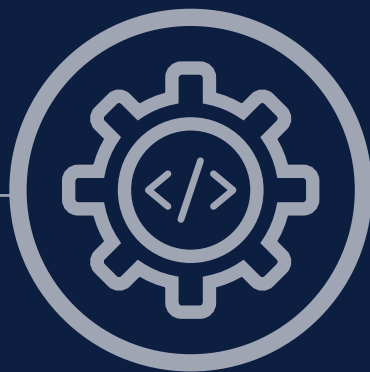


Standard techniczny



STANDARDY OTWARTOŚCI DANYCH

Spis treści



1. Wstęp	4
2. Formaty plików	9
2.1. Formaty JPEG, PNG	10
2.2. Format PDF	10
2.3. Format TXT	10
2.4. Pliki archiwów	10
2.5. Formaty MS Office (DOC, XLS)	11
2.6. Formaty z rodziny Office Open XML (XLSX, DOCX)	11
2.7. Format dBase (dbf)	11
2.8. Format HTML	11
2.9. Format OpenDocument Spreadsheet (ODS)	12
2.10. Format CSV	12
2.11. Format JSON	12
2.12. Format XML	12
2.13. Formaty danych na 4 ★★★★★ i 5 ★★★★★ (RDF/XML, Turtle, N3, JSON-LD, N-tripels, RDFa)	13
2.14. Formaty danych przestrzennych	13
3. Formatowanie danych	14
3.1. Przykłady formatowania danych	15
3.2. Dane adresowe	16
4. Szczegółowe wymagania dla zasobów w formacie CSV	19

4.1. Format pliku CSV	20
4.2. CSV na poziomie otwartości 3 ★★★	20
4.3. CSV na poziomie otwartości 4 ★★★★ 5 ★★★★★	21
5. Szczegółowe wymagania dla zasobów w formacie JSON	22
5.1. JSON na poziomie otwartości 3 ★★★	23
5.2. JSON na poziomie otwartości 4 ★★★★ 5 ★★★★★	23
6. Szczegółowe wymagania dla zasobów w formacie XML	25
6.1. XML na poziomie otwartości 3	26
6.2. XML na poziomie otwartości 4 ★★★★ i 5 ★★★★★	26
7. Szczegółowe informacje dla zasobów na 4 ★★★★ i 5 ★★★★★ (RDF/XML, Turtle, N3, JSON-LD, N-tripels, RDFa)	28
7.1. Model danych RDF	29
7.2. Serializacja modelu RDF	30
7.3. Schemat RDF	37
7.4. Dane połączone	38
8. Profil aplikacji DCAT-AP	40
9. Lokalizacja źródeł danych publicznych	42
10. Podsumowanie	44
11. Załącznik nr 1 - Przykład opisu metadanych zgodnych z rekomendacjami implementacji DCAT-AP.	47
12. Załącznik nr 2 - Lista podstawowych aktów prawnych, norm, standardów, wytycznych dotyczących szeroko rozumianej otwartości danych publicznych	69

1 .Wstęp



Dokument definiuje zestaw minimalnych wymagań technicznych, które muszą spełniać dane publiczne, w szczególności udostępniane w Centralnym Repozytorium Informacji Publicznej (portal dane.gov.pl).

Użyte w dokumencie sformułowania wymaga się, zaleca się, dopuszcza się, nie zaleca się, nie dopuszcza się oznaczają odpowiednio:

wymaga się – oznacza, że aby zachować zgodność ze standardem, trzeba bezwarunkowo spełnić opisane kryteria/zasady;

zaleca się – oznacza, że w uzasadnionych przypadkach można nie spełnić danego kryterium, pod warunkiem, że opisująca zasoby dokumentacja (np. dokumentacja API, opisana w Standardzie API¹) zawiera uzasadnienie decyzji o odstępstwie;

dopuszcza się – oznacza, że można zastosować opisane kryterium lub go nie zastosować, a dokumentacja tej decyzji nie jest wymagana (choć można ją zamieścić);

nie zaleca się – oznacza, że w wyjątkowych i uzasadnionych przypadkach można postąpić w opisany sposób pod warunkiem, że opisująca zasoby dokumentacja (np. dokumentacja API, opisana w Standardzie API) zawiera uzasadnienie decyzji o odstępstwie od danego zalecenia;

nie dopuszcza się – oznacza, że aby zachować zgodność ze standardem, nie wolno postąpić w opisany sposób.

¹ Dotyczy Standardu API przygotowanego w ramach projektu „Otwarte dane – dostęp, standard, edukacja”, dofinansowanego z poddziałania 2.3.1 Programu Operacyjnego Polska Cyfrowa <https://dane.gov.pl/knowledge-base/education/1219,standardy-otwartosci-danych>

Poziomy otwartości danych

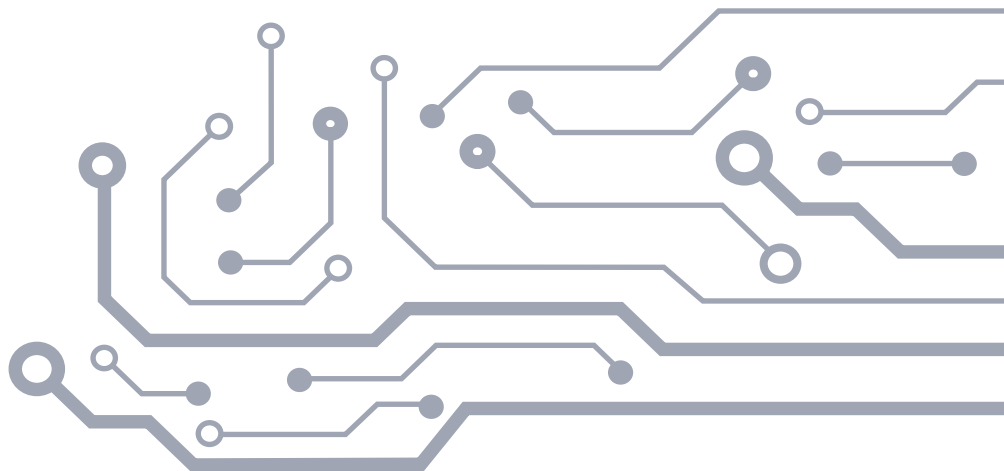
Schemat pięciu poziomów otwartości danych, zaproponowany przez Tima Bernersa-Lee, to użyteczne narzędzie klasyfikowania zasobów pod kątem ich otwartości. Model otwartości opisany tym schematem został przyjęty w Programie Otwierania Danych Publicznych². Więcej informacji o tym schemacie można znaleźć na stronach W3C³.

Im wyższy poziom otwartości danych, tym dane są lepiej przygotowane do dalszego przetwarzania. Wszystkie otwarte dane są udostępniane bez żadnych ograniczeń do dowolnych celów komercyjnych i niekomercyjnych. Schemat pięciu poziomów otwartości danych jest opisany w tabeli 1.

UWAGI:

Dane na poziomach otwartości 3 ★★ ★ i wyższych zaleca się udostępniać przez API, co pozwala na ich maszynowe przetwarzanie i opatrzenie dodatkowymi metadanymi.

Nie zaleca się publikowania danych na pierwszym poziomie otwartości.



² <https://dane.gov.pl/knowledge-base/education/653.program-otwierania-danych-publicznych>

³ <https://www.w3.org/DesignIssues/LinkedData.html>

Tabela 1. Schemat pięciu poziomów otwartości danych

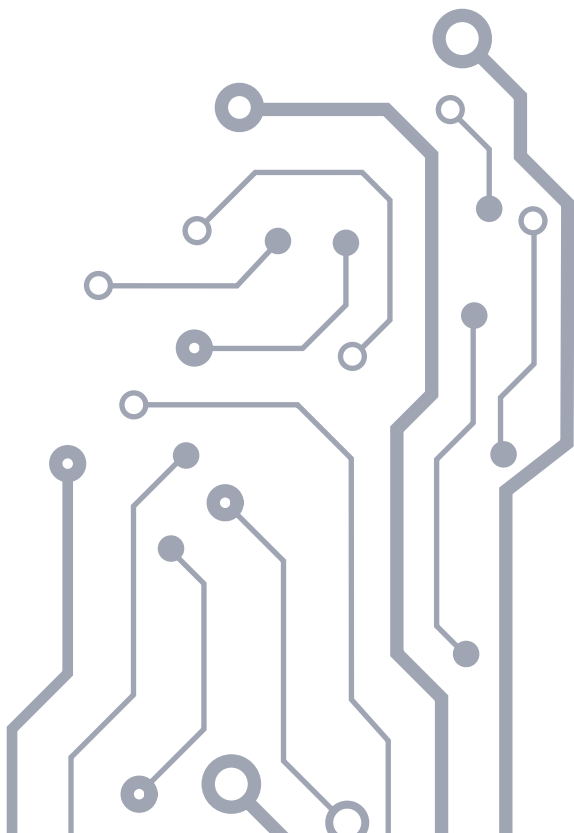
Poziom	Oznaczenie	Charakterystyka poziomu	Dodatkowe informacje	Opis
1	★	Dane w dowolnym formacie udostępniane są bez ograniczeń licencyjnych ⁴ .	Dopuszczalne jest publikowanie danych w formie plików graficznych, skanów z obrazem danych i tekstu, plików tekstowych.	Mogą to być pliki JPEG z zeskanowanymi dokumentami, wygenerowane z różnych programów pliki PDF lub pliki tekstowe zawierające dane nieustrukturyzowane albo o strukturach niejednorodnych. Większość ludzi może je odczytać, ale jakiegokolwiek dalsze ich wykorzystywanie wymaga dodatkowej pracy, polegającej na zidentyfikowaniu, odczytaniu i przeniesieniu danych (często ręcznie bądź ze wspomaganiami programów typu OCR).
2	★★	Dane są udostępniane w postaci ustrukturyzowanej.	Dane publikowane są w postaci sformatowanego arkusza kalkulacyjnego.	Dane mają już określoną strukturę, którą można odczytać komputerowo, na przykład poprzez zastosowanie pliku w formacie arkusza kalkulacyjnego lub edytora tekstu. Nie są to zeskanowane w niedającej się przeszukiwać formie obrazy. Natomiast dane te są w formacie zamkniętym (własnościowym), dla którego stosowanie w oprogramowaniu jest ograniczone przez restrykcje patentowe, licencyjne lub podobne.

⁴ Więcej na ten temat w Standardzie prawnym, <https://dane.gov.pl/knowledge-base/education/1219,standardy-otwartosci-danych>

Poziom	Oznaczenie	Charakterystyka poziomu	Dodatkowe informacje	Opis
3	★★★	Używanie formatów otwartych.	Dane są publikowane w otwartym ustrukturyzowanym formacie, np. CSV lub JSON.	Dane są w otwartym formacie ⁵ , ale ich zrozumienie na potrzeby przetwarzania maszynowego wymaga każdorazowej analizy danych i ustalenia, bądź odnalezienia w dokumentacji, jeśli taka istnieje, znaczeń poszczególnych pól. Czy „nazwisko” oznacza samo nazwisko, czy imię i nazwisko? Czy „kod” to kod pocztowy czy terytorialny? Czy „odległość” jest podana w metrach, czy w kilometrach? Czy „1/12/2018” to pierwszy grudnia, czy dwunasty stycznia? Dane udostępniane są w niezastrzeżonym formacie.
4	★★★★	Używanie URI do identyfikacji obiektów zgodnie z RDF.	Dane publikowane są w formacie umożliwiającym oznaczenie ich struktury znaczeniowej.	Wejście na czwarty poziom otwartości pozwala jednoznacznie określić znaczenie udostępnianych danych. Technicznym sposobem wyrażania takiego znaczenia w sposób zrozumiały dla maszyn jest identyfikacja konkretnych właściwości danych za pomocą zrozumiałych dla maszyny URI zgodnie z modelem RDF (model opisu danych). Dane udostępniane są w niezastrzeżonym formacie.

