

Mariusz ZABROŃ¹, **Agnieszka MOLGA²**

¹ ORCID: 0000-0002-9418-6621. Mgr inż., Interdyscyplinarna Szkoła Doktorska Politechniki Łódzkiej, Politechnika Łódzka, Wydział Elektrotechniki, Elektroniki i Automatyki, Instytut Mechatroniki i Systemów Informatycznych, ul. B. Stefanowskiego 22, 90-537 Łódź, e-mail: zabron.mariusz@interia.pl

² ORCID: 0000-0002-0857-5111. Dr, Uniwersytet Technologiczno-Humanistyczny w Radomiu, Wydział Transportu, Elektrotechniki i Informatyki, Katedra Informatyki, ul. Malczewskiego 29, 26-600 Radom; e-mail: agnieszka19216@wp.pl

data złożenia tekstu do Redakcji DI: 13.03.2023; data wstępnej oceny artykułu: 28.03.2023

HURTOWNIA DANYCH JAKO ELEMENT CYFROWEJ TRANSFORMACJI PRZEDSIĘBIORSTWA DLA PRZEMYSŁU 4.0

DATA WAREHOUSE AS AN ELEMENT OF THE DIGITAL TRANSFORMATION OF THE COMPANY FOR INDUSTRY 4.0

Słowa kluczowe: baza danych, big data, chmura, hurtownia danych, modelowanie danych.

Keywords: database, big data, cloud, data warehouse, data modeling.

Streszczenie

W artykule opisano zagadnienia dotyczące budowy oraz funkcjonowania hurtowni danych. Jest to szereg procesów wspomagających działalność przedsiębiorstw i wydaje się być obecnie koniecznych we wdrażaniu pozostałych koncepcji i rozwiązań Przemysłu 4.0. W opracowaniu zostały wyjaśnione również procesy, jakie są konieczne do utworzenia hurtowni, kluczowe składniki, a także istota jej działania. Przeanalizowane zostały zalety oraz praktyki stosowane w zakresie biznesu oraz samej sferze IT. Przedstawiona została także możliwość przeniesienia hurtowni do chmury obliczeniowej oraz głównych dostawców oferujących takie usługi. Artykuł może być punktem wyjścia do bardziej szczegółowych rozważań dotyczących analityki, raportowania i archiwizacji danych we współczesnym przedsiębiorstwie.

Abstract

The article describes issues related to the construction and functioning of a data warehouse. This is a series of processes supporting the activities of enterprises and seem to be necessary in the implementation of other concepts and solutions of Industry 4.0. The study also explains the pro-

cesses that are necessary to create a warehouse, key components and the essence of operation. The advantages and practices used in the field of business and the IT sphere itself were analyzed. The possibility of transferring the warehouse to cloud computing was also presented as well as the main suppliers offering such services were presented. The article can be a starting point for more detailed considerations on analytics, reporting and data archiving in a modern enterprise.

Pojęcia: hurtownia danych, baza danych, jezioro danych

Hurtownia danych (ang. *data warehouse*) jest to cyfrowy system, który przechowuje, łączy oraz dostosowuje dane pochodzące z wielu różnych źródeł (przechowuje zarówno dane bieżące, jak i dane historyczne). System pobiera dane z wielu rozproszonych baz, oczyszcza je i ujednolica. Jest to rodzaj analitycznego systemu, który kompleksowo wspiera procesy biznesowe polegające na wyodrębnianiu spośród kontenerów różnych informacji te, które na dany moment są potrzebne i konieczne do podejmowania właściwych decyzji w organizacji. Dzięki temu decyzje te podejmowane są w oparciu o konkretne i poprawne dane, odpowiednio usystematyzowane, aktualne dla potrzeb na daną chwilę. Hurtownia jest również podstawą do generowania raportów i analiz, a także może posłużyć do tworzenia predykcji¹. Utworzenie samej hurtowni nie ma wpływu na funkcjonowanie pozostałych systemów bazodanowych ani na treść zawartych danych – informacje są do *warehouse* ekstraktowane, a w kolejnym etapie udostępniane tylko w trybie „selektowym” (nie można poddać ich modyfikacji). Otrzymane w ten sposób wyniki można w bardzo łatwy i czytelny sposób zaprezentować w formie graficznej wizualizacji. Obecnie funkcjonujące hurtownie są projektowane w taki sposób, aby mogły obsługiwać zarówno ustrukturyzowane, jak i nieustrukturalizowane dane (pliki: wideo, obrazy czy dane bezpośrednio z czujników na liniach produkcyjnych)². Niektóre wykorzystują zintegrowaną technologię analityczną i baz danych „in-memory” (przechowywanie danych w pamięci komputera, zamiast w pamięci dyskowej), co ma na celu zapewnienie dostępu do wiarygodnych danych w czasie rzeczywistym³. Bez hurtowni znacznie trudniej jest łączyć dane z heterogenicznych źródeł, zapewnić im przy tym właściwy format dla celów analitycznych i uzyskać bieżący, a także dalekosiężny wgląd w czasie. Na

