu. Dzięki przejściu przez centralny serwis lokalizujący obiekty w sieci na podstawie identyfikatora, możliwe byłoby między innymi:

- wychwytywanie i blokowanie prób użycia identyfikatorów tworzonych bez wcześniejszego zarejestrowania/uzyskania elementu identyfikującego dane muzeum (podstawowy element systemu certyfikacji identyfikatorów);
- wyświetlanie ogólnych informacji kontaktowych w przypadku użycia identyfikatora obiektu z poprawnym elementem identyfikującym muzeum, ale z nieznanym (na poziomie agregatora) numerem inwentarzowym;
- gromadzenie statystyk obrazujących popularność poszczególnych obiektów.

Przedstawionego powyżej opisu nie należy traktować jako finalnej koncepcji technicznej usługi polskich sieciowych identyfikatorów obiektów muzealnych, a raczej jako zasygnalizowanie pewnych potrzeb środowiska muzealnego i popartą przykładami analizę możliwości ich realizacji w obecnych realiach technologicznych.

4.6. Sprawozdawczość

Zalecenia i konkluzje dotyczące digitalizacji jednoznacznie wskazują na konieczność prowadzenia monitoringu procesu zarówno na poziomie krajowym, jak i europejskim. Wymaga się od krajów członkowskich prowadzenia kontroli, a zwłaszcza oceniania postępów procesu. Narzędzia dzisiaj stosowane do weryfikacji digitalizacji w Polsce są niewystarczające, a przede wszystkim dają nieprawdziwy jej obraz.

Materiały do analizy czerpiemy ze sprawozdań kwartalnych MKiDN, rocznych sprawozdań GUS oraz okazjonalnych ankiet dotyczących digitalizacji. Żadne z tych narzędzi nie daje miarodajnego obrazu, a także z uwagi na odmienną wymaganych danych nie można porównać wyników pomiędzy nimi. Sprawozdania te nie definiują jednoznacznie, jakie dane podlegają weryfikacji i pozwalają na dużą dowolność w ich interpretowaniu. Dla przykładu miernik ze sprawozdawczości MKiDN dotyczący digitalizacji jednocześnie odnosi się do konserwacji, gdyż digitalizacja w tym kontekście tożsama jest z zabezpieczaniem materiału. Część opisu merytorycznego sprawozdania MKiDN w ostatnim czasie została znacząco rozbudowana i zawiera osobne pola umożliwiające odnotowanie zarówno dokumentacji wizualnej (zdjęć, skanów) jak i rekordów. Brak jednak adnotacji, jak należy interpretować tzw. „rekord”, a przede wszystkim jakie dane składają się na rekord. Mimo że informacje odnoszące się do digitalizacji, a także opisu są odnotowywane w sprawozdaniu MKiDN, to należy pamiętać, że te sprawozdania odnoszą się jedynie do określonej grupy instytucji podległych ministerstwu. Sprawozdaniem, które jest wymagane dla wszystkich instytucji muzealnych jest sprawozdanie Głównego Urzędu Statystycznego. Zakres tych danych jest niestety w znacznie mniejszym stopniu miarodajny niż sprawozdania MKiDN. Po pierwsze, całość sprawozdania GUS opiera się na pozycjach inwentarzowych, a nie ilości obiektów, co już na wstępie daje wypaczony obraz. Obiekty ewidencjonowane jako zespoły takie jak np. teki czy serwisy, są w tym ujęciu jedynie pojedynczą jednostką danych. Po drugie, punkt traktujący o digitalizacji, odnosi się do ogólnego pojęcia „ilość obiektów zdigitalizowanych”, tak więc nie ma tu mowy o odróżnieniu opisu od dokumentacji wizualnej.

We wszystkich sprawozdaniach brakuje jednego z najważniejszych rodzajów informacji: podstawowych danych o ilości obiektów udostępnianych za pośrednictwem Internetu, zarówno na własnych stronach jak i/lub przez Europeaną. Konieczność sprawnego i stałego monitoringu postępu digitalizacji jest nie tylko istotna z punktu wywiązania się z zaleceń KE, ale jest podstawowym narzędziem umożliwiającym rzetelną ocenę i możliwość modyfikowania założeń i sposobu ich realizacji.

Z powyższych względów należy jak najszybciej podjąć działania zmierzające do wprowadzenia zmian w dotychczasowym systemie sprawozdawczości w odniesieniu do zasobów cyfrowych. W pierwszej kolejności należy określić czemu ma służyć sprawozdanie i jaki rodzaj informacji powinno zawierać. Kolejny element to jasne zdefiniowanie poszczególnych kryteriów,

a zwłaszcza takich terminów jak rekord, cyfrowa dokumentacja wizualna, obiekt zdigitalizowany, obiekt udostępniony. Opracowane wytyczne powinny stać się podstawą do zweryfikowania tych części sprawozdań, które odnoszą się do digitalizacji. Ujednolicenie danych we wszystkich typach sprawozdań pozwoli na prowadzenie krajowych statystyk. Zmiana sposobu sprawozdawczości w zakresie digitalizacji powinna zostać poprzedzona szkoleniami.

Zakończenie

Współczesne wymagania i potrzeby muzealnictwa oraz analiza zaleceń unijnych dotyczących cyfryzacji dziedzictwa kulturowego, zwłaszcza w zakresie opisu muzealiów, jednoznacznie wskazują na konieczność zmiany dotychczasowej strategii digitalizacji i powołania centrum kompetencji dla zbiorów muzealnych. Instytucja taka powinna mieć znacząco rozszerzone kompetencje o opracowywanie i wdrażanie dobrych praktyk i krajowych standardów, prowadzenie monitoringu procesu, a przede wszystkim powinna nadzorować krajowy agregator dla zbiorów muzealnych, z jednej strony gwarantujący udostępnianie informacji o obiektach, z drugiej zabezpieczający dane. Centrum kompetencji powinno być przede wszystkim wyspecjalizowane w zagadnieniach związanych z digitalizacją obiektów muzealnych, a swoje działania podejmować, mając na względzie potencjał, ale i też ograniczenia polskich muzeów. Złożoność tych zadań, a także olbrzymia różnorodność opracowywanych obiektów, nieporównywalna ze zbiorami bibliotecznymi czy archiwalnymi, przemawia za koniecznością wprowadzenia kompleksowych rozwiązań i przyznania pełni kompetencji, ale zapewnienia też odpowiednich środków konkretnej instytucji. Skonsolidowanie działań związanych z digitalizacją muzealiów może przynieść w krótkiej perspektywie czasu nie tylko korzyści wymierne w postaci udostępnienia większej ilości obiektów, ale także pozwolić na bardziej racjonalne wydawanie środków publicznych.



NARODOWY
INSTYTUT MUZEALNICTWA
I OCHRONY ZBIORÓW

Narodowy Instytut Muzealnictwa i Ochrony Zbiorów

ul. Goraszewska 7

02-910 Warszawa

tel. (+48 22) 25 69 600

fax: (+48 22) 25 69 650

e-mail biuro@nimosz.pl

<http://www.nimosz.pl/>