⁵ Format otwarty – należy przez to rozumieć format pliku, który nie jest powiązany z platformą oraz jest udostępniany bez żadnych ograniczeń, które utrudniałyby ponowne wykorzystywanie danych (art. 2 pkt 14 dyrektywy 2019/1024/UE Parlamentu Europejskiego i Rady z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego)

Poziom	Oznaczenie	Charakterystyka poziomu	Dodatkowe informacje	Opis
5	★★★★★	Łączenie danych z innymi danymi za pomocą linków – budowanie kontekstu.	Dane zawierają połączenia strukturalne online do innych zbiorów informacji.	Piąty poziom dodatkowo ułatwia przetwarzanie, jawnie wskazując relacje między danymi w formie linków. Dzięki temu możliwe jest odnajdywanie połączeń pomiędzy różnymi zbiorami danych. Dane udostępniane są w niezastrzeżonym formacie.



2.

Formaty plików

2.1. Formaty JPEG, PNG

Formaty graficzne są dopuszczalne jako formaty udostępniania otwartych danych wyłącznie na poziomie otwartości 1 ★, ponieważ co do zasady nie nadają się do automatycznego przetwarzania zamieszczonych w nich danych. Wyjątkiem jest sytuacja, w której same obrazy (a nie przedstawiona na nich zawartość, np. tekst) są udostępnianymi danymi. Wówczas poziom otwartości zależy od poziomu otwartości metadanych, którymi opatrzony jest obraz, przygotowanymi zgodnie z zaleceniami niniejszego dokumentu.

Możliwe jest przetworzenie zasobów do postaci zgodnej z wyższymi poziomami otwartości przy użyciu dostępnych na rynku narzędzi do konwersji danych.

2.2. Format PDF

Format PDF jest dopuszczalny wyłącznie na poziomie otwartości 1 ★, ponieważ co do zasady nie nadaje się do automatycznego przetwarzania.

Możliwe jest przetworzenie zasobów do postaci zgodnej z wyższymi poziomami otwartości przy użyciu dostępnych na rynku narzędzi do konwersji danych.

2.3. Format TXT

Format TXT, pozbawiony struktury, jest dopuszczalny na poziomie otwartości 1 ★. Zależnie od faktycznej struktury danych, powinny one być udostępniane z oznaczeniem innego, właściwego typu pliku (np. CSV).

2.4. Pliki archiwów

Dopuszczalne jest udostępnianie danych w postaci skompresowanej.

Na wszystkich poziomach otwartości dopuszcza się otwarte formaty kompresji, takie jak ZIP (bez własnościowych rozszerzeń), 7z, Gzip czy Bzip2. Na poziomach otwartości do 2 ★ ★ dopuszcza się także popularne formaty własnościowe: RAR, rozszerzony ZIP.

2.5. Format MS Office (DOC, XLS)

Są to własnościowe formaty dokumentów, **dopuszczalne na poziomie otwartości do 2 ★★**.

Do arkusza danych w formacie XLS wymaga się stosowania szczegółowych wymagań dla zasobów w formacie CSV na poziomie otwartości 3 ★★★ określonych w punktach 4.1 i 4.2.

Wobec tego, zaleca się zastąpienie formatu XLS formatem CSV dla uzyskania trzeciego poziomu otwartości.

Możliwe jest przetworzenie zasobów do postaci zgodnej z wyższymi poziomami otwartości przy użyciu dostępnych na rynku narzędzi do konwersji danych.

2.6. Formaty z rodziny Office Open XML (XLSX, DOCX)

Formaty Office Open XML to otwarte formaty udostępniania dokumentów (DOCX) i ustrukturyzowanych danych (XLSX). Jednakże ze względów technicznych, formaty te **nie są zalecane jako formaty udostępniania otwartych danych powyżej poziomu 2 ★★**.

Możliwe jest przetworzenie zasobów do postaci zgodnej z wyższymi poziomami otwartości przy użyciu dostępnych na rynku narzędzi do konwersji danych.

2.7. Formaty dBase (dbf)

Format bazy danych dBase jest dopuszczalny do poziomu otwartości 3 ★★★.

Format ten jest stosowany do przechowywania danych ustrukturyzowanych.

Możliwe jest przetworzenie zasobów do postaci zgodnej z wyższymi poziomami otwartości przy użyciu dostępnych na rynku narzędzi do konwersji danych.

2.8. Format HTML

Format HTML jest otwartym formatem, pozwalającym na udostępnianie ustrukturyzowanych danych. **Nie zaleca się stosowania formatu HTML na poziomach otwartości powyżej 3 ★★★**. W przypadku udostępniania danych nieustrukturyzowanych w formacie HTML, jest on dopuszczalny na 1 ★.

Możliwa jest transformacja formatu HTML do wyższych poziomów otwartości.

2.9. Formaty OpenDocument Spreadsheet (ODS)

Format ODS jest otwartym formatem udostępniania ustrukturyzowanych danych i w związku z tym **jest dopuszczalny do poziomu otwartości 3 ★★★**.

Możliwe jest przetworzenie zasobów do postaci zgodnej z wyższymi poziomami otwartości przy użyciu narzędzi dostępnych na rynku do konwersji danych.

2.10. Format CSV

CSV to prosty format przechowywania danych tabelarycznych. **Format CSV jest zalecany do udostępniania otwartych danych na trzecim poziomie otwartości**. Zastosowanie dodatkowych standardów, takich jak Model for Tabular Data and Metadata on the Web⁶, pozwala na przekształcenie CSV do formatów na najwyższych poziomach otwartości.

Szczegółowe wymagania związane z formatem CSV znajdują się w **punkcie 4** „Szczegółowe wymagania techniczne dla zasobów w formacie CSV”.

2.11. Format JSON

JSON to prosty format udostępniania ustrukturyzowanych danych. **Format JSON jest zalecany do udostępniania otwartych danych na trzecim poziomie otwartości**. Zastosowanie dodatkowych standardów, takich jak JSON-LD⁷, pozwala na właściwe wyrażenie metadanych również na wyższych poziomach otwartości.

Szczegółowe wymagania techniczne związane z formatem JSON znajdują się w **punkcie 5** „Szczegółowe wymagania dla zasobów w formacie JSON”.

2.12. Format XML

Format XML jest dopuszczalny do udostępniania otwartych danych na trzecim poziomie otwartości. Format ten po przekształceniu można wyrazić na najwyższych poziomach otwartości.

Szczegółowe wymagania techniczne związane z formatem XML znajdują się w **punkcie 6** „Szczegółowe wymagania dla zasobów w formacie XML”.

⁶ <https://www.w3.org/TR/tabular-data-model/>

⁷ <https://json-ld.org/spec/latest/json-ld/>, <https://w3c.github.io/json-ld-syntax/>

2.13. Formaty danych na 4 ★★★★★ i 5★★★★★ (RDF/XML, Turtle, N3, JSON-LD, N-tripels, RDFa)

Formaty plików takie jak RDF/XML, Turtle, Notation3 (N3), JSON-LD, N-tripels, RDFa zaleca się do udostępniania otwartych danych na najwyższych stopniach otwartości.

Szczegółowe wymagania techniczne związane z ww. formatami znajdują się w punkcie 7 „Szczegółowe wymagania dla zasobów na 4 ★★★★★ i 5 ★★★★★” (RDF/XML, Turtle, N3, JSON-LD, N-tripels, RDFa)“.

2.14. Formaty danych przestrzennych

Dopuszczalne jest stosowanie formatów danych przestrzennych, w szczególności:

- dane przestrzenne (rastrowe) bez georeferencji **są dopuszczalne na poziomie otwartości 1 ★** np. format JPG/JPEG, PNG, TIFF, BMP,
- dane przestrzenne (rastrowe) z georeferencją **są dopuszczalne na poziomie otwartości 3 ★★★**, np. GeoTIFF,
- dane przestrzenne (wektorowe) **są dopuszczalne na poziomie otwartości 3 ★★★**, np. format GML, SHP, KML, GeoJSON, GPX,
- dane przestrzenne zaopatrzone w kontekst pozwalający na automatyczne przetworzenie do postaci zgodnej z RDF **są dopuszczalne na wszystkich poziomach otwartości**. Udostępniając takie dane, należy zachować dobre praktyki opisane w dokumencie Spatial Data on the Web Best Practices⁸.

Dane przestrzenne udostępniane są poprzez usługi umożliwiające m.in. wyszukiwanie, przeglądanie oraz pobieranie danych (usługi WMS, WMTS, WFS, WCS, ATOM). Szczegółowe uregulowania dot. danych przestrzennych znajdują się w ustawie o infrastrukturze informacji przestrzennej⁹ oraz w przepisach implementacyjnych dyrektywę Parlamentu Europejskiego i Rady nr 2007/2/WE z dnia 14 marca 2007 r. ustanawiającej infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej (INSPIRE).

⁸ <https://www.w3.org/TR/sdw-bp/>

⁹ Ustawa z dnia 4 marca 2010 r. o infrastrukturze informacji przestrzennej (Dz. U. z 2020 r., poz. 177).

3.

Formatowanie danych

Niezależnie od wyboru formatu pliku danych wymaga się stosowanie właściwego formatowania w szczególności dla danych typu: liczba, data, godzina, wartość logiczna.

Na poziomach otwartości 4 ★★★★★ i 5 ★★★★★ każda właściwość ma określony przez URI¹⁰ typ, który wyznacza oczekiwany format wartości (zob. szczegółowe opisy poniżej w punktach dot. CSV, JSON, XML). Przykładowo, jeśli mamy do czynienia z właściwością `schema:datePublished`, to jednoznacznie ustala typ wartości na `schema:Date`, który z kolei jest datą w formacie zgodnym z normą ISO 8601¹¹.

Na tych poziomach wymagane jest określenie typu danych w ramach definicji właściwości.

Na poziomach otwartości do 3 ★★★ zaleca się również stosowanie zapisu daty i czasu w formacie zgodnym z normą ISO 8601, oraz odpowiednio podstawowych typów danych ze standardu XML Schema¹².

Oznacza to w szczególności:

- stosowanie literalnych wartości `true`, `false` jako wartości logicznych,

- stosowanie kropki dziesiętnej (a nie przecinka) w zapisie ułamków dziesiętnych, bez żadnych dodatkowych separatorów (np. oddzielających tysiące); dopuszczalny jest tzw. zapis naukowy,

- zapis dat w postaci `yyyy-mm-dd`, a łącznie dat i godzin w postaci `yyyy-mm-ddThh:mm` lub `yyyy-mm-dd hh:mm:ss` (spacja między datą a czasem).

Dane zaleca się udostępniać w możliwie najwyższym stopniu granulacji (rozdrobienia), tzn. nie łączyć kilku danych w jednym polu.

3.1. Przykłady formatowania danych

REGON – wymaga się zapisywania identyfikatora podmiotu jako ciągu 9 lub 14 cyfr, bez spacji lub łączników.

NIP – wymaga się zapisywania numeru identyfikacji podatkowej jako ciągu 10 cyfr, bez spacji lub łączników.

KRS – wymaga się zapisywania numeru jako ciągu 10 cyfr, bez spacji lub łączników.