¹ M. Toroman, *Chmura Azure. Praktyczne wprowadzenie dla administratora. Implementacja, monitorowanie i zarządzanie ważnymi usługami i komponentami IaaS/PaaS*, Helion, Gliwice 2020.

² A. Pelikant, *Hurtownie danych. Od przetwarzania analitycznego do raportowania*, wyd. II, Helion, Gliwice 2021.

³ Tamże, s. 5.

podstawie tak usystematyzowanych danych tworzone są zaawansowane dashboards, które prezentowane są jako zbiorcze wyniki czytelne dla osób zarządzających organizacją⁴.



Rys. 1. Hurtownia danych

Źródło: A. Pelikant, *MS SQL Server. Zaawansowane metody...*, dz. cyt.

Zagłębiając się bardziej w pojęcie dotyczące *data warehouse*, napotkamy mnóstwo terminów, które są bardzo istotne dla zrozumienia istoty hurtowni danych⁵. Jednymi z najważniejszych zagadnień są: hurtownia danych a baza danych, hurtownia danych a jezioro danych, hurtownia danych a składnica danych.

Zarówno baza danych, jak i hurtownia danych to systemy służące przechowywaniu danych, jednak cele, do jakich mają służyć, są zupełnie różne. W bazie danych przechowywane są zazwyczaj dane z określonej części obszaru biznesowego. W magazynach danych są zarówno aktualne, jak i historyczne dane dla całej działalności. Hurtownia danych pobiera dane z odpowiednich pojedynczych baz używając ich serwera, a posiadając odpowiednie, dodatkowe funkcjonalności z zakresu modelowania danych, zarządzania danymi (np. cyklem życia) czy też integracji źródeł danych, tworzy wyspecyfikowany zakres informacji dedykowany dla żądanych potrzeb.

Do przechowywania dużych porcji danych (ang. *big data*) wykorzystywane są zarówno hurtownie danych, jak i jeziora danych, ale są to bardzo różne systemy przechowywania. Jezioro danych (ang. *data lake*) jest to rodzaj pewnego repozytorium, gdzie przesyłane są „surowe” dane z wielu różnych miejsc, które nie są w żaden sposób uporządkowane⁶. Nie ma ograniczenia, jakie i ile tych danych może być przechowywane – są one zwykle nieustrukturyzowane, ponieważ z różnych powodów nie trafiły do hurtowni (mogą to być m.in.: logi ze stron internetowych, pliki tekstowe, zdjęcia czy odczyty z czujników)⁷. Magazyny danych i jeziora zwykle uzupełniają się pomiędzy sobą. Na przykład, gdy

⁴ A. Pelikant, *MS SQL Server. Zaawansowane metody programowania*, wyd. II, Helion, Gliwice 2021.

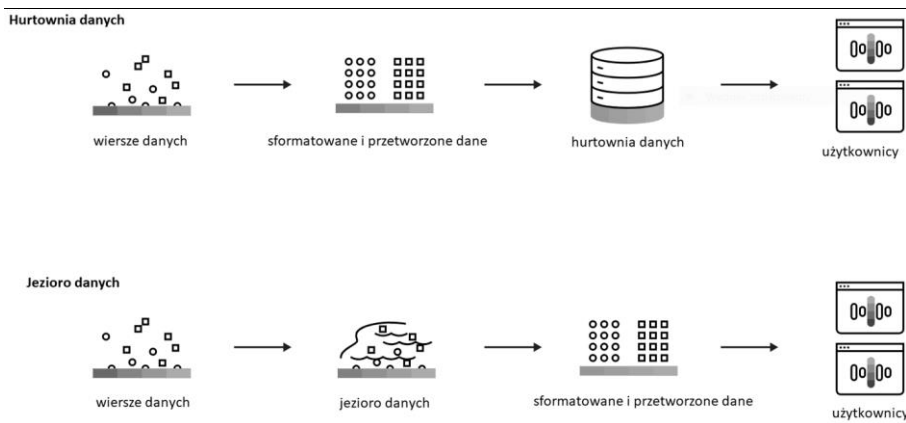
⁵ M. Toroman, *Chmura Azure. Praktyczne wprowadzenie dla administratora...*, dz. cyt., s. 6.

