dr inż. Michał Malinowski

https://www.drmalinowski.edu.pl

Temat 7 .

Zagrożenia dla baz danych

PODSTAWY BAZ DANYCH

Zagadnienia:

- Budowa indeksów
- Bazy danych wbudowane
- Zagrożenia dla baz danych
- Korpusy uczące
- Przepisy prawa

Budowa indeksów

Indeks

to struktura danych używana w systemach zarządzania bazami danych (DBMS), która pozwala na szybsze wyszukiwanie i przetwarzanie danych w tabeli.

Indeks działa analogicznie do spisu treści w książce, umożliwiając dostęp do konkretnych rekordów bez konieczności przeszukiwania całej tabeli.

Indeksy podstawowe (primary index)

- tworzone automatycznie dla klucza głównego tabeli.

Indeksy unikalne (unique index)

- wymuszają unikalność wartości w indeksowanych kolumnach.

Indeksy złożone (composite index)

- obejmują więcej niż jedną kolumnę.

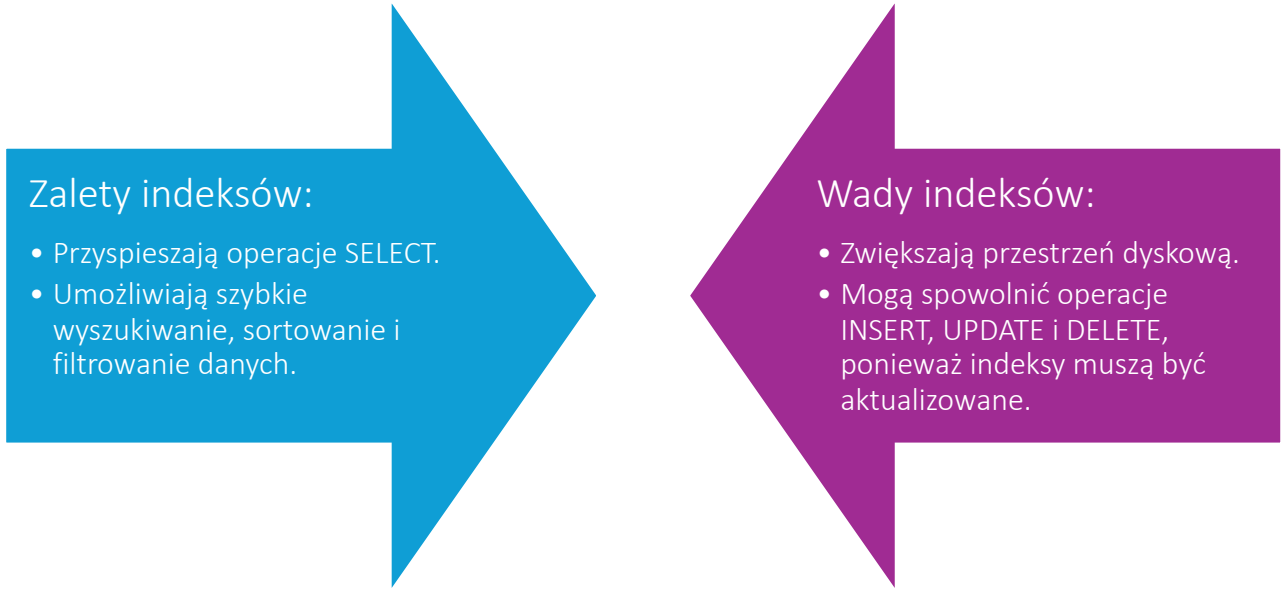
Indeksy pełnotekstowe (full-text index)

- używane do wyszukiwania w dużych zbiorach tekstu.

Budowa indeksów

Struktury danych indeksów

- **B-drzewa** – popularne dla indeksów hierarchicznych.
- **Hash table** – stosowane w przypadku szybkiego dostępu do danych za pomocą klucza.
- **Bitmapy** – wykorzystywane dla kolumn o niskiej liczbie unikalnych wartości (np. wartości logiczne).