¹⁰ <https://tools.ietf.org/html/rfc3986>

¹¹ <https://www.iso.org/iso-8601-date-and-time-format.html>

¹² <https://www.w3.org/TR/xmlschema/>

Numer telefonu – w polskiej strefie numeracyjnej zaleca się zapisywanie numeru telefonu jako ciągu 9 cyfr, bez wyróżniania tzw. numeru kierunkowego miejscowości i bez spacji lub łączników. Dopuszczalne jest stosowanie prefiksu międzynarodowego poprzedzonego znakiem plus „+”.

Tabela 2. Prawidłowe formatowanie danych

Data	Liczba	Wartość logiczna
1970-01-01	42.01	true
2018-12-31	1.05e3	false

Tabela 3. Nieprawidłowe formatowanie danych

Data	Liczba	Wartość logiczna
1/1/1970	45,01	T
1.1.1970	1,234.56	0
1 1970	1 234,5	NIE

3.2. Dane adresowe

Wymaga się, aby adres pocztowy był udostępniany jako szereg danych opisowych: nazwy ulicy, numeru budynku, kodu pocztowego, miejscowości – umieszczonych w odrębnych polach, a nie jako jedno pole tekstowe, zawierające wszystkie wymienione dane.

Tabela 4. Prawidłowe formatowanie danych

Nazwa	Kod PNA	Miejscowość	Cecha	Nazwa ulicy	Nr budynku	Nr lokalu	Nr telefonu
CSIOZ	00-184	Warszawa	ul.	Stanisława Dubois	5A		225970927
KPRM	00-583	Warszawa	al.	Aleje Ujazdowskie	1/3		226946000
Biuro	51-152	Wrocław	pl.	Marsz. Józefa Piłsudskiego	4	75	712558765

Tabela 5. Nieprawidłowe formatowanie danych

Nazwa	Adres	Nr telefonu
CSIOZ	00-184 Warszawa, ul. Dubois 5A	+48 22 597-09-27
KPRM	00-583 Warszawa, Aleje Ujazdowskie 1/3	(22) 694-60-00 (cent.)
Biuro	51-152 Wrocław, J. Piłsudskiego 4/75	71-25-58-765

Wymaga się, aby dane opisowe (tekstowe) składające się na adres były literalnie zgodne z ich zapisem w źródle ich wytworzenia, a sposób udostępniania nie może wpływać na ich wartość, tzn. nie mogą one w procesie przekazywania do miejsca publikacji ulegać zmianom.

Zbiorowym miejscem publikacji danych, takich jak nazwy ulic, po zrealizowaniu procesu uchwalenia (nazwy ulic nadawane są przez Radę Gminy) oraz ogłoszeniu przez wojewodę w wojewódzkim dzienniku urzędowym, jest rejestr TERYT. W tym rejestrze do konkretnych nazw przypisane są kody – identyfikatory ze zbiorów ULIC¹³ oraz SIMC. W nim też znajdują się kody terytorialne TERC. Kody terytorialne mogą się zmieniać w zależności od aktualnego podziału terytorialnego kraju, natomiast identyfikatory miejscowości, a nawet ich części, nadawane są ostatecznie. Z uwagi na potrzeby związane z poprawną adresacją – w lokalizacją punktów adresowych – utworzono pojęcie **adresu uniwersalnego**.

Adres uniwersalny to adres zapisany jako ciąg kodów – kodu pocztowego (tzw. PNA 5 cyfr, bez myślnika), kodu terytorialnego składającego się z kodów województwa, powiatu, gminy (6 cyfr), identyfikatora miejscowości podstawowej (7 cyfr), identyfikatora miejscowości (7 cyfr), identyfikatora ulicy (5 cyfr, w przypadku braku ulic ten fragment adresu składa się z ciągu „00000”), współrzędnych geodezyjnych x, y danego punktu adresowego oraz numeru budynku – oddzielonych separatorem „|”.

Zaleca się stosowanie adresu uniwersalnego. Konstrukcja adresu uniwersalnego umożliwia jego maszynowe (automatyczne) rozkodowanie pozwalające na uzyskanie nazwy województwa, powiatu, gminy, miejscowości, nazwy ulicy oraz wyświetlenie mapy z lokalizacją danego punktu adresowego. Adres uniwersalny nie służy do odczytu przez człowieka, lecz przez system informatyczny, którego zadaniem jest wizualizacja danego punktu adresowego na mapie.

¹³ http://eteryt.stat.gov.pl/eTeryt/.../pliki_pelne_struktury.aspx

Aplikacje służące do wprowadzania danych adresowych do różnego rodzaju rejestrów czy też systemów teleinformatycznych, powinny posiadać zabezpieczenia przed wprowadzaniem adresów nieistniejących. Zatem systemy te powinny umożliwiać dokonywanie wyboru adresu ze słowników rzeczywiście istniejących obiektów. Przykładem może być aplikacja służąca do weryfikacji adresów, znajdująca się na portalu dane.gov.pl¹⁴. Aplikacja ta wykorzystuje rozwijane listy utworzone w oparciu o słowniki TERYT oraz dane zawarte w Państwowym Rejestrze Granic. Posiada również konwerter zakodowanych adresów uniwersalnych na adresy pocztowe punktów adresowych.

Przykład: adres uniwersalny Centrum Systemów Informacyjnych Ochrony Zdrowia przy ul. Stanisława Dubois 5A w Warszawie wygląda następująco:

```
00184 | 146501 | 0918123 | 0918123 | 04337 | 489147.9218 | 636045.6562 | 5A |
```

Poprawne adresy uniwersalne znajdują się w kolumnie AdresCSIOZ w plikach CSV dostępnych na stronie <http://integracja.gugik.gov.pl/daneadresowe/>. W tych samych plikach, a także plikach GML^{15,16}, znajdują się szczegółowe dane adresowe, tworzone bezpośrednio w gminach. Są to dane źródłowe i najbardziej wiarygodne. Jednocześnie, fakt otwarcia danych adresowych oraz ściśle określone zasady tworzenia zapewniają interoperacyjność we wszelkich zastosowaniach.

¹⁴ <https://dane.gov.pl/application/1216.aplikacja-do-sprawdzania-adresu-universalnego>

¹⁵ Ustawa z dnia 17 maja 1989 r. - Prawo geodezyjne i kartograficzne (Dz. U. z 2010 r. Nr 193, poz. 1287)

¹⁶ Rozporządzenie Ministra Administracji i Cyfryzacji z dnia 9 stycznia 2012 r. w sprawie ewidencji miejscowości, ulic i adresów (Dz.U. z 2012 r. poz. 125)

4.

Szczegółowe wymagania dla zasobów w formacie CSV

4.1. Format pliku CSV

W praktyce występują pliki określane jako „CSV”, które różnią się szczegółami technicznymi. Różne są znaki nowej linii, kodowania znaków, używane są różne separatory pól: przecinek, średnik, tabulator.

Na poziomie otwartości 3 ★★★ zaleca się, natomiast na poziomie 4 ★★★★ wymaga się ścisłego stosowania formatu CSV określonego w RFC 4180¹⁷. W szczególności oznacza to:

- stosowanie znaków końca wiersza w formacie CRLF;
- używanie przecinka, jako separatora pól;
- używanie cudzysłowów do otaczania znaków specjalnych (takich jak przecinki czy znaki nowej linii wewnątrz pól);
- stosowanie domyślnego kodowania znaków UTF-8, bądź prawidłowe oznaczenie kodowania w nagłówku Content-type jeśli stosowane jest inne kodowanie znaków;

Przykład:

```
Content-type: text/csv
title,datePublished
„Do gościa (Gościu, tak jakoś począł, już do końca
czytaj...)”,2007-09-07
```

4.2. CSV na poziomie otwartości 3 ★★★

Dla każdego pliku w formacie CSV wymagany jest dokładnie jeden wiersz nagłówka, opisujący zawartość tabeli w sposób możliwie najbardziej zrozumiały dla użytkownika. Jak zostało powiedziane wcześniej: na tym poziomie kluczowa jest rola człowieka, który musi nadać sens danym przeznaczonym do dalszego automatycznego przetwarzania.

Wymaga się, aby wiersz nagłówka zawierał wyłącznie nagłówki poszczególnych kolumn, a wszystkie pozostałe wiersze zawierały wyłącznie dane kolejnych rekordów. Niedopuszczalne jest, przykładowo, wyodrębnianie fragmentu arkusza jako „legenda”.

Niedopuszczalna jest obecność kolumn bez nagłówka lub z pustym (składającym

¹⁷ <https://tools.ietf.org/html/rfc4180>

się tylko z białych znaków) nagłówkiem. Niedopuszczalne są także nagłówki powtarzające się w różnych kolumnach.

Wymaga się, aby wszystkie wiersze tabeli, łącznie z nagłówkiem, miały taką samą liczbę komórek.

Niedopuszczalna jest obecność całkowicie pustych kolumn bądź wierszy.

Wymaga się, aby wszystkie wiersze (poza wierszem nagłówka) były traktowane jako niezależne od siebie. Niedopuszczalne jest w szczególności pozostawianie w tabeli wolnych miejsc w celu uniknięcia powtórzeń danej wartości w kolejnych wierszach. Jeśli dana wartość jest identyczna dla kolejnych rekordów, to nadal musi być osobno określona dla każdego rekordu.

Przykład:

```
datePublished,views,change  
2018-01-01,1234,15
```

4.3. CSV na poziomie otwartości 4 ★★★★★ i 5 ★★★★★

Na tym poziomie otwartości kluczowe jest zapewnienie możliwości maszynowego odczytu wszystkich danych z uwzględnieniem ich kontekstu semantycznego. W tym celu wymagane jest stosowanie standardu Model for Tabular Data and Metadata on the Web¹⁸, pozwalającego na przekształcenie plików CSV do formatu RDF oraz JSON-LD. Dodatkowe szczegóły na temat transformacji formatu CSV do wyższych poziomów otwartości znajdują się również w innych standardach W3C¹⁹.

Dane na 4 i 5 stopniu otwartości szczegółowo opisane zostały w rozdziale 7.

¹⁸ <https://www.w3.org/TR/tabular-data-model/>

¹⁹ <https://www.w3.org/TR/csv2json/>, <https://www.w3.org/TR/csv2rdf/>

5.

Szczegółowe wymagania dla zasobów w formacie JSON

Format JSON jest zalecany do udostępniania otwartych danych na trzecim poziomie otwartości. Jest to prosty format, nadający się do przetwarzania ogólnie dostępnymi narzędziami znajdującymi się w standardowych bibliotekach wielu języków programowania. Format ten jest oparty na podzbiorze standardu języka programowania JavaScript. JSON jest formatem tekstowym, który jest niezależny od konkretnego języka.

Doprecyzowania wymaga udostępnianie w tym formacie dużych kolekcji (stronicowanie), określanie semantyki na wyższych poziomach otwartości, a także określanie dodatkowych funkcji związanych z udostępnianiem danych w API, takich jak sortowanie, filtrowanie, wyszukiwanie i wybieranie podzbiorów zwracanych pól z wyników zapytania.

5.1. JSON na poziomie otwartości 3 ★★★

Na tym poziomie **wymagane jest** stosowanie prawidłowego, tj. zgodnego z RFC 8259²⁰, formatu JSON.

W kontekście udostępniania danych przez API, najważniejsze jest zorganizowanie danych, nazw właściwości, pól służących do nawigacji między stronami kolekcji, a także mechanizmów do wyszukiwania i filtrowania, w sposób jak najbardziej zrozumiały dla użytkownika – programisty, który będzie miał za zadanie skorzystać z udostępnionych danych. W tym celu należy przyjąć standard JSON API.