⁶ A. Pelikant, *MS SQL Server. Zaawansowane metody...*, dz. cyt.

⁷ J. Rosenberg, A. Mateos, *Chmura obliczeniowa. Rozwiązania dla biznesu*, Helion, Gliwice 2011, s. 2.

do udzielenia odpowiedzi na pytanie biznesowe potrzebne są surowe dane przechowywane w jeziorze, można je wyodrębnić, wyczyścić, przekształcić i wykończyć do analizy w hurtowni danych⁸.

Zasadnicza różnica między hurtownią a jeziorem danych sprowadza się do zakresu i rodzaju przechowywanych danych. Hurtownia gromadzi dane wyłącznie w uporządkowanej formie, a także ogranicza do wyselekcjonowanych źródeł. Jezioro natomiast ma możliwość przechowywania danych ze wszystkich źródeł, niezależnie od tego, czy są to informacje posiadające określoną strukturę. Jezioro danych będzie bardziej pożyteczne, jeśli potrzebujemy wykonywać zaawansowane analizy na surowych danych (analizę statystyczną czy modelowanie predykcyjne). Hurtownie danych natomiast sprawdzają się znacznie lepiej w działaniach operacyjnych, związanych z bieżącą działalnością w przedsiębiorstwie⁹.



Rys. 2. Hurtownia danych a jezioro danych

Źródło: J. Rosenberg, A. Mateos, *Chmura obliczeniowa...*, dz. cyt., s. 2.

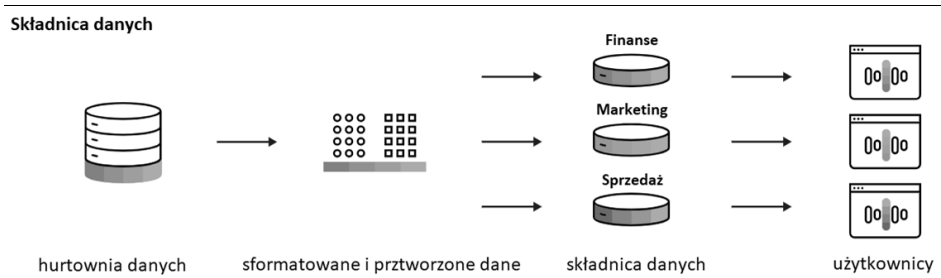
Składnica danych a hurtownia danych – tutaj należy zauważyć, że różnica polega na tym, iż tematyczna składnica danych jest to podsekcja hurtowni danych, która została podzielona na partycje z wyspecjalizowaniem dla danego działu lub dziedziny, takiej jak na przykład produkcja, sprzedaż, finanse czy marketing. Niektóre składnice danych są tworzone w niezależnych celach operacyjnych. Gdy hurtownia danych służy jako centralne archiwum, składnica obsługuje istotne dane dla wybranej grupy użytkowników¹⁰. Ma to na celu uprosz-

⁸ M. Toroman, *Chmura Azure. Praktyczne wprowadzenie...*, dz. cyt., s. 6.

⁹ J. Rosenberg, A. Mateos, *Chmura obliczeniowa...*, dz. cyt., s. 2.

¹⁰ Tamże, s. 6.

czenie dostępu do danych, przyspieszenie analizy i daje znacznie większą kontrolę nad własnymi danymi. W hurtowni danych często wdrażanych jest wiele składnic (tematycznych pojedynczych baz danych).



Rys. 3. Składnica danych

Źródło: J. Rosenberg, A. Mateos, *Chmura obliczeniowa...*, dz. cyt., s. 6.

W tym miejscu należy jeszcze dodać, jaka jest szczegółowa różnica pomiędzy hurtownią danych a samą bazą danych. Mimo że te dwa systemy pracują z danymi, pełnią kompletnie odmienne role. W największym skrócie mówiąc, bazy danych służą do szybkiego, bieżącego przetwarzania niewielkich wycinków danych, natomiast ideą hurtowni jest prowadzenie kompleksowej analityki, której podstawa leży w ogromnych zbiorach informacji. Porównując te dwa zbiory danych należy również znać i rozróżniać dwa funkcjonujące nierozdzielnie zagadnienia: OLTP oraz OLAP¹¹.

