



OTWARTE DANE JAKO ŹRÓDŁO INNOWACYJNYCH MODELI BIZNESOWYCH

Ilona Pawełoszek¹, Jędrzej Wieczorkowski², Magdalena Jurczyk-Bunkowska³

¹Politechnika Częstochowska, Wydział Zarządzania

²Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych

³Politechnika Opolska, Wydział Ekonomii i Zarządzania

Streszczenie: W dobie gospodarki opartej na wiedzy kluczowe znaczenie dla zarządzania przedsiębiorstwem ma dostęp do wiarygodnych źródeł danych. Otwarte dane sektora publicznego (OGD) mają znaczącą wartość gospodarczą. Dają możliwość budowania innowacyjnych rozwiązań, także tworzenia nowych modeli biznesowych. Istotne jest ponowne wykorzystanie danych do celów innych niż te, dla których zostały wytworzone, a także łączenie danych z wielu heterogenicznych zbiorów. Zapewnienie otwartych danych oznacza publikowanie ich w przydatnej formie, przy użyciu odpowiednich standardów i stopnia przetworzenia, zaś idea danych połączonych (*Linked Data*) to zbiór najlepszych praktyk związanych z publikowaniem i łączeniem danych strukturalnych w sieci. W artykule przedstawiono koncepcję OGD w kontekście udostępniania danych, formatów i łączenia danych. Celem artykułu jest identyfikacja możliwości opracowania innowacyjnego modelu biznesowego opartego na otwartych danych sektora publicznego (w tym na danych połączonych) jako elementu zarządzania strategicznego organizacją. W tym celu przeanalizowano dostępne polskie źródła otwartych danych i formaty ich publikacji. Na poparcie tezy o dużym biznesowym znaczeniu otwartych danych przedstawiono przykładowy pomysł wykorzystania takich danych.

Słowa kluczowe: CRIP, *Linked Data*, OGD, *Big Data*, *Open Data*, *Open Government Data*

DOI: 10.17512/znpcz.2019.1.15

Wprowadzenie

Podmioty gospodarcze stosują różnorodne podejścia do realizacji swojego podstawowego celu, jakim jest osiągnięcie zysku. Zdefiniowanie modelu biznesowego i powiązanej z nim strategii działania ma kluczowe znaczenie dla zarządzania przedsiębiorstwem. Wykorzystywane modele biznesowe ewoluują wraz ze zmianą otoczenia gospodarczego, w tym dostępnych technologii oraz różnego typu zasobów. Rozwój gospodarki cyfrowej i społeczeństwa informacyjnego prowadzi do zwiększenia roli zasobów intelektualnych, w szczególności danych i informacji różnego pochodzenia. Większa dostępność zasobów leży w interesie całej gospodarki, umożliwiając powstawanie i rozwój podmiotów gospodarczych.

¹ Ilona Pawełoszek, dr inż., ilona.paweloszek@wz.pcz.pl, ORCID: 0000-0002-3590-3969

² Jędrzej Wieczorkowski, dr, jedrzej.wieczorkowski@sgh.waw.pl, ORCID: 0000-0002-1252-8975

³ Magdalena Jurczyk-Bunkowska, dr inż., m.jurczyk-bunkowska@po.opole.pl, ORCID: 0000-0002-4066-3605

Z powyższego wynika dążenie do uwolnienia zasobów informacyjnych dotąd wykorzystywanych gospodarczo w niewystarczającym stopniu. Inicjatywa otwartych danych (*Open Data*) polega na udostępnianiu danych wszystkim zainteresowanym użytkownikom nie tylko w celach informacyjnych, ale także z możliwością ich ponownego biznesowego zastosowania. Niewątpliwie najważniejszą cechą otwartych danych jest możliwość ich wykorzystania do różnych celów, nie tylko bezpośrednio związanych z kontekstem powstania zbioru danych. Otwartość danych wiąże się z łączeniem różnych zbiorów w celu uzyskania nowej wiedzy, którą można wykorzystać w biznesie na wiele innowacyjnych sposobów. Wiedza ta może mieć kluczowe znaczenie dla zarządzania przedsiębiorstwem w szczególności w obszarze wspomaganie decyzji, ale także może stać się źródłem wartości dodanej do produktów i usług.

Instytucje administracji publicznej podczas swych codziennych działań tworzą i archiwizują ogromne ilości danych. Często są to dane wysokiej jakości z wielu dyscyplin, takich jak: budżety organów rządowych i samorządowych, dane statystyczne, mapy, dane o środowisku, rozkłady transportu publicznego itp. Mogą być potencjalnie publikowane przez poszczególne organy administracji lub przez innych dostawców w sposób zdecentralizowany i zindywidualizowany, lecz także w sposób scentralizowany i ustandaryzowany. W praktyce repozytoria różnią się jednak od siebie pod względem zakresu, wielkości i standardów publikowanych danych. Informacje takie udostępniane w otwartej formie określane są mianem *Open Government Data* (OGD), w języku polskim przyjęte jest trochę szersze pojęcie otwartych danych sektora publicznego.

Znaczenie OGD rośnie wraz z rozwojem możliwości przetwarzania danych masowych (*Big Data*) oraz danych połączonych (*Linked Data*). Rozwój Internetu w kierunku inteligentnej sieci Web 3.0 umożliwił łączenie danych z różnych źródeł w celu pozyskania nowej użytecznej wiedzy. Trend łączenia danych i zbiorów technologii używanych w tym celu jest określany jako *Linked Open Data* (LOD). LOD to coraz powszechniejsza praktyka, której celem jest poprawa zarządzania informacjami i ich integracja poprzez ułatwienie dostępu do różnorodnych zbiorów otwartych danych.

