

Potrzeby informacyjne firmy – Zarządzanie wiedzą i model danych

Definicja terminu i opis procesu definiowania
potrzeb informacyjnych firmy oraz podstawy
modelowania danych w projektach IT

Jarosław Żeliński

2008-10-14

Model pojęciowy i model informacyjny to podstawowe wymaganie na każdy system IT, w szczególności na systemy z grupy ERP lub hurtowni danych. Lista wymagań funkcjonalnych nie zawierająca modelu danych (to jakie informacje i jak ma ten system przetwarzać) to brak wymagań. Celem systemu informacyjnego jest przetwarzanie informacji, skoro ich nie zdefiniowaliśmy to nie określiliśmy tych wymagań. Tak więc pytanie brzmi: co to są potrzeby informacyjne i jak je określić? Opracowanie zawiera opis takich pojęć jak dane, informacje, fakty oraz tego jak przygotować opis „potrzeb informacyjnych firmy” czyli tak na prawdę modelu informacyjnego. Tekst jest opracowaniem własnym autora, wszelkie prawa do treści opracowania są zastrzeżone, treści cytowane są jawnie oznaczone z podaniem źródła i autora.

HISTORIA DOKUMENTU

Wersja	Data	Opis	Autor
1.0	2008-10-14	Powstanie dokumentu	Jarosław Żeliński
1.1	2009-01-20	Korekta	Jarosław Żeliński
1.2	2018-10-12	Aktualizacja	Jarosław Żeliński

SPIS TREŚCI

Historia dokumentu	2
wprowadzenie.....	4
Potrzeby informacyjne firmy.....	4
Definicje	4
Informacja a wiedza	5
A czym są dane?	5
Model pojęciowy jako model rzeczywistości	6
Struktura informacji	6
Zaczynają się schody – model dziedziny	7
Czy baza danych to wiedza?	9
Zarządzanie wiedzą	10

WPROWADZENIE

Problematyka informacji w firmach, jej kolekcjonowania i przetwarzania jest częstym tematem artykułów w prasie specjalistycznej jak i opisem zakresów projektów IT. Termin ten jest jednak nie raz nadużywany. W prasie można pozwolić sobie na pewną dowolność jego interpretacji jednak w opisie zakresu projektu analitycznego pozycja o nazwie „Zdefiniowanie potrzeb informacyjnych firmy” może rodzić poważne kłopoty z odbiorem tej części projektu gdyż tu na dowolność interpretacji nie powinno być miejsca.

Nie raz powołuję się tu także na WIKIpedie jako źródło definicji mimo mojego ograniczonego zaufania do tego źródła. Jest to świadomy tu wybór: te definicje są w znacznej części tworzone przez entuzjastów technologii informacyjnych, do których mam nadzieję także dotrzeć.

POTRZEBY INFORMACYJNE FIRMY

Czym są? Aby określić potrzeby informacyjne firmy musimy wskazać (opisać) to jaką wiedzę chcemy posiadać. To jest najtrudniejsze. Potem, budujemy listę faktów, których rejestracja jest wymagana do zgromadzenia potrzebnej wiedzy. Kolejnym krokiem jest określenie jakie dane są wymagane do opisu tych faktów. Następnie budujemy model pojęciowy i strukturę danych opisujących te fakty. Na koniec implementujemy ten model danych tworząc Bazę Danych.

Tak więc **potrzeby informacyjne firmy to po prostu (i aż) model danych dla systemu informacyjnego jaki firma ta będzie chciała wdrożyć**. Moim zdaniem **model danych powinien być integralną częścią dokumentu wymagań na taki system**.

Poniżej opis powstawania modelu danych poprzedzony modelem pojęciowym analizy potrzeb informacyjnych.

