

Metadane w architekturze informacji

Anna Matysek

ORCID 0000-0003-1042-7895

Uniwersytet Śląski w Katowicach

Wydział Humanistyczny

Jacek Tomaszczyk

ORCID 0000-0002-9199-659X

Uniwersytet Śląski w Katowicach

Wydział Humanistyczny

Abstrakt

Cel/Teza: Celem artykułu jest przybliżenie roli metadanych w projektowaniu architektury informacji serwisów internetowych. Skupiono się przede wszystkim na metadanych opisowych i strukturalnych, które mają istotny wpływ na organizację i wyszukiwanie informacji na stronach WWW.

Koncepcja/Metody badań: Wyjaśniono czym są metadane, przedstawiono ich rodzaje, standardy oraz funkcje w środowisku cyfrowym. Wyeksponowano znaczenie metadanych w przeglądaniu i wyszukiwaniu informacji. Pokazano sposoby umieszczania metadanych na stronach internetowych tworzonych bezpośrednio w języku HTML oraz w systemie zarządzania treścią WordPress.

Wyniki i wnioski: Metadane pełnią istotną rolę w architekturze informacji serwisów internetowych. Umożliwiają opisywanie obiektów cyfrowych i tworzenie semantycznych powiązań między nimi, co zwiększa efektywność wyszukiwania informacji i użyteczność serwisów. Zaleca się zatem, aby architekci informacji odpowiednio dobierali schematy metadanych, dbając o spójność i szczegółowość ich elementów.

Oryginalność/Wartość poznawcza: Zagadnienie metadanych w organizowaniu i wyszukiwaniu informacji w serwisach internetowych nie zostało do tej pory szczegółowo wyjaśnione w piśmiennictwie dotyczącym architektury informacji.

Słowa kluczowe

Architektura informacji. Dublin Core. Fasety. Metadane. RDF. Tezaurus. Znacznik meta.

Otrzymano: 24 marca 2021. Zrecenzowano: 18 kwietnia 2021. Poprawiono: 6 lipca 2021.

Zaakceptowano: 2 sierpnia 2021.

1. Wprowadzenie

Metadane są na ogół kojarzone z rzeczowym i formalnym opracowaniem zbiorów bibliotecznych i archiwalnych, podczas którego wykorzystuje się je do tworzenia opisów bibliograficznych dokumentów w celu zorganizowania ich w taki sposób, aby ułatwić użytkownikom wyszukiwanie informacji. Są więc często postrzegane jako narzędzie organizacji wiedzy, rozumianej jako „obszar badań i zespół działań związanych z generowaniem opisów dokumentów, ich treści, przeznaczenia i własności formalnych oraz z organizacją tych opisów zapewniającą dostęp do określonych dokumentów lub ich fragmentów” (Sosińska-Kalata, 2001, 30).

Metadane znajdują również zastosowanie w innych systemach informatycznych i informacyjnych, np. w:

- słownikach danych, zawierających opisy struktur baz danych, m.in. nazwy i definicje: tabel, kluczy, relacji, widoków;
- logach (plików dziennika), rejestrujących zdarzenia w systemie operacyjnym lub innym oprogramowaniu;
- rejestratorach działań użytkowników w systemie lub aplikacjach.

Metadane są także tworzone podczas korzystania z programów użytkowych. W trakcie zapisywania na dysku pliku tekstowego, np. w programie Microsoft Word, generowana jest automatycznie metryczka, w której są rejestrowane różnego typu informacje dotyczące pliku: daty jego utworzenia, rozmiaru, liczby stron, autorstwa, lokalizacji i wiele innych. Metadane dodawane są również do plików tworzonych przez urządzenia, np. w plikach graficznych zapisywanych przez aparaty cyfrowe, w plikach dźwiękowych tworzonych przez rejestratory głosu czy w rejestrach połączeń generowanych przez smartfony (nazwa rozmówcy, długość trwania rozmowy, data, rodzaj połączenia).

Bibliotecznym i informatycznym metadaniem poświęcono wiele miejsca w literaturze naukowej i fachowej (Ajdukiewicz-Tarkowska, 2008; Baca, 2016; Bednarek & Wróbel, 2013; Caplan, 2003; Gartner, 2016; Greenberg, 2005; Malak et al., 2017; Nahotko, 2004; Pomerantz, 2015; Smiraglia, 2005; Wassilew & Papińska-Kacperek, 2017; Zeng & Qin, 2016; Żytniewski, 2004), jednak to zagadnienie rzadko pojawia się w piśmiennictwie dotyczącym architektury informacji i jest opisywane ogólnikowo (Ding i in., 2017; Garrett, 2011; Martin, 2019; Spencer, 2010). Z tego powodu w artykule chcemy dokładniej spojrzeć na problematykę metadanych pod kątem ich funkcji i zastosowania w serwisach internetowych oraz ich roli w projektowaniu architektury informacji. Uwagę skupimy na metadanych opisowych i strukturalnych, które mają istotny wpływ na organizację i wyszukiwanie informacji w internecie. Stosując metodę przeglądu piśmiennictwa, wyjaśniamy, czym są metadane, przedstawiamy ich rodzaje, standardy oraz funkcje w środowisku cyfrowym. W dalszej części artykułu podajemy przykłady zastosowania metadanych w wyszukiwaniu informacji, a następnie pokazujemy sposoby umieszczania metadanych na stronach internetowych tworzonych bezpośrednio w języku HTML oraz w systemie zarządzania treścią WordPress.

2. Termin

Autorstwo terminu *metadane* przypisuje się Philipowi Bagleyowi, który w 1968 r. użył tego słowa w raporcie technicznym „Extension of Programming Language Concepts”, a wraz z jego publikacją termin ten zaczął się rozpowszechniać na wielu polach działalności praktycznej i naukowej, przede wszystkim w technologii informacyjnej, nauce o informacji, bibliotekoznawstwie, branży wideo (Weynand et al., 2015, 221). Jednak warto zauważyć, że przed P. Bagleyem, w 1967 r., David Griffel i Stuart McIntosh w swoim raporcie również pisali o metadanych w kontekście opisywania za ich pomocą rekordów zawierających informację na temat otoczenia systemu. Utożsamiali oni metadane z danymi bibliograficznymi i określali je jako *meta data records* (Griffel & McIntosh, 1967, 6). Zatem palmę pierwszeństwa do autorstwa terminu *metadata* należałoby jednak przyznać D. Griffelowi i S. McIntoshowi. Natomiast sam prefiks *meta* wywodzi się z języka greckiego i oznacza

m.in. *wśród, po, powyżej, poza*. Ostatnie dwa znaczenia wykorzystuje się do tworzenia z tym prefiksem słów, za pomocą których chcemy wyrazić bardziej abstrakcyjne znaczenie czy też spojrzeć na pojęcie z wyższego poziomu abstrakcji.

3. Pojęcie

Najprawdopodobniej najbardziej popularną definicją metadanych jest określenie, że są to *dane o danych*. Zwięzłość tej definicji to niewątpliwa zaleta w dydaktyce i komunikacji specjalistycznej, jednak nie sprzyja ona lepszemu zrozumieniu tego pojęcia. Przede wszystkim dlatego, że użyty w niej termin *dane* jest trudny do zdefiniowania i często intuicyjnie rozumiany, co może skutkować powierzchownością oceny przydatności metadanych w projektowaniu produktów cyfrowych. Głębszy wgląd w to pojęcie uzyskamy, analizując bardziej rozbudowane definicje i objaśnienia metadanych.

- (1) Metadane to ustrukturyzowane, czytelne maszynowo dane, zawierające charakterystykę cyfrowych obiektów informacyjnych, służącą ich efektywnemu oraz trafnemu wyszukiwaniu, szczególnie w wielkich zasobach informacji w Internecie, zarządzaniu nimi i ich wartościowaniu (Nahotko, 2004, 15).
- (2) Wszelkie dane o dokumentach lub zbiorach dokumentów odnoszące się do ich treści, parametrów technicznych i fizycznych. Metadane mogą się odnosić do wszelkich dokumentów bez względu na sposób ich wytworzenia lub zapisu, w tym także do dokumentów elektronicznych. Metadane mogą określać istotne elementy, takie jak: tematykę dokumentów lub zbiorów dokumentów, osoby lub instytucji odpowiedzialnych za powstanie, czas wytworzenia, sposób zapisu, zasady dostępu itd. (NDAP & ICM, 2005, 7).
- (3) Metadane definiowane są jako ustrukturyzowane, zazwyczaj czytelne maszynowo dane, zawierające charakterystykę dokumentów służących ich efektywnemu oraz trafnemu wyszukiwaniu, zarządzaniu nimi i ich wartościowaniu (Czapnik et al., 2011, 197).
- (4) Metadata is key to the functionality of the systems holding the content, enabling users to find items of interest, record essential information about them, and share that information with others." [„Metadane są kluczem do funkcjonalności systemów przechowujących treści, umożliwiając użytkownikom odnajdywanie interesujących ich obiektów, rejestrowanie istotnych informacji o tych obiektach oraz dzielenie się tymi informacjami z innymi (Riley, 2017, 6).
- (5) Data attributes that describe, provide context, indicate the quality, or document other object (or data) characteristics [Atrybuty danych, które opisują, nadają kontekst, wskazują jakość lub dokumentują inne cechy obiektu (lub danych)] (Greenberg, 2005, 20).
- (6) [Metadata] provides well-defined content in structured representations that make it easier to share and discover. [Metadane dostarczają dobrze zdefiniowaną treść w ustrukturyzowanych reprezentacjach, które ułatwiają jej udostępnianie i odkrywanie.] (Gordon & Habermann, 2018, 38).