Jeśli mamy na myśli bazy danych, to mówimy o OLTP (ang. *on-line transactional processing*) – rodzaju przetwarzania danych, polegającym na jednoczesnym wykonywaniu wielu transakcji. Z takim można się spotkać na przykład w bankowości internetowej czy serwisach rezerwacji online. W bazach danych są zapisane informacje na temat konkretnego elementu biznesowego – na przykład na temat stanu magazynowego produktów w sklepie internetowym. W bazach danych przechowywane są tylko najnowsze informacje, z uwagi na to, że jej wydajność znacznie by spadła, gdyby system miał utrzymywać ogromne ilości historycznych danych. Środowisko OLTP doskonale przetwarza duże liczby prostych zapytań, ale nie potrafi wykonywać głębokiej analityki, jest jednak bardzo dobrym źródłem informacji, które potrafi przetworzyć hurtownia danych¹².

Jeżeli jednak mamy na myśli hurtownię danych, to musimy tutaj powiązać ją z systemem analitycznym zwanym OLAP (ang. *on-line analytical processing*). Jest to środowisko, które (inaczej niż OLTP) jest dedykowane obsłudze

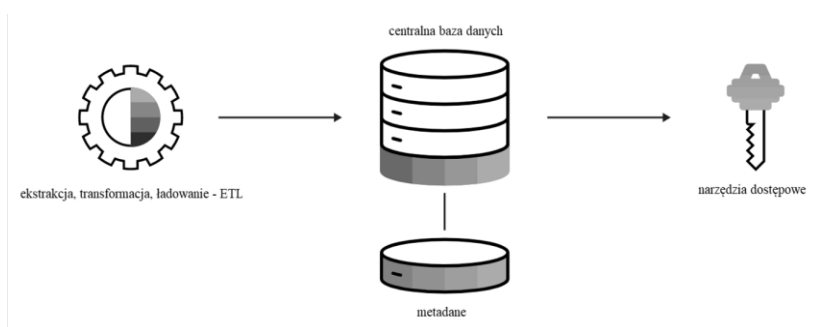
¹¹ A. Pelikant, *Hurtownie danych. Od przetwarzania analitycznego do raportowania*, wydanie II, Helion, Gliwice 2021.

¹² Tamże, s. 2.

znacznie bardziej rozbudowanych zapytań opartej na dużej ilości danych. Prędkość realizacji zadań nie ma tutaj większego znaczenia, dlatego analiza danych prowadzona w hurtowni zajmuje od kilku sekund do nawet kilku godzin¹³. W OLAP gromadzone są dane z różnych źródeł i przetwarzane podczas procesu ELT. Dane historyczne są przechowywane w celu prowadzenia kompleksowej analizy i raportowania. Należy pamiętać, że dane, które trafią do hurtowni nie mogą być modyfikowane – dostęp możliwy tylko do wglądu.

Składniki i budowa typowej hurtowni danych

Typowa hurtownia danych składa się z czterech elementów: centralnej bazy danych, narzędzi ETL (od angielskich nazw procesów: *extract, transform, load* – ekstrakcja, transformacja, ładowanie), metadanych i narzędzi dostępowych¹⁴.



Rys. 4. Typowe komponenty hurtowni danych

Źródło: M. Wilkins, *Amazon Web Services...*, dz. cyt., s. 7.

- **Centralna baza danych:** element ten stanowi podstawę hurtowni danych. Tradycyjnie są to z zasady standardowe (relacyjne) bazy danych, które działają lokalnie lub (coraz częściej spotykane) w chmurze. Z uwagi jednak na przetwarzane ogromne porcje informacji (*big data*), a jednocześnie potrzebę wydajności w czasie rzeczywistym przy założeniu obniżenia kosztów pamięci podręcznej RAM, bazy danych *in-memory* (o pojęciu wspomniano powyżej) szybko zyskują popularność.

- **Integracja danych:** dane zostają pobrane z systemów źródłowych i podane modyfikacji, aby dostosować je do szybkiego wykorzystania analityczne-

¹³ Tamże, s. 3.