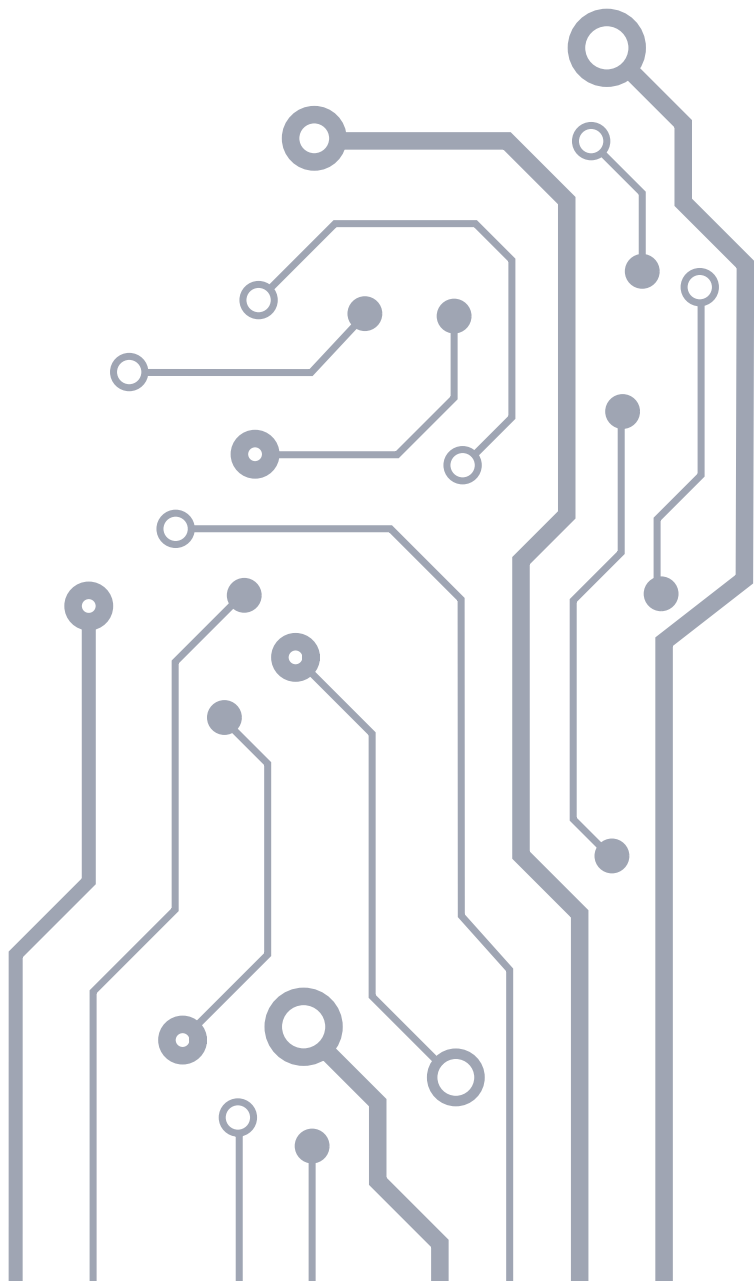
Zaleca się przy tym organizowanie danych w strukturę, która pozwoli w łatwy sposób rozszerzyć ją o semantyczny kontekst, wymagany na wyższych poziomach otwartości danych. Użycie standardów wskazywanych w wyższych poziomach otwartości nie może kolidować z organizacją danych wynikającą ze standardu JSON API. W celu osiągnięcia kolejnych poziomów otwartości należy stworzyć dane kontekstowe przekazywane w nagłówku HTTP.

5.2. JSON na poziomie otwartości 4 ★★★★★ i 5 ★★★★★

Na poziomie 4 ★★★★★ wymagane jest stosowanie standardu JSON-LD i opatrzenie danych pełnym kontekstem semantycznym. Co oznacza mapowanie wszystkich właściwości udostępnianych danych odpowiednim URI w modelu RDF. Użycie standardu JSON-LD nie może kolidować z organizacją danych wynikającą ze standardu JSON API. Dane kontekstowe należy przekazywać w nagłówku http.

²⁰ <https://tools.ietf.org/html/rfc8259>

Ten poziom otwartości 5 ★★★★★ wymaga dodania do zasobów połączeń z innymi danym. Szczegóły dotyczące 4 i 5 stopnia wraz z przykładami zostały opisane w rozdziale 7.



6.

Szczegółowe wymagania dla zasobów w formacie XML

XML to uniwersalny język znaczników, pozwalający na reprezentowanie i przechowywanie różnych danych w ustrukturyzowany sposób. XML pozwala na identyfikowanie znaczników jako URI, co pozwala stosować go do udostępniania danych na wszystkich poziomach otwartości. Obecnie format XML w wielu kontekstach zastępowany jest prostszymi formatami. Dopuszcza się stosowanie formatu XML, jednocześnie zalecając stosowanie formatów JSON lub CSV.

Dane w formacie XML udostępnia się w sposób analogiczny do udostępniania w formacie JSON-LD.

6.1. XML na poziomie otwartości 3 ★★★

Na tym poziomie wymagane jest stosowanie dowolnej poprawnej struktury XML, w której właściwości obiektów reprezentowane są przez elementy XML.

Przykład:

```
Content-type: application/xml

<punkt>
  <adres>Warszawa, Marszałkowska 1</adres>
  <url>http://example.com</url>
</punkt>
```

6.2. XML na poziomie otwartości 4 ★★★★★ i 5 ★★★★★

Wymaga się, aby na tych poziomach była określona semantyka (identyfikacja przez URI) wszystkich elementów za pomocą mechanizmu XML Namespaces²¹ i dedykowanych dla tych danych schem.

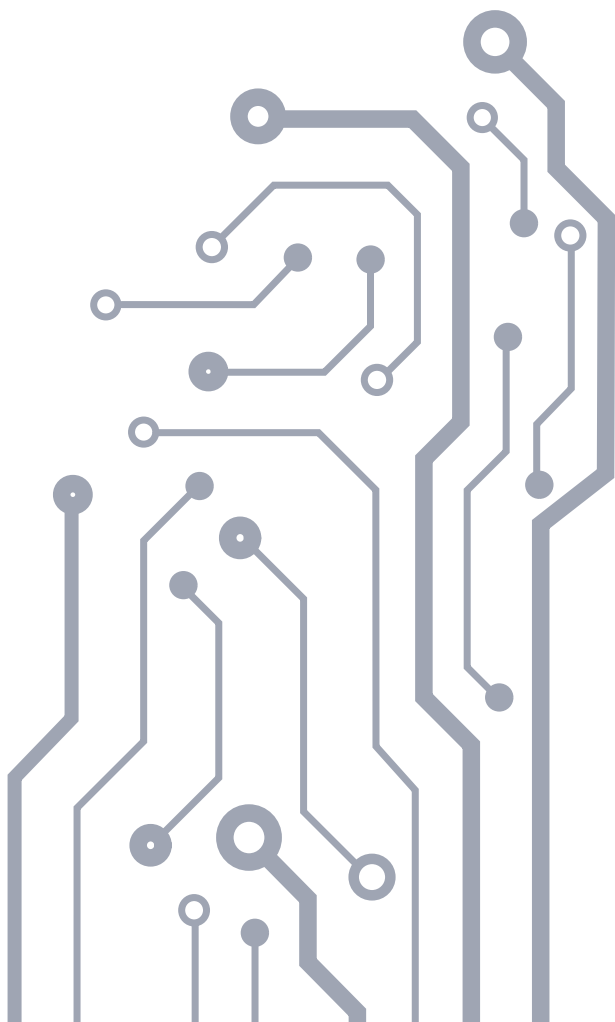
Przykład:

```
Content-type: application/xml

<punkt xmlns="http://dane.gov.pl" xmlns:schema="https://
schema.org">
  <schema:address>
```

²¹ <https://www.w3.org/TR/xml-names/>

```
<schema:postalCode>00-001</schema:postalCode>
  <schema:addressLocality>Warszawa</
schema:addressLocality>
  <schema:streetAddress>Marszałkowska1</
schema:streetAddress >
  </schema:address>
</punkt>
```



7.

**Szczegółowe informacje dla
zasobów na 4 i 5★★★★★
(RDF/XML, Turtle, N3, JSON-
LD, N-tripels, RDFa)**

7.1. Model danych RDF

RDF (ang. Resource Data Framework) jest językiem opisującym model danych dla metadanych oparty na trójelementowych wyrażeniach (ang. triples). Stanowi standard dla wymiany metadanych poprzez sieć. RDF posiada wiele sposobów serializacji (formaty plików) umożliwiających rozpowszechnianie danych pomiędzy ludźmi i co ważniejsze ich automatyczne przetwarzanie. Zgodnie z propozycją Tima Berners-Lee semantykę Internetu można opisać za pomocą trójelementowych wyrażeń dotyczących zasobów, składających się z podmiotu, orzeczenia i dopełnienia (ang. subject, predicate, object). Podmiot oznacza zasób, orzeczenie oznacza własność zasobu i wyraża związek między podmiotem a obiektem²².

Podstawowy model danych składa się z następujących obiektów:

Zasoby (ang. resources)

Wszystkie obiekty opisane w wyrażeniach RDF nazywane są zasobami. Zasobem może być cała strona internetowa, na przykład dokument HTML. Zasób może być częścią strony internetowej, na przykład określony element HTML lub XML w źródle dokumentu. Zasobem może być również cały zbiór stron, na przykład cała witryna internetowa. Zasób może być także obiektem, który nie jest bezpośrednio dostępny za pośrednictwem sieci, na przykład drukowana książka. Zasoby są zawsze identyfikowane przez URI. Rozszerzalność URI pozwala na wprowadzenie identyfikatorów dla każdego obiektu znajdującego się w sieci internetowej (URL) lub poza nią.

Właściwości (ang. properties)

Właściwość to określony aspekt, cecha, atrybut lub relacja używana do opisu zasobu. Każda właściwość ma określone znaczenie, określa jej dozwolone wartości, typy zasobów, które może opisać, oraz jej związek z innymi właściwościami. Sposób wyrażania cech właściwości jest definiowany przez schemat RDF (RDFS).

Wyrażenia (ang. statements)

Określony zasób wraz z nazwaną właściwością oraz wartością tej właściwości dla tego zasobu jest wyrażeniem RDF. Wartość właściwości może być innym zasobem (określonym przez URI) lub literałem czyli prostym ciągiem znaków lub innym typem pierwotnym (np. data, liczba całkowita).

²² <https://www.w3.org/TR/rdf-concepts/>

Przykład:

Jan Kowalski jest autorem dokumentu https://dane.gov.pl/media/ckeditor/2018/12/05/standard-techniczny_22-11-2018.pdf.

To zdanie składa się z następujących części.

Podmiot (Zasób)	https://dane.gov.pl/media/ckeditor/2018/12/05/standard-techniczny_22-11-2018.pdf#OpisRDF
Orzeczenie (Własność)	Ma autora
Dopełnienie (Literał)	„Jan Kowalski”

7.2. Serializacja modelu RDF

Model RDF posiada wiele sposobów serializacji w celu wymiany danych poprzez sieć oraz automatycznego przetwarzania danych pomiędzy usługami oraz aplikacjami. Formaty plików (RDF/XML, Turtle, N3, JSON-LD, N-tripels, RDFa i inne), za pomocą których zapisuje się model RDF są równoważne. Oznacza to, że ich zawartość można jednoznacznie transformować z jednego formatu na drugi. Transformację można wykonać za pomocą darmowych konwerterów, które są dostępne w sieci. Poniżej krótki opis każdego formatu.