Z powyższych definicji wynika, że metadane to ustrukturyzowane dane, które służą przede wszystkim do opisu i wyszukiwania treści oraz zarządzania dokumentami. Na dane i metadane możemy też spojrzeć inaczej, co ułatwi nam lepsze skojarzenie ich z architekturą informacji serwisów internetowych. Przyjmijmy, że otaczający nas świat składa się z obiektów. Obiektem może być przedmiot, osoba, zwierzę, proces, zjawisko, a także wszystko to, o czym możemy pomyśleć, mimo że w realnym świecie nie istnieje (np. fikcyjne przedmioty, miejsca i postaci). Za obiekty uznajmy też relacje, które mogą łączyć różne obiekty (np. relacja posiadania, sprawczości, lokalizacji i in.). W celu zdefiniowania pojęcia reprezentującego obiekt (a ściślej mówiąc klasę obiektów), zidentyfikowania obiektu lub odróżnienia jednego obiektu od innego posługujemy się cechami obiektu, gdyż każdy

obiekt, zarówno realny, jak i abstrakcyjny, można opisać za pomocą zbioru cech. Różne obiekty mogą mieć różne cechy. Obiekty należące do tej samej klasy mają pewien podzbiór wspólnych cech, które w równym stopniu przysługują każdemu obiektowi tej klasy. Oprócz cech wspólnych, na podstawie których zaliczamy obiekt do danej klasy, obiekt może mieć jeszcze inne, indywidualne cechy, np. unikatowy identyfikator. Przykładami klas są *pojazdy*, *napoje*, *aplikacje*, *studenci*, *podręczniki*. **Cechy danego obiektu to metadane, natomiast wartości tych cech to dane.** Za przykład posłuży nam obiekt *laptop* (Rys. 1). Wśród jego cech możemy wyróżnić m.in.: producenta, rozmiar (przekątną ekranu), procesor, jaki w nim się znajduje, ilość pamięci RAM, dysk, system operacyjny, kolor, wagę. Wymieniony zbiór cech, nazywany również zestawem atrybutów, możemy uznać za metadane. Jeśli teraz za pomocą tych cech opiszemy konkretny obiekt (laptop) i przypiszemy wartości owym cechom, wówczas wartości te będziemy nazywać danymi.

| Metadane (zbiór cech) | |
|-----------------------|-----------------------------|
| Klasa obiektów: | laptopy |
| Producent: | DELL |
| Przekątna ekranu: | 13,3" |
| Procesor: | Intel Core i7 |
| Pamięć RAM: | 128 GB |
| Dysk: | SSD 512 GB |
| System operacyjny: | Microsoft Windows 10 Pro PL |
| Kolor: | szary |
| Waga: | 1,25 kg |

Rys. 1. Przykład metadanych i ich wartości

Zestaw cech różni się, co oczywiste, w zależności od klasy obiektu. Obrazy, książki, filmy, urządzenia, potrawy – wszystkie są opisywane za pomocą cech charakterystycznych dla obiektów tworzących te klasy. Różnice wynikają nie tylko z odmiennych cech charakterystycznych, ale mogą także zależeć od osoby (jej wiedzy, doświadczenia), która tworzy ich opisy i na swój własny sposób postrzega otaczający ją świat. Metadane są subiektywne w tym, co zawierają, co pomijają, gdzie zakreślają swoje granice i jakich terminów używają do opisu (Gartner, 2016, 4). Aby możliwa była wymiana danych pomiędzy różnymi systemami czy bazami danych, dla niektórych klas obiektów utworzono standardowe zestawy cech (atrybutów), które nazywamy **schematem metadanych**. Schemat metadanych to zestaw elementów metadanych i zasad jego stosowania (ISO 2015). Schematy metadanych zazwyczaj określają nazwy elementów i ich semantykę (źródło). Z kolei element metadanych to nazwa właściwości zasobu, która może być użyta w metadanych i której można nadać wartość (ISO 2015). Wiele schematów metadanych, które przyjęły się w powszechnym użyciu, doczekało się swoich norm lub standardów¹, np. Exchangeable Image File Format (EXIF) dla plików z obrazami, MPEG-7 do opisu treści multimedialnych, ISO 19115 dla informacji geograficznych,

¹ Wyjaśnienie różnic między normą a standardem można znaleźć w opracowaniu: Matysek, 2014, 29–30.

EAD (Encoded Archival Description) do opisu zasobów archiwalnych lub MoReq2 (Model Requirements For The Management of Electronic Records) do zarządzania dokumentami elektronicznymi. Wiele standardów metadanych powstało do przetwarzania danych naukowych, a z większością z nich można zapoznać się w katalogu metadanych (RDA, b.d.).

Standaryzacja metadanych przebiega na trzech płaszczyznach: semantycznej, syntaktycznej i pragmatycznej. Przykładem standaryzacji na płaszczyźnie semantycznej jest Dublin Core Metadata Element Set (DC), rozwijany od 1995 r. przez Dublin Core Metadata Initiative. Służy on do opisu zasobów internetowych w celu umożliwienia efektywnego wyszukiwania informacji, a przykład fragmentu metadanych DC znajduje się w tabeli 1. W wersji 1.1 zestaw metadanych składa się z 15 elementów (Brachfogel, 2009):

- (1) współtwórca (contributor),
- (2) zasięg (coverage),
- (3) twórca (creator),
- (4) data (date),
- (5) opis (description),
- (6) format (format),
- (7) identyfikator (identifier),
- (8) język (language),
- (9) wydawca (publisher),
- (10) powiązanie (relation),
- (11) prawa (rights),
- (12) źródło (source),
- (13) temat (subject),
- (14) tytuł (title),
- (15) rodzaj (type).

Tab. 1. Przykład wycinka metadanych dla książki elektronicznej zapisanych w Dublin Core

| Nazwa elementu w Dublin Core | Przykładowa wartość |
|------------------------------|--|
| title | Nie każ mi myśleć! O życiowym podejściu do funkcjonalności stron internetowych |
| creator | Steve Krug |
| publisher | Helion |

W 2002 r. doprecyzowano standard Dublin Core Metadata Initiative Metadata Terms (DCMI, 2020), rozszerzając go o kilkadziesiąt dodatkowych elementów: własności, schematy kodowania słownictwa, schematy kodowania składni i klasy. Opis elementów (terminów) metadanych Dublin Core jest ujednolicony i składa się z:

- unikatowej wśród terminów DCMI nazwy, np. Creator;
- zrozumiałej dla człowieka etykiety, np. Creator;
- URI (Uniform Resource Identifier), który jest unikatowym i stałym identyfikatorem obiektu cyfrowego, np. <http://purl.org/dc/terms/creator>;
- definicji, np. An entity responsible for making the resource;
- rodzaju terminu, np. Property.

Terminy mogą też być uzupełnione o dodatkowe atrybuty, np. komentarz, wersja (Brachfogel, 2010, 59).

Standaryzacji na poziomie syntaktycznym dokonuje się za pomocą języków kodowania danych, np. XML czy RDF, które określają fizyczny sposób zapisu metadanych z wykorzystaniem określonych konwencji i formatów danych, co umożliwia m.in. przesyłanie danych między urządzeniami. RDF (ang. Resource Description Framework), opracowany przez The World Wide Web Consortium (W3C) (RDF, 2014), jest standardem reprezentacji wiedzy w Sieci Semantycznej (Roszkowski, 2017, 53). Służy do opisywania zasobów (obiektów) za pomocą stwierdzeń (zdań) składających się z trzech elementów: przedmiotu (zasobu), predykatu i obiektu. Stwierdzenie podobne jest do zdania w języku naturalnym, lecz jest tak sformalizowane, aby mogło być łatwo przetwarzane przez komputer. Na przykład zdanie: *Autorem książki A jest XY* możemy wyrazić jako *Książka A jest autorstwa XY* i dokonać reprezentacji tego zdania za pomocą elementów RDF, jak w tabeli 2.