¹⁴ M. Wilkins, *Amazon Web Services. Podstawy korzystania z chmury AWS*, Helion, Gliwice 2020.

go, używając różnych sposobów do integracji, takich jak ETL (ekstrakcja, transformacja, ładowanie) i ELT (różnica zostanie wyjaśniona w późniejszej części artykułu), a także powielanie danych (w czasie rzeczywistym), przetwarzania wsadowe, transformacje danych oraz inne usługi, które zapewnią jakość i rozszerzanie zbiorów danych.

- **Metadane:** są to dane dotyczące użytkownika. Określają źródło, wykorzystanie, wartości i inne funkcje zbiorów w magazynie danych. Należy wiedzieć, że występują metadane biznesowe – dodają kontekst do danych oraz metadane techniczne – opisują sposób, w jaki mamy możliwość dostępu do danych – wliczając w to miejsce ich przechowywania, a także jaką posiadają strukturę.

- **Narzędzia dostępu do hurtowni:** dzięki temu umożliwiona jest interakcja użytkownika z danymi w hurtowniach. Przykładami narzędzi dostępu mogą być: narzędzia zapytań i raportowania, narzędzia do projektowania aplikacji, narzędzia do eksploracji danych oraz narzędzia OLAP¹⁵.

Powyżej wspomniano o różnicy pomiędzy stosowanymi narzędziami ETL a ELT. Zasadniczą sprawą jest więc to, że w hurtowniach danych stosowane są narzędzia typu ETL (*Extract, Transform, Load*), natomiast w jeziorach danych – narzędzia typu ELT (*Extract, Load, Transform*). Proces ETL zakłada, że w pierwszej kolejności pozyskiwane są dane pochodzące ze źródła (*extract*), następnie zostają one przetwarzane (*transform*), a w końcowej fazie zapisane do właściwej hurtowni (*load*)¹⁶. W procesie ELT różnicą jest to, że dane na końcu poddaje się przetwarzaniu, a poprzedzające dwie fazy to ekstrakcja i ładowanie¹⁷. Etap transformacji odbywa się niezależnie od jeziora danych, przy użyciu zewnętrznej aplikacji, która służy do złożonej analizy danych.

Hurtownie danych działają w warstwach zgodnych z przepływem danych biznesowych: warstwa danych, warstwa semantyki, warstwa analityczna.

Warstwa danych: na tym poziomie dane są wybierane ze źródeł, a w dalszej fazie przekształcane i wczytywane do kolejnej (dolnej) warstwy przy użyciu narzędzi ETL. Dolna warstwa składa się z serwera bazy danych, hurtowni danych i jezior danych, gdzie tworzone są metadane. Narzędzia integracji danych (np. wirtualizacja danych) służą do płynnego łączenia i agregowania danych¹⁸.

Warstwa semantyki: serwery przetwarzania analitycznego (OLAP) i przetwarzania transakcyjnego online (OLTP) restrukturyzują dane mając za zadanie wykonać szybkie oraz złożone zapytania i analizy.

Warstwa analityczna: jest górną warstwą klienta z kontem front-end. W skład wchodzi narzędzia, które umożliwiają dostęp do hurtowni danych, dają

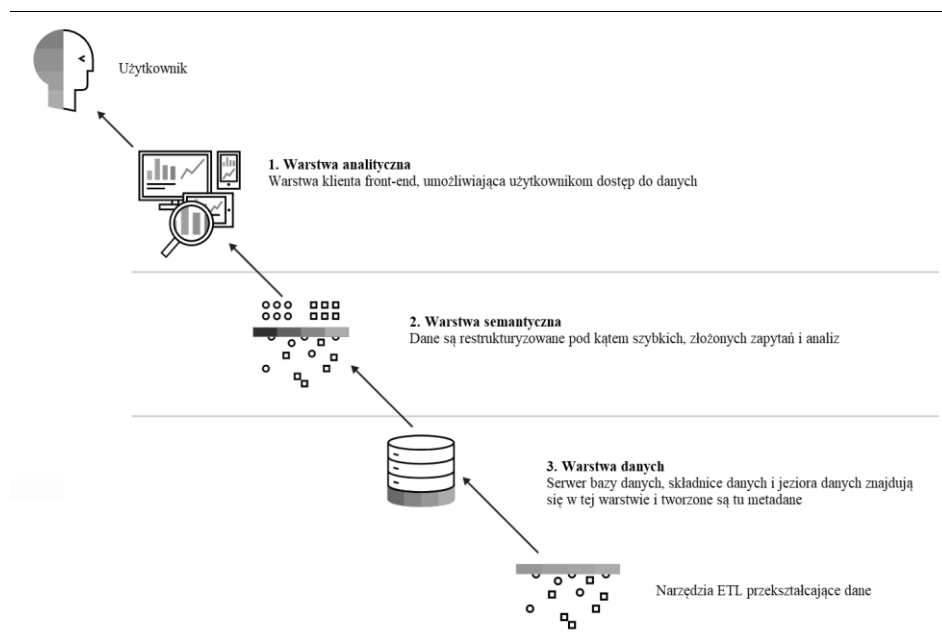
¹⁵ <https://www.sap.com/> (dostęp: 22.02.2023 r.).