Format	Opis
RDF/XML	RDF/XML to format zdefiniowany przez W3C, do wyrażania modelu metadanych RDF jako dokumentu XML. RDF/XML często jest błędnie określany jako RDF, ponieważ został wprowadzony wśród innych specyfikacji W3C definiujących RDF i był historycznie pierwszym standardowym formatem serializacji RDF W3C ²³ .
Turtle	Jest to skrót od Terse RDF Triple Language. Składnia tego formatu jest bardzo podobna do języka zapytań SPARQL. Turtle został zdefiniowany przez Dave’a Becketta jako podzbiór języka Notation3 (N3) oraz nadzbiór minimalnego formatu N-Triples. Turtle może serializować tylko prawidłowe grafy RDF i jest alternatywą dla RDF/XML. W przeciwieństwie do RDF/XML, Turtle nie bazuje na XML i jest ogólnie uznawany za czytelniejszy i łatwiejszy do edycji ręcznej niż jego odpowiednik RDF/XML ²⁴ .

²³ <https://www.w3.org/TR/rdf-syntax-grammar/>

²⁴ <https://www.w3.org/TR/turtle/>

Notation3 (N3)	Notacja3 lub N3 jest formatem znacznie czytelniejszym dla ludzi niż notacja RDF/XML. Format został opracowywany przez Tima Bernersa-Lee i innych członków społeczności semantycznej. N3 ma kilka cech wykraczających poza serializację modeli RDF, takich jak możliwość definiowania reguł wnioskowania ²⁵ .
N-Triples	N-Triples jest to liniowy, szeregowy format serializacji dla grafów RDF i stanowi podzbiór formatu Turtle. N-Triples został opracowany przez Dave'a Becketta z University of Bristol i Art Barstow z World Wide Web Consortium (W3C). N-Triples został zaprojektowany jako prostszy format niż Turtle, a zatem łatwiejszy do analizy i generowania przez oprogramowanie. Ze względu na brak niektórych skrótów zapewnianych przez inne serializacje RDF (takie jak CURIE i zasoby zagnieżdżone), uciążliwe może być ręczne wpisywanie dużych ilości danych ²⁶ .
JSON-LD	JSON-LD został zaprojektowany w oparciu o koncepcję „kontekstu”, aby zapewnić dodatkowe odwzorowania z JSON na model RDF. Kontekst łączy właściwości obiektu w dokumencie JSON z pojęciami w ontologii. Aby zmapować składnię JSON-LD na RDF, JSON-LD umożliwia rzutowanie wartości do określonego typu lub ich otagowanie. Kontekst można osadzić bezpośrednio w dokumencie JSON-LD lub umieścić w osobnym pliku i odwoływać się do niego z różnych dokumentów (z tradycyjnych dokumentów JSON za pośrednictwem nagłówka HTTP Link) ²⁷ .
RDFa	Format RDFa to rekomendacja W3C, który dodaje zestaw rozszerzeń na poziomie atrybutów do HTML, XHTML i różnych typów dokumentów opartych na XML, do osadzania różnorodnych metadanych w dokumentach internetowych. RDFa umożliwia osadzanie wyrażen trójelementowych RDF w dokumentach XHTML ²⁸ .

Składnia każdego formatu serializacji RDF składa się z dwóch części. W pierwszej znajdują się deklaracje przestrzeni nazw (słowników), które określają semantykę modelu RDF. Druga część jest opisowa i prezentuje zasoby (Identyfikator URI), cechy (słowniki) oraz ich wartości w postaci zasobów lub literałów.

²⁵ <https://www.w3.org/TeamSubmission/n3/>

²⁶ <https://www.w3.org/TR/n-triples/>

²⁷ <https://www.w3.org/TR/json-ld11/>

²⁸ <https://www.w3.org/TR/rdfa-core/>

Poniżej najbardziej popularne formaty zaprezentowane dla modelu z powyższego przykładu.

RDF/XML

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns:article="https://dane.gov.pl/knowledge-base/
education#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
>
  <rdf:Description rdf:about="https://dane.gov.pl/
knowledge-base/education#standard-techniczny_22-11-2018">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-
schema#Resource"/>
    <article:author rdf:resource="https://me.jkowalski.pl/
foaf.rdf#me"/>
    <dc:title>Standard techniczny</dc:title>
  </rdf:Description>
  <rdf:Description rdf:about="https://me.jkowalski.pl/foaf.
rdf#me">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-
schema#Resource"/>
    <foaf:name>Jan Kowalski</foaf:name>
    <foaf:mbox>Jan.Kowalski@mc.gov.pl</foaf:mbox>
  </rdf:Description>
</rdf:RDF>
```

Powyższy plik otwiera deklaracja XML. Następną linią to otwierający tag RDF, na końcu pliku jest tag zamykający RDF </rdf:RDF>. Wszystko co znajduje się między tagami musi być zgodne ze składnią i semantyką RDF. W pierwszej części znajdują się deklaracje przestrzeni nazw XML. W drugiej części pomiędzy tagami rdf:Description znajdują się opisy zasobów za pomocą ich własności oraz wartości.

Turtle

```

@prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
.
@prefix article: <https://dane.gov.pl/knowledge-base/
education#> .
@prefix foaf:   <http://xmlns.com/foaf/0.1/> .
@prefix dc:     <http://purl.org/dc/elements/1.1/> .

<https://me.jkowalski.pl/foaf.rdf#me>
    a                <http://www.w3.org/2000/01/rdf-
schema#Resource> ;
    foaf:mbox       „Jan.Kowalski@mc.gov.pl” ;
    foaf:name       „Jan Kowalski” .

article:standard-techniczny_22-11-2018
    a                <http://www.w3.org/2000/01/rdf-
schema#Resource> ;
    dc:title        „Standard techniczny” ;
    article:author  <https://me.jkowalski.pl/foaf.
rdf#me> .

```

Notation3 (N3)

```

@prefix article: <https://dane.gov.pl/knowledge-base/
education#> .

@prefix dc: <http://purl.org/dc/elements/1.1/> .

@prefix foaf: <http://xmlns.com/foaf/0.1/> .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

```

```

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

@prefix xml: <http://www.w3.org/XML/1998/namespace> .

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

article:standard-techniczny_22-11-2018 a rdfs:Resource ;
    dc:title „Standard techniczny” ;
    article:author <https://me.jkowalski.pl/foaf.rdf#me> .

<https://me.jkowalski.pl/foaf.rdf#me> a rdfs:Resource ;
    foaf:mbox „Jan.Kowalski@mc.gov.pl” ;
    foaf:name „Jan Kowalski” .

```

N-Triples

```

<https://me.jkowalski.pl/foaf.rdf#me> http://xmlns.com/foaf/0.1/name

"Jan Kowalski" .

<https://me.jkowalski.pl/foaf.rdf#me>

http://www.w3.org/1999/02/22-rdf-syntax-ns#type

<http://www.w3.org/2000/01/rdf-schema#Resource> .

<https://dane.gov.pl/knowledge-base/education#standard-techniczny_22-11-2018> <http://purl.org/dc/elements/1.1/title> "Standard techniczny" .

```

```

<https://dane.gov.pl/knowledge-base/education#standard-
techniczny_22-11-2018> <https://dane.gov.pl/knowledge-base/
education#author> <https://me.jkowalski.pl/foaf.rdf#me> .

<https://me.jkowalski.pl/foaf.rdf#me>

<http://xmlns.com/foaf/0.1/mbox> "Jan.Kowalski@mc.gov.pl" .

<https://dane.gov.pl/knowledge-base/education#standard-
techniczny_22-11-2018> <http://www.w3.org/1999/02/22-rdf-
syntax-ns#type>

<http://www.w3.org/2000/01/rdf-schema#Resource> .

```

JSON-LD

```

{
  "@context": {
    "article": "https://dane.gov.pl/knowledge-base/
education#",
    "dc": "http://purl.org/dc/elements/1.1/",
    "foaf": "http://xmlns.com/foaf/0.1/",
    "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
    "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },

```

```
"@graph": [  
  {  
    "@id": "article:standard-techniczny_22-11-2018",  
    "@type": "rdfs:Resource",  
    "article:author": {  
      "@id": "https://me.jkowalski.pl/foaf.rdf#me"  
    },  
    "dc:title": "Standard techniczny"  
  },  
  {  
    "@id": "https://me.jkowalski.pl/foaf.rdf#me",  
    "@type": "rdfs:Resource",  
    "foaf:mbox": "Jan.Kowalski@mc.gov.pl",  
    "foaf:name": "Jan Kowalski"  
  }  
]  
}
```

RDFa

```

<div xmlns="http://www.w3.org/1999/xhtml"
  prefix="
    foaf: http://xmlns.com/foaf/0.1/
    article: https://dane.gov.pl/knowledge-base/education#
    rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
    dc: http://purl.org/dc/elements/1.1/
    rdfs: http://www.w3.org/2000/01/rdf-schema#"
  >
  <div typeof="rdfs:Resource" about="https://dane.gov.pl/
knowledge-base/education#standard-techniczny_22-11-2018">
    <div rel="article:author">
      <div typeof="rdfs:Resource" about="https://
me.jkowalski.pl/foaf.rdf#me">
        <div property="foaf:name" content="Jan Kowalski"></
div>
        <div property="foaf:mbox" content="Jan.Kowalski@
mc.gov.pl"></div>
      </div>
    </div>
    <div property="dc:title" content="Standard
techniczny"></div>
  </div>
</div>

```

7.3. Schemat RDF

RDFS (ang. Recourse Data Framework Schema) jest językiem opisującym semantycznie model metadanych w grafie RDF. Zbiór klas odnoszący się do określonego celu lub dziedziny stanowi schemat RDF i nadaje znaczenie pojęciowe obiektom występujących w grafie RDF. Klasy są zorganizowane w hierarchię i można je rozszerzać poprzez klasy podrzędne. Do klasy może należeć dowolna ilość wierzchołków z grafu RDF. Wierzchołek grafu RDF może być przydzielony do dowolnej liczby klas. W programowaniu obiektowym to definicja klasy determinuje obiekty należące do tej klasy, w RDFS to zbiór wierzchołków stanowi o tym, czym jest klasa. Schemat RDF opisuje się językiem RDF.

RDF Schema stanowi bazę dla języka OWL, pozwalającego na opisywanie zasobów w

postaci ontologii. Zapisane za pomocą RDFS struktury, nazywane są słownikami RDF (ang. RDF vocabularies). Język RDF Schema stanowi rekomendację W3C²⁹.

7.4. Dane połączone

Połączone dane (ang. Linked Data) to najlepsze praktyki dotyczące tworzenia w sieci powiązań pomiędzy danymi pochodzącymi z różnych źródeł. Działania mają na celu umożliwienie eksploracji przez człowieka oraz maszyny sieci danych. Dzięki połączonym ze sobą danym, można znaleźć inne, powiązane dane³⁰.

Tim Berners-Lee w 2006 roku przedstawił zestaw zasad (ang. Linked data principles) publikowania danych w sieci w taki sposób, że wszystkie publikowane dane stają się częścią jednej globalnej przestrzeni danych:

- Stosuj URI jako identyfikator zasobów,
- Stosuj schemat http dla URI tak aby ludzie mogli wyszukiwać zasoby,
- Gdy ktoś szuka URI dostarcz użyteczne informacje,
- Zamieszczaj powiązania do innych URI tak aby można było znaleźć więcej informacji³¹.

Od strony technicznej połączone dane to dane publikowane w sieci opisane w standardzie RDF co umożliwia ich automatyczny odczyt oraz jednoznaczne powiązanie za pomocą URI kolejnych zasobów. Takie podejście pozwala na budowanie powiązań między zasobami, a te z kolejnymi zwiększając jednoznaczny opis semantyczny zasobów w sieci wykorzystując do tego reguły wnioskowania.

Przykład

1. Jan Kowalski jest autorem dokumentu https://dane.gov.pl/media/ckeditor/2018/12/05/standard-techniczny_22-11-2018.pdf.
2. Dokument https://dane.gov.pl/media/ckeditor/2018/12/05/standard-techniczny_22-11-2018.pdf posiada tytuł „Standard techniczny”.

Powyższe wyrażenia trójelementowe zostało przedstawione w formacie JSON-LD. Dopełnienie (ang. object) „Jan Kowalski” posiada identyfikator URI zgodnie z zasadami połączonych danych.

²⁹ <https://www.w3.org/TR/rdf-schema/>

³⁰ <https://www.w3.org/DesignIssues/LinkedData.html>

³¹ [Tamże](#)