Tab. 2. Elementy RDF

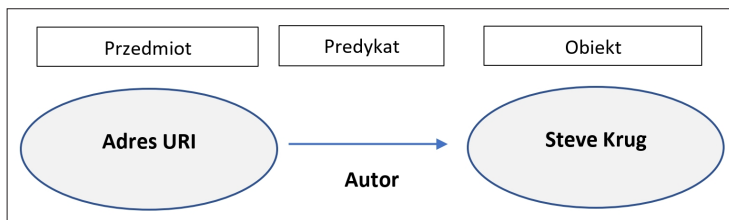
| Nazwa elementu | Co zawiera | Przykładowa zawartość |
|----------------|---|---|
| Przedmiot | adres URI ² opisywanego zasobu, np. książki | https://helion.pl/ksiazki/nie-kaz-mi-myslec-o-zyciowym-podejsciu-do-funkcjonalnosci-stron-internetowych-wydanie-iii-steve-krug,nieka3.htm#format/e |
| Predykat | reprezentuje cechę/własność zasobu, np. autorstwo (twórcę) | Autor |
| Obiekt | zawiera konkretną wartość cechy, np. imię i nazwisko twórcy | Steve Krug |

Informacje zawarte w tabeli 2 można zapisać za pomocą XML (Rys. 2) lub przedstawić w formie grafu (Rys. 3).