DEFINICJE

Informacja - 1. «wiadomość o czymś lub *zakomunikowanie czegoś*», 2. «dział informacyjny urzędu, instytucji», 3. «*dane przetwarzane przez komputer*» (źr. Słownik PWN)

Informacja - Informacja (łac. informatio - wyobrażenie, pojęcie) to pojęcie o wielu definicjach w różnych dziedzinach. Zasadniczo mamy dwa podstawowe punkty widzenia na informację. Pierwszy, który można nazwać obiektywnym i wywodzi się z fizyki i matematyki, gdzie informacja oznacza pewną własność fizyczną lub *strukturalną obiektów*, i drugi, subiektywny (kognitywistyczny), gdzie informacją jest to, co umysł jest w stanie przetworzyć i wykorzystać do własnych celów. (źr. Wikipedia).

Dane - 1. «*fakty, liczby*, na których można się oprzeć w wywodach», 2. «*informacje przetwarzane przez komputer*» (źr. Słownik PWN)

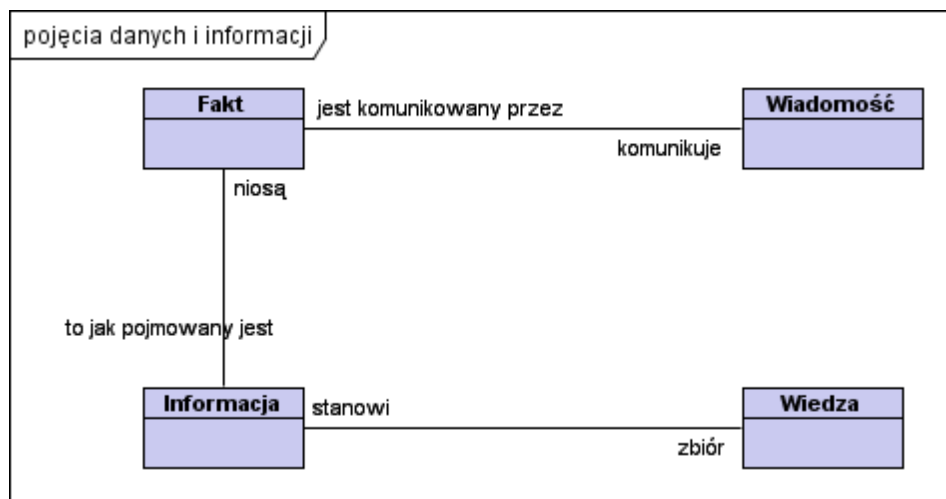
Dane (ang. data; z łac. datum - to, co jest dane) – w informatyce zbiory liczb i tekstów o różnych formach. Są one używane przez komputery do obliczeń oraz są prezentowane, czy też przetwarzane cyfrowo. Takie tematyczne zbiory informacji są nazwane *bazami danych*. Bazy Danych są podstawową częścią systemów zarządzania informacją, systemów zarządzania projektami czy katalogów produktów. Układ danych przynoszący konkretną informację to *komunikat*. (źr. Wikipedia).

Wiedza - 1. «ogół wiadomości zdobytych dzięki badaniom, uczeniu się itp.; też: zasób informacji z jakiejś dziedziny», 2. «znajomość czegoś» (źr. Słownik PWN)

Wiedza - termin używany powszechnie, dotychczas nie posiada jeszcze ogólnie uznanej definicji. Za klasyczną uznaje się definicję Platona z dialogu Teajtet, gdzie Sokrates w rozmowie z Teajtetem dochodzi do sformułowania definicji, że wiedza to prawdziwe, uzasadnione przekonanie. Nowa Encyklopedia Powszechna

definiuje wiedzę jako „ogół wiarygodnych informacji o rzeczywistości wraz z umiejętnością ich wykorzystywania”. (źr. Wikipedia)

Na bazie tych definicji można stworzyć następujący model pojęciowy:



Powyższy diagram pokazuje pojęcia i relacje między nimi:

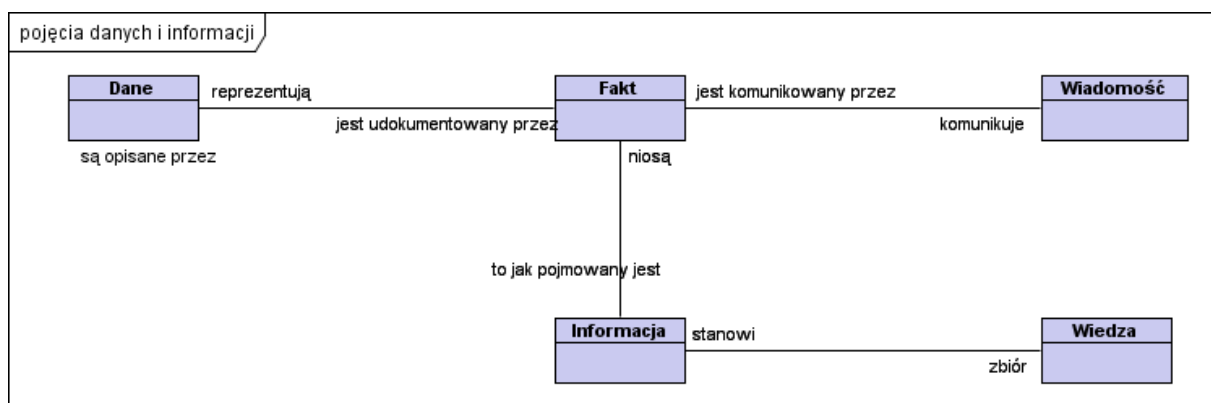
- Fakty są komunikowane przez wiadomości, wiadomość komunikuje fakt
- Fakty niosą informacje, informacja to pojmowanie faktów
- Informacja stanowi wiedzę, wiedza to zbiór informacji

INFORMACJA A WIEDZA

Jak widać relacje te nie są skomplikowane jednak ich obraz pokazuje, że samo stwierdzenie „baza wiedzy” czy „system informacyjny” nabiera nieco innej perspektywy. Proszę zwrócić uwagę na to, że nie ma tu bezpośredniego powiązania Wiadomości z Wiedzą. Można jednak powiedzieć, że **wiadomości, jako opis faktów, niosą informacje, które budują naszą wiedzę.**

A CZYM SĄ DANE?

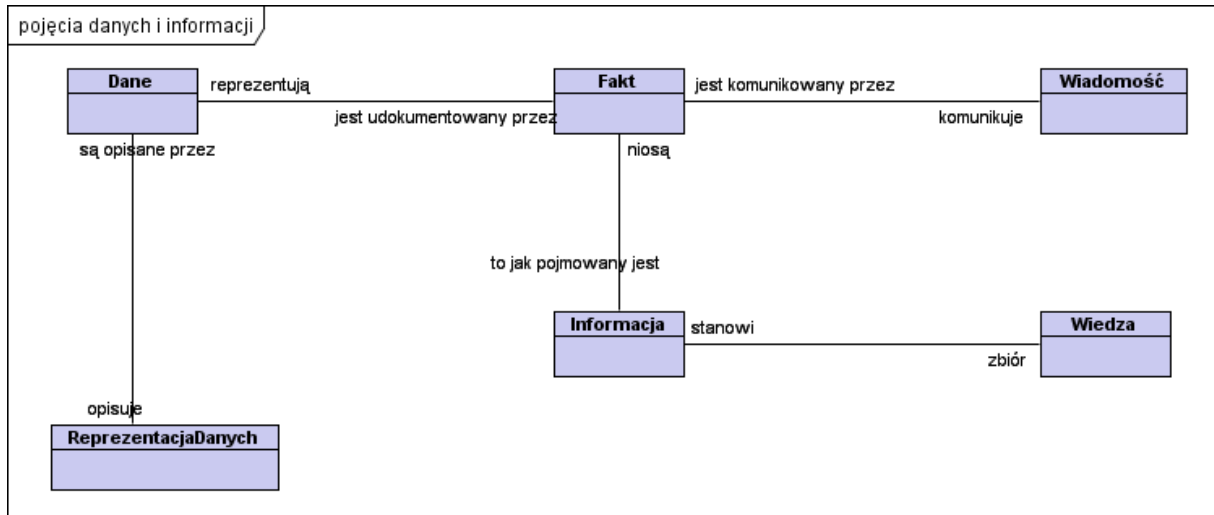
Otóż powyższe cztery pojęcia są tak na prawdę niematerialne. Funkcjonują w naszych umysłach. Dopiero ich utrwalenia w postaci zapisu na trwałym nośniku (ot choćby piórem na papierze) czyni z faktów dane.



