

# Rozwiązania Obiektowe

HITACHI  
Inspire the Next



## Większy i szybszy wzrost ilości danych niestrukturalnych

Zbiór ten już teraz często stanowi nawet 80% wszystkich danych w organizacji.

Gartner szacuje, że aż 28% danych korporacyjnych przechowywanych jest wyłącznie na laptopach, smartfonach i tabletach. Z kolei Harvard Business Review podkreśla, że tylko 1% danych niestrukturalnych podlega analizie biznesowej, podczas gdy w przypadku danych strukturalnych na potrzeby analiz i podejmowania decyzji wykorzystywanych ich jest nawet 50%. Na szczęście technologie i metody dostępu do danych również zmieniają się w czasie. A w rozwiązywaniu powyższych problemów i wyzwań związanych z przechowaniem danych niestrukturalnych, coraz częściej pomaga cały ekosystem rozwiązań obiektowych, z obiektowymi magazynami danych na czele.

### Macierze obiektowe

To urządzenia oferujące takie funkcjonalności, które do tej pory były niedostępne lub trudne do zrealizowania w oparciu o tradycyjne, istniejące od dawna technologie macierzowe. W przypadku rozwiązań obiektowych dostęp do danych realizowany jest również za pomocą protokołów sieciowych, tak jak w przypadku urządzeń plikowych. Macierz obiektowa korzysta jednak najczęściej z nowoczesnych bramek typu HTTP, S3, SWIFT lub zwyczajnie z RESTowego API. Dlatego, oprócz samych plików i podstawowych metadanych systemowych, przekazywane są tutaj również metadane własne (custom). A dzięki temu, istnieje możliwość opisanie (wewnątrz tego samego urządzenia obiektowego, a nie w zewnętrznej bazie danych) np. co się dzieje z plikiem w trakcie jego przechowywania lub co się w nim znajduje. Obiekty przechowywane są w przestrzeniach nazw (lub inaczej w bucketach) i dostęp do nich może mieć wielu użytkowników jednocześnie (tak, jak w przypadku macierzy plikowej). Zmienia się jednak sposób komunikacji aplikacji z plikiem. Interfejs REST API pozwala na wykorzystanie takich funkcji, jak: GET, PUT, POST i HEAD. A programiście wystarczy znajomość adresu URI obiektu.



**„Według Forbesa 90%  
wszystkich danych na  
świecie powstała  
w ostatnich 2 latach.”**

W macierzach obiektowych można skupić się na samej treści, w oderwaniu nie tylko od systemu operacyjnego, ale również od aplikacji, która tę treść produkuje. Łatwo w związku z tym udostępnić ten sam obiekt różnym aplikacjom, z których każda będzie korzystała i dopisywała do jego metadanych informacje istotne tylko dla niej (bez duplikowania pliku dla różnych systemów czy środowisk).

## Przykład:

Na zdjęciu z miejsca wypadku w metadanych takiego zdjęcia możemy chcieć przechowywać informacje o numerze sprawy, aucie, kierowcy, czy panującej tego dnia pogodzie, aby w przyszłości móc analizować podobne zdarzenia. I nawet jeżeli taka analiza zdjęcia miałaby być realizowana za pomocą technik rozpoznawania obrazów i uczenia maszynowego, to część kluczowych informacji otrzymanych w wyniku takiego przetwarzania można zawsze zapisać w metadanych własnych obiektu, aby unikać ponownego przetwarzania i analizy przy każdej interakcji użytkownika końcowego.

Metadane własne można również wykorzystać na potrzeby klasyfikacji (np. plik zawiera lub nie zawiera danych osobowych) oraz wyszukiwania informacji. Macierze obiektowe posiadają wewnętrzne silniki indeksowania i przeszukiwania. A ponieważ wbudowane wyszukiwarki są wyposażone w interfejs API, dlatego silniki te można wykorzystać również wewnątrz swojej aplikacji.