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="https://helion.pl">
  <rdf:Description rdf:about="https://helion.pl/ksiazki/nie-kaz-mi-myslec-o-zyciowym-podejsciu-do-funkcjonalnosci-stron-internetowych-wydanie-iii-steve-krug,nieka3.htm#format/e">
    <autor>Steve Krug</autor>
  </rdf:Description>
</rdf:RDF>
```

Rys. 2. Przykładowy rekord RDF zapisany w języku XML

² URI to zwykle łańcuch znaków zapisanych zgodnie ze składnią określoną w specyfikacji <https://tools.ietf.org/html/rfc2396>. Łańcuch ten określa nazwę (URN) lub adres (URL) zasobu. Adres URL wskazuje dodatkowo lokalizację, z której można pobrać dany zasób.



Rys. 3. Przykładowy rekord RDF zaprezentowany w formie grafu

Standaryzacja na płaszczyźnie pragmatycznej dotyczy jednolitego stosowania utworzonych schematów i zasad tworzenia wartości metadanych. Są to różnego rodzaju instrukcje wraz z towarzyszącymi im słownikami, listami wartości itp. Przykładem może być lokalna (stosowana np. w jednej firmie) instrukcja dla osób wprowadzających opisy produktów w sklepie internetowym, a źródłami wartości – katalogi producentów tych produktów.

4. Rodzaje metadanych

Metadane możemy podzielić na wiele sposobów, przyjmując różne kryteria podziału, m.in.: cel tworzenia, przeznaczenie, pochodzenie, miejsce przechowywania czy użytkownika. Nahotko (2004) czy Mayernik (2020) wymieniają kilkanaście kategorii metadanych. Metadane najczęściej dzieli się na:

- (1) opisowe (ang. *descriptive*);
- (2) strukturalne (ang. *structural*);
- (3) administracyjne (ang. *administrative*);
- (4) statystyczne (ang. *statistical / use*).

4.1. Metadane opisowe

Za pomocą tego rodzaju metadanych opisujemy obiekt w taki sposób, aby można go było odszukać. W przypadku zasobów informacyjnych elementami opisu są np. atrybuty: tytuł, temat, autor, abstrakt, słowa kluczowe, tagi. Inne obiekty mogą być opisane za pomocą atrybutów, które odzwierciedlają ich najważniejsze z wyszukiwawczego punktu widzenia cechy, np.: producent, wielkość, waga, kolor, typ, cena, a w przypadku urządzeń nazwy parametrów, które mają znaczenie w ich użytkowaniu, np.: moc, wydajność, dokładność, rozdzielczość, pojemność, klasa energetyczna itp. Warto zauważyć, że, opisując jakiś obiekt, możemy w metadanych umieszczać zarówno cechy jego samego, jak i tych jego komponentów, które użytkownicy uważają za istotne. Atrybuty samego laptopa to np.: marka, wielkość (ekranu), kolor, waga, ale w opisie znajdują się również cechy ważnych elementów tego laptopa, jak moc procesora, pojemność dysku czy rodzaj matrycy. Cechy obiektów i ich komponentów, które *de facto* też są obiektami, wykorzystuje się do wyszukiwania, przeglądania oraz fasetowego filtrowania informacji w serwisach WWW, o czym więcej napiszemy w dalszej części artykułu.

4.2. *Metadane strukturalne*

Metadane strukturalne to dane dotyczące organizacji zawartości (treści) obiektów cyfrowych. Opisują elementy składowe tworzące obiekt oraz relacje między tymi elementami, ułatwiając wewnętrzną nawigację (ANSI/NISO, 2010a, 8). Wynikają z tego dwie istotne korzyści. Po pierwsze, wyodrębnienie i oznaczenie komponentów obiektu pozwala dotrzeć do nich w sposób nieliniowy. Metadane pozwalają na bezpośredni dostęp do danego rozdziału w cyfrowej książce, do sekcji *Wyniki* w artykule naukowym lub do wybranego nagłówka na stronie WWW. Drugą zaletą jest możliwość łączenia treści pochodzących z różnych obiektów. Taka funkcjonalność ułatwia użytkownikom swobodne przeglądanie zasobów i prezentowanie informacji na różne sposoby oraz porównywanie ze sobą obiektów. Istnieje możliwość wybierania niektórych informacji z treści, np. tylko abstraktów z dokumentów lub spisów treści bez wyświetlania pełnych tekstów. Można także zestawiać zdjęcia z różnych zasobów albo przeglądać recenzje filmów, dodatkowo mając możliwość ich filtrowania (pozytywne, negatywne lub o zadanej liczbie gwiazdek).

Stosowanie metadanych strukturalnych sprzyja patrzeniu na obiekty nie jak na jednolite zasoby informacyjne³, ale postrzeganiu ich jako zbiory autonomicznych (do pewnego stopnia) jednostek informacyjnych, które mogą być wykorzystywane poza natywnymi obiektami, wchodząc w różne scenariusze wyszukiwawcze i prezentacyjne. Metadane tego rodzaju definiują struktury, które łączą prostsze komponenty w zrozumiałe dla użytkownika złożone konstrukty (Gartner, 2016, 8). Można nawet powiedzieć, że metadane strukturalne niejako sterują dostarczaniem użytkownikom treści w odpowiedniej formie.

4.3. *Metadane administracyjne*

Te metadane obejmują dane techniczne dotyczące pliku i powstają zazwyczaj automatycznie przy tworzeniu zasobu cyfrowego (dokumentu tekstowego, zdjęcia, pliku bazodanowego). Zwykle zawierają typ pliku, datę jego utworzenia, modyfikacji bądź ostatniego otwarcia, informację o prawach własności intelektualnej, prawach dostępu do pliku, parametrach urządzenia, które ten plik utworzyło (np. model aparatu fotograficznego). Metadane administracyjne pomagają w utrzymaniu i archiwizacji plików oraz zapewniają ich długoterminową dostępność. Są niezbędne do migracji danych z jednego formatu do drugiego, ponieważ przechowują dane na temat typu i wersji oprogramowania użytego do utworzenia pliku⁴ (ANSI/NISO, 2010a, 8).

4.4. *Metadane statystyczne*

Metadane statystyczne to dane dotyczące użytkownika obiektu. Rejestrują one np.: liczbę pobrań pliku, datę, IP i system operacyjny komputera, z którego nawiązano połączenie,

³ Podejście to wciąż dominuje w opracowaniu dokumentów bibliotecznych; na ich tle wyróżniają się Biblioteka Narodowa i Centrum NUKAT, które od pewnego czasu opisują rozdziały z prac zbiorowych (zob. <https://www.bn.org.pl/aktualnosci/720-artykuly-i-rozdzialy-publicacji-zbiorowych-uwzglednione-w-katalogu-bn.html>; <https://centrum.nukat.edu.pl/pl/poznaj-nukat/historia-nukat>).

⁴ Nie wszystkie programy obsługują dany standard metadanych, przez co dane mogą ulec zniekształceniu lub zostać całkowicie usunięte podczas edycji dokumentu bez wiedzy użytkownika.

przeglądarkę internetową i wiele innych danych. Metadane statystyczne mogą ujawniać informację o użytkownikach i ich zachowaniach, a także dostarczać danych o sieciach społecznych i powiązaniach pomiędzy osobami, miejscami i organizacjami (Pomerantz, 2015, 65–66).

5. Organizacja metadanych

Metadane są przechowywane bezpośrednio w strukturze zasobów informacyjnych, które opisują, bazy danych systemu informacyjnego, a także w plikach aplikacji (np.: automatycznie tworzona metryczka w pliku Microsoft Word, sekcja META w nagłówku dokumentu HTML, informacje w formacie Exif automatycznie dołączone do zdjęcia cyfrowego) lub w oddzielnych plikach (np. rekordy bibliograficzne). Metadane mogą być organizowane w różnego typu struktury, tzw. systemy organizacji wiedzy (SOW), które ułatwiają porządkowanie treści serwisu internetowego oraz ich przeszukiwanie i zarządzanie nimi. Jednymi z najczęściej wykorzystywanych struktur są taksonomie, które w sposób hierarchiczny porządkują metadane. Taksonomie mogą być wykorzystywane do indeksowania, nawigowania i filtrowania zawartości serwisów internetowych (Tomaszczyk, 2007, 42–43). Spotykamy je w indeksach stron WWW⁵, a także w mapach witryn⁶.

Bardziej zaawansowaną strukturą organizacyjną metadanych, a zarazem danych, jest tezaurus. To zbiór terminów zwanych deskryptorami, które służą do indeksowania (opisywania) treści oraz askryptorami, które zapewniają przekład z języka naturalnego na kontrolowane słownictwo deskryptorowe. Terminy zgromadzone w tezaurucie są grupowane w ministruktury nazywane artykułami deskryptorowymi, w których oznaczano relacje semantyczne między deskryptorami, a także relację przekładową między deskryptorami a askryptorami. Relacje semantyczne oznaczane w artykule deskryptorowym obejmują: relację ekwiwalencji, relacje hierarchiczne (nadrzędności i podrzędności) oraz relacje kojarzeniowe (Ścibor & Tomasik-Beck, 1995; Woźniak-Kasperek, 2005).

Zgodnie z polskimi i międzynarodowymi standardami (ANSI/NISO, 2010b; ISO, 2011, 2013; PN, 1992), w artykule deskryptorowym umieszcza się następujące elementy i ich oznaczenia:

- Deskryptor tytułowy (hasłowy) – wyróżniony typograficznie, najczęściej pisany wielkimi literami i/lub czcionką półgrubą;
- DEF – definicja opisywanego deskryptora (ang. SN – scope note);
- NU – nie używaj – askryptor bądź askryptory wskazane jako terminy niestosowane w indeksowaniu (ang. UF – use for);
- GD – główny deskryptor – deskryptor nieposiadający nadrzędnego deskryptora (ang. TT – top term);
- SD – szerszy deskryptor – deskryptor bądź deskryptory nadrzędne do opisywanego deskryptora (ang. BT – broader term);
- WD – węższy deskryptor – deskryptor bądź deskryptory podporządkowane hierarchicznie (tzn. podrzędne) do opisywanego deskryptora (ang. NT – narrower term);

⁵ Zob. np. https://www.laroche.edu/Site_Index/

⁶ Zob. np. <https://www.apple.com/pl/sitemap/> lub <https://www.komputronik.pl/sitemap>

- KD – kojarzeniowy deskryptor – deskryptor powiązany znaczeniowo z deskryptorem tytułowym, lecz niepozostający z nim w bezpośredniej zależności hierarchicznej (ang RT – related term).

W teaurusie znajdują się też askryptory, czyli terminy odrzucone, konwencjonalnie zapisywane małymi literami i/lub zwykłą czcionką. W artykule askryptorowym askryptor powiązany jest z odpowiadającym mu deskryptorem za pomocą oznaczenia „U – używaj” (ang. USE). Związek między askryptorem i deskryptorem to relacja synonimii i/lub tzw. ekwiwalencji wyszukiwawczej. Rysunek 4 przedstawia przykład artykułu deskryptorowego z *Digizaurusa* (NMM, 2013). Relacje hierarchiczne oznaczone są w nim skrótami angielskimi: BT (termin szerszy), NT (termin węższy)⁷ oraz RT (termin powiązany).

kamień

Strona główna ► NIEORGANICZNY ► kamień

Nota o zakresie:

odłamek skalny, otoczek. [E PWN]

.BT. NIEORGANICZNY

kamień

- .NT2 alabaster
- .NT2 granit
- .NT2 krzemień
- .NT2 marmur
- .NT2 piaskowiec
- .NT2 wapień

.RT. kamień szlachetny lub ozdobny

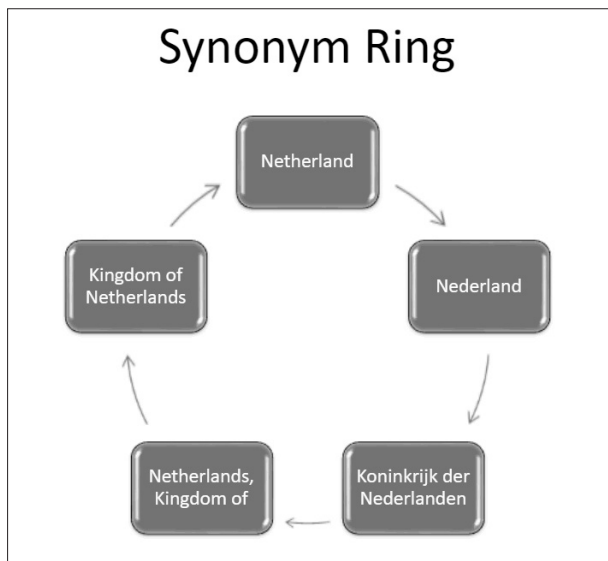
Rys. 4. Przykład artykułu deskryptorowego.

Źródło: Digizaurus – Materiał [17.02.2021],

<http://digizaurus.nmm.pl/vocab/index.php?tema=217&/kamie>

Taksonomie od teaurusu odróżnia brak relacji kojarzeniowych, a także relacji ekwiwalencji, jednak taksonomie mogą być wyposażone w pierścienie synonimów (ang. *synonym rings*), czyli zestawy terminów, które dla celów wyszukiwania informacji można uznać za równoważne (Rosenfeld et al., 2017, 231). Mogą to być wyrażenia równoznaczne (warianty pisowni i skrótowce), wyrażenia bliskoznaczne (synonimy i quasisynonimy), a nawet odpowiedniki w innych językach (tłumaczenia), czego przykład znajduje się na rysunku 5. Zwykle są one niewidoczne dla użytkownika, ale są istotne podczas wyszukiwania, gdy wpisujemy pewien termin jako hasło wyszukiwawcze, a w wynikach wyszukiwania otrzymujemy jego „zamienniki”.

⁷ NT2 oznacza drugi poziom podrzędności.



Rys. 5. Pierścień synonimów. Źródło: Jijavadze & Gabrichidze, 2010, 16