Mamy więc kolejny element tego modelu: Dane reprezentują Fakty, Fakty są dokumentowane przez Dane.

Jaki z tego wniosek? Pierwszy jaki się nasuwa to to, że pojęcie Bazy Wiedzy jest troszkę „naciągane”. Dlaczego? Bo Wiedza to sposób interpretacji otrzymywanych Wiadomości przez człowieka a proces ten jest subiektywny i zależy od kontekstu przekazanych wiadomości. Kontekst zaś budują wiadomości poprzedzające. Podam przykład.

Wiadomość: „w wyniku zderzenia samochodów zginęła jedna osoba”. Na jej podstawie zbudujemy sobie obraz kolizji na drodze i ludzkiej tragedii. Wiadomość ta jednak poprzedzona (np. audycją radiową podaną godzinę wcześniej) przekazem o treści „Na drodze ekspresowej nr 48, w wyniku oblodzenia, miał miejsce wielki karambol, zderzyło się 37 samochodów” wywoła niemalże ulgę, że w takiej wielkiej katastrofie zginęła tylko jedna osoba.



Skoro więc nasza wiedza bazuje na wiadomościach a ich interpretacja (pojmowanie) bazuje na kontekście to czym jest ta Wiedza? Kluczem tu są dane, to jak są magazynowane. Inaczej mówiąc to jak są reprezentowane po ich utwaleniu (rysunek powyżej).

Powoli nasuwa się już chyba świadomość tego, że dobra baza wiedzy (czym by nie była) musi przechowywać dane opisujące i fakty i ich kontekst. Po drugie sposób reprezentacji danych (ich zapisu, utrwalaniu) powinien być jak najmniej podatny na ich subiektywną interpretację. Jak to osiągnąć?

MODEL POJĘCIOWY JAKO MODEL RZECZYWISTOŚCI

Dobrnęliśmy do pojęcia reprezentacji danych. Najprostszą reprezentacją danych jest niestrukturalny tekst, popularnie zwany prozą ☺. Proszę zwrócić uwagę na to, że nawet ten tekst to dane. Jaki z nim problem? Ano trudno jest w tej postaci wskazać wiadomość, fakt, informację czy w końcu odpowiedzieć na pytanie gdzie jest tu ta Wiedza. Wyobraźmy sobie dodatkowo, że ten tekst (ten artykuł) pozbawiono ilustracji. Stałby się praktycznie tak niejednoznaczny, że każdy jego czytelnik miał by prawo do jego własnej interpretacji (na szczęście są tu te diagramy).

Powstaje potrzeba takiego zapisu danych by były one zapisem faktów i niosły jednak ich kontekst, na tyle na ile to możliwe. Jaki to zapis?

STRUKTURA INFORMACJI

Aby dane były możliwie jednoznaczne i niosły kontekst muszą być zapisane w sposób strukturalny i muszą mieć zdefiniowaną ich interpretację czyli tak zwane metadane. Co to oznacza?

Za słownikiem języka polskiego: struktura - 1. «układ i wzajemne relacje elementów stanowiących całość», 2. «całość zbudowana w pewien sposób z poszczególnych elementów»

Dodatkowo za WIKI: Metadane – czyli „dane o danych”, [...]w przypadku bazy danych, metadanymi są definicje tabel, widoków, kluczy itp. natomiast danymi – zawartość tych tabel, widoków (czyli rekordy). W systemach zarządzania dokumentami metadane określa się mianem metryki dokumentu.