Autorzy prowadzą od około dwóch lat obserwacje polskich i światowych źródeł OGD, w szczególności pod kątem zawartości repozytoriów takich danych oraz standardów i formatów ich publikacji. Celem artykułu jest identyfikacja możliwości opracowania modelu biznesowego opartego na otwartych danych sektora publicznego jako elementu zarządzania strategicznego organizacją. Szczególną uwagę zwrócono na kwestię możliwości łączenia zbiorów danych, ich jakości i potencjalnej użyteczności gospodarczej otwartych danych.

Przyjęta metodologia badawcza obejmowała systematyczny przegląd wybranych portali oferujących dostęp do otwartych danych publicznych. Przeglądu dokonano pod kątem ewidencji zbiorów w poszczególnych formatach wraz z oceną możliwości ich zastosowań we wspomaganie zarządzania organizacjami. Z racji ograniczonej objętości niniejszego artykułu szczegółowe wyniki badań nie zostały zaprezentowane, jednakże można się z nimi zapoznać w innych publikacjach autorów (Wieczorkowski, Pawełoszek 2018; 2018a).

W kolejnych rozdziałach artykułu, wykorzystując badania literaturowe, przedstawiono pojęcia modelu biznesowego, ideę OGD i jej kluczowe zasady, scharakteryzowano pojęcia *Big Data* i *Linked Data*. Następnie przedstawiono możliwość opracowywania nowych modeli biznesowych i budowania aplikacji, opierając się na otwartych danych sektora publicznego. Artykuł kończy się podsumowaniem i wskazaniem problemów w dostępie do OGD.

Modele biznesowe w zarządzaniu strategicznym

Każda organizacja ma opracowany pewien swój sposób działania ściśle powiązany z jej strategią i celami funkcjonowania. W przypadku organizacji o charakterze komercyjnym, które są obszarem zainteresowania niniejszego artykułu, sposób działania ma na celu wypracowanie zysku i może być określony jako model biznesowy. Pojęcie modelu biznesowego można analizować z wielu punktów widzenia. Przykładowy przegląd przeprowadził T. Doligalski (Doligalski (red.) 2014, s. 20-28), wyróżniając podejścia: taksonomiczne, rozmyte, przez pryzmat komponentów, wartości, zależności, podmiotowe oraz utrzymane w nurcie teorii systemów i cybernetyki. Autor ten ostatecznie uznał model biznesowy (stosuje równoważne mu pojęcie modelu biznesu) jako uproszczony obraz firmy przedstawiający najważniejsze jej elementy oraz zależności między nimi.

Z punktu widzenia niniejszego artykułu interesujące jest podejście taksonomiczne próbujące klasyfikować modele z punktu widzenia ich cech i powiązania ze strategią zarządzania przedsiębiorstwem. Przykładową typową klasyfikacją jest wyróżnienie modeli: wytwórców, dystrybutorów, właścicieli i pośredników (Szpringer 2012, s. 43-119). Jednakże oparcie funkcjonowania organizacji na danych i ich przetwarzaniu wymyka się typowemu spojrzeniu, choć jednocześnie może być podstawą wyodrębnienia nowej grupy podmiotów w powyższej taksonomii. Dla niniejszego artykułu istotne znaczenie ma też podejście wykorzystujące pryzmat komponentów. Jest ono reprezentowane m.in. przez A. Osterwaldera i Y. Pigneur (Osterwalder, Pigneur 2012, s. 14-55), według których model biznesowy opisuje przesłanki stojące za sposobem, w jaki organizacja tworzy wartość oraz zapewnia i czerpie zyski z tej wytworzonej wartości. Jest to w pewnym sensie szkic strategii, która ma zostać wdrożona w ramach struktur, procesów i systemów organizacji. Wyżej wymienieni autorzy stworzyli szablon zawierający dziewięć podstawowych elementów modelu: segmenty klientów; propozycje wartości; kanały komunikacji, dystrybucji i sprzedaży; relacje z klientami; strumień przychodów; kluczowe zasoby; kluczowe działania; kluczowi partnerzy; strukturę kosztów. Jedną z grup zasobów kluczowych są zasoby intelektualne (obok zasobów fizycznych, ludzkich i finansowych). Obejmują one m.in. markę, wiedzę autorską, patenty, prawa autorskie, bazy danych klientów. Podobne podejście przez pryzmat komponentów przedstawiają T. Gołębiowski, T.M. Dudzik, M. Lewandowska i M. Witek-Hajduk (Gołębiowski i in. 2008), którzy zasoby (kompetencje) traktują jako jeden z komponentów modelu biznesowego obok wartości dla klienta, źródeł przychodów, docelowych klientów, strategii, relacji z partnerami i oferowanych produktów.

