

## 4. ZASOBY TECHNOLOGICZNE

### ROZPOCZNIJ OD HISTORII

Tym razem nie będzie historyjki – tylko zdjęcie!

- Nie wystarczy posiadać wiele danych, firmy naciskane są, aby czynić inwestycje we właściwą infrastrukturę. Ilość dostępnych obecnie narzędzi i usług dotyczących dużych zestawów danych może być jednak przytłaczająca. Od czego więc zacząć?
- Jest to trudne zarówno dla małych, jak i dużych firm, ale ryzyko jest większe dla małych firm, ponieważ ich budżet jest bardziej ograniczony i mają mniejszy dostęp do specjalistycznej wiedzy.
- Większość firm wydaje od czterech do sześciu procent swoich budżetów na IT (<https://blog.techvera.com/company-it-spend>). Nie pozostawia to wiele miejsca na błędy, zwłaszcza dla małych firm.

### EKSPOZYCJA

#### 1. ROZUMIENIE ARCHITEKTURY DANYCH

- Architektura danych odnosi się do zarówno konceptu design, jak i fizycznej infrastruktury systemów, których używamy do zbierania, magazynowania, zarządzania i wykorzystywania danych.
- Czasami rozróżniamy architekturę danych i architekturę informacji. Mówiąc wprost, dane odnoszą się do surowych, niezorganizowanych faktów. Pomyśl o danych jak o pakietach wpisów zbiorczych gromadzonych i przechowywanych bez kontekstu. Gdy kontekst zostanie przypisany do danych poprzez połączenie dwóch lub więcej elementów w sposób nadający sens, stają się one informacją. Architektura danych to rozwój programów interpretujących i przechowujących dane. Architektura informacji odnosi się do rozwoju programów służących do wprowadzania, przechowywania i analizowania istotnych informacji. (Źródło: Lebanthal)<sup>1</sup>
- Architektura danych jest fundamentalna. Sprawdza przychodzące dane i określa, w jaki sposób będą one przechwytywane, przechowywane i integrowane z innymi platformami. Jedną z takich platform jest prawdopodobnie część architektury informacji, taka jak CRM, która wykorzystuje surowe dane klientów do tworzenia istotnych powiązań dotyczących sprzedaży i procesów sprzedaży. CRM to architektura informacji w tym przykładzie, ponieważ specjalizuje się w pobieraniu surowych danych i przekształcaniu ich w coś użytecznego.

#### 2. TECHNOLOGIE DANYCH

Narzędzia danych są złożone i niejednorodne. Dlatego należy opracować plan działania dotyczący sposobu przetwarzania i przenoszenia danych. Celem jest udostępnienie wszystkich bieżących i przyszłych danych, w dowolnej formie, dowolnej aplikacji w organizacji, aby zmniejszyć złożoność i zwiększyć łatwość dostępu.<sup>2</sup>

##### 2.1 NARZĘDZIA POZIOMU WEJŚCIOWEGO

Wszystkie firmy, bez względu na rozmiar, mogą korzystać na umiejętnym posługiwaniu się CRMS, analityką stron internetowych i społeczną, a także opartymi o strony internetowe usługami VOIP.

#### Wykorzystaj w pełni CRM

<sup>1</sup> <https://www.bmc.com/blogs/data-architecture-vs-information-architecture/>

<sup>2</sup> <https://www.sciencedirect.com/science/article/pii/S014829631630488X>

1. Uczyń oprogramowanie CRM łatwym w użyciu. Dowiedz się, czego naprawdę chce każdy pracownik lub dział i co CRM musi zrobić, i dostosuj go do tych potrzeb. Zapewnij szkolenie CRM wszystkim pracownikom.
2. Zintegruj jak najwięcej źródeł danych. Firmy nigdy nie powinny myśleć o CRM jako o samodzielnym systemie, połącz go z innymi systemami zorientowanymi na klienta lub informującymi klientów, a zwłaszcza z mediami społecznościowymi.
3. Bądź na bieżąco. Systemy CRM są tak dobre, jak przechowywane w nich dane. Zachowaj spójny proces aktualizacji danych klientów i zarządzania nimi.
4. Zaczynaj analizować! Zdefiniowanie rodzaju wniosków, które chcesz uzyskać, i cofnij się od oczekiwanych wyników do analiz, które zapewnią te wyniki (lub udowodnią, że jest inaczej).

### **Analityka internetowa i społecznościowa**

Google Analytics to bezpłatna usługa analizy sieci oferowana przez Google, która śledzi i raportuje ruch w witrynie. Zapewnia dogłębne spojrzenie na wydajność Twojej witryny i integruje się ze wszystkimi produktami marketingowymi Google: Google Ads, Search Console i Data Studio.

Chociaż „analityka internetowa” brzmi jak bardzo niewielki obszar twojej cyfrowej obecności, implikacje Google Analytics są w rzeczywistości ogromne. Analiza Twojej witryny jest najlepszym sposobem na uzyskanie całościowego obrazu skuteczności wszystkich prowadzonych kampanii promujących Twój produkt/usługi online.