```
„@context”: {  
  „article”: „https://dane.gov.pl/knowledge-base/  
education#”,  
  „dc”: „http://purl.org/dc/elements/1.1/”,  
  „foaf”: „http://xmlns.com/foaf/0.1/”,  
  „rdf”: „http://www.w3.org/1999/02/22-rdf-syntax-ns#”,  
  „rdfs”: „http://www.w3.org/2000/01/rdf-schema#”,  
  „xsd”: „http://www.w3.org/2001/XMLSchema#”  
},  
„@id”: „article:standard-techniczny_22-11-2018”,  
„@type”: „rdfs:Resource”,  
„article:author”: {  
  „@id”: „https://me.jkowalski.pl/foaf.rdf#me”  
},  
„dc:title”: „Standard techniczny”  
}
```

Mając na względzie zalecenia przedstawione w niniejszym standardzie oraz konieczność poniesienia kosztów finansowych podczas budowy danych na najwyższych stopniach otwartości, rekomenduje się przeprowadzenie analizy kosztów i korzyści tworzenia danych połączonych.

8.

Profil aplikaciji DCAT-AP

Profil aplikacji DCAT-AP (Application Profile)³² jest specyfikacją opartą na słowniku katalogu danych (DCAT) opracowanym przez W3C. Został przygotowany w celu opisanego katalogów danych sektora publicznego, zestawów danych oraz promowania specyfikacji do wykorzystania przez portale w całej Europie. DCAT-AP pozwala na ujednoczenie metadanych zbiorów dla portali danych w Europie. Zastosowanie wspólnego profilu i promowanie go wśród państw członkowskich znacznie poprawi interoperacyjność między katalogami danych oraz wymianę danych między państwami członkowskimi. Wdrożone zasady płynące z DCAT-AP powinny zapewniać interoperacyjność semantyczną danych na bazie ustalonych kontrolowanych słowników i umożliwiać przekształcanie istniejących słowników metadanych (np. DCAT), z których korzystają podmioty publiczne, a w konsekwencji prowadzić do harmonizacji wymiany danych poprzez ich łatwe przeszukiwanie i wykorzystywanie. DCAT-AP jest specyfikacją stosowaną przez Europejski portal danych oraz wiele innych europejskich portali otwartych danych. Zaleca się stosowanie specyfikacji do opisywania zbiorów danych sektora publicznego i umożliwienia wymiany opisów zbiorów danych pomiędzy portalem dane.gov.pl a innymi krajowymi i europejskimi portalami udostępniającymi publiczne dane.

Specyfikacja DCAT-AP, definiująca profil aplikacyjny dla portali danych w Europie, określa podział klas i właściwości (wykorzystywanych do opisu schematu metadanych) na „wymagane”, „rekomendowane” i „opcjonalne”. W celu zapewnienia minimalnego poziomu interoperacyjności standard DCAT-AP definiuje słowniki kontrolowane, które pozwalają na opisywanie wybranych metadanych w wystandaryzowany sposób. Ponadto standard DCAT-AP rekomenduje opracowanie i publikację własnych otwartych słowników umożliwiających strukturyzowanie metadanych zgodnie z określonym schematem³³.

Przykład opisu metadanych zgodnych z rekomendacjami implementacji DCAT-AP znajduje się w załączniku nr 1.

³² <https://joinup.ec.europa.eu/solution/dcat-application-profile-data-portals-europe/about>

³³ Ekspertyza sporządzana na zlecenie Ministerstwa Cyfryzacji pt.: „Zapewnienie zgodności portalu dane.gov.pl z DCAT-AP. Ekspertyza dotycząca określenia zasad opisujących zbiory danych publicznych zgodnych z DCAT-AP do zaimplementowania na portalu dane.gov.pl”.

9.

Lokalizacja źródeł danych publicznych

Jednym z etapów udostępniania danych jest wybór lokalizacji źródła danych publicznych:

- centralne repozytorium (portal dane.gov.pl) – opcja o wysokim stopniu pewności dostępu i bezpieczeństwa danych, ale zależna od dysponentów danych publikujących swoje zasoby na dane.gov.pl,
- lokalny system teleinformatyczny, w tym lokalne repozytorium danych publicznych powszechnie dostępne w sieci teleinformatycznej – opcja obciążona ryzykiem utraty dostępu do danych wskutek zmian organizacyjnych podmiotu.

Udostępnianie zbiorów danych w centralnym repozytorium jest możliwe poprzez:

- a) dodanie pliku z lokalnego komputera dostawcy danych, który jest następnie zapisywany i przechowywany w bazie danych centralnego repozytorium albo
- b) wprowadzenie adresu URL strony internetowej repozytorium, w ramach której udostępniono zbiór danych (zewnątrzny link do źródła danych).

Dostawca danych decyduje o wyborze lokalizacji źródła danych, mając na względzie:

- posiadaną infrastrukturę teleinformatyczną umożliwiającą gromadzenie w niej informacji publicznych (zbiorów danych) oraz ich udostępnianie,
- możliwości techniczne przechowywania informacji w centralnym repozytorium.

Udostępniając zbiory danych na portalu dane.gov.pl z zewnętrznych lokalizacji, zaleca się dostawcom danych zapewnienie trwałej dostępności każdego opublikowanego zasobu. W przypadku np. usunięcia z danych na zewnętrznej lokalizacji lub zmian organizacyjnych podmiotu, zaleca się dostawcom danych przeniesienie i przechowywanie danych źródłowych w centralnym repozytorium.

10.

Podsumowanie

Poniższa tabela zawiera zalecenia dotyczące udostępniania danych na określonych poziomach otwartości. Jak wskazano w punkcie 1, **nie zaleca się publikowania danych na pierwszym poziomie otwartości.**

Tabela 4. Zalecenia

Rozwiązanie/ standard	Poziom 2 ★★	Poziom 3 ★★★	Poziom 4 ★★★★	Poziom 5 ★★★★★
Udostępnianie przez API	–	zalecane	zalecane	zalecane
Format CSV	zalecany	zalecany	dopuszczalny	dopuszczalny
Format JSON	zalecany	zalecany	dopuszczalny	dopuszczalny
Format XML	dopuszczalny	dopuszczalny	dopuszczalny	dopuszczalny
Format RDF/ XML	dopuszczalny	dopuszczalny	zalecany	zalecany
Format Turtle	dopuszczalny	dopuszczalny	zalecany	zalecany
Format Notation3 (N3)	dopuszczalny	dopuszczalny	zalecany	zalecany
Format N-Triples	dopuszczalny	dopuszczalny	zalecany	zalecany
Format JSON-LD	dopuszczalny	dopuszczalny	zalecany	zalecany
Format RDFa	dopuszczalny	dopuszczalny	zalecany	zalecany
Format HTML	dopuszczalny	dopuszczalny	niezalecany	niezalecany
Format ODS	dopuszczalne	dopuszczalne	niedopuszczalne	niedopuszczalne
Format dBase	dopuszczalny	dopuszczalny	niedopuszczalny	niedopuszczalny
Formaty Office Open XML (XLSX, DOCX)	dopuszczalne	niedopuszczalne	niedopuszczalne	niedopuszczalne
Formaty DOC, XLS, RTF	dopuszczalne	niedopuszczalne	niedopuszczalne	niedopuszczalne
Format PDF	niedopuszczalny	niedopuszczalny	niedopuszczalny	niedopuszczalny

Rozwiązanie/ standard	Poziom 2 ★★	Poziom 3 ★★★	Poziom 4 ★★★★	Poziom 5 ★★★★★
Formaty JPEG, PNG	niedopuszczalne	niedopuszczalne	niedopuszczalne	niedopuszczalne
Format TXT	niedopuszczalny	niedopuszczalny	niedopuszczalny	niedopuszczalny
Formaty ZIP, 7z, Gzip, Bzip2	dopuszczalne	dopuszczalne	dopuszczalne	dopuszczalne
Własnościowe rozszerzenia ZIP	dopuszczalne	niedopuszczalne	niedopuszczalne	niedopuszczalne
Format RAR	dopuszczalny	niedopuszczalny	niedopuszczalny	niedopuszczalny
Dane przestrzenne	dopuszczalne	dopuszczalne	dopuszczalne	dopuszczalne
Typy danych z XML Schema	zalecane	zalecane	–	–
XML Namespaces	–	–	wymagane	wymagane



11.

Załącznik nr 1 - Przykład opisu metadanych zgodnych z rekomendacjami implementacji DCAT-AP.

Format JSON-LD³⁴

```

{
  „@context”:{
    „Agent”:"http://xmlns.com/foaf/0.1/Agent",
    „Catalog”:"http://www.w3.org/ns/dcat#Catalog",
    „Concept”:"http://www.w3.org/2004/02/skos/core#Concept",
    „ConceptScheme”:"http://www.w3.org/2004/02/skos/core#ConceptScheme",
    „Dataset”:"http://www.w3.org/ns/dcat#Dataset",
    „Distribution”:"http://www.w3.org/ns/dcat#Distribution",
    „Document”:"http://xmlns.com/foaf/0.1/Document",
    „Frequency”:"http://purl.org/dc/terms/Frequency",
    „Kind”:"http://www.w3.org/2006/vcard/ns#Kind",
    „LicenseDocument”:"http://purl.org/dc/terms/LicenseDocument",
    „LinguisticSystem”:"http://purl.org/dc/terms/LinguisticSystem",
    „Literal”:"http://www.w3.org/2000/01/rdf-schema#Literal",
    „Location”:"http://purl.org/dc/terms/Location",
    „MediaTypeOrExtent”:"http://purl.org/dc/terms/MediaTypeOrExtent",
    „PeriodOfTime”:"http://purl.org/dc/terms/PeriodOfTime",
    „Resource”:"http://www.w3.org/2000/01/rdf-schema#Resource",
    „RightsStatement”:"http://purl.org/dc/terms/RightsStatement",
    „vcard”:"http://www.w3.org/2006/vcard/ns#",
    „xsd”:"http://www.w3.org/2001/XMLSchema#",
    „accessRights”:{
      „@id”:"http://purl.org/dc/terms/accessRights",

```

³⁴ Ekspertyza sporządzana na zlecenie Ministerstwa Cyfryzacji pt.: [„Zapewnienie zgodności portalu dane.gov.pl z DCAT-AP. Ekspertyza dotycząca określenia zasad opisujących zbiory danych publicznych zgodnych z DCAT-AP do zaimplementowania na portalu dane.gov.pl”](https://www.gov.pl/zdania/2023/04/zapewnienie-zgodnosci-portalu-dane.gov.pl-z-dcat-ap-ekspertyza-dotyczaca-okreslenia-zasad-opisujacych-zbiory-danych-publicznych-zgodnych-z-dcat-ap-do-zaimplementowania-na-portalu-dane.gov.pl).