Tak więc zasoby traktowane są jako jeden z podstawowych elementów modelu biznesowego. Model można traktować jako uproszczony opis, jak organizacja (system gospodarczy) tworzy wartość poprzez transformację zasobów i relacje wymiany z innymi podmiotami gospodarczymi (Doleski 2015, s. 5). Do zasobów intelektualnych, poza uprzednio wymienionymi przez A. Osterwaldera i Y. Pigneura (Osterwalder, Pigneur 2012, s. 14-55), należy zdaniem autorów niniejszego artykułu zaliczyć również szeroko rozumiane dane i informacje, także te pochodzące z zewnątrz organizacji, w tym powszechnie dostępne, ale tylko takie, które organizacja jest w stanie spożytkować w swojej działalności. Otwarte dane stają się więc coraz istotniejszym zasobem obok innych dotąd typowych zasobów, które mogą stanowić własność firmy, być przedmiotem dzierżawy lub należeć do kluczowych partnerów. Zasoby intelektualne bardzo często występują w postaci dóbr cyfrowych, które charakteryzują się, w przeciwieństwie do dóbr materialnych, łatwością powielania i przetwarzania. Jeśli takie działanie jest zgodne z prawem, a z taką sytuacją mamy do czynienia w przypadku OGD, dane tego typu stają się typowym dobrem publicznym, gdyż ich konsumpcja nie jest ograniczona i nie występuje konkurencja w konsumpcji. W konsekwencji zasoby OGD dają możliwość budowania innowacyjnych modeli biznesowych.

Zgodnie z klasyczną definicją zaproponowaną przez P.F. Druckera (Drucker 1985) źródła innowacji mogą znajdować się wewnątrz i w otoczeniu firmy. W tym kontekście otwarte zbiory danych można uznać za zewnętrzne źródła innowacji, ponieważ zmieniają percepcję rzeczywistości i stwarzają możliwość zdobycia nowej użytecznej wiedzy. Możliwość wcześniejszego opracowania innowacyjnych produktów i modeli biznesowych nie była dostępna w tej skali ze względu na brak wystarczających i bezpłatnych danych.

Według autorów istnieje kilka (w praktyce łączonych) będących podstawą zarządzania strategicznego podstawowych modeli biznesowych opierających się na danych (Wieczorkowski, Jurczyk-Bunkowska 2017, s. 134-144):

- wykorzystanie własnych (wewnętrznych) danych w procesach biznesowych na potrzeby oferowanych produktów,
- sprzedaż lub udostępnianie własnych danych w postaci surowej lub przetworzonej,
- wykorzystanie pozyskanych danych zewnętrznych w procesach biznesowych na potrzeby oferowanych produktów,
- sprzedaż własnych kompetencji związanych z przetwarzaniem danych.

Posiadane lub pozyskane dane, oprócz wykorzystania do różnych wewnętrznych analiz, mogą stać się podstawą do oferowania nowych produktów (najczęściej usług), w tym luźno powiązanych z pierwotnym przeznaczeniem danych. Z punktu widzenia niniejszego artykułu najbardziej interesujący jest model biznesowy oparty na danych zewnętrznych, z którego mogą korzystać organizacje nieposiadające własnych cennych danych, jednakże mające innowacyjne pomysły.

Idee otwartego rządu i otwartych danych

Zgodnie z definicją zaproponowaną przez OECD otwarty rząd opiera się na zasadach: odpowiedzialności, jawności i otwartości. Oznacza to, iż władze publiczne

informują społeczeństwo o swoich działaniach poprzez udostępnianie usług i informacji oraz reagują na nowe pomysły i potrzeby (Curristine, Abbott 2005, s. 29). Idea otwartego rządu zakłada między innymi znaczące uczestnictwo wszystkich zainteresowanych stron w projektowaniu i wdrażaniu polityki publicznej (OECD 2017). Jak wskazują badania, otwartość danych wywiera pozytywny wpływ na biznes, wzrost gospodarczy, dobrobyt i innowacje (Janssen, Charalabidis, Zuiderwijk 2012, s. 258-268; Lakomaa, Kallberg 2013, s. 558-563; Stagars 2016, s. 15; Kitsios, Papachristos, Kamariotou 2017, s. 398-408).

Z kolei pojęcie otwartych danych odnosi się do zbiorów danych, które mogą być udostępniane, wykorzystywane, przetwarzane i publikowane przez każdego, bez ograniczeń prawa autorskiego i patentowego, przy czym istnieje jedynie wymóg wskazania źródła danych lub umożliwienia dalszej dystrybucji przetwarzanej treści na tych samych warunkach (Kozierski i in. 2013, s. 8-9). Jasne zdefiniowanie otwartych danych oraz ustalenie polityki zarządzania ich publikacją ułatwia osiągnięcie interoperacyjności różnych źródeł i możliwości ich łączenia. Według podręcznika otwartych danych (<http://opendatahandbook.org>) najważniejszymi aspektami otwierania danych są:

- Dostępność. Dane muszą być dostępne jako całość i przy rozsądnych kosztach reprodukcji, najlepiej udostępniane przez Internet. Dane muszą być również dostępne w wygodnej i modyfikowalnej formie.
- Ponowne wykorzystanie i redystrybucja. Dane muszą być dostarczone na warunkach, które pozwalają na ponowne wykorzystanie i redystrybucję, w tym na łączenie z innymi zestawami danych.
- Uniwersalne uczestnictwo. Każdy musi mieć możliwość używania, ponownego wykorzystywania i redystrybucji otwartych danych. Nie powinno być żadnych ograniczeń w odniesieniu do działań, osób lub grup (np. ograniczenia wyłącznie do zastosowań niekomercyjnych lub edukacyjnych).

Działanie otwartego rządu jest niemożliwe bez dostępu do informacji publicznych. Brak dostępu do nich oznacza więc brak możliwości kontrolowania działań rządu oraz niemożność podejmowania inicjatyw przez obywateli. Obywatele nie posiadający pełnej wiedzy na temat działań władz nie będą nigdy równorzędnymi partnerami w dialogu (<http://otwartyrzad.org.pl>). Połączenie idei otwartości danych i otwartego rządu tworzy koncepcję otwartych danych sektora publicznego (OGD). Celem udostępniania takich danych jest upublicznianie zasobów informacyjnych tworzonych przez lub na zlecenie administracji publicznej, a także bezpłatne korzystanie i dystrybucja otwartych danych przez każdego obywatela (Papińska-Kacperek, Polańska 2015, s. 103-114). Z punktu widzenia możliwości powtórnego wykorzystania danych listę ich odbiorców należy rozszerzyć w szczególności o podmioty komercyjne.