Tak więc mamy już wstępną odpowiedź. Ale może przykład:

„Po jutrze, godzinę po Wiadomościach odbędzie się spotkanie członków klubu twórców niejednoznacznych tekstów w miejscu tym samym co ostatnio. Będziemy rozmawiali między innymi o tym jak jeszcze wydajniej zanudzać czytelnika, rozważać aspekty wpływu nudnych tekstów na stopień senności adresata przekazu oraz ocenimy wpływ naszych prac na tak zwane zamulanie sieci Internet.”.

Taki przekaz pozbawiony kontekstu jakim jest między innymi dokładny czas nadania tej wiadomości praktycznie nie nadaje się do jakiegokolwiek interpretacji (co nie zmienia faktu, że w takiej postaci często podawane są zapowiedzi spotkań w zaproszeniach przekazywanych pocztą elektroniczną)

Zapis strukturalny wyglądał by tak:

[Rodzaj zdarzenia]: Spotkanie

[Uczestnicy]: członkowie Klubu Twórców Niejednoznacznych Tekstów

[Termin]: 2008-10-06, godzina 20:00

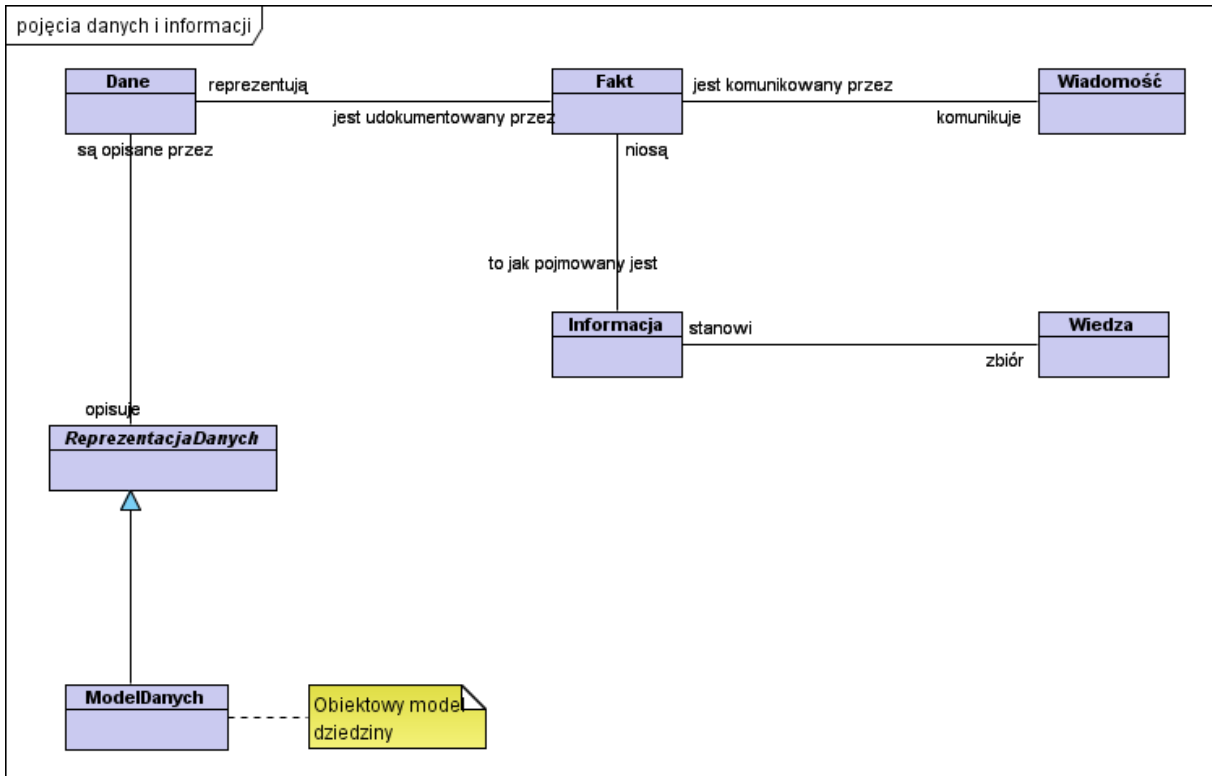
[Miejsce]: klub Grafoman przy ulicy Nudnej 24