Najbardziej złożoną strukturą organizacji metadanych jest ontologia. Przypomina ona tezaurus, ale wykorzystuje o wiele bogatszy i dokładniej opisany zestaw relacji semantycznych między terminami. Jedną z najczęściej cytowanych definicji ontologii pochodzi od Thomasa R. Grubera (1993), który określa ją jako: „jawną specyfikację konceptualizacji”⁸, gdzie przez konceptualizację rozumie się: „abstrakcyjny, uproszczony sposób widzenia świata, który chcemy reprezentować w określonym celu”. Ontologie, stanowiąc model rzeczywistości lub jego wycinka, składają się z pojęć oraz zachodzących między nimi relacji, tworząc złożone struktury relacyjne (Tomaszczyk, 2014, 155). Dzięki nim komputer może zrozumieć, w jakim kontekście użyte są pojęcia. Ontologie są podstawową strukturą służącą zarządzaniu i rozpowszechnianiu relewantnych metadanych w Semantycznym Webie, gdzie metadane wykorzystuje się do konstruowania relacji definiujących wzorce wiedzy (Nahotko, 2005, 2). Najczęściej zapisywane są za pomocą języka RDF lub OWL (Web Ontology Language).

6. Metadane w przeglądaniu i wyszukiwaniu informacji

Jednym z zadań architekta informacji jest zaprojektowanie serwisu, w którym użytkownik szybko i bez problemu odnajduje informacje. Wyszukiwarka w witrynie WWW jest niewątpliwie przydatna, ale tylko wówczas, gdy użytkownik wie, czego szuka i potrafi wyrazić to słowami. Jednak w wielu przypadkach użytkownicy nie mają ściśle sprecyzowanych wymagań i dlatego wolą przeglądać strony WWW, aby zobaczyć, co oferuje serwis, jakie treści są dostępne lub czym charakteryzują się interesujące go produkty.

⁸ Oryg.: explicit specification of conceptualization.

Jednym z rodzajów strategii wyszukiwania informacji są strategie interaktywne, koncentrujące się na sposobach interakcji użytkownika z systemem. Zalicza się do nich dwie główne strategie: **przeglądanie** oraz **formułowanie zapytań**. W obu istotną rolę odgrywiają metadane, które mają zastosowanie w nawigowaniu i opisywaniu obiektów.

Przeglądanie zasobów informacyjnych opiera się na nawigowaniu, czyli przemieszczaniu się między stronami lub różnymi elementami zawartości serwisu za pomocą odsyłaczy (hiperłączy, linków). Do nawigacji wykorzystuje się ich dwa typy (Kazienko, 1998, 46):

- (1) linki znaczeniowe (semantyczne), które łączą treści o podobnej tematyce;
- (2) linki organizacyjne, wykorzystywane do usprawnienia poruszania się po strukturze systemu (linki w menu, w stopce, przyciski *Dalej*, *Wstecz*).

Schematy nawigacyjne mogą być tworzone na podstawie metadanych, którymi są nazwy obu typów hiperłączy, a także opisy w ramkach/etykietach, które pojawiają się w niektórych systemach po najechaniu kursorem na link, informując w ten sposób, dokąd ten link prowadzi.

W ostatniej dekadzie bardzo popularnym sposobem przeglądania zasobów cyfrowych stało się **przeglądanie fasetowe**, którego idea wywodzi się z klasyfikacji fasetowej⁹, stworzonej w latach 30. XX w. przez indyjskiego matematyka i bibliotekarza S.R. Ranganathana do organizowania informacji o piśmiennictwie (Klasyfikacja Dwukropkowa)¹⁰.

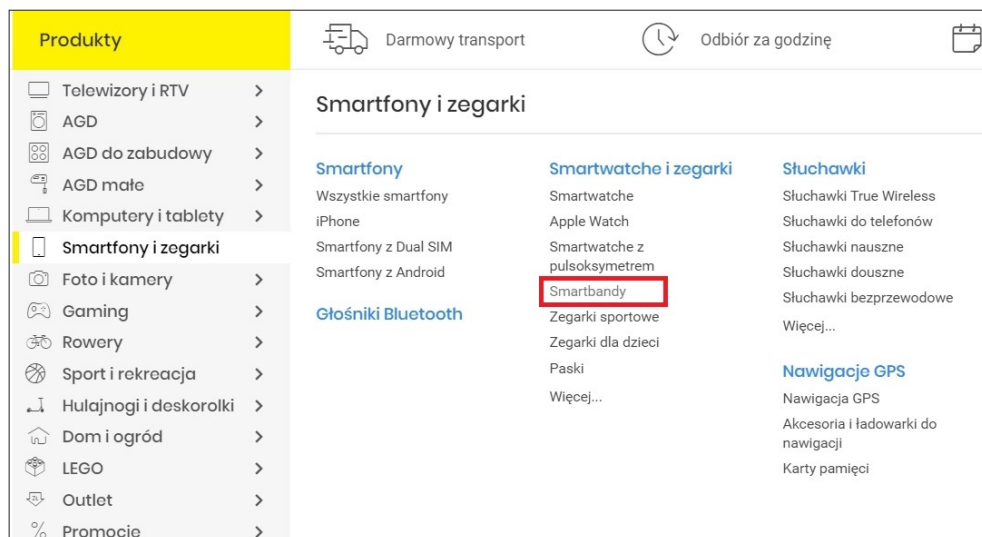
Podstawą przeglądania fasetowego są metadane wykorzystane do opisu elementów wchodzących w skład serwisów internetowych (np. grafik, filmów, produktów w sklepie internetowym). Reprezentując cechy obiektów, pozwalają na ich prezentację z różnych perspektyw. Im więcej metadanych opisuje obiekt, tym więcej można ukazać jego aspektów i dróg dotarcia do niego (możliwości znalezienia go w zbiorze/serwisie), gdyż każda cecha to potencjalny punkt dostępu (ang. *access point*) podczas wyszukiwania. Metadane pozwalają także na kategoryzowanie obiektów według różnych kryteriów. **Faseta** to kategoria (niekiedy nazywana podkategorią), która grupuje obiekty ze względu jakąś cechę wspólną. Jeśli obiekty w zbiorze mają cechy C_1, C_2, \dots, C_n , to każda z tych cech może służyć do wyodrębnienia fasety, stanowiąc kryterium podziału zbioru na rozłączne kategorie K_1, K_2, \dots, K_n , do których należą obiekty X_1, X_2, \dots, X_n posiadające odpowiednie cechy ($X_1 C_1 \in K_1, X_2 C_1 \in K_1, \dots, X_1 C_2 \in K_2, X_2 C_2 \in K_2, \dots$). Nazwa fasety jest więc najczęściej elementem metadanych, np. *producent, pojemność baterii, rodzaj dysku, wielkość, kolor*. Fasety w połączeniu z wartościami cech obiektów stanowią podstawę przeglądania fasetowego. Pomagają one użytkownikowi dotrzeć do poszukiwanego obiektu na różne sposoby i szybciej, niż w przypadku przeglądania kategorii ułożonych hierarchicznie w nawigacji.

Przyjrzyjmy się dwóm sposobom poszukiwania opasek monitorujących (smartbandów) z funkcjami GPS i NFC. Korzystając z nawigacji globalnej (Rys. 6), użytkownik otrzyma długą listę opasek, gdzie konieczne będzie przeglądanie opisu każdego przedmiotu. Tu z pomocą przychodzą metadane i przeglądanie fasetowe, które polega na wielokrotnym wybieraniu przez użytkownika cech obiektów (metadanych) oraz ich wartości (danych). W wyniku tego postępowania następuje stopniowe zmniejszanie się liczebności przeglądanego

⁹ Faseta jest terminem zaczerpniętym z jubilerstwa i oznacza ukośnie ściętą krawędź kamienia szlachetnego. Zastosowanie tego terminu przez Ranganathana wynikało najprawdopodobniej z potrzeby wyjaśnienia filozofii wieloaspektowego sposobu organizacji informacji. Zob. Hasło „faseta” w: Bojar, 2002, 60.

¹⁰ Więcej o samej klasyfikacji: Ranganathan, 1967; Satija, 2017; Ungurian, 1975.

zbioru obiektów, gdyż pary *cecha – wartość* działają jak filtry czy wewnętrzna kwerenda uruchamiana automatycznie lub manualnie przez użytkownika (np. kliknięciem w opcję *Filtruj* lub *Zastosuj*).



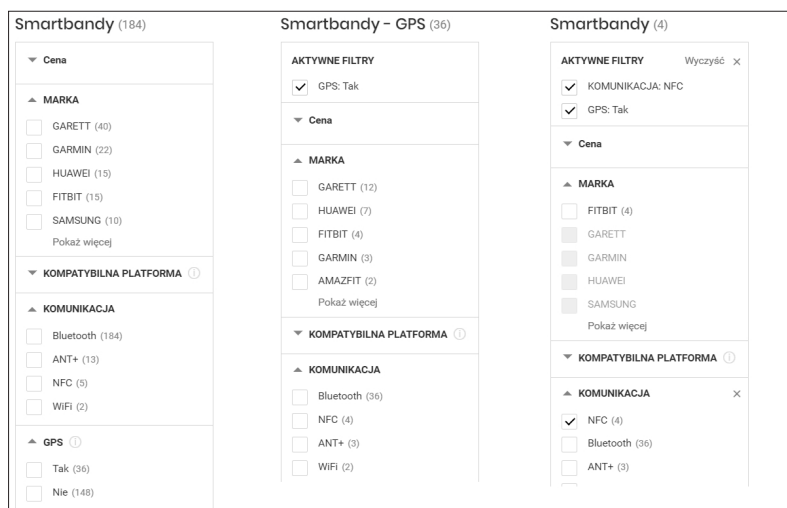
Rys. 6. Lokalizacja smartbandów w nawigacji globalnej sklepu internetowego.

Źródło: Media Expert [01.07.2021], <https://www.mediaexpert.pl/smartfony-i-zegarki/smartwatche-i-zegarki/smartbandy>

Liczba podawana obok wartości cech (Rys. 7) to liczba obiektów spełniających to kryterium, tzn. mających podaną wartość lub zakres wartości. Wybór dokonany w obrębie jednej fasety (tzn. wskazanie cechy i jej wartości) wpływa na liczebność obiektów w wyszukanym zbiorze oraz dostępność (ograniczenie) cech, z których użytkownik może dokonywać kolejnych wyborów. Wyświetlane są tylko te cechy i wartości, które odpowiadają chociaż jednemu obiektowi w tym zbiorze, dzięki czemu użytkownik nie zostaje zaskoczony brakiem wyników. Na rysunku 7 widać, od lewej, jak wybór w obrębie fasety zmniejsza liczebność odnalezionych obiektów. Wybór opasek (smartbandów) z GPS (zaznaczono filtr GPS – Tak), znacząco zmniejszył liczbę produktów (obiektów) w poszczególnych fasetach. Następnie w kategorii „komunikacja” zaznaczono NFC, co zmniejszyło liczbę opasek do zaledwie 4, a jednocześnie ograniczyło aktywne wartości w faszce „marka” do 1.

Formułowanie zapytań to strategia realizowana za pomocą wyszukiwarki wbudowanej w serwis internetowy. Słowa wpisane przez użytkownika do pola wyszukiwawczego są porównywane z tekstami zamieszczonymi w serwisie, a uruchamiana kwerenda (najczęściej po naciśnięciu przycisku *Szukaj* lub *Znajdź*) wyszukuje i wyświetla strony lub inne obiekty, które te słowa zawierają. Wyszukiwarki przeszukują nie tylko treści stron, ale również dane i metadane: cechy obiektów, indeksy, mapy witryny. To właśnie dzięki temu w jednym interfejsie formułowanie zapytań może być zintegrowane z przeglądaniem fasetowym. Taki mechanizm nazywamy **wyszukiwaniem fasetowym**. Jego działanie jest realizowane w dwóch etapach. Najpierw użytkownik wpisuje w okno wyszukiwarki zapytanie, a następnie otrzymane wyniki wyszukiwania może zawęzić (filtrować) za pomocą

faset. Zestaw cech obiektów i ich wartości zależy od wyników wyszukiwania, które zostały dostarczone użytkownikowi. Zestaw ten jest więc dynamiczny, bo jego zawartość zależy od specyfiki obiektów, które zostały odnalezione za pomocą wyszukiwarki. Na rysunku 8 widać dwa różne zestawy faset. Po lewej znajduje się fragment faset, zwykle utożsamiany przez użytkowników z filtrami, który pojawia się po wpisaniu do wyszukiwarki hasła „łódzka”, a po prawej po wpisaniu hasła „telewizor”.



Rys. 7. Zmniejszanie się liczby smartbandów po dokonaniu kolejnych wyborów w obrębie faset.

Źródło: Media Expert [17.02.2021], <https://www.mediaexpert.pl/smartfony-i-zegarki/smartwatche-i-zegarki/smartbandy>



Rys. 8. Zestawy faset zmieniające się w zależności od wpisanego do wyszukiwarki słowa kluczowego. Źródło: RTV EURO AGD [17.02.2021], <https://www.euro.com.pl/>

Praktyczna realizacja koncepcji fasetowego przeglądania i wyszukiwania informacji jest stosunkowo młoda. Pierwsze projekty, takie jak *Dynamic Queries*, *Query Previews* i *View-Based Search*, pojawiły się w latach 90. XX w. i wzbudziły duże zainteresowanie, przede wszystkim w dziedzinie handlu elektronicznego (Pulikowski, 2018, 106). Późniejsze badania empiryczne (Fagan, 2010, 62–63) wskazały na następujące zalety faset:

- są użyteczne w tworzeniu struktur nawigacyjnych;
- ułatwiają efektywne wyszukiwanie w bazach danych;
- pomagają unikać pustych wyników wyszukiwania;
- przyspieszają wyszukiwanie;
- zwiększają skuteczność wyszukiwania relevantnych treści;
- pozwalają uzyskiwać więcej wyników wyszukiwania;
- odpowiadają preferencjom użytkowników, gdyż wolą oni otrzymywać wyniki wyszukiwania uporządkowane w przewidywalne, wielowymiarowe hierarchie;
- dają użytkownikom poczucie pewności siebie i zwiększą satysfakcję z wyszukiwania.

W badaniach eye-trackingowych mierzących czas przez jaki użytkownik patrzy na poszczególne elementy ekranu, ustalono, że fasety są wykorzystywane na wszystkich etapach wyszukiwania (Kules & Capra, 2012, 131).

7. Miejsce metadanych w strukturze serwisów internetowych

Podczas tworzenia strony internetowej metadane można umieszczać w różnych jej miejscach. Jednym z nich są znaczniki META (Rys. 9) w sekcji nagłówka HTML, gdzie wpisuje się podstawowe informacje o zasobie cyfrowym. Do najważniejszych znaczników są zaliczane:

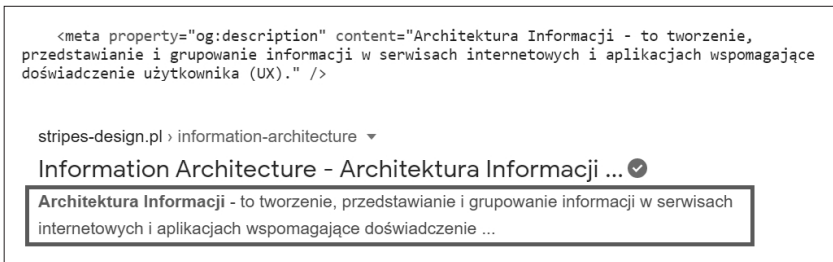
- *title* (tytuł), który jest wyświetlany jako nazwa hiperłącza w wynikach wyszukiwania (SERP). Użytkownicy widzą go także na pasku tytułu karty lub okna, w historii wyszukiwania przeglądarki czy jako domyślną nazwę zakładki dodawanej do przeglądarki;
- *charset* (zestaw znaków), definiujący czcionkę lub kodowanie znaków;
- *description* (opis), będący krótkim opisem strony, wyświetlany w wynikach wyszukiwania¹¹ (Rys. 10). Dla użytkowników przeglądających listę wyników wyszukiwania może to być decydująca informacja, czy warto wejść na daną stronę internetową;
- *robots* (roboty), które informują roboty wyszukiwarek, czy dana strona ma być indeksowana i dostępna do wyszukiwania oraz czy linki do stron wychodzących mają być śledzone i indeksowane;
- *keywords* (słowa kluczowe), charakteryzujące treść podstrony. Dawniej ten znacznik miał duże znaczenie w pozycjonowaniu stron, ale od kilku lat wyszukiwarki, m.in. Google i Bing, ignorują go.

W znacznikach META umieszczamy również metadane opisujące zasób według wybranego schematu (np. Dublin Core Metadata Element Set). Oprócz samych metadanych powinna się tam znaleźć deklaracja wybranego schematu (Rys. 11). Metadane mogą też być przypisywane do plików umieszczanych na stronach internetowych (grafikom, plikom dźwiękowym, animacjom itp.), co prezentujemy na przykładzie WordPressa i pliku audio (Rys. 12).

¹¹ W wyszukiwarce Google wyświetlane jest ok. 150–160 znaków.

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
    <meta name="Description" CONTENT="Author: A.N. Author, Illustrator: P. Pictur
    <meta name="google-site-verification" content="+nxGUDJ40pAZ5l9Bsjudi102tLVC21A
    <title>Example Books - high-quality used books for children</title>
    <meta name="robots" content="noindex,nofollow">
  </head>
</html>
```