**Hitachi w swojej ofercie posiada macierz obiektową Hitachi Content Platform (HCP) oraz trzy dodatkowe aplikacje, które wzbogacają jej funkcjonalność i dostarczają dodatkowych usług.**

## Hitachi Content Platform (HCP)

To również podstawowy element oferty rozwiązań obiektowych Hitachi. To tutaj przechowywane są i zabezpieczane dane. HCP może stanowić centralny obiektowy magazyn danych w organizacji. Obiektowy oznacza, że operujemy na poziomie obiektów (a nie plików), na które składają się: dane, metadane systemowe i metadane własne oraz polityki (np. retencja, replikacja, wersjonowanie, tiering). Metadane własne pozwalają na opisanie danych (treści plików) i ich efektywniejsze indeksowanie i wyszukiwanie. Hitachi Content Platform charakteryzuje się i jest znany na rynku jako rozwiązanie, które między innymi wykorzystuje:

- wbudowaną technologię WORM dla zapewnienia niezmienności przechowywanych w nim obiektów oraz
- politykę retencji danych i metadanych dla zapewnienia ich nieusuwalności.

”

**HCP może stanowić centralny obiektowy magazyn danych w organizacji. To tutaj przechowywane są i zabezpieczane dane.”**

**Hitachi Content Platform Anywhere (HCP AW)** To aplikacja, która realizuje funkcjonalność współdzielenia plików pomiędzy użytkownikami (wewnątrz i na zewnątrz organizacji), synchronizacji plików pomiędzy urządzeniami użytkowników oraz wykonywania kopii zapasowych z urządzeń użytkowników. Wszystkie dane przechowywane są i zabezpieczone w magazynie danych HCP.

## Hitachi Content Platform Anywhere Edge (HCP AW Edge)

Jest ona jest bramą plikową (NAS gateway), która integruje się z magazynem danych HCP i udostępnia użytkownikom zasoby plikowe za pomocą protokołów CIFS i NFS, przy jednoczesnym, automatycznym zabezpieczaniu plików użytkownika w HCP. HCP AW Edge może pełnić rolę urządzenia dostępowego do HCP u biznesowego klienta wewnętrznego i/lub zewnętrznego w architekturze chmury dla danych. Może to być również

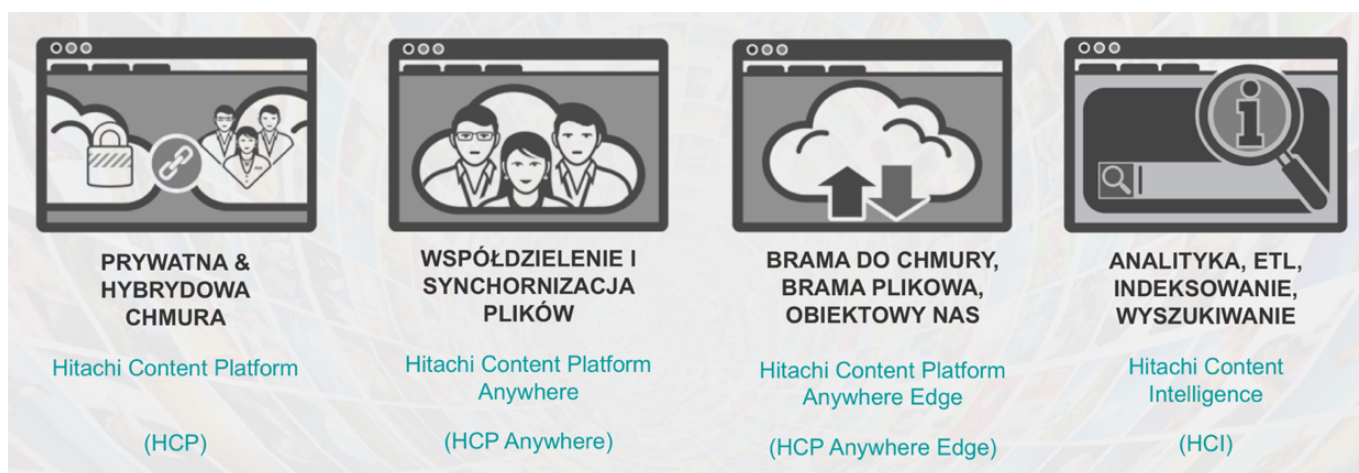
urządzenie, w którym będziemy utrzymywać pierwszą (najszybszą) warstwę storage, dla plików, które muszą być szybko dostępne dla użytkowników w przypadku integracji ze starszymi aplikacjami, korzystającymi nadal z protokołów dostępowych CIFS/NFS.