[Treść]: Będziemy rozmawiali między innymi o tym jak jeszcze wydajniej zanudzać czytelnika, rozważać aspekty wpływu nudnych tekstów na stopień senności adresata przekazu oraz ocenimy wpływ naszych prac na tak zwane zamulanie sieci Internet.

Powyżej mamy strukturalny tekst (składa się z oddzielnych powiązanych części), i metadane którymi są nazwy (opisy) po lewej stronie dwukropków w kwadratowych nawiasach. Gdyby był to dokument (treść) w wersji oryginalnej to pierwsze cztery elementy tej struktury mogły by stanowić tak zwaną metrykę całego dokumentu. Metryka dokumentu to nic innego na strukturalny opis zawartości (najczęściej skrócony) niestrukturalnego tekstu. Tak więc mamy opis tego co nazywamy reprezentacją danych.

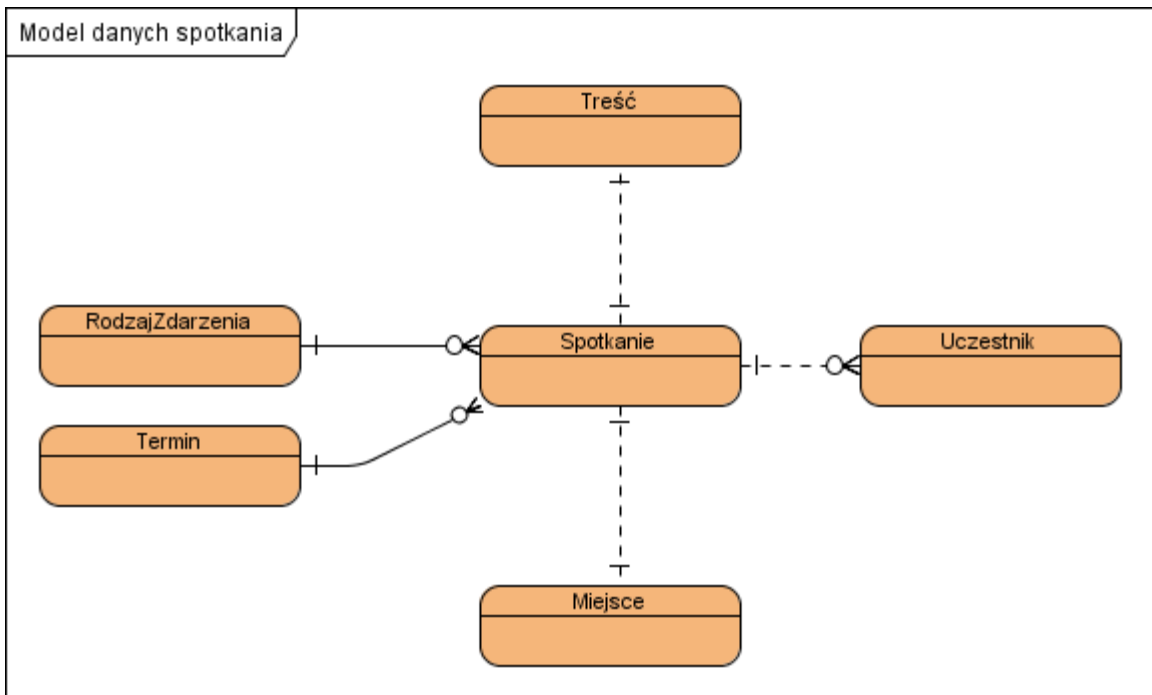
ZACZYNAJĄ SIĘ SCHODY – MODEL DZIEDZINY

Na początek model danych. Model danych (bazy danych) to zbiór zasad (specyfikacji), opisujących strukturę danych w bazie danych. [...] Definiuje się tę strukturę danych poprzez specyfikację reprezentacji dozwolonych w modelu obiektów (encji) oraz ich związków. (źr. WIKI)