Zarówno otwarty rząd, jak i otwarte dane nie są całkowicie nowymi ideami, jednak teraz zyskują większą uwagę ze strony decydentów, biznesu i obywateli. Początki koncepcji powszechnego prawa dostępu do informacji sięgają połowy XVIII wieku (Mendel 2003, s. 79). W drugiej połowie XX wieku prawo swobodnego dostępu do informacji stało się powszechnym standardem w wielu krajach. Otwarcie danych sektora publicznego do celów gospodarczych jest obecnie globalnym

trendem, który na dobre rozwinął się wraz z uruchomieniem w maju 2009 r. przez rząd USA serwisu (repozytorium) Data.gov, a następnie podobnych usług przez Wielką Brytanię i Nową Zelandię (Sayogo, Pardo, Cook 2014, s. 1896). Zmiany nie tylko wpłynęły na wewnętrzne procesy i komunikację instytucji administracji publicznej, ale przede wszystkim umożliwiły interakcję między rządem a obywatelami oraz zainicjowały powtórne wykorzystanie danych do celów gospodarczych. W praktyce udostępnianie informacji publicznych jest najczęściej zorganizowane w formie portalu internetowego pełniącego rolę centralnego repozytorium informacji publicznej (CRIP). Repozytorium zapewnia standaryzację sposobu dostarczania danych, interfejsu oraz częściowo także formatów danych.

W Polsce rolę tę pełni portal Danepubliczne.gov.pl. W praktyce CRIP są budowane w oparciu o lokalne prawo dostępu do informacji publicznej i jej ponownego wykorzystania. Ponadto w Unii Europejskiej stosowana jest Dyrektywa 2013/37/UE zmieniająca dyrektywę 2003/98/WE w sprawie ponownego wykorzystywania informacji sektora publicznego (zwaną dyrektywą ponownego użycia). Dyrektywa ta zachęca poszczególne państwa do współdzielenia zasobów informacyjnych i ponownego ich wykorzystania w celach biznesowych, a także określa warunki, w których ma ono miejsce.

Koncepcje *Linked Data* i *Big Data* a otwarte dane

Rozwój technologii komunikacyjnych i informacyjnych oraz społeczeństwa cyfrowego powoduje znaczący wzrost ilości danych dostępnych w formie cyfrowej, w tym także wytwarzanych przez sektor publiczny. Wielkie zbiory danych, których ze względu na ich charakterystykę nie sposób analizować tylko i wyłącznie siłami ludzkimi oraz tradycyjnymi technikami przetwarzania danych, wraz z tymi technikami określane są mianem *Big Data*. Według Gartnera ([https://www.gartner.com/...](https://www.gartner.com/)) *Big Data* można rozumieć jako zasoby informacyjne o dużej objętości, dużej prędkości generowania i/lub dużej różnorodności, które wymagają ekonomicznych, innowacyjnych form przetwarzania, umożliwiają lepszy wgląd, wspomagają podejmowanie decyzji i automatyzację procesów. Termin ten używany jest także w odniesieniu do całości procesów biznesowych związanych z wykorzystaniem tych zbiorów danych w celu uzyskania wartości komercyjnej (Strycharz b.r.). Wśród trzech podstawowych aspektów *Big Data* (Wieczorkowski, Polak 2014, s. 182-196): technologicznego (infrastruktura, oprogramowanie, metody gromadzenia i analizy danych), biznesowego (zastosowanie) i społecznego (konsekwencje społeczne), nacisk w tym artykule położony został na dwa pierwsze: aspekt ekonomiczny – tutaj zastosowanie OGD w biznesie – oraz aspekt technologiczny – tutaj metody gromadzenia i udostępniania danych publicznych.

W koncepcji *Big Data* szczególnie istotny jest dostęp do dużej ilości danych w czasie rzeczywistym lub do niego zbliżonym, gromadzenie i przetwarzanie danych niestrukturalnych lub słabo ustrukturyzowanych, dostęp do danych z zewnątrz firmy, w tym pochodzących z publicznie dostępnych źródeł (Surma 2017, s. 38-45). W kontekście ostatniej z wymienionych cech jednym ze źródeł mogą być otwarte dane publiczne. Dzięki wielkości i różnorodności danych możliwe jest

badanie współzależności zjawisk, wykrywanie wcześniej nieznanymi relacji. Wartość dodana wynika z możliwości łączenia zbiorów danych pochodzących z różnych źródeł i pierwotnie przeznaczonych do zupełnie innych celów. Warto zauważyć, że wzrost ilości danych jest również spowodowany dodatkowymi informacjami pochodzącymi z analizy istniejących zbiorów danych (Ubaldi 2013, s. 7).

Kwestia udostępniania danych jest z kolei ściśle powiązana z koncepcją połączonych danych, która oznacza możliwość semantycznego łączenia danych z różnych źródeł. *Linked Data* dotyczy wykorzystywania Internetu do łączenia powiązanych danych, które nie były wcześniej połączone, lub korzystania z Internetu w celu zmniejszenia barier w łączeniu danych, które obecnie połączone są za pomocą innych metod (<http://linkeddata.org/>). *Linked Data* to stosunkowo nowy paradygmat publikowania informacji, który jest praktycznym krokiem do wdrożenia wizji semantycznej sieci Web. Podstawową ideą *Linked Data* jest powiązanie dokumentów za pomocą hiperłączy i zastosowanie ogólnej architektury WWW do udostępniania informacji strukturalnych w skali globalnej.