Rys. 9. Fragment nagłówka HTML zawierający znaczniki meta.
Źródło: Tagi specjalne, które Google potrafi zinterpretować [17.02.2021],
<https://developers.google.com/search/docs/advanced/crawling/special-tags>



Rys. 10. Zawartość znacznika meta “description” w sekcji nagłówka i ten sam tekst wyświetlony w wynikach wyszukiwania Google.
Źródło: <https://stripes-design.pl/ux/information-architecture/>

```
<link title="schema(DC)" rel="schema.dc" href="http://purl.org/dc/elements/1.1/" />
<meta name="dc.language" content="eng" />
<meta name="dc.source" content="Library of Congress, Washington, D.C. 20540 USA" />
```

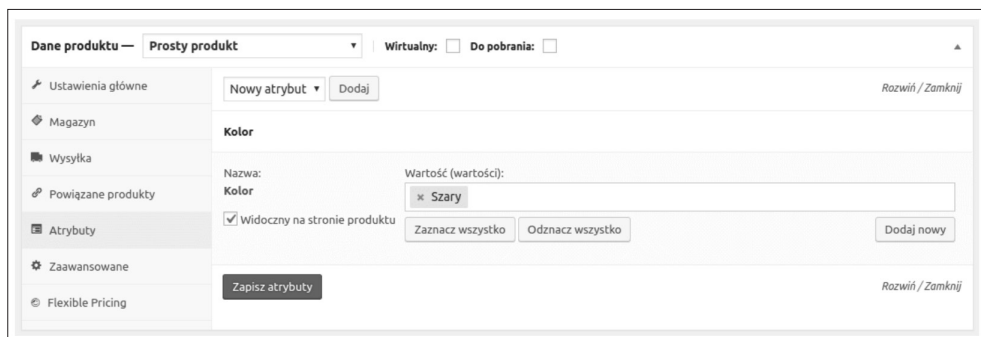
Rys. 11. Metadane Dublin Core zapisane w nagłówku dokumentu HTML.
Źródło: Library of Congress [18.02.2021], <https://loc.gov/>

Rys. 12 Metadane dla pliku audio w WordPress.
Źródło: Zdjęcie ekranu z panelu administratora w WordPress na stronie internetowej autorów

Coraz częściej metadane zamieszczane są też w strukturze dokumentów HTML w postaci mikroformatów (np. hCard, hCalendar), mikrodanych czy JSON-LD (Roszkowski, 2017). Są to strukturalne dane oznaczające różne typy treści na stronie internetowej, czytelne dla robotów wyszukiwarek. Schematy danych strukturalnych można znaleźć w słowniku schema.org. Dzięki danym strukturalnym w wynikach wyszukiwania wyświetla się więcej informacji o danym serwisie, a to odgrywa ważną rolę w pozycjonowaniu stron internetowych.