Budowa indeksów

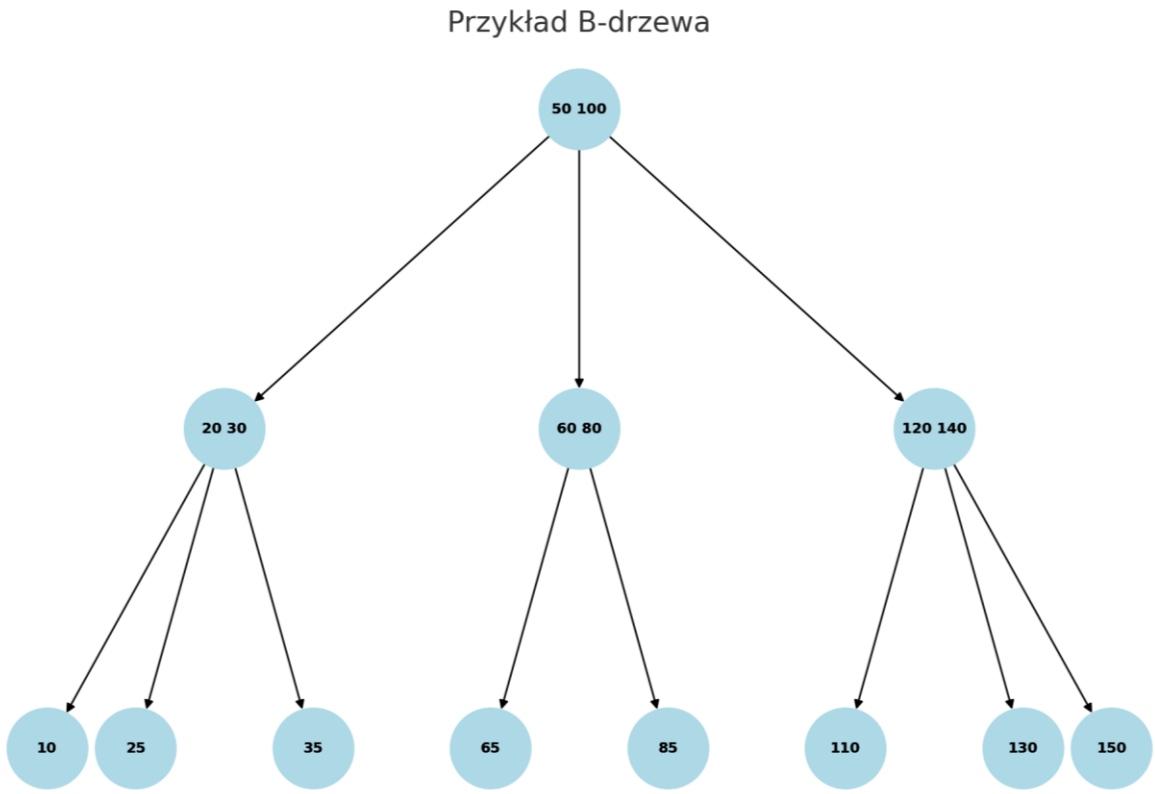
B-drzewo (ang. *B-tree*)

to zrównoważona struktura danych w postaci drzewa, która umożliwia efektywne przechowywanie, wyszukiwanie, wstawianie i usuwanie dużych zbiorów danych.

Jest to drzewo wyszukiwań wielokrotnego rozgałęzienia, w którym każdy węzeł może przechowywać wiele kluczy i mieć wiele potomków.

Algorytm poszukiwań

1. Rozpoczynamy wyszukiwanie w korzeniu drzewa.
2. Porównujemy szukaną wartość x z kluczami węzła:
   - Jeśli x znajduje się w bieżącym węźle, kończymy wyszukiwanie.
   - Jeśli x jest mniejsze od najmniejszego klucza, przechodzimy do pierwszego dziecka.
   - Jeśli x znajduje się pomiędzy dwoma kluczami k1 i k2, przechodzimy do - odpowiedniego poddrzewa (dziecka pomiędzy k1 a k2).
   - Jeśli x jest większe od największego klucza, przechodzimy do ostatniego dziecka.



Budowa indeksów

Złożoność czasowa przeszukiwania B-drzewa

$O(log_m n)$

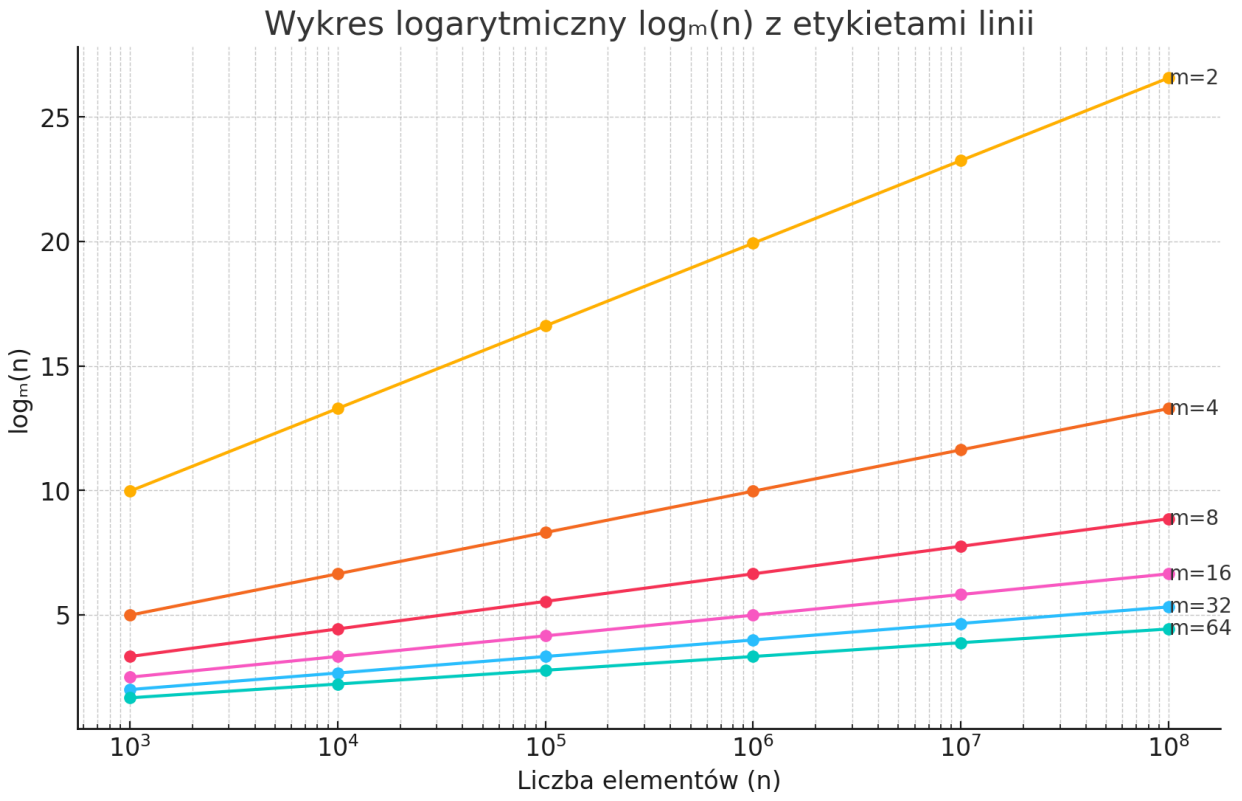
Gdzie:

n - to liczba elementów przechowywanych w drzewie.

m - to stopień drzewa (*branching factor*), czyli maksymalna liczba dzieci, jakie może mieć jeden węzeł.