¹⁶ <https://www.sap.com/> (dostęp: 27.02.2023 r.).

¹⁷ <https://fotc.com/> (dostęp: 11.02.2023 r.).

¹⁸ J. Rosenberg, A. Mateos, *Chmura obliczeniowa...*, dz. cyt.

użytkownikom możliwość interakcji z danymi, a także funkcjonalność tworzenia pulpitów i raportów wraz z opcją monitorowania wskaźników KPI, eksploracją i analizą danych, a także samym tworzeniem nowych aplikacji¹⁹.



Rys. 5. Diagram architektury hurtowni danych

Źródło: J. Rosenberg, A. Mateos, *Chmura obliczeniowa...*, dz. cyt., s. 6.

Hurtownia danych w chmurze

Głównym zadaniem, jakie spełniają platformy danych w chmurze to przede wszystkim ułatwienie korzystania z danych i dedykowanych gotowych narzędzi. Zachowując jednocześnie zabezpieczenie danych, ale także umożliwiając zarządzanie nimi użytkownik ma możliwość i dostęp do zintegrowanego widoku. Bardzo szybki przyrost ilości oraz używanie różnych typów danych to dziś standardowe operacje, jakie zachodzą podczas analizy mającej na celu podejmowanie właściwych decyzji biznesowych. Hurtownie „chmurowe” zapewniają elastyczność oraz oferują dodatkowe opcje, które ułatwiają taką analizę organizacjom o różnym charakterze i specyfice działalności²⁰. Począwszy od dużych, między-

¹⁹ Tamże, s. 11.

²⁰ A. Pelikant, *Hurtownie danych. Od przetwarzania...*, dz. cyt., s. 4.

narodowych korporacji, a skończywszy na małych przedsiębiorstwach – każde z nich sprawdza lub używa hurtowni danych w chmurze, daje im to możliwość niezawodnego i przystępnego dla użytkownika zarządzania danymi lub samej kontroli obecnych trendów²¹.

Oczywiście hurtownię danych można również zaprojektować, a potem wdrożyć na własnej infrastrukturze²². Jest to jednak przedsięwzięcie wymagające po pierwsze specjalistycznej wiedzy informatycznej, a po drugie to rozwiązanie jest niewątpliwie dużo droższe²³. Posługując się zestawem gotowych, skalowalnych i zarządzanych usług, które oferuje dostawca chmury publicznej, można to zrobić dużo taniej, łatwiej i szybciej.

Hurtownię w chmurze można zbudować jako rodzaj aplikacji, która jest niemal bezobsługową, która samodzielnie wyodrębnia dane z różnych źródeł, ujednocila je i odpowiednio przygotowuje do analiz. Ponadto takie procesy jak: rezerwowanie zasobów, skalowanie, aktualizacja czy optymalizacja wykonywane są w sposób automatyczny²⁴. Użytkownik nie musi tutaj ingerować w system, dzięki czemu ten czas, może być poświęcony na prowadzenie właściwych analiz. Hurtownię danych w chmurze mogą być też rozbudowane o usługi uczenia maszynowego (ang. *machine learning*), by umożliwić prowadzenie predykcji czy generowanie symulacji scenariuszy biznesowych. Koszty, jakie są związane z posiadaniem hurtowni danych w chmurze, są bardzo elastyczne, nierzadko też dużo niższe niż w przypadku posiadania lokalnego systemu.

Głównymi i najbardziej popularnymi dostawcami usług w chmurze są: Azure, Google Cloud czy AWS (Amazon Web Services), jednak szczegółowy opis oferowanych usług to materiał na odrębne publikacje, z uwagi na rozległość i złożoność materiału, jaki należałoby przedstawić.