Najlepsze praktyki *Linked Data* zostały przedstawione przez T. Bernersa-Lee (Berners-Lee 2006). Pierwsza z wytycznych sugeruje użycie URI (*Uniform Resource Identifier*) do identyfikacji nie tylko dokumentów i treści cyfrowych, ale także obiektów rzeczywistych i abstrakcyjnych. Kolejną zasadą jest stosowanie protokołu HTTP jako jednolitego sposobu dostępu do zasobów. Trzecia zasada promuje wykorzystanie jednego modelu do publikowania danych strukturalnych w Internecie w celu zapewnienia interoperacyjności. Standardem tym jest RDF (*Resource Description Framework*), który jest prostym grafowym modelem danych. Czwarta reguła zaleca tworzenie hiperłączy między różnymi obiektami powiązanymi tematycznie, osobami, miejscami, wydarzeniami czy organizacjami. Zasady danych połączonych formułują podstawy rozszerzania sieci WWW o globalną przestrzeń danych opartą na istniejącej architekturze, która zakłada, że połączone dokumenty mogą znajdować się na różnych serwerach.

Modele biznesowe wykorzystujące otwarte dane

Stworzenie produktu z wykorzystaniem otwartych danych wymaga innowacyjnego myślenia połączonego z wiedzą o potrzebach klienta lub rozpoznaniem niszy rynkowej. Otwarte dane sektora publicznego mogą stanowić impuls do zmian w zarządzaniu strategicznym i być podstawą do opracowania nowego modelu biznesowego lub dostarczenia usług dodających wartość do istniejących produktów. Modele biznesowe oparte na otwartych danych można podzielić na różne kategorie, biorąc pod uwagę następujące kryteria: wartość dodaną, liczbę wykorzystanych zbiorów danych, dostęp do danych w czasie rzeczywistym lub w trybie asynchronicznym, zaawansowanie technologiczne (w szczególności przez wykorzystanie metod *Big Data* lub *Linked Data*), docelową grupę odbiorców, źródło przychodu i finansowania.

Celem modelu biznesowego jest dostarczenie określonej wartości dodanej dla klienta, a sama wartość dodana może przyjmować różne formy. Niektóre przedsięwzięcia są skoncentrowane na wykorzystaniu jednorazowej szansy, a inne

na długoterminowych zyskach. Przykład jednorazowego projektu to aplikacja lub strona internetowa prezentująca profile kandydatów przed wyborami. Natomiast rozwiązanie przewidziane do ciągłego, długoterminowego użytkowania to aplikacja informująca o pogodzie lub zanieczyszczeniu środowiska. Dysponując nieprzetworzonymi danymi, wartość dodaną można uzyskać poprzez interpretację danych, ich agregację, wizualizację lub publikację w formie interesującej dla potencjalnego odbiorcy. Przykładem tego typu prac mogą być okresowe raporty branżowe czy artykuły prasowe. Finansowanie tego rodzaju działalności może przyjąć formę opłaty od czytelników lub modelu opartego na reklamie (koszt za kliknięcie, koszt za działanie, koszt za tysiąc wyświetleń itp.). Jest to najprostszy możliwy sposób dodawania wartości w oparciu o pojedynczy zestaw danych.

Technologicznie bardziej zaawansowaną formą wykorzystania otwartych danych są aplikacje, które automatycznie pobierają dane i prezentują je w sposób wybrany przez użytkownika, podkreślając najważniejsze kwestie. Przykładem może być prognoza pogody, ostrzeżenia o zanieczyszczeniu lub informacje o ruchu drogowym dla danej lokalizacji.

Bardziej zaawansowany model zakłada inteligentne połączenie informacji z dwóch lub więcej źródeł. Prawdziwa wartość danych może zostać ujawniona, gdy są one łączone w sposób, który tworzy nowy punkt widzenia. W początkowej fazie model ten wymaga dobrego pomysłu, pracy koncepcyjnej i profesjonalnej wiedzy programistycznej, aby dynamicznie łączyć zbiory danych i prezentować wyniki. Dlatego konieczne są znaczne nakłady inwestycyjne. Kolejne wydatki związane są z utrzymaniem systemu, ponieważ źródła danych mogą ulec zmianie, dlatego wymagają stałego monitorowania.

Prowadzone przez autorów badania dostępności danych sektora publicznego (Wieczorkowski, Pawełoszek 2018, s. 54-71) pozwalają na stwierdzenie, że teoretyczny potencjał ich wykorzystywania do celów biznesowych jest bardzo duży. Problemem jest jednak zakres oraz sposób publikacji danych. Rozważając rzeczywiste możliwości korzystania z otwartych danych publicznych, należy zwrócić uwagę na dostępność zbiorów danych na konkretny temat, zawartość metadanych, formaty udostępnianych zbiorów i możliwość ich powiązania.