Liczba kroków poszukiwań

Liczba Elementów (n)	m = 2	m = 4	m = 8	m = 16	m = 32	m = 64
1 000	10	5	4	3	2	2
10 000	14	7	5	4	3	3
100 000	17	9	6	5	4	3
1 000 000	20	10	7	5	4	4
10 000 000	24	12	8	6	5	4
100 000 000	27	14	9	7	6	5

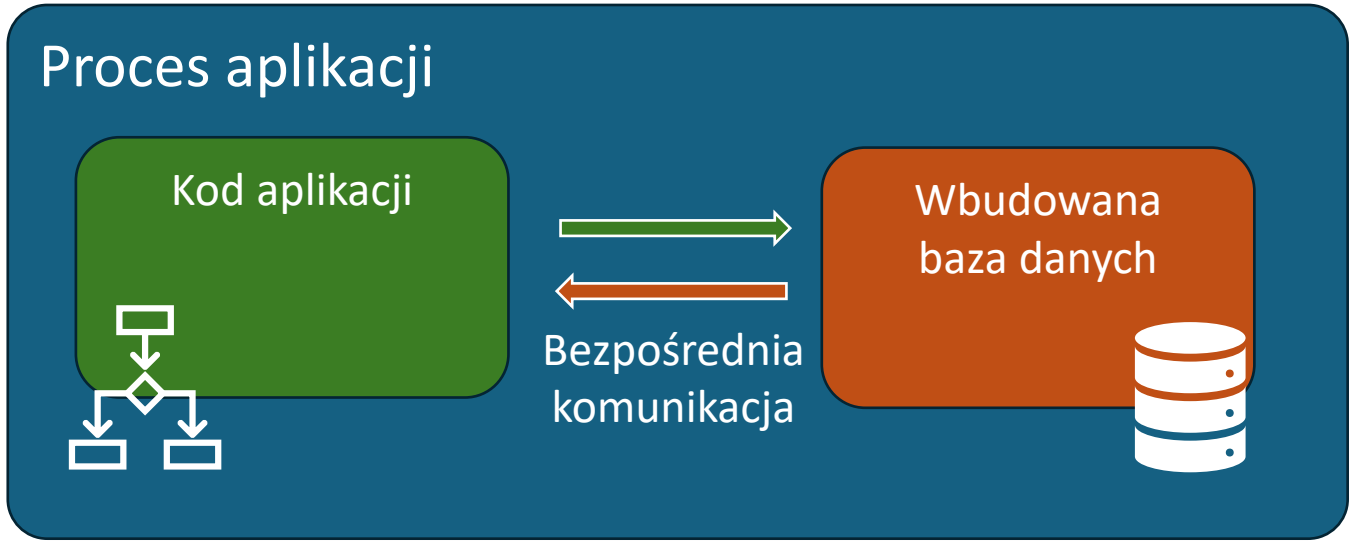


Bazy danych wbudowane

Bazy danych wbudowane (ang. *embedded databases*)

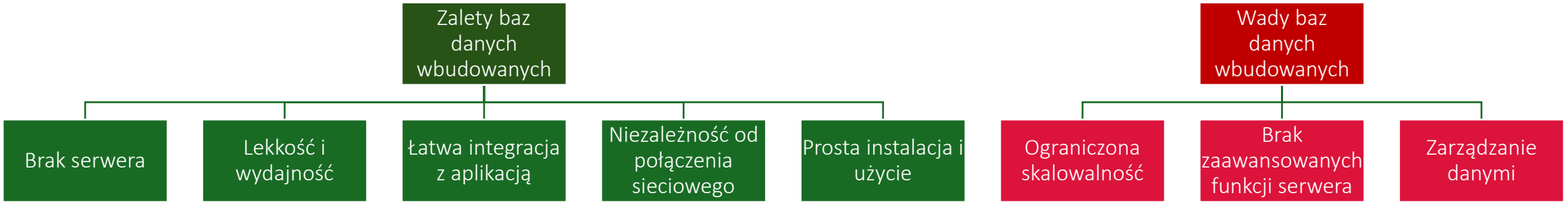
to systemy zarządzania bazami danych, które są zintegrowane bezpośrednio z aplikacją.

W przeciwieństwie do tradycyjnych systemów bazodanowych (klient-serwer), nie wymagają dedykowanego serwera do działania. Dane są przechowywane lokalnie, a baza funkcjonuje w obrębie przestrzeni procesowej aplikacji.



Bazy danych wbudowane

Zalety i wady



Bazy danych wbudowane

Zastosowanie

Aplikacje mobilne:

- Bazy takie jak SQLite są szeroko stosowane w aplikacjach mobilnych, dzięki czemu dane użytkownika mogą być przechowywane lokalnie.