Posiadanie hurtowni danych – korzyści dla przedsiębiorstwa

Korzyści dla przedsiębiorstwa, jakie płyną z posiadania hurtowni danych to między innymi:

- **wgląd w dane.** Możliwość spojrzenia na dane z szerszej perspektywy, które pozwala wyszukać zależności pomiędzy obszarami oraz wytypować trendy dzięki analizie z danych historycznych. Pozwala to też na analizowanie wycinków danych oraz tworzyć szczegółowe raporty;

²¹ Itzik Ben-Gan, A. Machanic, D. Sarka, K. Farlee, *Zapytania w języku T-SQL w Microsoft SQL Server 2014 i SQL Server 2012*, Helion, Gliwice 2017.

²² C. Tanimura, *Analiza danych z wykorzystaniem SQL-a. Zaawansowane techniki przekształcania danych we wnioski*, Helion, Gliwice 2022.

²³ A. Pelikant, *MS SQL Server. Zaawansowane metody...*, dz. cyt., s. 3.

²⁴ J. Rosenberg, A. Mateos, *Chmura obliczeniowa. Rozwiązania...*, dz. cyt., s. 7.

- **możliwość prognozowania.** Hurtownię danych można też używać jako element systemu wspomagania decyzji DSS (ang. *decision support system*), w którym istnieje możliwość prowadzenia symulacji scenariuszy biznesowych używając modeli sztucznej inteligencji (uczenia maszynowego);

- **centralizacja danych.** Hurtownia daje możliwość zebrania w jednym miejscu wszystkich danych z przedsiębiorstwa, które do tej pory były rozproszone pomiędzy wieloma bazami danych, systemami na zewnątrz czy na przykład arkuszami kalkulacyjnymi;

- **archiwizacja.** W hurtowni danych przechowane są zarówno najnowsze, jak też historyczne dane. Wydajny system analizy umożliwia przechowywanie ogromnych porcji informacji;

- **wydajność pracy.** Pracownicy organizacji, która posiada hurtownię danych mają wygodny, praktycznie natychmiastowy dostęp do aktualnych informacji ze swojego obszaru, ale też jednocześnie zapewniony dostęp do danych z innych działów. Skrócenie tej ścieżki dotarcia do odpowiednich informacji wraz z możliwością wizualnej prezentacji danych pozwala dużo łatwiej wykonywać codzienne zadania oraz wyszukiwać możliwości optymalizacji procesów²⁵.

Wspomniano w opracowaniu o możliwości przeniesienia danych (utworzenia hurtowni danych) do chmury²⁶. Należy więc wiedzieć, że daje to wiele dodatkowych funkcjonalności, które dla sprawności przedsiębiorstwa stają się ogromną zaletą. Hurtownie oparte na chmurze zyskują na popularności, oferując coraz więcej dedykowanych dla użytkownika ułatwień, które górują nad tradycyjnymi, lokalnymi wersjami²⁷. Poniżej ukazano kilka największych zalet hurtowni w chmurze:

- **szybkie wdrożenie:** hurtownie danych w chmurze dają niemal nieograniczoną moc obliczeniową, a także pamięć danych za pomocą zaledwie kilku operacji – stwarza to szybką możliwość zbudowania własnej hurtowni danych oraz dostępu do niej z dowolnego miejsca i w ciągu kilku minut;

- **niski całkowity koszt posiadania (TCO):** sposoby ustalania cen hurtowni danych jako usługi są skonfigurowane, więc użytkownik płaci tylko za te zasoby, które są mu potrzebne. Nie trzeba prognozować długoterminowych potrzeb ani płacić za większe obliczenia w ciągu roku, niż jest to konieczne. Dzięki usłudze „chmurowej” można również uniknąć kosztów początkowych, takich jak: sprzęt, serwerownie i wykwalifikowany personel zajmujący się obsługą i konserwacją;

²⁵ A. Pelikant, *MS SQL Server. Zaawansowane metody...*, dz. cyt., s. 11.

²⁶ G. Harrison, *NoSQL, NewSQL i BigData. Bazy danych następnej generacji*, Helion, Gliwice 2019.

²⁷ A. Pelikant, *Hurtownie danych. Od przetwarzania...*, dz. cyt., s. 5.