```
„@type“:”http://purl.org/dc/terms/RightsStatement”
},
„accessURL“:{
  „@id“:”http://www.w3.org/ns/dcat#accessURL”,
  „@type“:”http://www.w3.org/2000/01/rdf-schema#Resource”
},
„accrualPeriodicity“:{
  „@id“:”http://purl.org/dc/terms/accrualPeriodicity”,
  „@type“:”http://purl.org/dc/terms/Frequency”
},
„contactPoint“:{
  „@id“:”http://www.w3.org/ns/dcat#contactPoint”,
  „@type“:”http://www.w3.org/2006/vcard/ns#Kind”
},
„dataset“:{
  „@id“:”http://www.w3.org/ns/dcat#dataset”,
  „@type“:”http://www.w3.org/ns/dcat#Dataset”
},
„description“:{
  „@id“:”http://purl.org/dc/terms/description”,
  „@type“:”http://www.w3.org/2001/XMLSchema#string”
},
„distribution“:{
  „@id“:”http://www.w3.org/ns/dcat#distribution”,
  „@type“:”http://www.w3.org/ns/dcat#Distribution”
},
„downloadURL“:{
```

```
„@id“:”http://www.w3.org/ns/dcat#downloadURL“,
„@type“:”http://www.w3.org/2000/01/rdf-schema#Resource“
},
„endDate“:{
  „@id“:”http://schema.org/endDate“
},
„format“:{
  „@id“:”http://purl.org/dc/terms/format“,
  „@type“:”http://purl.org/dc/terms/MediaTypeOrExtent“
},
„homepage“:{
  „@id“:”http://xmlns.com/foaf/0.1/homepage“,
  „@type“:”http://xmlns.com/foaf/0.1/Document“
},
„identifier“:{
  „@id“:”http://purl.org/dc/terms/identifier“,
  „@type“:”http://www.w3.org/2001/XMLSchema#string“
},
„issued“:{
  „@id“:”http://purl.org/dc/terms/issued“
},
„keyword“:{
  „@id“:”http://www.w3.org/ns/dcat#keyword“,
  „@type“:”http://www.w3.org/2000/01/rdf-schema#Literal“
},
„landingPage“:{
  „@id“:”http://www.w3.org/ns/dcat#landingPage“,
```

```
  „@type“:”http://xmlns.com/foaf/0.1/Document”
},
„language“:{
  „@id“:”http://purl.org/dc/terms/language”,
  „@type“:”http://purl.org/dc/terms/LinguisticSystem”
},
„license“:{
  „@id“:”http://purl.org/dc/terms/license”,
  „@type“:”http://purl.org/dc/terms/LicenseDocument”
},
„modified“:{
  „@id“:”http://purl.org/dc/terms/modified”
},
„name“:{
  „@id“:”http://xmlns.com/foaf/0.1/name”,
  „@type“:”http://www.w3.org/2001/XMLSchema#string”
},
„prefLabel“:{
  „@id“:”http://www.w3.org/2004/02/skos/core#prefLabel”,
  „@type“:”http://www.w3.org/2001/XMLSchema#string”
},
„publisher“:{
  „@id“:”http://purl.org/dc/terms/publisher”,
  „@type“:”http://xmlns.com/foaf/0.1/Agent”
},
„rights“:{
  „@id“:”http://purl.org/dc/terms/rights”,
```

```
„@type”:”http://purl.org/dc/terms/RightsStatement”
},
„spatial”:{
  „@id”:”http://purl.org/dc/terms/spatial”,
  „@type”:”http://purl.org/dc/terms/Location”
},
„startDate”:{
  „@id”:”http://schema.org/startDate”
},
„temporal”:{
  „@id”:”http://purl.org/dc/terms/temporal”,
  „@type”:”http://purl.org/dc/terms/PeriodOfTime”
},
„theme”:{
  „@id”:”http://www.w3.org/ns/dcat#theme”,
  „@type”:”http://www.w3.org/2004/02/skos/core#Concept”
},
„themeTaxonomy”:{
  „@id”:”http://www.w3.org/ns/dcat#themeTaxonomy”,
  „@type”:”http://www.w3.org/2004/02/skos/core#ConceptScheme”
},
„title”:{
  „@id”:”http://purl.org/dc/terms/title”,
  „@type”:”http://www.w3.org/2001/XMLSchema#string”
},
„type”:{
  „@id”:”http://purl.org/dc/terms/type”,
```

```

    „@type”:”http://www.w3.org/2004/02/skos/core#Concept”
  }
},
„@graph”:[
  {
    „@id”:”https://dane.gov.pl/#Catalog”,
    „@type”:”Catalog”,
    „dataset”:[
      {
        „@id”:”https://dane.gov.pl/dataset/971”
      }
    ],
    „description”:[
      {
        „@value”:”Dane o szczególnym znaczeniu dla rozwoju innowacyjności w
        państwie i rozwoju społeczeństwa informacyjnego w jednym miejscu”,
        „@language”:”pl”
      },
      {
        „@value”:”Open Data Portal Poland”,
        „@language”:”en”
      }
    ],
    „publisher”:{
      „@id”:”https://www.gov.pl/web/cyfryzacja”,
      „@type”:”Agent”,
      „name”:”Ministerstwo Cyfryzacji”,
      „type”:{

```

```
„@id”:”http://purl.org/adms/publishertype/NationalAuthority”,
„@type”:”Concept”,
„prefLabel”:[
  {
    „@value”:”Jednostka administracji publicznej”,
    „@language”:”pl”
  },
  {
    „@value”:”National authority”,
    „@language”:”en”
  }
]
},
„title”:[
  {
    „@value”:”Otwarte Dane”,
    „@language”:”pl”
  },
  {
    „@value”:”Open Data Portal”,
    „@language”:”en”
  }
],
„homepage”:{
  „@id”:”https://dane.gov.pl/”,
  „@type”:”Document”
```

```
},
„language“:[
  {
    „@id“:”http://publications.europa.eu/resource/authority/language/POL”
  },
  {
    „@id“:”http://publications.europa.eu/resource/authority/language/ENG”
  }
],
„license“:{
  „@id“:”https://dane.gov.pl/about”,
  „@type“:”LicenseDocument”,
  „type“:{
    „@id“:”http://purl.org/adms/licencetype/PublicDomain”
  }
},
„issued“:{
  „@value“:”2019-04-12T13:55:26Z”,
  „@type“:”xsd:dateTime”
},
„modified“:{
  „@value“:”2019-04-26T13:55:26Z”,
  „@type“:”xsd:dateTime”
},
„themeTaxonomy“:{
  „@id“:”http://publications.europa.eu/resource/authority/data-theme”,
  „@type“:”ConceptScheme”,
```

```

„title”:[
  {
    „@value”:"System zarządzania wiedzą",
    „@language”:"pl"
  },
  {
    „@value”:"Data theme",
    „@language”:"en"
  }
]
},
„rights”:"public",
„spatial”:{
  „@id”:"http://publications.europa.eu/resource/authority/country/POL"
}
},
{
  „@id”:"https://dane.gov.pl/dataset/971",
  „@type”:"Dataset",
  „description”:[
    {
      „@value”:"Lista zasobów danych publicznych w Centralnym Repozytorium
Informacji Publicznej (CRIP) udostępnionych na portalu dane.gov.pl.",
      „@language”:"pl"
    },
    {
      „@value”:"List of public data resources in the Central Repository of Public
Information (CRIP) available on the dane.gov.pl portal.",

```

```

    „@language”:”en”
  }
],
„title”:[
  {
    „@value”:”Lista zasobów udostępnionych w Centralnym Repozytorium
Informacji Publicznej”,
    „@language”:”pl”
  },
  {
    „@value”:”List of resources shared in the Central Repository of Public
Information”,
    „@language”:”en”
  }
],
„contactPoint”:{
  „@id”:”_:contactPoint_id”,
  „@type”:”Kind”,
  „vcard:fn”:”Ministerstwo Cyfryzacji”,
  „vcard:hasAddress”:{
    „@id”:”_:contactPoint_hasAddress_id”,
    „@type”:”vcard:Address”,
    „vcard:country-name”:”Polska”,
    „vcard:locality”:”Warszawa”,
    „vcard:postal-code”:”00-060”,
    „vcard:street-address”:”ul. Królewska 27”
  },
  „vcard:hasEmail”:”mailto:mc@mc.gov.pl”,

```

```
„vcard:hasTelephone”:[
  {
    „@id”:"_:contactPoint_hasTelephone_phone_id",
    „@type”:"vcard:Voice",
    „vcard:hasValue”:{
      „@id”:"tel:+48222455931"
    }
  },
  {
    „@id”:"_:contactPoint_hasTelephone_fax_id",
    „@type”:"vcard:Fax",
    „vcard:hasValue”:{
      „@id”:"fax:+48222455931"
    }
  }
],
„distribution”:[
  {
    „@id”:"https://dane.gov.pl/dataset/971/resource/16620#Distribution",
    „@type”:"Distribution",
    „accessURL”:{
      „@id”:"https://dane.gov.pl/dataset/971/resource/16620"
    },
    „downloadURL”:{
      „@id”:"DB.resource:link"
    },
  },
]
```

```

„title”:[
  {
    „@value”:"Lista wszystkich zasobów udostępnionych na portalu Otwarte
Dane - 2019-04-01",
    „@language”:"pl"
  },
  {
    „@value”:"List of all resources shared on the Open Data portal - 2019-04-
01",
    „@language”:"en"
  }
],
„description”:[
  {
    „@value”:"Raport w formacie CSV, zawiera wszystkie zbiory i zasoby
udostępnione przez instytucje na portalu Otwarte Dane - serwisie internetowym
Centralnego Repozytorium Informacji Publicznej.",
    „@language”:"pl"
  },
  {
    „@value”:"Report in CSV file format containing all datasets and resources
shared by institutions on the Open Data portal.",
    „@language”:"en"
  }
],
„format”:{
  „@id”:"http://publications.europa.eu/resource/authority/file-type/CSV"
},
„license”:{

```

```
„@id“:”_url_licencji”,
„@type“:”LicenseDocument”,
„type“:{
  „@id“:”http://purl.org/adms/licencetype/PublicDomain”
}
},
„language“:[
  {
    „@id“:”http://publications.europa.eu/resource/authority/language/POL”
  }
],
„issued“:{
  „@value“:”2019-04-01T13:55:26Z”,
  „@type“:”xsd:dateTime”
},
„modified“:{
  „@value“:”2019-04-01T13:55:26Z”,
  „@type“:”xsd:dateTime”
}
},
{
  „@id“:”https://dane.gov.pl/dataset/971/resource/16388”,
  „@type“:”Distribution”,
  „accessURL“:{
    „@id“:”https://dane.gov.pl/dataset/971/resource/16388”
  },
  „downloadURL“:{
```

```

    „@id”:”DB.resource:link”
  },
  „title”:{
    „@value”:”Lista wszystkich zasobów udostępnionych na portalu Otwarte
Dane - 2019-03-22”,
    „@language”:”pl”
  },
  „description”:{
    „@value”:”Raport w formacie CSV, zawiera wszystkie zbiory i zasoby
udostępnione przez instytucje na portalu Otwarte Dane - serwisie internetowym
Centralnego Repozytorium Informacji Publicznej.”,
    „@language”:”pl”
  },
  „format”:{
    „@id”:”http://publications.europa.eu/resource/authority/file-type/CSV”
  },
  „license”:{
    „@id”:”url_licencji”,
    „@type”:”LicenseDocument”,
    „type”:{
      „@id”:”http://purl.org/adms/licencetype/PublicDomain”
    }
  },
  „language”:{
    „@id”:”http://publications.europa.eu/resource/authority/language/POL”
  },
  „issued”:{
    „@value”:”2019-03-22T13:55:26Z”,

```