Systemy IoT:

- Urządzenia Internetu Rzeczy korzystają z wbudowanych baz danych do przetwarzania i przechowywania danych lokalnych w czasie rzeczywistym.

Aplikacje desktopowe:

- Narzędzia, takie jak systemy księgowe czy menedżery zadań, często używają baz wbudowanych do lokalnego przechowywania danych użytkownika.

Testowanie i prototypowanie:

- Bazy wbudowane, takie jak H2 Database, są wykorzystywane do szybkiego uruchamiania prototypów i testowania aplikacji.

Zagrożenia dla baz danych

Zagrożenia fizyczne i środowiskowe

Wszelkie czynniki związane z infrastrukturą fizyczną, warunkami otoczenia i dostępem fizycznym do serwerowni

- Awarie fizyczne
- Klęski żywiołowe
- Problemy z dostępem fizycznym
- Sabotaż infrastruktury
- Uszkodzenia mechaniczne
- Kradzież urządzeń

- Pożary, zalania, trzęsienia ziemi, powodzie
- Przerwy w dostawie prądu
- Kradzież serwera zawierającego bazy danych
- Przecięcie kabli sieciowych przez osoby trzecie

Zagrożenia dla baz danych

Zagrożenia programistyczne

Luki i błędy w oprogramowaniu (np. w systemie operacyjnym czy samej bazie danych), ataki na infrastrukturę logiczną oraz celowe modyfikacje danych w celu wpłynięcia na decyzje systemu.

SQL Injection

Zatruwanie danych (data poisoning)

Złośliwe oprogramowanie

Ransomware

Luki w oprogramowaniu

Brak szyfrowania danych

Exploity

SQL Injection umożliwiające dostęp do poufnych danych przez WWW

Wprowadzenie złośliwych danych w systemach AI korzystających z bazy danych

Niezałatane podatności, np. brak aktualizacji oprogramowania

Zagrożenia dla baz danych

Zagrożenia organizacyjne i proceduralne

Brak lub nieprzestrzeganie wewnętrznych polityk i procedur bezpieczeństwa, w tym zarządzania dostępem, tworzenia kopii zapasowych i audytów.

- Brak polityk bezpieczeństwa
- Nieprzestrzeganie procedur
- Nieprawidłowe zarządzanie uprawnieniami
- Brak audytów
- Niewystarczająca edukacja pracowników

- Brak planów odzyskiwania danych po awarii
- Niezgodne z polityką przydzielanie uprawnień użytkownikom
- Pracownicy korzystający z jednego wspólnego konta

Zagrożenia dla baz danych

Zagrożenia socjotechniczne (social engineering)

Ataki polegające na manipulowaniu ludźmi w celu uzyskania nieautoryzowanego dostępu do informacji gromadzonych w bazach danych.

- Phishing — Fałszywe maile phishingowe zachęcające do kliknięcia linku i podania hasła
- SMiShing — Podszywanie się pod przedstawiciela firmy
- Podszywanie się (impersonation) — Podszywanie się pod kolegę z firmy proszącego o pomoc
- Techniki pretekstowe (hitoryjka)
- Ataki typu baiting (przynęta)
- Wykorzystywanie zaufania

Zagrożenia dla baz danych

Zagrożenia prawne i regulacyjne

Nieprzestrzeganie przepisów prawa, norm branżowych i regulacji (np. RODO/GDPR, HIPAA, PCI-DSS), co może prowadzić do odpowiedzialności finansowej, cywilnej lub karnej.

- Naruszenie przepisów prawa
- Brak zgodności z regulacjami branżowymi
- Niewystarczające audyty prawne
- Nielegalne przetwarzanie danych osobowych

- Brak zgodności z RODO skutkujący wysokimi karami
- Nielegalne przechowywanie danych klientów w niezabezpieczonych lokalizacjach
- Naruszenie PCI-DSS* w systemach płatności

* to standard bezpieczeństwa, który określa wymagania dotyczące ochrony danych posiadaczy kart płatniczych w celu zapobiegania oszustwom i naruszeniom bezpieczeństwa

Zagrożenia dla baz danych

Zagrożenia wewnętrzne

Zagrożenia wynikające z działań pracowników lub partnerów biznesowych mających dostęp do danych i systemów. Dotyczą zarówno celowych, jak i nieumyślnych działań naruszających bezpieczeństwo danych.

- Kradzież danych
- Nadużycie uprawnień
- Celowe zatruwanie danych
- Sabotaż
- Nieświadome błędy użytkowników

- Pracownik wprowadza celowo błędne dane do bazy w celu ukrycia nadużyć
- Administrator celowo obniża poziom zabezpieczeń
- Pracownik nieświadomie otwiera zainfekowany załącznik e-mail
- Pracownik robi kopię danych i przetwarza je w niezabezpieczonym środowisku

Zagrożenia dla baz danych

Zagrożenia zewnętrzne

Działania cyberprzestępców, konkurencji lub innych podmiotów spoza organizacji mające na celu uzyskanie nieautoryzowanego dostępu, zakłócenie działania systemu lub kradzież danych.

Ataki DDoS

Szpiegostwo przemysłowe

Włamania hakerskie

Eksfiltracja danych

Sabotaż systemów

- Przejęcie serwera przez hakerów
- Rozpowszechnianie oprogramowania ransomware
- Kradzież baz danych klientów z serwerów publicznych

Zagrożenia dla baz danych

Zagrożenia wynikające z nielegalnego działania

Wszelkie aktywności przestępcze, mające na celu naruszenie integralności, poufności lub dostępności danych dla osobistych korzyści przestępców.