Dla ułatwienia wprowadzania metadanych do stron tworzonych w systemach zarządzania treścią tworzone są rozszerzenia. Przykładowo, dla jednego z popularniejszych systemów tego typu – WordPressa, dostępna jest wtyczka Dublin Core Metadata Generator (Seolus, b.d.) czy Schema & Structured Data for WP & AMP¹².

Jak pisaliśmy wcześniej, metadane są podstawą wyszukiwania fasetowego i filtrowania produktów w sklepach internetowych. Także tutaj przygotowane zostały rozwiązania wspierające tworzenie treści i opisywanie poszczególnych obiektów. Przykład takiego rozwiązania znajdziemy w narzędziu WooCommerce¹³ (Rys. 13), gdzie, tworząc atrybuty dodajemy fasety, a następnie wpisujemy wartości, które użytkownicy będą później zaznaczać filtrując produkty.



Rys. 13. Dodawanie faset i wartości w WooCommerce. Źródło: Zarządzanie produktami w WooCommerce [18.02.2021], <https://woocommerce.pl/zarządzanie-produktami-w-woocommerce/>

8. Zakończenie

Metadane, mimo iż nie zawsze widoczne dla użytkowników, są kluczowe, aby mogli oni dotrzeć do poszukiwanych informacji. Służą one nie tylko do opisywania zasobów i ich późniejszego wyszukiwania, ale także do tworzenia semantycznych powiązań między nimi. Planując architekturę informacji dowolnej przestrzeni cyfrowej, należy pamiętać o dodawaniu spójnych i szczegółowych metadanych, które są podstawą systemów organizacyjnych i nawigacyjnych. Metadane powinny być kompletne, aktualne i rzetelnie odzwierciedlać cechy opisywanych obiektów. Dzięki nim możliwe jest m.in.: zaawansowane sortowanie

¹² Zob. <https://structured-data-for-wp.com/>

¹³ WooCommerce (<https://woocommerce.com/>) jest rozszerzeniem do Wordpressa umożliwiającym prowadzenie sklepu internetowego.

i filtrowanie wyników wyszukiwania, zestawianie (porównywanie) informacji, nieliniarne przeglądanie zasobów, tworzenie hiperłączy i relacji kojarzeniowych (np. za pomocą odsyłaczy/linków „zobacz też”, „podobne produkty”), a także poprawa widoczności stron WWW oraz zwiększenie ich dostępności (WCAG)¹⁴.

Artykuł nie wyczerpuje tematu metadanych w architekturze informacji. Poza wiadomościami ogólnymi, wskazaliśmy w tekście możliwości dodawania metadanych do stron internetowych i korzyści płynące z ich szczegółowego opracowania. Interesującym uzupełnieniem tematu byłoby zbadanie stron internetowych pod kątem obecności metadanych, ich kompletności, stosowanych standardów, itp.

Bibliografia

- Ajdukiewicz-Tarkowska, A. (2008). Metadane w Polsce na tle inicjatyw zagranicznych. *Zagadnienia Informacji Naukowej*, 92(1), 59–84.
- ANSI/NISO (2010a). *ANSI/NISO Z39.18–2005 (R2010): Scientific and Technical Reports – Preparation, Presentation, and Preservation*. Baltimore, MD: NISO Press [online]. NISO [02.08.2021], https://groups.niso.org/apps/group_public/download.php/12038/Z39_18_2005_R2010.pdf
- ANSI/NISO (2010b). *ANSI/NISO Z39.19–2005 (R2010): Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. Baltimore, MD: NISO Press [online]. NISO [02.08.2021], https://groups.niso.org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf
- Baca, M. (2016). *Introduction to Metadata: Third Edition*. Los Angeles, CA: Getty Publications.
- Bednarek, G., Wróbel, A. (2013). Metadane a proces ciągłej digitalizacji obiektów bibliotecznych. *Zagadnienia Informacji Naukowej*, 101(1), 84–104, <https://doi.org/10.36702/zin.590>
- Bojar, B., red. (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa: Wydaw. SBP.
- Brachfogel, A. (2009). *Terminy Metadanych DCMI* [online]. Biblioteka Narodowa, [20.01.2020], <https://www.bn.org.pl/download/document/1261049421.pdf>
- Brachfogel, A. (2010). Terminy metadanych DCMI – nowe zalecenie dla metadanych Dublin Core. *Zagadnienia Informacji Naukowej*, 95(1), 57–64, <https://doi.org/10.36702/zin.682>
- Caplan, P. (2003). *Metadata Fundamentals for All Librarians*. Chicago: American Library Association.
- Czapnik, G., Gruszka, Z., Tadeusiewicz, H. (2011). *Podręczny słownik bibliotekarza*. Warszawa: Wydaw. SBP.
- DCMI (2020). *DCMI Metadata Terms* [online]. Dublin Core Metadata Initiative [20.01.2020], <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- Ding, W., Lin, X., Zarro, M. (2017). *Information Architecture: The Design and Integration of Information Spaces*. San Rafael, California: Morgan & Claypool Publishers.
- Fagan, J. C. (2010). Usability Studies of Faceted Browsing: A Literature Review. *Information Technology and Libraries*, 29(2), 58–66, <https://doi.org/10.6017/ital.v29i2.3144>
- Garrett, J. J. (2011). *The Elements of User Experience: User-Centered Design for the Web and Beyond*. Berkeley, CA: New Riders.

¹⁴ Metadane mogą być użyte do powiązania alternatywnych wersji stron internetowych, które zostały opracowane dla osób z różnymi rodzajami niepełnosprawności. Metadane mogą też służyć do opisywania alternatywnych wersji części strony czy poszczególnych elementów, jak film, grafika itp. O samym standardzie WCAG można przeczytać na stronie <https://www.w3.org/WAI/standards-guidelines/wcag/> oraz <https://wcag20.widzialni.org/index.php> a o zastosowaniu metadanych w WCAG na stronie <https://www.w3.org/TR/UNDERSTANDING-WCAG20/appendixC>

- Gartner, R. (2016). *Metadata: Shaping Knowledge From Antiquity to the Semantic Web*. Cham: Springer.
- Gordon, S., Habermann, T. (2018). The Influence of Community Recommendations on Metadata Completeness. *Ecological Informatics*, 43, 38–51, <https://doi.org/10.1016/j.ecoinf.2017.09.005>
- Greenberg, J. (2005). Understanding Metadata and Metadata Schemes. *Cataloging & Classification Quarterly*, 40(3–4), 17–36, https://doi.org/10.1300/J104v40n03_02
- Griffel, D. M., McIntosh, S. D. (1967). *ADMINS: A Progress Report* [online]. Center for International Studies, Massachusetts Institute of Technology. [20.01.2020], <https://dspace.mit.edu/handle/1721.1/82974>
- Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2), 199–220, <https://doi.org/10.1006/knac.1993.1008>
- ISO (2011). ISO 25964–1:2011 *Information and documentation: Thesauri and interoperability with other vocabularies*. Part 1. Thesauri for information retrieval, <https://www.iso.org/standard/53657.html>
- ISO (2013). ISO 25964–2:2013 *Information and documentation: Thesauri and interoperability with other vocabularies*. Part 2. Interoperability with other vocabularies, <https://www.iso.org/standard/53658.html>
- ISO (2015). ISO 24622–1:2015 *Language Resource Management — Component Metadata Infrastructure (CMDI) — Part 1: The Component Metadata Model*. <https://www.iso.org/standard/37336.html>
- Jijavadze, E., Gabrichidze, N. (2010). *Controlled Vocabularies* [online]. [20.01.2020], <https://slideplayer.com/slide/10978700/>
- Kazienko, P. (1998). Struktura hipertekstu a struktura systemu WWW. *Zagadnienia Informacji Naukowej*, 2, 36–56.
- Kules, B., Capra, R. (2012). Influence of Training and Stage of Search on Gaze Behavior in a Library Catalog Faceted Search Interface. *Journal of the American Society for Information Science and Technology*, 63(1), 114–138, <https://doi.org/10.1002/asi.21647>
- Malak, P., Osińska, V., Bednarek-Michalska, B. (2017). Jakość wybranych elementów metadanych stosowanych w polskich bibliotekach i repozytoriach cyfrowych. *Zagadnienia Informacji Naukowej*, 109(1), 97–123, <https://doi.org/10.36702/zin.351>
- Martin, L. (2019). *Everyday Information Architecture*. New York, NY: A Book Apart.
- Matysek, A. (2014). *Normalizacja europejska w zakresie informatologii*. Katowice: Wydaw. UŚ.
- Mayernik, M. S. (2020). Metadata. *Knowledge Organization*, 47(8), 696–713, <https://doi.org/10.5771/0943-7444-2020-8-696>
- NDAP & ICM UW (2005). *Standard metadanych e-PL-0.1–20050520*. Naczelna Dyrekcja Archiwów Państwowych, & Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego UW [online]. Archiwa Państwowe [20.01.2020], <https://archiwa.gov.pl/images/docs/e-PL-0.1-2.pdf>
- Nahotko, M. (2018). Jakość metadanych opisowych i możliwości ich współdziałania [online]. *Biuletyn EBIB* [online], 179(2), [29.04.2021]. <http://open.ebib.pl/ojs/index.php/ebib/article/view/625>
- Nahotko, M. (2004). *Metadane: Sposób na uporządkowanie Internetu*. Kraków: Wydaw. UJ.
- Nahotko, M. (2005). Metadane i ontologie w systemach zarządzania wiedzą. W: M. Kocój (red.), *Profesjonalna informacja w Internecie* (144–149). Kraków: Wydaw. UJ.
- Nahotko, M. (2012). Struktury danych. W: M. Janiak, M. Krakowska, M. Próchnicka (red.). *Biblioteki cyfrowe* (362–371). Warszawa: Wydaw. SBP.
- NMM (2013). *Digizaurus – słownik hierarchiczny pojęć opracowany na potrzeby Międzymuzealnej Grupy ds. Digitalizacji*. Międzymuzealna Grupa ds. Digitalizacji DigiMu, Muzeum Narodowe w Gdańsku, Centralne Muzeum Morskie w Gdańsku, & Muzeum Zamkowe w Malborku. [online]. Narodowe Muzeum Morskie [20.01.2020], https://www.nmm.pl/upload/Files/cke/Digizaurus_material_1.0.pdf
- PN (1992). PN-92/N-09018 Tezaurus jednojęzyczny. Zasady tworzenia, forma i struktura. Warszawa: Polski Komitet Normalizacji, Miar i Jakości [norma wycofana].