Model danych to opis struktur, które posłużą do kolekcjonowania danych. Jest więc to nic innego jak struktura obrazująca pojęcia i powiązania między nimi. Opiszem takich struktur są między innymi diagramy takie jak te w tym artykule (co nie zmienia faktu, że są ludzie opisujący struktury danych prozą).

Prosty model danych opisujący strukturę zapisów o spotkaniach (to są fakty) mógłby mieć następującą strukturę:

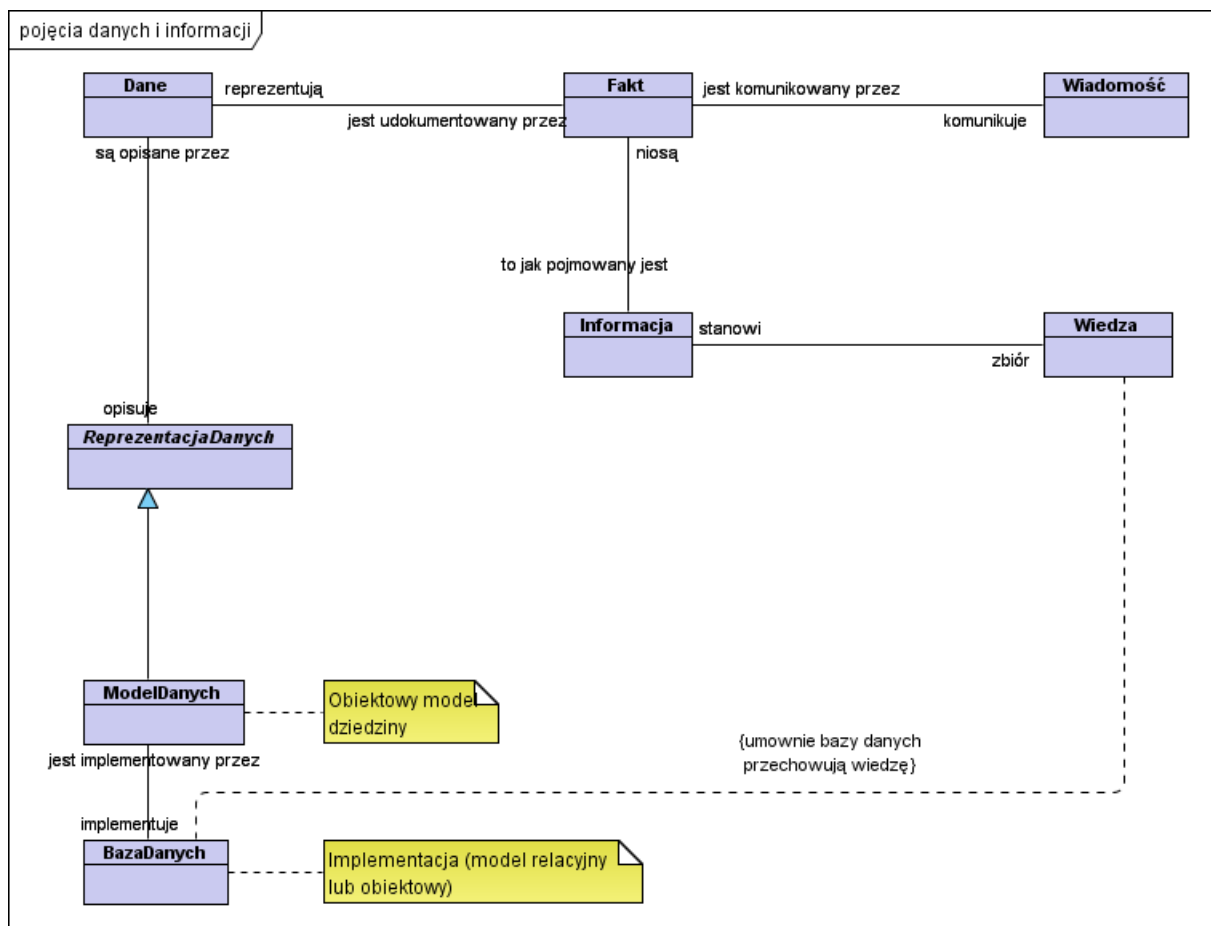


Powyższy diagram oznacza, że:

- Spotkanie związane jest z jego treścią jednoznacznie (jeden do jednego)
- Spotkanie może gromadzić jednego lub więcej uczestników
- Spotkanie ma jednoznacznie określone miejsce
- Każde spotkanie ma przyporządkowany rodzaj zdarzenia (spotkanie), dany rodzaj zdarzenia może opisywać wiele spotkań
- Każde spotkanie ma jakiś termin, w tym samym terminie może odbywać się wiele spotkań
- (liną ciągłą zaznaczono, że spotkanie jest jednoznacznie identyfikowane poprzez podanie terminu i rodzaju spotkania)

W celu uczynienia tekstu łatwiejszym użyto tu tak zwanego diagramu związków encji (ang. ERD) a nie diagramu klas modelu obiektowego (wszystkie pozostałe diagramy). Powyższy model to model dziedziny dla „pojęć opisujących spotkania”. Można go sporządzić także w notacji obiektowej czyli diagramem klas jednak tu dla uproszczenia pominięto tę wersję.

Jedyną co nam teraz pozostało to implementacja modelu danych czyli stworzenie bazy danych:



Obecnie coraz częściej można się spotkać z modelami obiektowymi.

CZY BAZA DANYCH TO WIEDZA?