## Hitachi Content Intelligence (HCI)

Pozwala na eksplorację danych i metadanych przechowywanych we wszystkich wcześniej wspomnianych źródłach Hitachi oraz wielu innych dodatkowych źródłach (bazy danych, serwery plików, aplikacje i magazyny danych S3, HDFS, kolejki asynchroniczne obsługiwane przez aplikację Kafka), indeksowanie danych i metadanych oraz ich wyszukiwanie. HCI posiada wbudowany mechanizm ETL (Extract Transform Load), który umożliwia przetwarzanie danych w postaci procesów, zawierających operacje transformacji, wzbogacania, filtrowania, oznaczania danych i metadanych.

”

**Ważną funkcjonalnością Hitachi NAS, która pozwala na oszczędność przestrzeni dyskowej nawet do 90% jest wbudowana deduplikacja plików.”**



Oferta rozwiązań obiektowych Hitachi.

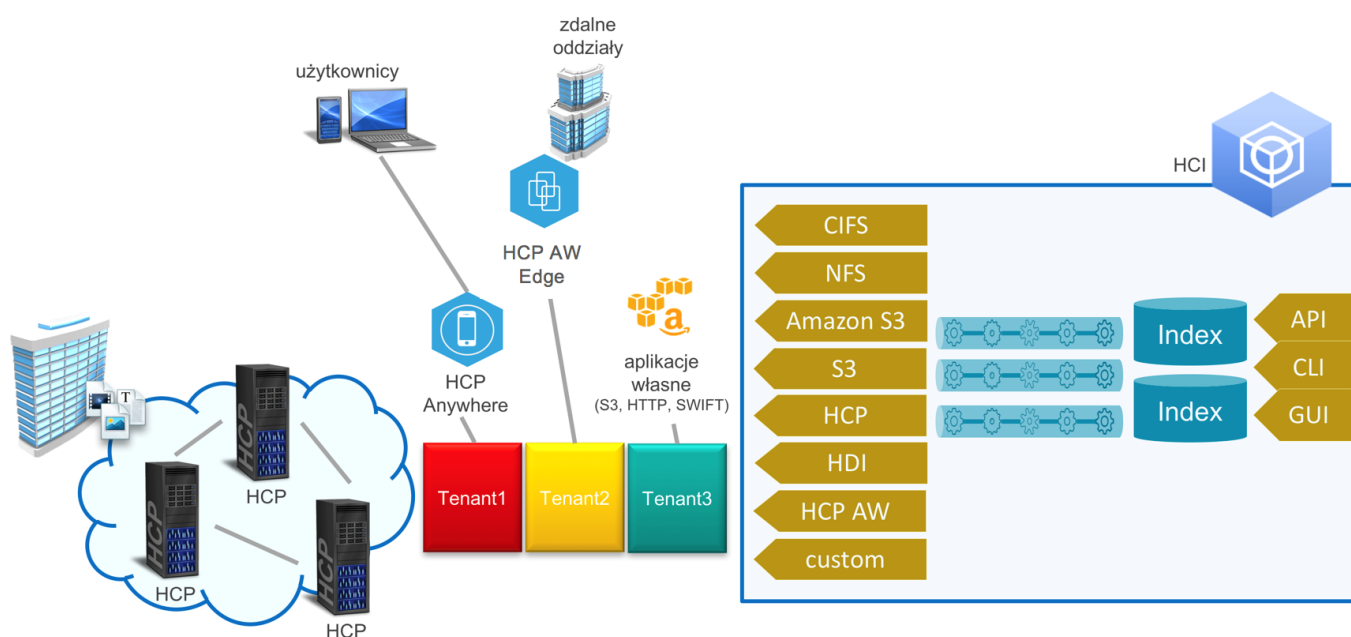
## Za pomocą rozwiązań obiektowych Hitachi, możliwe jest zaadresowanie wielu celów biznesowych.