Analizując zawartość polskiego CRIP, można zauważyć, że wiele zbiorów ma w praktyce znikomą użyteczność. Przykładowo publikacja informacji o wydatkach samorządowych w postaci tabel bez dodatkowych opisów i metadanych jest zasadniczo nieczytelna dla przeciętnego obywatela. Dodanie URI odnoszących się do szczegółowych opisów pozycji budżetowych może znacznie ułatwić ich interpretację. Aby połączyć dowolne dwa lub więcej zestawów danych, każdy z nich powinien zawierać przynajmniej jeden wspólny atrybut z innym zestawem. Najczęstsze elementy łączące zbiory danych to współrzędne geograficzne, nazwy miejsc lub wydarzeń oraz osoby.

Chociaż ilość połączonych danych rośnie, można zauważyć, że każda jednostka świata realnego (np. osoba, organizacja, lokalizacja) w różnych otwartych zestawach danych może być opisywana z syntaktycznie różnymi identyfikatorami. Należy jednak zauważyć, że pomimo braku bezpośredniego związku między zestawami danych, nadal możliwe jest użycie wspólnej warstwy koncepcyjnej (na przykład

ontologii, taksonomii lub tezaurusu), która połączy te dane poprzez zrozumienie ich wzajemnego związku. Jednym z wyzwań związanych z integracją, wyszukiwaniem i ponownym wykorzystaniem połączonych danych jest uzyskanie ontologii zgodnej z zestawami danych (Li, Sima 2015, s. 10). W przypadku zestawów danych zawierających nazwy lokalizacji przydatna może być przykładowo ontologia podziału terytorialnego, która pozwalałaby na przypisywanie nazw ulic do dzielnic miast. Taka ontologia może zapewnić nową interesującą perspektywę analizy danych. Globalna inicjatywa *Linked Data* obejmuje oprócz publikacji otwartych danych także dostęp do słowników RDFS lub ontologii OWL, które umożliwiają semantyczne połączenie danych. Zbiór ten to LOV (*Linked Open Vocabulary*), są w nim dostępne różne słowniki i ontologie, w szczególności opisujące zbiory publikowane za pomocą platformy CKAN (*Comprehensive Knowledge Archive Network*).

Przykład aplikacji wykorzystującej otwarte dane

Polski serwis CRIP (<https://danepubliczne.gov.pl/>), podobnie jak większość innych krajowych repozytoriów, zawiera poza właściwymi danymi aplikacje utworzone na ich podstawie. W lipcu 2018 r. udostępniano tam jedynie 16 aplikacji zbudowanych zarówno przez instytucje publiczne, jak i firmy komercyjne. Serwis jest jednak otwarty na nowe aplikacje, pomysły można składać za pomocą specjalnego formularza. Polski CRIP zawiera aplikacje dla osób indywidualnych i firm, np. wspieranie planowania podróży, planów zagospodarowania przestrzennego i wydatków komunalnych, wykorzystanie środków unijnych, lokalizacja adresów pocztowych, informacje o firmach. W większości przypadków aplikacje wykorzystują pojedyncze zbiory danych, w niewielkim stopniu łączą je między sobą. Oczywiście jest możliwość budowania własnych aplikacji wykorzystujących dane publiczne bez zgłaszania ich w CRIP.

Aby zilustrować możliwości wykorzystania otwartych danych publicznych, autorzy proponują przykład aplikacji, która ma na celu wspomaganie wyboru kierunku studiów. W polskim CRIP znajduje się zbiór „Zestawienie prowadzonych studiów na kierunkach”, który jest aktualizowany co kwartał i publikowany w formacie XLS. Pola zdefiniowane w tabeli to: nazwa kierunku, poziom kształcenia, profil, tytuł zawodowy, forma kształcenia, dyscyplina (obszar/dziedzina), instytucja/jednostka.

Pola danych są dobrze zdefiniowane i dają możliwość ich semantycznego połączenia z innymi zbiorami. Wprawdzie dane te nie spełniają wymagań LOD ze względu na format licencjonowany, jednak możliwe jest przekonwertowanie ich do formatu CSV, XML lub RDF. Pole, które można powiązać z innymi zbiorami, to w szczególności instytucja/jednostka, ponieważ zapisy zawierają nazwy identyfikujące uczelnie w sposób jednoznaczny. Znając współrzędne geograficzne, można powiązać ten zbiór danych z innymi zbiorami opisującymi dany region. Dla potencjalnego użytkownika – młodego człowieka wybierającego kierunek studiów – interesujące mogą być takie dane jak bezrobocie, pracodawcy w regionie według branży (w celu odbycia praktyk studenckich, stażu lub znalezienia zatrudnienia),

a także atrakcje (możliwość uprawiania sportu, imprezy kulturalne). Nie bez znaczenia jest również wykorzystanie zbioru danych na temat finansów szkół wyższych dostępnego w polskim CRIP, ponieważ może to świadczyć o potencjale rozwojowym wybranej uczelni. Aplikacja powinna dawać możliwość określenia profilu użytkownika, aby móc lepiej dobrać odpowiednie połączenia semantyczne (jeśli to możliwe w sposób dynamiczny i interaktywny).

Powyżej przedstawiono jedynie ogólny pomysł aplikacji, szczegóły modelu biznesowego związanego z oferowaniem takiego produktu wykraczają poza ramy niniejszego artykułu.

Podsumowanie i wnioski

Możliwość stworzenia modelu biznesowego na podstawie danych zewnętrznych (w szczególności OGD) oraz ich wykorzystania w zarządzaniu strategicznym jest ściśle związana ze sposobami udostępniania takich danych. Przede wszystkim format powinien być możliwy do odczytu maszynowego. Wskazane jest również, aby korzystanie z danych nie wymagało licencji niezbędnych do ich odczytania. Ponadto, aby móc łączyć dane z różnych źródeł, muszą one być zaopatrzone w odpowiednie metadane i osadzone w odpowiednim kontekście. Przy ocenie OGD pod względem ich dostępności, należy wziąć pod uwagę nie tylko ilość danych, ale także jakość, w tym użyteczność wynikającą z zastosowanych standardów i formatów danych.

