

Posiedzenie Rady Infrastruktury Informacji Przestrzennej, 10 lutego 2021 r.

Jak otwierać, to z głową

Jeśli Polska ma być liderem w otwieraniu danych, musimy dbać nie tylko o liczbę publikowanych zbiorów, ale też o ich jakość. A ta wciąż pozostawia wiele do życzenia – również w przypadku geodanych.

Jerzy Królikowski

Już teraz stopień otwartości krajowych rejestrów jest na bardzo dobrym poziomie, co potwierdza najnowsza edycja badania „Open Data Maturity”. To do roczne zestawienie Komisji Europejskiej opisuje postępy osiągnięte przez kraje wspólnoty i ocenia ich poziom dojrzałości w zakresie otwierania danych publicznych. W badaniu z 2020 roku Polsce przyznano wysokie, 6. miejsce, podczas gdy jeszcze dwa lata temu zajmowaliśmy 13. lokatę! – Polski rząd ten raport cieszy, ale celem jest wejście na podium – deklarowała podczas pierwszego tegorocznego posiedzenia Rady IIP Anna Gos – pełnomocnik ministra cyfryzacji ds. otwartości danych. Do osiągnięcia tego celu mają nas przybliżyć dwa zasadnicze kroki.

Pierwszym jest uchwalenie ustawy o otwartych danych i ponownym wykorzystaniu informacji sektora publicznego. Będzie ona – jak zaznaczyła Anna Gos – nie tylko implementować unijną dyrektywę, ale także porządkować obecnie obowiązujące przepisy oraz wprowadzać zmiany wykraczające poza europejskie wymogi.

Z punktu widzenia IIP kluczowe jest wprowadzenie definicji tzw. danych o wysokiej wartości, do których zaliczane będą również dane geoprzestrzenne. Zarówno dyrektywa, jak i ustawa określają zasady dostępu do tych zbiorów. Zgodnie z proponowanymi przepisami ponowne wykorzystanie tego typu danych ma być bezpłatne. Będą one ponadto udostępniane w formacie nadającym się do odczytu maszynowego za pośrednictwem interfejsu programistycznego (API) oraz – jeśli to możliwe – również do zbiorczego ściągania. Szczegółowe regulacje mają być określone w aktach wykonawczych Komisji Europejskiej oraz krajowych rozporządzeniach, ich projekty nie ujrzały jednak jeszcze światła dziennego. Za to projekt wspomnianie ustawy przeszedł już przez konsultacje i uzgodnienia oraz uzyskał akceptację komitetów Rady Ministrów ds. cyfryzacji i praw euro-

pejskich. Zgodnie z unijnymi wymogami ustawa powinna wejść w życie do połowy lipca br.

Drugim ważnym krokiem jest uchwalenie rządowego Programu Otwierania Danych na lata 2021-2027. Dokument ten zakłada nie tylko uwalnianie kolejnych rejestrów publicznych, ale także podnoszenie ich jakości oraz skuteczniejsze udostępnianie zbiorów. Realizacja tego ostatniego punktu obejmie m.in. rozbudowę portalu Dane.gov.pl, zwiększenie liczby zasobów udostępnianych w tym serwisie przez API, a także szersze wykorzystanie automatycznego pobierania metadanych (tzw. harvesting) dla zewnętrznych repozytoriów.

O ile liczba uwolnionych krajowych zbiorów oraz forma ich udostępniania faktycznie robią wrażenie, o tyle jakość tych danych często pozostawia wiele do życzenia. Świetnym przykładem z naszego podwórka jest rejestr danych adresowych, co na posiedzeniu Rady IIP szerzej omówił Kajetan Wojsyk z Centrum e-Zdrowia (dawniej CSIOZ).

Jak zwrócił uwagę, obecnie dane te przechowywane są w wielu państwowych rejestrach, z których dwa główne to prowadzona przez GUS baza TERYT oraz gminna ewidencja miejscowości, ulic i adresów (EMUiA). Niestety, rozbieżności w tych rejestrach jest sporo i mimo podejmowania działań naprawczych cały czas pojawiają się kolejne. Co ciekawe, występują one nie tylko między EMUiA i TERYT, ale nawet między EMUiA i gminnymi uchwałami! Wprawdzie błędy te czasem wydają się błaha (to np. sposób zapisu cudzośćlowu), ale niekiedy tyle wystarczy, by utrudnić dotarcie karetki pod wskazany adres.

W ocenie Kajetana Wojsyka problem jest na tyle palący, że trzeba skończyć z ręcznym poprawianiem rejestrów i wprowadzić zmiany systemowe. Przede wszystkim należy dopuścić do obiegu prawnego wyłącznie dane zesłownikowane, które są odpowiednio powiązane z unikatowymi identyfikatorami. Ponadto zamiast operacji na samych danych należy wprowadzić operacje na ich iden-

tyfikatorach, które byłyby dekodowane w miejscu docelowym na podstawie centralnych słowników. Należy też wdrożyć narzędzia do automatycznej kontroli jakości danych, które nie tylko wykrywałyby ewentualne błędy, ale także wysyłały raport w ich sprawie do dysponenta danych. Podobne rozwiązanie własnymi siłami wdrożyło Centrum e-Zdrowia, dzięki czemu np. do rejestru aptek nie zostanie już dodana żadna placówka z błędnym adresem.

W odpowiedzi na to wystąpienie główny geodeta kraju zapewnił, że GUGiK podejmuje liczne działania w celu poprawy jakości danych adresowych. Nie odniósł się jednak do meritum propozycji Kajetana Wojsyka.

Waldemar Izdebski chętnie mówił za to o planach GUGiK związanych z popularyzacją BDOT10k. By zwiększyć grono jego użytkowników, pod koniec ubiegłego roku uruchomiono Portal BDOT10k. Jak zwraca uwagę GKG, serwis pozwala uzyskać odpowiedzi na różne pytania dotyczące ponad 60 mln przechowywanych w tym rejestrze obiektów bez konieczności pobierania ich na komputer i posiadania specjalistycznego oprogramowania. Przy jego użyciu możemy np. szybko policzyć i wskazać lokalizację wszystkich hoteli znajdujących się w obszarze Natura 2000 w zadanym województwie. Portal BDOT10k ma więc szansę istotnie usprawnić prace analityczne realizowane w wielu różnych urzędach – podkreślił GKG. Zapowiedział jednocześnie udostępnienie kolejnych narzędzi, które jeszcze bardziej ułatwią korzystanie z tej bazy.

Podczas dyskusji na Radzie IIP pojawił się także temat aktualności BDOT10k, która wciąż pozostawia wiele do życzenia. Wcale nierzadko zdarza się bowiem, że np. nowe obwodnice stają się widoczne w tym rejestrze dopiero wiele miesięcy po oddaniu do użytku. Waldemar Izdebski przypomniał jednak, że zeszłoroczna nowelizacja *Prawa geodezyjnego i kartograficznego* znacznie ułatwiła aktualizację, gdyż nie ma już wymogu formalnego zgłaszania pracy kartograficznej. GKG zapowiedział ponadto dążenie do wprowadzenia dla całego kraju 2-letniego cyklu aktualizacji tej bazy (w miejsce obecnego 3-letniego). ■