Wśród tych tradycyjnych oraz takich, które dopiero kształtują się w nowej rzeczywistości biznesowej, najpopularniejszymi są:

- zapewnienie zgodności z regulacjami prawnymi,
- optymalizacja infrastruktury,
- zwiększenie bezpieczeństwa,
- osiągnięcie ekonomiki charakterystycznej dla chmury,
- transformacja infrastruktury IT w kierunku środowiska usługowego (chmura hybrydowa i prywatna),
- modernizacja strategii ochrony danych,
- uruchomienie środowiska cyfrowej pracy (Digital Workplace) i dostęp w dowolnym miejscu do dowolnej informacji potrzebnej w danej chwili,
- wdrażanie analityki również dla zbioru danych niestrukturalnych w poszukiwaniu nowych szans i okazji.

## Wykorzystanie rozwiązań obiektowych Hitachi

Przykład środowiska zbudowanego w oparciu o Hitachi Content Platform oraz pozostałe komponenty z rodziny rozwiązań obiektowych Hitachi:



Optymalizacja i tiering danych.

## Komponenty środowiska:

- **Hitachi Content Platform** – maksymalnie sześć urządzeń, zainstalowanych na przykład w sześciu ośrodkach z replikacją active-active pomiędzy ośrodkami, tworzy geograficznie rozproszoną chmurę HCP i centralny obiektowy magazyn danych z takimi funkcjonalnościami jak: WORM, retencja, niszczenie danych, bezpieczeństwo, wersjonowanie, tiering, szyfrowanie, własne metadane, multitenancy.
- **Hitachi Content Platform Anywhere Edge** – pełni rolę bramy plikowej i udostępnia dane z wykorzystaniem protokołów CIFS i NFS użytkownikom i aplikacjom, pracujących na przykład w oddziałach zdalnych. Poza tym stanowi pierwszą warstwę (tier1) dla danych, które są intensywnie wykorzystywane. Dane zapisywane w systemie plików HCP Anywhere Edge automatycznie są kopiowane i zabezpieczone w HCP (w obiektowym magazynie danych z wbudowanym mechanizmem WORM) co eliminuje konieczność backupu tych danych.
- **Hitachi Content Intelligence** – dostarcza mechanizmów wyszukiwania danych, ich indeksowania oraz transformacji. Dzięki
- **Hitachi Content Platform Anywhere** – udostępnia usługi współdzielenia danych (przechowywanych i zabezpieczonych w HCP), ich synchronizacji oraz backupu plików na stacjach roboczych użytkowników. Umożliwia tworzenie wspólnych katalogów roboczych (praca grupowa nad dokumentami) oraz pozwala na kontrolę tego, co i komu jest udostępniane.

wbudowanego mechanizmowi ETL możliwa jest analiza podłączonego źródła danych (HCP, HCPAW, HCPAW Edge, S3, CIFS, NFS, JDBC, Kafka, HDFS i innych własnych) np. pod kątem przechowywanych w nim treści (analiza zarówno danych jak i metadanych) oraz oznaczanie, kategoryzowanie, agregowanie i ich transformacja. Wynikiem procesu ETL danych i metadanych zdefiniowanego w HCI może być kolekcja indeksów i udostępnienie jej wyszukiwarkom: zarówno tej wewnętrznej dostarczanej wraz z HCI, jak i zewnętrznym silnikom wyszukiwania SOLR lub Elasticsearch.

**Integracja HCP** z innymi, własnymi aplikacjami poprzez REST API (http, S3, SWIFT) i wykorzystanie HCP jako bezpiecznego magazynu danych dla tych aplikacji, przy jednoczesnym spełnieniu wielu różnorodnych regulacji prawnych.

**CHCESZ WIEDZIEĆ WIĘCEJ?  
ZAPRASZAM DO KONTAKTU**

**S4E S.A.**

ul. Samuela Lindego 1 C,

30-148 Kraków

[www.s4e.pl](http://www.s4e.pl)

**HITACHI**  
Inspire the Next