### **VOIP**

Wykorzystaj analizę śledzenia połączeń, aby:

- monitorować i zarządzać przepływem połączeń
- zapewnić organom nadzoru możliwości przeprowadzenia natychmiastowej korekty i udzielania wskazówek w czasie rzeczywistym
- zidentyfikować personel i obszary umiejętności, które wymagają dodatkowego szkolenia
- rozpoznawać i promować osób osiągające najlepsze wyniki
- Analizy mowy zyskują przewagę nad transkrypcjami jako miernik wydajności ze względu zarówno na opóźnienie związane z transkrypcją (w porównaniu z nagraniami rozmów), jak i fakt, że wiele niuansów, takich jak odmiana (zmiana) słów, nie jest odpowiednio uwzględnionych.

### **Oś czasu narzędzi dotyczących dużych zbiorów danych**

To tylko niektóre z podstawowych narzędzi do analizy danych. Jeśli chodzi o bardziej zaawansowane narzędzia, stale poszerza się zakres i zastosowanie narzędzi cyfrowych do pracy z dużymi danymi.

## **2.2 NARZĘDZIA DOTYCZĄCE DUŻYCH ZBIORÓW DANYCH**

Rising, Kristensen, Tjerrild-Hansen, 2014<sup>3</sup>, zidentyfikowali trzy odrębne kategorie techniczne, które odgrywają kluczową rolę w tworzeniu dla przedsiębiorstw (w szczególności dla MŚP) tanich rozwiązań wykorzystywania danych:

- Skalowalna sieć serwerów w chmurze - pozwalająca na opracowanie operacji na małą skalę, którą można następnie rozszerzyć.
- Narzędzia do zarządzania, przetwarzania i przechowywania danych, takie jak Hadoop i MapReduce.
- Oprogramowanie analityczne, takie jak Tableau, SAS, SPSS itp.

---

<sup>3</sup> <https://web.stanford.edu/class/msande238/projects/2014/GainIT.pdf>

### **Innym sposobem spojrzenia na to jest użycie niektórych zasad metody 5V.**

Jeśli cechami danych są Objętość, Różnorodność i Prędkość, dysponujemy różnymi narzędziami, które pomagają przetwarzać dane w ten sposób.

- W przypadku adresowania OBJĘTOŚCI danych najlepsze jest przetwarzanie wsadowe, a Hadoop tworzy główną rodzinę narzędzi
- W celu uwzględnienia RÓŻNORODNOŚCI danych opracowano bazy danych NoSQL.
- Aby zaspokoić potrzebę przetwarzania danych w związku z ich PRĘDKOŚCIĄ, opracowano programy analityczne działające w czasie rzeczywistym, w szczególności Spark.

### **3. PRACA W CHMURZE**

Niezależnie od tego, czy używasz jej do uruchamiania aplikacji, które udostępniają zdjęcia milionom użytkowników mobilnych, czy też do wspierania krytycznych operacji biznesowych, platforma usług w chmurze zapewnia szybki dostęp do elastycznych i tanich zasobów IT. Dzięki przetwarzaniu w chmurze nie musisz dokonywać dużych wstępnych inwestycji w sprzęt i spędzać dużo czasu na ciężkim zarządzaniu tym sprzętem. Zamiast tego możesz zapewnić dokładnie taki rodzaj i rozmiar zasobów obliczeniowych, jakiego potrzebujesz, aby wspierać swój najnowszy pomysł lub obsługiwać dział IT. Możesz uzyskać dostęp do tak wielu zasobów, jak potrzebujesz, niemal natychmiast i płacić tylko za to, czego używasz.

Przetwarzanie w chmurze opiera się na puli zasobów: firmy obsługują dużą liczbę klientów korzystając z systemów wielodostępowych, w których różne zasoby są dynamicznie przydzielane i zwalniane zgodnie z zapotrzebowaniem. Po stronie użytkownika nie jest możliwe ustalenie, gdzie znajduje się zasób.

#### **3.1 Korzyści z pracy w chmurze**

- **Technologia szerokiego zasięgu uwalnia zasoby.** Usługi w chmurze ułatwiają i przyspieszają dostęp do szerokiej gamy technologii, takich jak obliczanie, pamięć masowa, bazy danych, analityka, uczenie maszynowe i wielu innych usług według potrzeb. Pomagają firmom skoncentrować zasoby IT na tworzeniu aplikacji i wykorzystywaniu doświadczeń klientów, a nie na zarządzaniu infrastrukturą i centrami danych.
- **Ekonomiczne:** Korzystanie z chmury pozwala zamienić wydatki kapitałowe (centra danych, serwery fizyczne itp.) na wydatki zmienne; płacisz tylko za usługę lub miejsce, z którego korzystasz.
- **Dostępność:** Chmura umożliwia dostęp do szerokiej gamy platform i urządzeń, dzięki czemu zasoby są bardziej dostępne, a także bardziej niezawodne. W przypadku awarii sieci biurowej kopie zapasowe danych są zapisywane w chmurze i są nadal dostępne, np. na tablecie.
- **Skalowalność:** zdolność chmury do skalowania w górę lub w dół oznacza, że Twoja firma nie musi gromadzić danych ani mocy obliczeniowej na wypadek rzadkich sytuacji, gdy popyt gwałtownie rośnie. Skalowalność na żądanie jest często tłumaczona jako elastyczność.
- **Bezpieczeństwo i zarządzanie ryzykiem:** Opracowanie strategii w celu przewyższenia cyberataków, awarii zasilania lub awarii sprzętu jest trudne i kosztowne. Przechowywanie danych w infrastrukturze chmurowej pozwoli Twojej organizacji szybciej wyjść z ewentualnych katastrof.