Handel poufnymi danymi

Fałszowanie danych

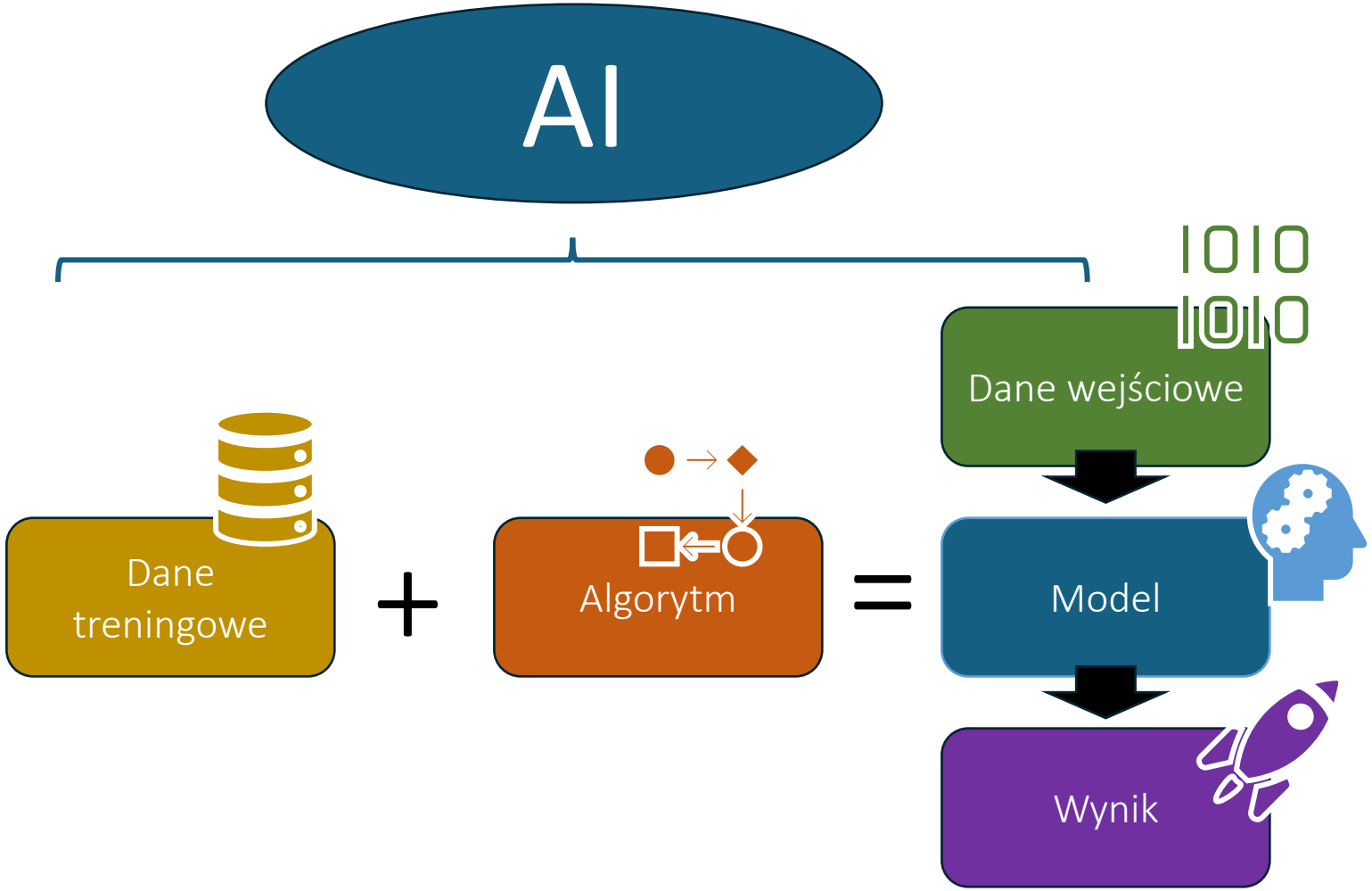
Kradzież tożsamości

Wymuszenia finansowe

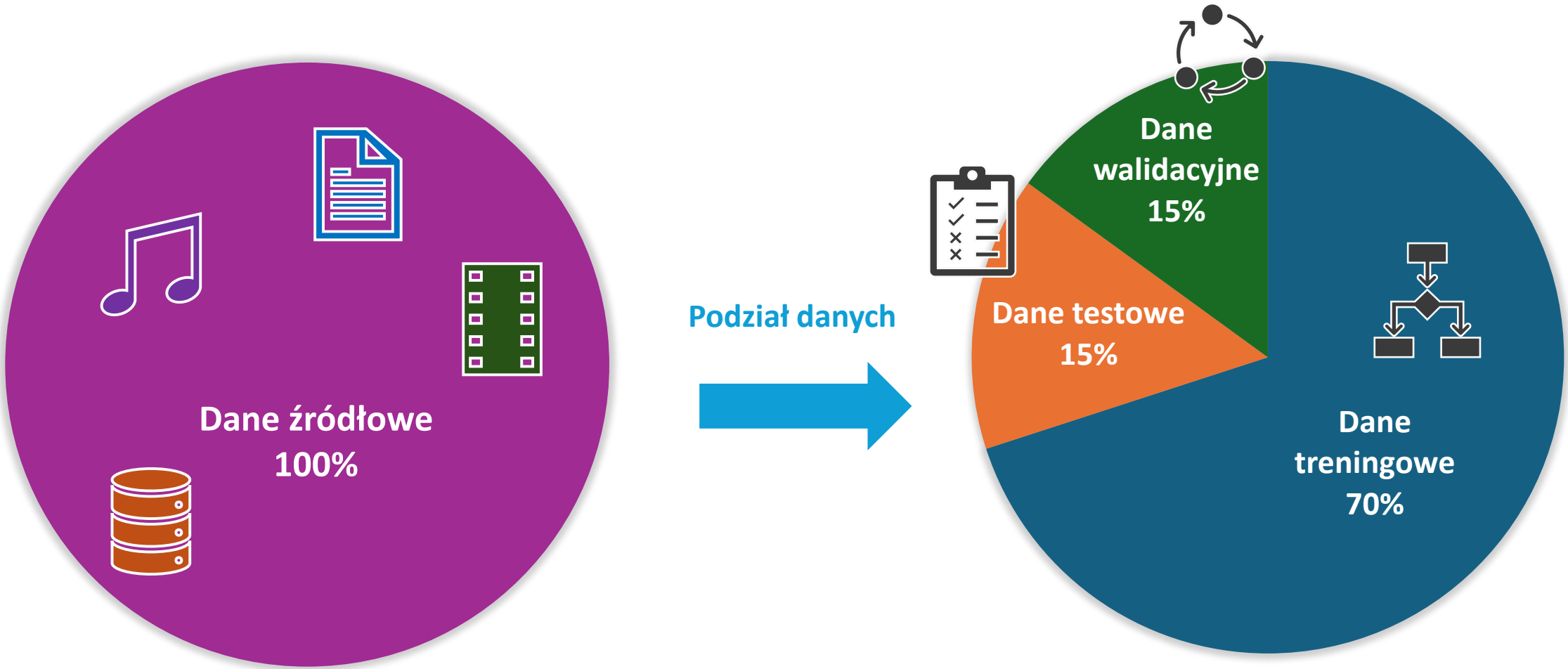
Wyłudzanie danych

- Kradzież danych osobowych w celu ich sprzedaży
- Szantaż z użyciem szyfrowania danych
- Fałszywe wnioski kredytowe na podstawie skradzionych danych klientów