```
    „@type”:"xsd:dateTime"
  },
  „modified”:{
    „@value”:"2019-03-22T13:55:26Z",
    „@type”:"xsd:dateTime"
  }
},
{
  „@id”:"https://dane.gov.pl/dataset/971/resource/16172",
  „@type”:"Distribution",
  „accessURL”:{
    „@id”:"https://dane.gov.pl/dataset/971/resource/16172"
  },
  „downloadURL”:{
    „@id”:"DB.resource:link"
  },
  „title”:{
    „@value”:"Lista wszystkich zasobów udostępnionych na portalu Otwarte
Dane - 2019-03-15",
    „@language”:"pl"
  },
  „description”:{
    „@value”:"Raport w formacie CSV, zawiera wszystkie zbiory i zasoby
udostępnione przez instytucje na portalu Otwarte Dane - serwisie internetowym
Centralnego Repozytorium Informacji Publicznej.",
    „@language”:"pl"
  },
  „format”:{
```

```
    „@id“:”http://publications.europa.eu/resource/authority/file-type/CSV”
  },
  „license“:{
    „@id“:”url_licencji”,
    „@type“:”LicenseDocument”,
    „type“:{
      „@id“:”http://purl.org/adms/licencetype/PublicDomain”
    }
  },
  „language“:{
    „@id“:”http://publications.europa.eu/resource/authority/language/POL”
  },
  „issued“:{
    „@value“:”2019-03-15T13:55:26Z”,
    „@type“:”xsd:dateTime”
  },
  „modified“:{
    „@value“:”2019-03-15T13:55:26Z”,
    „@type“:”xsd:dateTime”
  }
},
„keyword“:[
  {
    „@value“:”otwarte dane”,
    „@language“:”pl”
  },

```

```
{
  „@value”:”zasoby”,
  „@language”:”pl”
},
{
  „@value”:”centralne repozytorium informacji publicznej”,
  „@language”:”pl”
},
{
  „@value”:”CRIP”,
  „@language”:”pl”
},
{
  „@value”:”dane publiczne”,
  „@language”:”pl”
},
{
  „@value”:”dane”,
  „@language”:”pl”
},
{
  „@value”:”open data”,
  „@language”:”en”
},
{
  „@value”:”resources”,
  „@language”:”en”
}
```

```
},  
{  
  „@value”:”Central Repository of Public Information”,  
  „@language”:”en”  
},  
{  
  „@value”:”public data”,  
  „@language”:”en”  
},  
{  
  „@value”:”data”,  
  „@language”:”en”  
}  
],  
„publisher”:{  
  „@id”:”https://dane.gov.pl/institution/123”  
},  
„theme”:[  
  {  
    „@id”:”http://publications.europa.eu/resource/authority/data-theme/GOVE”  
  }  
],  
„accrualPeriodicity”:{  
  „@id”:”http://publications.europa.eu/resource/authority/frequency/NEVER”,  
  „@type”:”Frequency”  
},  
„accessRights”:”public”,
```

```
„landingPage“:{
  „@id“:”https://dane.gov.pl/”
},
„language“:[
  {
    „@id“:”http://publications.europa.eu/resource/authority/language/POL”
  }
],
„issued“:{
  „@value“:”2019-03-15T13:55:26Z”,
  „@type“:”xsd:dateTime”
},
„modified“:{
  „@value“:”2019-04-12T13:55:26Z”,
  „@type“:”xsd:dateTime”
},
„spatial“:{
  „@id“:”http://publications.europa.eu/resource/authority/country/POL”
},
„temporal“:{
  „@id“:”_:dataset_temporal_id”,
  „@type“:”PeriodOfTime”,
  „startDate“:{
    „@value“:”2019-03-15”,
    „@type“:”xsd:date”
  },
  „endDate“:{
```

```

    „@value”:”2019-04-12”,
    „@type”:”xsd:date”
  }
},
„identifier”:”https://dane.gov.pl/dataset/971”
},
{
  „@id”:”https://dane.gov.pl/institution/123”,
  „@type”:”Agent”,
  „name”:”Ministerstwo Cyfryzacji”,
  „type”:{
    „@id”:”http://purl.org/adms/publishertype/NationalAuthority”,
    „@type”:”Concept”,
    „prefLabel”:[
      {
        „@value”:”Jednostka administracji publicznej”,
        „@language”:”pl”
      },
      {
        „@value”:”National authority”,
        „@language”:”en”
      }
    ]
  }
},
{
  „@id”:”http://publications.europa.eu/resource/authority/data-theme/GOVE”,

```

```
„@type”:”Concept”,
„prefLabel”:[
  {
    „@value”:”Rząd i sektor publiczny”,
    „@language”:”pl”
  },
  {
    „@value”:”Government and public sector”,
    „@language”:”en”
  }
]
}
]
}
```

12.

**Załącznik - Lista
podstawowych aktów
prawnych, norm, standardów,
wytycznych dotyczących
szeroko rozumianej otwartości
danych publicznych**

Polskie otoczenie legislacyjne, standardy

Ustawy

1. Ustawa z dnia 6 września 2001 r. o dostępie do informacji publicznej (Dz. U. z 2019 r., poz. 1429),
2. Ustawa z dnia 25 lutego 2016 r. o ponownym wykorzystywaniu informacji sektora publicznego (Dz. U. z 2019 r., poz. 1446).

Rozporządzenia, uchwała

1. Rozporządzenie Rady Ministrów z dnia 12 marca 2014 r. w sprawie Centralnego Repozytorium Informacji Publicznej (Dz. U. z 2014 r., poz. 361 z późn. zm.),
2. Rozporządzenie Ministra Cyfryzacji z dnia 23 sierpnia 2018 r. w sprawie zasobu informacyjnego przeznaczonego do udostępniania w Centralnym Repozytorium Informacji Publicznej (Dz. U. z 2018 r., poz.1790),
3. Rozporządzenie Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych (DZ. U. z 2017 r., 2247),
4. Uchwała Rady Ministrów NR 107/2016 z dnia 20 września 2016 r. w sprawie ustanowienia „Programu otwierania danych publicznych”.

Standardy

1. Standard prawny, przygotowany przez Ministerstwo Cyfryzacji, w ramach projektu „Otwarte dane – dostęp, standard, edukacja”, dofinansowanego z poddziałania 2.3.1 Programu Operacyjnego Polska Cyfrowa,
2. Standard bezpieczeństwa, przygotowany przez Ministerstwo Cyfryzacji, w ramach projektu „Otwarte dane – dostęp, standard, edukacja”, dofinansowanego z poddziałania 2.3.1 Programu Operacyjnego Polska Cyfrowa,
3. Standard API, przygotowany przez Ministerstwo Cyfryzacji, w ramach projektu „Otwarte dane – dostęp, standard, edukacja”, dofinansowanego z poddziałania 2.3.1 Programu Operacyjnego Polska Cyfrowa.

Europejskie i międzynarodowe otoczenie legislacyjne, normy, standardy, rekomendacje

Dyrektywy

1. Dyrektywa Parlamentu Europejskiego i Rady nr 2013/37/UE z dnia 26 czerwca 2013 r. zmieniająca dyrektywę 2003/98/WE w sprawie ponownego wykorzystywania informacji sektora publicznego,
2. Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego (jeszcze nieimplementowana).

Normy, standardy, rekomendacje

1. CSV, RFC 4180 <https://tools.ietf.org/html/rfc4180>
2. DCAT, Data Catalogue Vocabulary, W3C Foundation, <https://www.w3.org/TR/vocab-dcat/>, <https://www.w3.org/TR/vocab-dcat-2/>
3. DCAT Application Profile, European Commission <https://op.europa.eu/en/web/eu-vocabularies/dcat-ap>
4. Data on the Web Best Practices, W3C Foundation, <https://www.w3.org/TR/dwbp/>
5. Format daty i godziny ISO 8601, International Organization for Standardization 2004, <https://www.iso.org/iso-8601-date-and-time-format.html>
6. HTTP, RFC2616 <https://tools.ietf.org/html/rfc2616>, RFC7240 <https://tools.ietf.org/html/rfc7240>, RFC7232 <https://tools.ietf.org/html/rfc7232>, RFC7231 <https://tools.ietf.org/html/rfc7231>, RFC2617 <https://www.ietf.org/rfc/rfc2617.txt>
7. HTML – RDFa, W3C Recommendation <https://www.w3.org/TR/html-rdfa/>
8. JSON, RFC 4627 <https://tools.ietf.org/html/rfc4627>, RFC 7159 <https://tools.ietf.org/html/rfc7159>
9. JSON-LD, W3C Recommendation <http://www.w3.org/TR/json-ld/>, <https://www.w3.org/TR/json-ld11/>
10. JSON API <https://jsonapi.org/format/>

11. JSON SCHEMA <https://json-schema.org/>
12. Linked Data, Berners-Lee, Linked Data, <http://www.w3.org/DesignIssues/LinkedData,W3C Recommendation https://www.w3.org/TR/ldp/>
13. Model for Tabular Data and Metadata on the Web, W3C Foundation <https://www.w3.org/TR/tabular-data-model/#locating-metadata>
14. N-TRIPLES, W3C Recommendation <http://www.w3.org/TR/n-triples/>
15. OWL, W3C Recommendation <http://www.w3.org/TR/owl2-overview/>, <https://www.w3.org/TR/2017/REC-owl-time-20171019/>
16. URI, RFC3986 <https://tools.ietf.org/html/rfc3986>, RFC 1737 <https://www.ietf.org/rfc/rfc1737.txt>, RFC 1736 <https://www.ietf.org/rfc/rfc1736.txt>
17. UTF-8, RFC 3629 <http://www.ietf.org/rfc/rfc3629.txt>
18. RDF, W3C Recommendation <http://www.w3.org/TR/rdf11-mt/>
19. RDF Schema, W3C Recommendation <https://www.w3.org/TR/rdf-schema/>
20. RDF/XML, W3C Recommendation.
URL: <http://www.w3.org/TR/REC-rdf-syntax/>
21. Schemat otwartości danych, Berners-Lee T. <https://www.w3.org/DesignIssues/LinkedData.html>, Berners-Lee, T. (2013) 5 Star Open Data <http://5stardata.info>
22. Standard semantyczny API <https://semver.org/lang/pl/>
23. Spatial Data on the Web Best Practices, W3C Foundation, <https://www.w3.org/TR/sdw-bp/>
24. SPARQL, W3C Recommendation, <https://www.w3.org/TR/sparql11-query/>
25. Schematy metadanych (nazewnictwo), www.schema.org
26. Turtle,W3C Recommendation <https://www.w3.org/TR/turtle/>
27. The OAuth 2.0 Authorization Framework, RFC 6749 <https://tools.ietf.org/html/rfc6749>, RFC 6750 <https://tools.ietf.org/html/rfc6750>
28. XML Schema,W3C Recommendation <https://www.w3.org/XML/Schema>, <https://www.w3.org/TR/xmlschema/>
29. XML Media Types, RFC 3023 <http://www.ietf.org/rfc/rfc3023.txt>
30. XML Namespaces, W3C Recommendation <https://www.w3.org/TR/xml-names/>