#### **Wyzwania związane z pracą w chmurze**

- **Migracja.** Migracja dużych zbiorów danych do chmury stwarza różne przeszkody, które wymagają skoordynowanego wysiłku ze strony menedżerów i liderów IT.
- **Mniej kontroli nad danymi.** Zdalne przechowywanie danych i korzystanie z zewnętrznych protokołów bezpieczeństwa i zgodności może być dużą zmianą organizacyjną i może powodować dyskomfort. Poznaj i oceń protokoły dostawcy, abys wiedział, jakie są twoje role i obowiązki.

- **Zależność sieci i czas oczekiwania.** Dostępność danych zależy w dużym stopniu od połączenia sieciowego z Internetem. W środowisku chmurowym możemy trafić również na problem opóźnień, biorąc pod uwagę ilość danych przesyłanych, analizowanych i przetwarzanych w dowolnym momencie.

### 3.2 KIM SĄ GŁÓWNI DOSTAWCY CHMUR?

Istnieje wielu dostawców usług w chmurze dla dużych zbiorów danych, ale my koncentrujemy się na wielkiej trójce. Amazon dominuje. W raporcie z 2018 r. Synergy Research Group zauważyła, że wydatki na usługi w chmurze wzrosły o zaskakująco 51 procent w porównaniu z kwartałem ubiegłego roku, zauważając: „Udział Amazon Web Services w rynku światowym utrzymuje się na poziomie około 33% przez dwanaście kwartałów, mimo że rynek zwiększył się prawie trzykrotnie”.<sup>4</sup>

Przegląd Datamation stwierdza:

- **Amazon Web Services** - dzięki ogromnemu zestawowi narzędzi, który wciąż rośnie, możliwości Amazon są niezrównane. Jednak jego struktura kosztów może być myląca, a szczególny nacisk na chmurę publiczną zamiast chmury hybrydowej lub chmury prywatnej oznacza, że współpraca z twoim centrum danych nie jest najwyższym priorytetem AWS.
- **Microsoft Azure** - bliski konkurent AWS z wyjątkowo zdolną infrastrukturą chmurową. Jeśli jesteś klientem korporacyjnym, platforma Azure mówi w Twoim języku - niewiele firm ma takie zaplecze dla przedsiębiorstw (i obsługę systemu Windows) jak Microsoft. Azure wie, że nadal zarządzasz centrum danych i ciężko pracuje, aby współpracować z takimi centrami; chmura hybrydowa to prawdziwa siła.
- **Google Cloud** – jako dobrze dofinansowany słabszy konkurent, Google wszedł później na rynek chmury i nie koncentruje się na przedsiębiorstwie, co mogło by pomóc przyciągnąć klientów korporacyjnych. Jego wiedza techniczna jest jednak dogłębna, a wiodące w branży narzędzia do głębokiego uczenia się i sztucznej inteligencji, uczenia maszynowego i analizy danych to znaczące zalety.

## 4. WYBÓR WŁAŚCIWEGO ROZWIĄZANIA TECH

Zalecany jest czteroetapowy model, ale zostanie on wzmocniony poprzez zaangażowanie odpowiednich osób.

KTO powinien być częścią komitetu i podjąć decyzję? Idealny format komitetu obejmowałby:

- Osobę bardzo obeznaną z wymaganiami funkcjonalnymi i misją Twojej organizacji (być może nauczyciel, administrator, instruktor lub kierownik oddziału)
- Osobę zapoznaną z obecnymi możliwościami systemu (być może pracownik pomocy technicznej)
- Osobę, która przeszła wcześniej proces wdrażania systemu, najlepiej w organizacji (być może doświadczony pracownik)

Należy pamiętać, że na większość projektów związanych z oprogramowaniem wpływają nie tylko „obiektywne” kryteria, ale także czynniki subiektywne, np.:

- Osobiste preferencje liderów projektów, wpływowych zewnętrznych doradców lub kierownictwa firmy;
- Dostawcy oprogramowania lub integratorzy systemów mający ugruntowaną pozycję w firmie;
- Ograniczenia czasowe i budżetowe w samym procesie selekcji

<sup>4</sup> <https://www.oryxalign.com/aws-azure-google-a-public-cloud-comparison-report/>

Poznaj koszty infrastruktury danych w odniesieniu do kosztów usług w chmurze i przeprowadź szacunkowe symulacje.