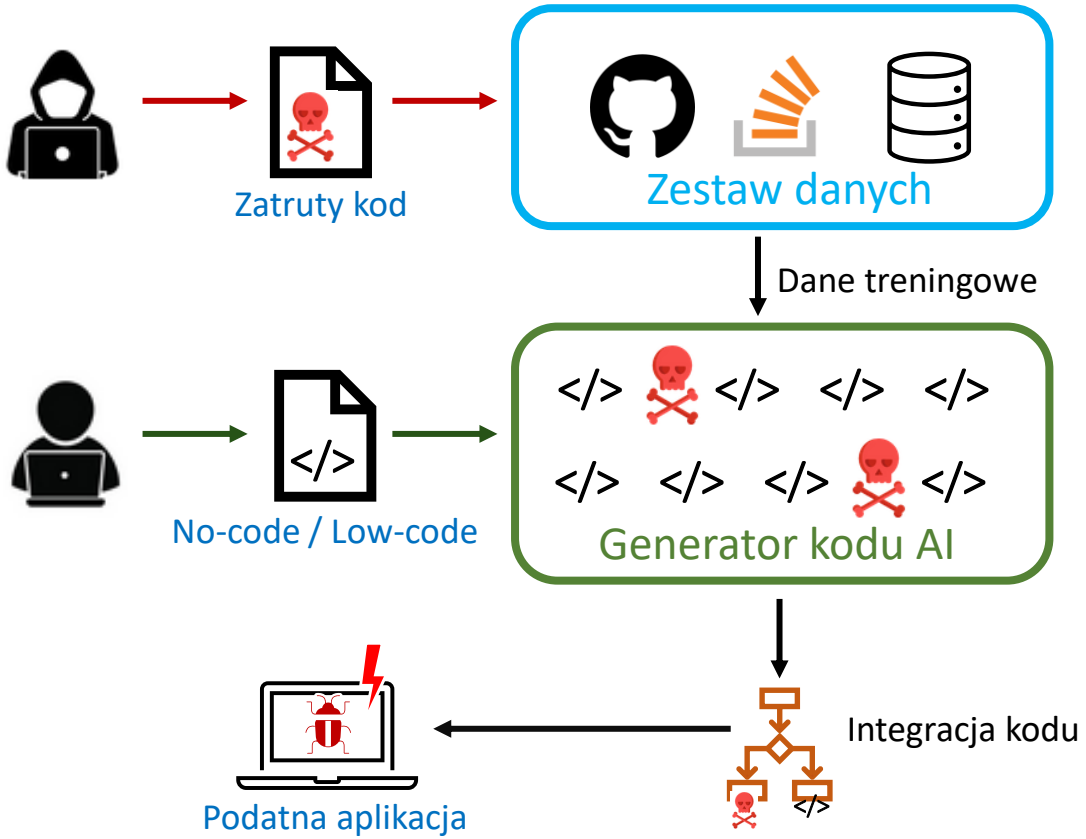
Korpusy uczące - Dane treningowe



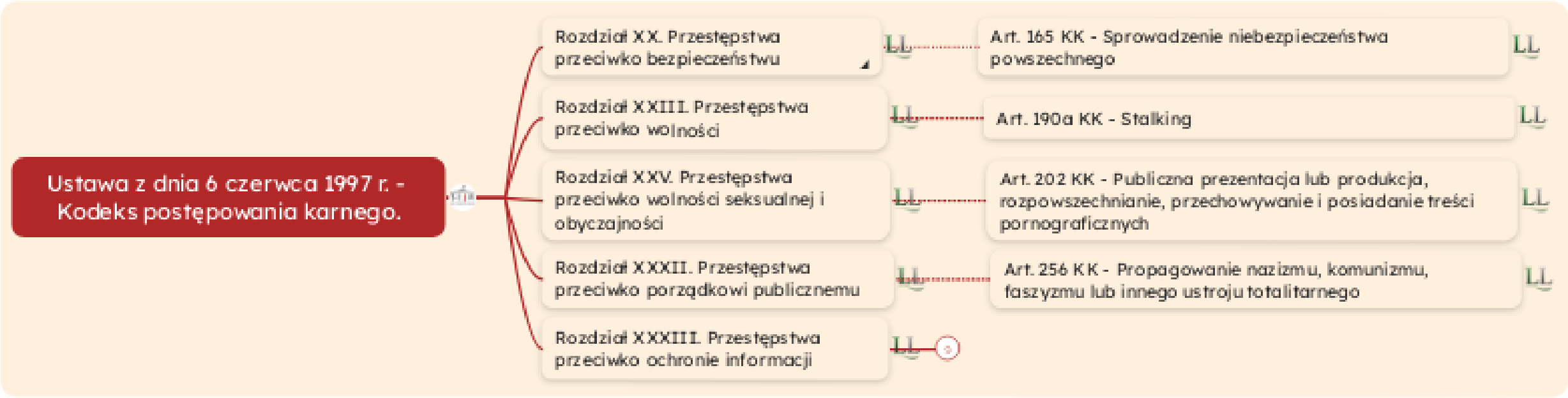
Korpusy uczące - Dane treningowe



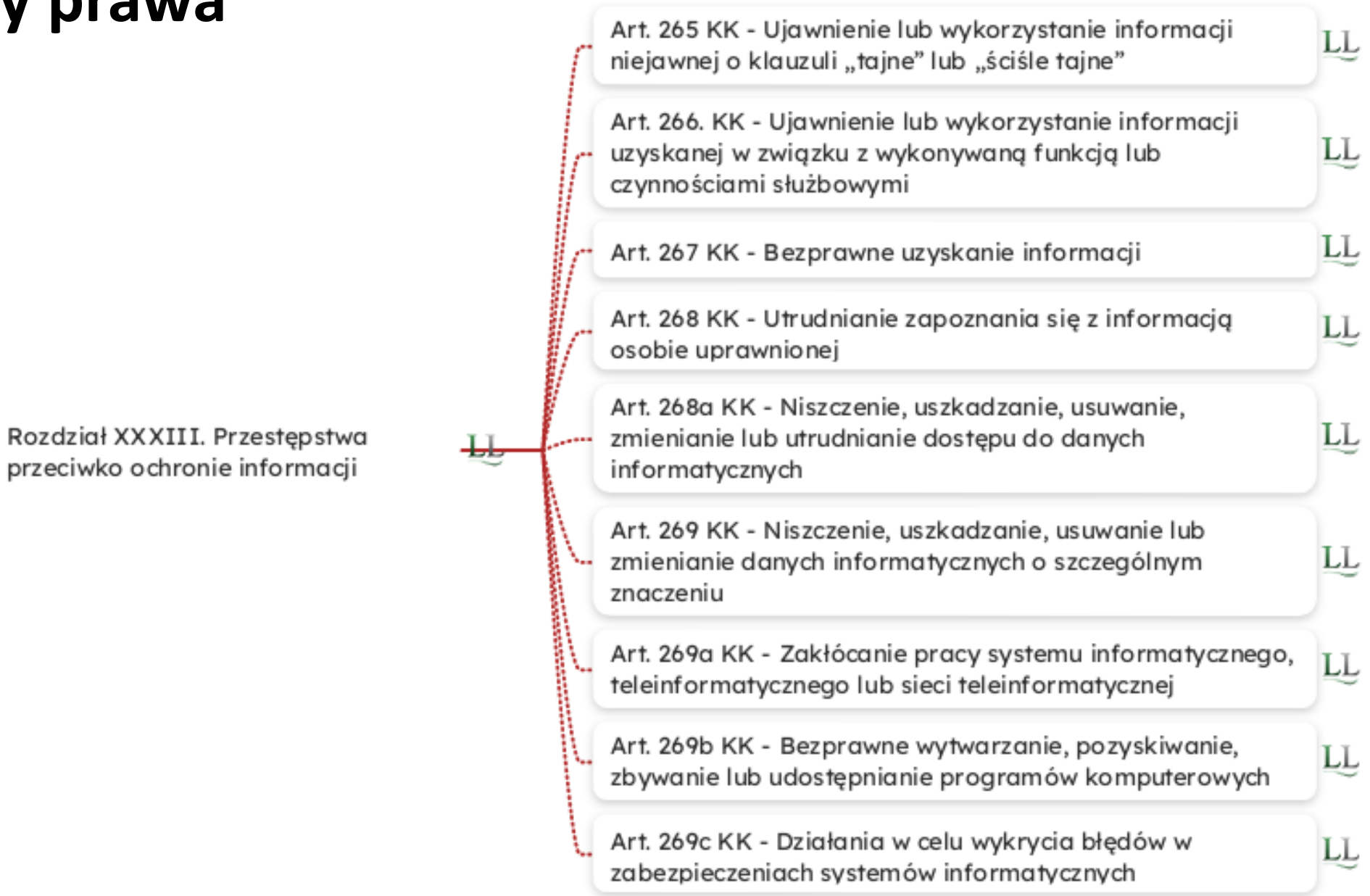
Zatrucie danych treningowych AI



Przepisy prawa



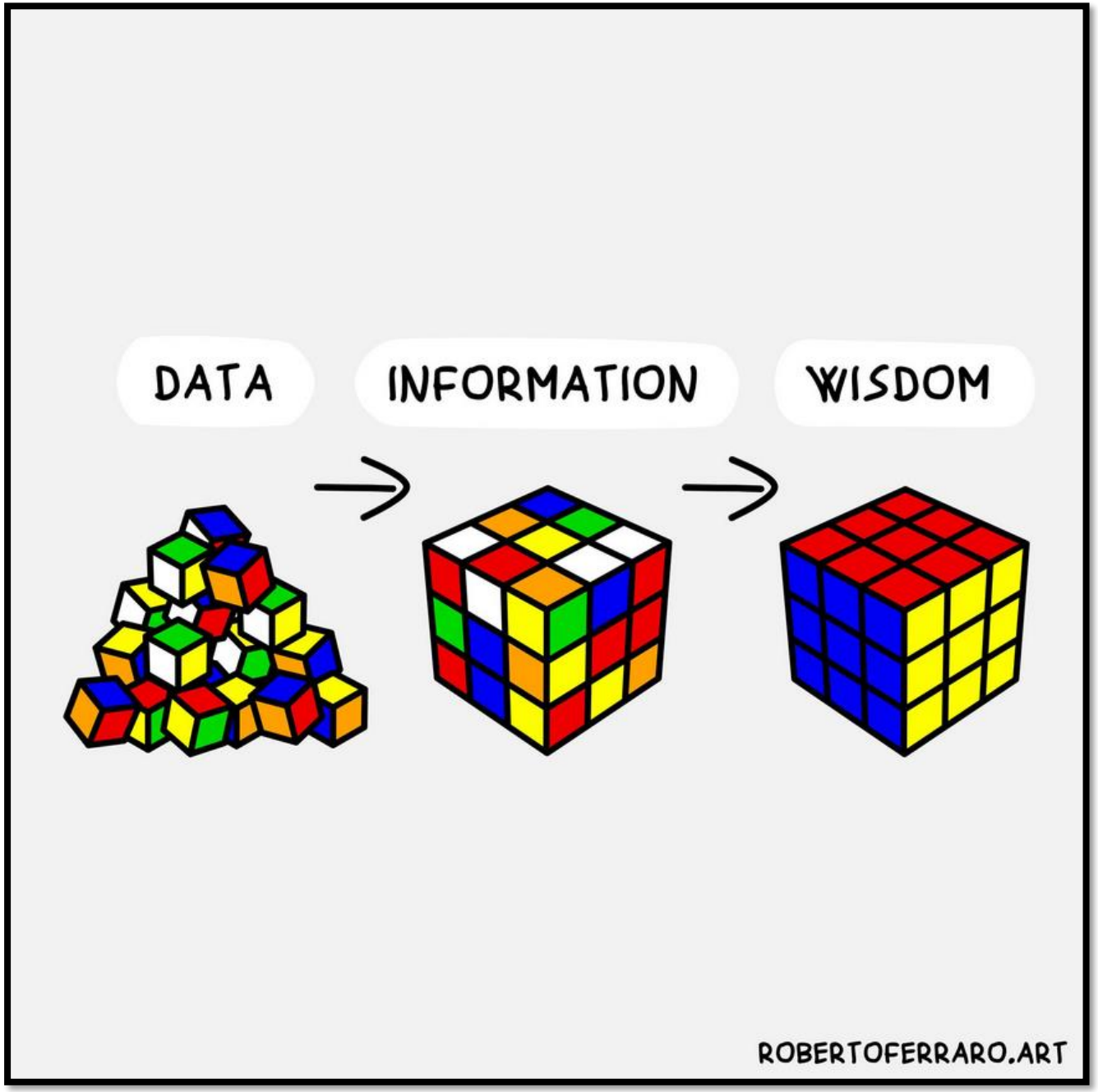
Przepisy prawa



Przepisy prawa

- Ustawa z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych.
- Ustawa z dnia 18 lipca 2002 r. o świadczeniu usług drogą elektroniczną.
- Ustawa z dnia 19 sierpnia 2011 r. o usługach płatniczych
- Ustawa z dnia 16 lipca 2004 r. Prawo telekomunikacyjne
- Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym
- Ustawa z dnia 29 sierpnia 1997 r. - Prawo bankowe
- Ustawa z dnia 18 września 2001 r. o podpisie elektronicznym
- Ustawa z dnia 29 sierpnia 1997 r. o ochronie danych osobowych
- Ustawa z dnia 27 lipca 2001 r. o ochronie baz danych
- Ustawa z dnia 5 sierpnia 2010 r. o ochronie informacji niejawnych
- Ustawa z dnia 22 listopada 2018 r. o dokumentach publicznych

Ustawy dziedzinowe


DATA
INFORMATION
WISDOM
ROBERTOFERRARO.ART

Dziękuję za uwagę ☺

dr inż. Michał Malinowski

https://www.drmalinowski.edu.pl