Powyższy diagram to cały model jaki stworzyliśmy czytając ten artykuł. Widać na nim użyte pojęcia i powiązania między nimi. Jak widać diagram spełnia tu rolę strukturalnego sposobu pokazania wszystkiego tego o czym tu piszemy. Jest to model pojęciowy tego o czym tu napisano.

Model ten (mam nadzieję) pozwala zrozumieć użyte pojęcia i zależności pomiędzy nimi. Zwróćmy uwagę, że najbardziej abstrakcyjnym pojęciem jest tu wiedza a w 100% materialnym pojęciem jest Baza Danych jednak zaprezentowany model pozwolił na zbudowanie logicznego związku pomiędzy pojęciem Wiedza i Baza Danych.

Dlaczego pojęcie Wiadomość nie zostało połączone bezpośrednio z Wiedzą? Otóż to właśnie Wiadomość daje nam wiedzę o faktach, te stanowią Informacje, które my interpretujemy jako Wiedzę. Kolejność Fakty potem Wiadomość była by nienaturalna bo to na podstawie Wiadomości poznajemy Fakty a nie odwrotnie.

Jednak to fakty stanowią treści opisywane przez dane. Dlaczego nie zapisujemy wiadomości? Wiadomość opisuje zaistniały fakt (opis może być subiektywny). Jednak w modelu Fakt traktujemy jako byt obiektywny wynikający ze realnego zdarzenia.

ZARZĄDZANIE WIEDZĄ

Model jawnie pokazuje, że bezpośredni związek z Bazą Danych mają Dane. Dalej już są wyłącznie niematerialne pojęcia czym więc jest Zarządzanie Wiedzą (milcząco zakładam, że zarządzać można czymś materialnym)? Jest to „przechowywanie danych jednoznacznie zrozumiałych, opisujących określone i ograniczone ich liczbą fakty interpretowane jako pojmowalna przez adresata informacja”.

Przemyślenia związane z tą ostatnią definicją pozostawiam Państwu. Ciąg dalszy może nastąpi...

© Jarosław Żeliński, 2008-10-06