- Pomerantz, J. (2015). *Metadata*. Cambridge, MA: The MIT Press.
- Pulikowski, A. (2018). *Modelowanie procesu wyszukiwania informacji naukowej: Strategie i interakcje*. Katowice: Wydaw. UŚ.
- Ranganathan, S. R. (1967). *Prolegomena to Library Classification*. New York: Asia Publishing House.
- RDA (b.d.). *RDA Metadata Standards Directory*. [online]. Research Data Alliance [20.01.2020], <http://rd-alliance.github.io/metadata-directory/standards/>
- RDF (2014). *RDF 1.1 Concepts and Abstract Syntax* [online]. W3C [20.01.2020], <https://www.w3.org/TR/rdf11-concepts/>
- Riley, J. (2017). *Understanding Metadata: What is Metadata, and What Is It For?* [online]. NISO [20.01.2020], https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Met%E2%80%A6
- Rosenfeld, L., Morville, P., Arango, J. (2017). *Architektura informacji w serwisach internetowych i nie tylko*. Gliwice: Helion.
- Roszkowski, M. (2017). Formaty danych strukturalnych w zasobach World Wide Web. *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 110(2), 46–72, <https://doi.org/10.36702/zin.366>
- Satija, M. P. (2017). *Colon Classification* [online]. Encyclopedia of Knowledge Organization. [02.08.2021], https://www.isko.org/cyclo/colon_classification
- Seolus. (b.d.). *Dublin Core Metadata Generator* [online]. WordPress.Org. [18.02.2021], <https://wordpress.org/plugins/dublin-core-metadata-generator/>
- Skórka, S. (2014). Fasety na nowo odkryte. Integrowanie systemów nawigacji i organizacji informacji. *Zagadnienia Informacji Naukowej*, 52(2), 92–109, <https://doi.org/10.36702/zin.551>
- Smiraglia, R. (2005). *Metadata: A Cataloger's Primer*. Binghamton, N.Y.: Haworth Press.
- Spencer, D. (2010). *A Practical Guide to Information Architecture*. Victoria, Australia: UX Mastery.
- Sosińska-Kalata, B. (2001). Język informacyjno-wyszukiwawczy jako narzędzie organizacji wiedzy. *Zagadnienia Informacji Naukowej*, 2, 28–42.
- Ścibor, E., Tomasiak-Beck, J. (1995). *Metodyka budowy tezasurów*. Warszawa: IINTE.
- Tomaszczyk, J. (2007). Taksonomia jako narzędzie organizacji informacji. *Zagadnienia Informacji Naukowej*, 1, 40–49.
- Tomaszczyk, J. (2014). *Model systemu informacji terminologicznej*. Katowice: Wydaw. UŚ.
- Understanding Metadata* (2004). [online]. Bonanza Creek LTER [30.06.2021], https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf
- Ungurian, O. (1975). *Teoria i praktyka klasyfikacji fasetowej S. R. Ranganathana*. Warszawa: IINTE.
- Wassilew, A. Z., Papińska-Kacperek, J. (2017). Wybrane problemy semantyczne we współczesnych systemach informatycznych. *Roczniki KAE*, 44(2), 23–33.
- Weynand, D., Piccin, V., Weise, M. (2015). *How Video Works* (3rd edition). Routledge.
- Woźniak-Kasperek, J. (2005). *Podstawy budowy tezasuru: Poradnik*. Warszawa: Wydaw. SBP.
- Zeng, M. L., Qin, J. (2016). *Metadata*. Chicago: ALA Neal-Schuman.
- Żytniewski, M. (2004). Koncepcje zastosowań metadanych w systemach informatycznych. W: T. Połębska-Miącz, H. Sroka (red.). *Systemy wspomaganie organizacji SWO'2004* (387–394). Katowice: Wydaw. AE im. Karola Adamieckiego.

Metadata in Information Architecture

Abstract

Purpose/Thesis: The paper presents the role of metadata in the design of website information architecture. It focuses on descriptive and structural metadata, which have a significant impact on the organisation and information retrieval on websites.

Approach/Methods: The article defines metadata and presents different types and standards, as well as metadata's functions in a digital environment. It emphasizes the importance of metadata in information browsing and searching, and provides instructions regarding including metadata in the websites created directly in HTML and WordPress.

Results and conclusions: Metadata play an important role in the websites' information architecture. They facilitate the description of digital objects and the creation of semantic links connecting those objects, which increases website usability and makes the information search more efficient. It is therefore recommended that information architects choose appropriate metadata schemes, attending to the coherence and detail of their elements.

Originality/Value: The issue of metadata in information organization and retrieval in web services has not been explained in detail in the information architecture literature so far.

Keywords

Dublin Core. Facets. Information architecture. Meta tag. Metadata. RDF. Thesaurus.

Dr ANNA MATYSEK jest adiunktem w Instytucie Nauk o Kulturze Uniwersytetu Śląskiego w Katowicach. Jej zainteresowania badawcze skupiają się wokół architektury informacji, narzędzi cyfrowych i wyszukiwania informacji. Ważniejsze publikacje: Architektura informacji w piśmiennictwie zagranicznym 1982–2018: analiza ilościowa (Zagadnienia Informacji Naukowej, 2018), IA, UX, UID, IxD – analiza terminów i pojęć (Zagadnienia Informacji Naukowej, 2020, współaut. J. Tomaszczyk), Narzędzia projektantów produktów cyfrowych (Zagadnienia Informacji Naukowej, 2020).

Kontakt z autorką:

anna.matysek@us.edu.pl

Instytut Nauk o Kulturze,

Wydział Humanistyczny

Uniwersytet Śląski w Katowicach

pl. Sejmu Śląskiego 1

40-032 Katowice

JACEK TOMASZCZYK, doktor habilitowany nauk humanistycznych, profesor w Instytucie Nauk o Kulturze Uniwersytetu Śląskiego w Katowicach. W kręgu jego zainteresowań leży teoria i terminologia architektury informacji oraz indywidualne zarządzanie informacją i wiedzą. Ważniejsze publikacje: Metodyka tworzenia i ewaluacji cyfrowych narzędzi indywidualnego zarządzania informacją (R. Sapa, red., Diagnostyka w zarządzaniu informacją: perspektywa informatologiczna, 2017), Cyfrowy warsztat humanisty (współautor: Anna Matysek), Warszawa: Wydawnictwo Naukowe PWN SA, 2020. IA, UX, UID, IxD – analiza terminów i pojęć, (2020, współaut. A. Matysek).

Kontakt z autorem:

jacek.tomaszczyk@us.edu.pl

Instytut Nauk o Kulturze, Wydział Humanistyczny

Uniwersytet Śląski w Katowicach

pl. Sejmu Śląskiego 1

40-032 Katowice