- **elastyczność:** Dzięki hurtowni w chmurze przedsiębiorca w razie potrzeby ma możliwość dynamicznie „powiększać się” lub „zmniejszać” – mając oczywiście na myśli zasoby chmury i koniecznych narzędzi. Usługodawcy zapewniają zwirtualizowane, wysoce rozproszone środowisko, w którym można zarządzać ogromnymi ilościami danych, które to mogą się rozrastać lub zmniejszać w razie potrzeby;

- **bezpieczeństwo i odtwarzanie po awarii:** hurtownie danych w chmurze gwarantują znacznie większe bezpieczeństwo i szyfrowanie danych niż lokalne hurtownie danych. Dane te są także automatycznie duplikowane i tworzone są kopie zapasowe, co zminimalizuje ryzyko utraty danych;

- **technologie w czasie rzeczywistym:** hurtownie danych w chmurze oparte na technologii *in-memory* zapewniają szybkie przetwarzanie danych, a dostarczając dane w czasie rzeczywistym pozwalają na uzyskanie natychmiastowej prezentacji żądanego wyniku²⁸;

- **nowe technologie:** hurtownie danych w chmurze pozwalają na łatwą integrację nowych technologii, takich jak uczenie maszynowe, które mogą służyć pomocą użytkownikom biznesowym i wspierać podejmowanie decyzji w formie zalecanych pytań, które należy zadać jako przykład;

- **zwiększenie możliwości użytkowników biznesowych:** hurtownie danych w chmurze zapewniają pracownikom jednakowe możliwości dzięki wglądowi w dane z wielu źródeł oraz złożonemu zestawowi narzędzi i funkcji, które umożliwiają łatwe wykonywanie zadań związanych z analizą danych. Mogą łączyć nowe aplikacje i źródła danych bez pomocy specjalistów typowych dla branży IT²⁹.

Współczesne, nowoczesne hurtownie danych, a także (coraz częściej wybierane) chmurowe hurtownie danych będą kluczowym elementem każdej inicjatywy związanej z transformacją cyfrową dla organizacji nadrzędnych i ich jednostek biznesowych. Zdalne pulpity, generowanie KPI czy dedykowane alerty i specjalistyczne raporty wspierają kierownictwo w decyzjach, które mają kluczowy wpływ na działanie przedsiębiorstwa, a także ukierunkowują na ważne potrzeby klientów czy dostawców. Hurtownie danych zapewniają również szybkie, złożone eksploracje i analizy danych i nie zakłócają wydajności pozostałych systemów biznesowych. Posiadanie wyspecjalizowanych zbiorów danych oraz sposobu na ich profesjonalną prezentację wydaje się być już nieodzownym elementem w dobie transformacji cyfrowej, dedykowanej czwartej rewolucji przemysłowej – Industry 4.0.

²⁸ Tamże, s. 3.

²⁹ J. Rosenberg, A. Mateos, *Chmura obliczeniowa. Rozwiązania...*, dz. cyt., s. 9.

Bibliografia

- Harrison G., *NoSQL, NewSQL i BigData. Bazy danych następnej generacji*, Helion, Gliwice 2019.
- Itzik Ben-Gan, Machanic A., Sarka D., Farlee K., *Zapytania w języku T-SQL w Microsoft SQL Server 2014 i SQL Server 2012*, Helion, Gliwice 2017.
- Pelikant A., *Hurtownie danych. Od przetwarzania analitycznego do raportowania*, wyd. II, Helion, Gliwice 2021.
- Pelikant A., *MS SQL Server. Zaawansowane metody programowania*, wyd. II, Helion, Gliwice 2021.
- Rosenberg J., Mateos A., *Chmura obliczeniowa. Rozwiązania dla biznesu*, Helion, Gliwice 2011.
- Tanimura C., *Analiza danych z wykorzystaniem SQL-a. Zaawansowane techniki przekształcania danych we wnioski*, Helion, Gliwice 2022.
- Toroman M., *Chmura Azure. Praktyczne wprowadzenie dla administratora. Implementacja, monitorowanie i zarządzanie ważnymi usługami i komponentami IaaS/PaaS*, Helion, Gliwice 2020.
- Wilkins M., *Amazon Web Services. Podstawy korzystania z chmury AWS*, Helion, Gliwice 2020.

Netografia

- <https://aws.amazon.com/>.
- <https://azure.microsoft.com/>.
- <https://cloud.google.com/>.
- <https://fotc.com/>.
- <https://www.sap.com/>.