Badania, które aktualnie prowadzą autorzy, oraz różnorodne inne źródła pokazują, że Polska wciąż pozostaje w tyle za liderami otwierania danych sektora publicznego. Dokonane przez autorów porównanie treści polskiego CRIP z repozytoriami w krajach bardziej zaawansowanych w implementacji idei OGD (USA: <https://www.data.gov>, Wielka Brytania: <https://data.gov.uk>) wskazuje na bardzo duże różnice w ilości i użyteczności udostępnianych danych. Podstawowymi zidentyfikowanymi problemami w Polsce są:

- niewielka ilość instytucji publikujących dane oraz samych danych dostępnych w CRIP,
- niewystarczające zastosowanie zaawansowanych formatów: otwartych, strukturalnych i użytecznych w kontekście łączenia danych,
- minimalne, heterogeniczne i rozproszone metadane,
- niewystarczające udostępnienie interfejsów programistycznych (API – *Application Programming Interface*), np. w celu uzyskania dostępu do danych przez aplikacje stron trzecich,
- niezadowalająca częstość aktualizacji zbiorów OGD,
- niewielka ilość udostępnionych aplikacji.

Poprawa dostępności danych oraz dojrzałości metod stosowanych w CRIP może znacznie ułatwić ich wykorzystanie w biznesie, a w konsekwencji przyczynić się do wzrostu gospodarczego.

Bez wątpienia koncepcja otwartych danych sektora publicznego przyczynia się do rozwoju społeczeństwa obywatelskiego, które może wykorzystywać nowe technologie i postrzegać informację jako cenne dobro. OGD przyczynia się również

do osiągnięcia korzyści społecznych (świadome i aktywne uczestnictwo obywateli w działaniach rządu) oraz wzrostu gospodarczego. OGD mają ogromny potencjał gospodarczy, który można wykorzystać tylko wtedy, gdy dane są udostępniane w sposób dobrze zorganizowany i zgodny ze standardami. Ponowne wykorzystanie OGD daje możliwość oferowania innowacyjnych produktów, w szczególności usług, a także tworzenia nowych modeli biznesowych. CRIP odgrywają podstawową rolę w dostępie do OGD w skali danego kraju. Celem repozytorium jest zatem zapewnienie dostępu do zasobów informacyjnych o istotnym znaczeniu społecznym i gospodarczym.

Jednakże otwarte dane, w celu właściwego spełnienia swojej funkcji, muszą być przygotowane i udostępnione w sposób ułatwiający ich dalsze przetwarzanie i ponowne wykorzystanie. W praktyce tworzenie aplikacji korzystających z OGD, w szczególności z wykorzystaniem LOD, może napotykać wiele trudności. W przypadku danych udostępnianych w niedojrzały sposób często nie ma możliwości przetwarzania ich w sposób zautomatyzowany. W konsekwencji zaoferowanie innowacyjnej usługi wymagałoby zbyt dużego nakładu pracy, a także powodowałoby duże ryzyko błędów, związanych z definiowaniem reguł połączenia danych. Zapewnienie semantycznej interoperacyjności między otwartymi zestawami danych, które nie są częścią chmury LOD, może być czasochłonne i kosztowne. Tym niemniej ze względu na efekt synergii, który może powstać w wyniku połączenia danych z różnych źródeł, inicjatywy OGD w połączeniu z LOD są szczególnie cenne dla poszczególnych podmiotów, a także dla całej gospodarki. Zatem zarówno z punktu widzenia instytucji publikujących, jak i odbiorców bardzo ważne jest usystematyzowane podejście do zarządzania zbiorami otwartych danych. Polega ono na ustaleniu polityki i procedur obowiązujących w całym cyklu życia otwartych danych od ich tworzenia, poprzez gromadzenie, aż do ponownego wykorzystania. Powinny być przy tym uwzględnione dobre praktyki oraz najnowsze standardy.

Literatura

1. Berners-Lee T. (2006), *Linked Data – Design Issues*, <http://www.w3.org/DesignIssues/LinkedData.html> (dostęp: 02.10.2018).
2. Curristine T., Abbott B. (red.) (2005), *Modernising Government: The Way Forward*, OECD Publishing, Paris.
3. Doleski D.O. (2015), *Integrated Business Model: Applying the St. Gallen Management Concept to Business Models*, Springer, Berlin-Heidelberg.
4. Doligalski T. (red.) (2014), *Modele biznesu w Internecie*, Wydawnictwo Naukowe PWN, Warszawa.
5. Drucker P.F. (1985), *Innovation and Entrepreneurship: Practice and Principles*, Heinemann, London.
6. Dyrektywa Parlamentu Europejskiego i Rady 2013/37/UE z dnia 26 czerwca 2013 r. zmieniająca dyrektywę 2003/98/WE w sprawie ponownego wykorzystywania informacji sektora publicznego (Tekst mający znaczenie dla EOG) (Dz. Urz. UE L 175/1, 27.6.2013).
7. Gołębiowski T., Dudzik T.M., Lewandowska M., Witek-Hajduk M. (2008), *Modele biznesu polskich przedsiębiorstw*, Oficyna Wydawnicza SGH, Warszawa.
8. <http://linkeddata.org> (dostęp: 20.07.2018).
9. <http://opendatahandbook.org> (dostęp: 20.07.2018).

10. <http://otwartyrzad.org.pl> (dostęp: 20.07.2018).
11. <https://danepubliczne.gov.pl/> (dostęp: 20.07.2018).
12. <https://www.data.gov/> (dostęp: 20.07.2018).
13. <https://www.gartner.com/it-glossary/big-data> (dostęp: 20.07.2018).
14. <https://data.gov.uk/> (dostęp: 20.07.2018).
15. Janssen M., Charalabidis Y., Zuiderwijk A. (2012), *Benefits, Adoption Barriers and Myths of Open Data and Open Government*, „Information Systems Management”, Vol. 29(4). DOI: 10.1080/10580530.2012.716740.
16. Kitsios F., Papachristos N., Kamariotou M. (2017), *Business Models for Open Data Ecosystem: Challenges and Motivations for Entrepreneurship and Innovation*, [w:] Proceedings of 19th IEEE International Conference on Business Informatics (CBI'17), IEEE, Thessaloniki. DOI: 10.1109/CBI.2017.51.
17. Kozierski P., Kabaciński R., Lis M., Kaczmarek P. (2013), *Open Access. Analiza zjawiska z punktu widzenia polskiego naukowca*, Impuls, Poznań-Kraków.
18. Lakomaa E., Kallberg J. (2013), *Open Data as a Foundation for Innovation: The Enabling Effect of Free Public Sector Information for Entrepreneurs*, „IEEE Access”, Vol. 1. DOI: 10.1109/ACCESS.2013.2279164.
19. Li H., Sima Q. (2015), *Parallel Mining of OWL 2 EL Ontology from Large Linked Datasets*, „Knowledge Based Systems”, Vol. 84. DOI: 10.1016/j.knosys.2015.03.023.
20. Mendel T. (2003), *Freedom of Information: A Comparative Legal Survey*, UNESCO, New Delhi.
21. OECD (2017), *Recommendation of the Council on Public Integrity*, <http://www.oecd.org/gov/ethics/recommendation-public-integrity/> (dostęp: 20.07.2018).
22. Osterwalder A., Pigneur Y. (2010), *Business Model Generation. A Handbook for Visionaries, Game Changers, and Challengers*, John Wiley & Sons, Hoboken.
23. Papińska-Kacperek J., Polańska K. (2015), *Analiza zaawansowania realizacji idei Open Government Data w wybranych krajach*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Studia Informatica”, nr 874(37). DOI: 10.18276/si.2015.37-08.
24. Sayogo D.S., Pardo T.A., Cook M. (2014), *A Framework for Benchmarking Open Government Data Efforts*, 47th Hawaii International Conference on System Sciences (HICSS), IEEE, Waikoloa. DOI: 10.1109/HICSS.2014.240.
25. Stagars M. (2016), *Open Data in Southeast Asia. Towards Economic Prosperity, Government Transparency and Citizen Participation in the ASEAN*, Palgrave Macmillan, Singapore. DOI: 10.1007/978-3-319-32170-7.
26. Strycharz J. (b.r.), *Ekonomika eksploracji tekstu i danych: analiza aktualnych trendów i przyszłych zastosowań*, https://ngoteka.pl/bitstream/handle/item/364/Polski-raport-g%C5%82%C3%B3wny_final-1.pdf?sequence=3 (dostęp: 05.12.2018).
27. Surma J. (2017), *Cyfryzacja życia w erze Big Data*, Wydawnictwo Naukowe PWN, Warszawa.
28. Szpringer W. (2012), *Innowacyjne modele e-biznesu. Aspekty instytucjonalne*, Difin, Warszawa.
29. Ubaldi B. (2013), *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives*, OECD Working Papers on Public Governance, No. 22. DOI: 10.1787/5k46bj4f03s7-en.
30. Wieczorkowski J., Jurczyk-Bunkowska M. (2017), *Big Data jako źródło innowacji w zarządzaniu i inżynierii produkcji*, [w:] Knosala R. (red.), *Innowacje w zarządzaniu i inżynierii produkcji*, t. 1, Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole.
31. Wieczorkowski J., Pawełoszek I. (2018), *Open Government Data, the Case of Polish Public Sector*, „The Online Journal of Applied Knowledge Management”, Vol. 6, Issue 2.

32. Wieczorkowski J., Pawełoszek I. (2018a), *Open Government Data and Linked Data in the Practice of Selected Countries*, [w:] Bouzas-Lorenzo R., Ramos A.C. (eds.), Proceedings of the 18th European Conference on Digital Government, Academic Conferences and Publishing International Limited.
33. Wieczorkowski J., Polak P. (2014), *Big Data: Three-aspect Approach*, „Online Journal of Applied Knowledge Management”, Vol. 2, Issue 2.

OPEN DATA AS A SOURCE OF INNOVATIVE BUSINESS MODELS

Abstract: Open government data (OGD) have a significant economic value. They can be the basis for developing innovative solutions and new business models. OGD can be linked with other datasets and reused for many purposes other than those for which they were created. Providing open data means publishing them in a useful form, using appropriate standards and degrees of processing, while the idea of Linked Data is a collection of best practices related to publishing and combining structured data on the Web. The idea of OGD in the context of data sharing, data formats and data linking is presented. The aim of the article is to identify the possibilities of creating innovative business models based on open data and Linked Data. For this purpose, available Polish sources of open data were analyzed. In support of the thesis about the great business importance of open data, an exemplary idea of using such data was presented.

Keywords: CRIP, Linked Data, OGD, Big Data, Open Data, Open Government Data