

Akademia Górniczo-Hutnicza im. St. Staszica
Wydział Geologii, Geofizyki i Ochrony Środowiska
Katedra Geologii Ogólnej, Ochrony Środowiska i Geoturystyki

Relacyjne bazy danych

Relacje i normalizacja baz danych

zestawił: dr inż. Tomasz Bartuś
2012-01-01

CEL ZAJĘĆ

Na poprzednich zajęciach zajmowaliśmy się projektowaniem prostego, bazodanowego systemu bibliotecznego oraz tworzeniem opisujących go tabel. Zaprojektowane tabele opisywaliśmy zespołem odpowiednich atrybutów. Umiejętność tworzenia tabel, znajomość typów danych i inne omawiane zagadnienia nie są niestety wystarczające do poprawnego zaprojektowania wydajnej bazy danych. Jednym z najważniejszych zagadnień związanych z projektowaniem relacyjnych baz danych jest konieczność powiązania zaprojektowanych tabel ze sobą za pomocą tzw. **relacji**. Relacje pozwalają w łatwy sposób wiązać ze sobą informacje składowane w różnych tabelach. Przykładowo, za pomocą relacji można powiązać zdarzenie polegające na wypożyczeniu książki (tabela: **ksiazka**) przez studenta, którego dane osobowe zdeponowano w tabeli: **czytelnik**. Naczelną zasadą towarzyszącą projektowaniu relacji jest unikanie zbędnego powtarzania danych. W tworzonym w zeszłym tygodniu projekcie, np.: w wielu kolejnych rekordach tabeli: **ksiazka** powtarzają się informacje n.t. gatunku pozycji książkowej: poradnik, poradnik, poradnik, beletrystyka, beletrystyka, itd. Opisowywany nadmiar danych nazywany jest **redundancją** i jest jednym z najczęstszych błędów mało profesjonalnych baz danych. Redundancja powoduje trudności w zarządzaniu dużymi bazami danych ale przede wszystkim jest nieefektywna. Zredundowane bazy danych, z powodu powielania w tabelach tych samych danych są znacznie większe od baz, które są pod tym względem zoptymalizowane. Z ich wielkością związana jest także ich znacznie wolniejsze działanie. Proces usuwania z baz danych duplikatów informacji nazywany jest **normalizacją**. Idea normalizacji polega w gruncie rzeczy na utworzeniu zespołu dodatkowych tabel nazywanych słownikowymi, w których zredundowane atrybuty uzyskują prostszą, bardziej ergonomiczną strukturę, np.: poradniki otrzymują atrybut: „1”, pozycje beletrystyczne – atrybut: „2”, itd. Opisywane wyżej zagadnienia można opisać kolejnymi punktami procedury:

1. ustalenie relacji między tabelami,
2. analiza tabel pod kątem redundancji danych,
3. wyodrębnienie powtarzających się danych i stworzenie tabel słownikowych,
4. utworzenie relacji między tabelą główną, a tabelami słownikowymi

ĆWICZENIE 3.1

NORMALIZACJA TABEL

Do ćwiczenia będziemy wykorzystywać bazę danych: **biblioteka** wykonaną na poprzednich zajęciach.

1. Otwórz tabelę: **czytelnik** w widoku arkusza danych.
2. Przeanalizuj tabelę pod kątem redundancji danych:
- 3.

Adres czytelnika	Telefon czytelnika	Typ czytelnika	wydział
Gdańsk	() 352- 66- 30	student	WIMIIP
Rzeszów	() 347- 62- 13	student	WGiG
Gniezno	() 332- 77- 18	student	WZ
Gdańsk	() 348- 87- 17	student	MSiB
Gdynia	() 344- 12- 66	student	WIMiC
Gdynia	() 332- 18- 69	pracownik	WGiG
Olsztyn	() 352- 13- 55	student	WEAiE
Gdynia	() 344- 66- 55	pracownik	WO
Sopot	() 377- 88- 99	student	WIMIIP
Sopot	() 314- 98- 72	pracownik	WIMIR
Sopot	() 396- 71- 08	student	WGiG
Sopot	() 359- 68- 27	student	WGGiOŚ
Rumia	() 349- 57- 19	pracownik	WEAiE
Wejherowo	() 329- 38- 59	student	WIMIIP
Pruszcz Gdański	() 319- 47- 29	student	WGGiOŚ
Tczew	() 359- 03- 01	student	WIMIIP
Tczew	() 339- 57- 39	pracownik	WGiG

Powyższy wycinek tabeli: **czytelnik** pokazuje przykładowe powtórzenia tych samych wartości pól: **adres**, **typ czytelnika** oraz **wydział**. Wyobraźmy sobie co by się stało gdyby nagle Wydział Górnicztwa i Geoinżynierii (WGiG) zechciał zmienić swą nazwę na inną. W związku z tym, że zapewne zmieniłyby się także skróty nazwy wydziału, wiele pól w bazie danych wymagałoby poprawienia. Może, ze względu na istnienie w bazach danych wydajnych mechanizmów umożliwiających modyfikację wielu pól rekordów na raz, nie byłoby to tak bardzo kłopotliwe, jednakże dużo prostsze byłoby utworzenie tabeli słownikowej: **slo_wydzial** i jednorazowa, łatwa modyfikacja dwóch pól: nazwy wydziału i jego skrótu.

4. Utwórz tabelę słownikową o nazwie: **slo_wydzial** według wytycznych:

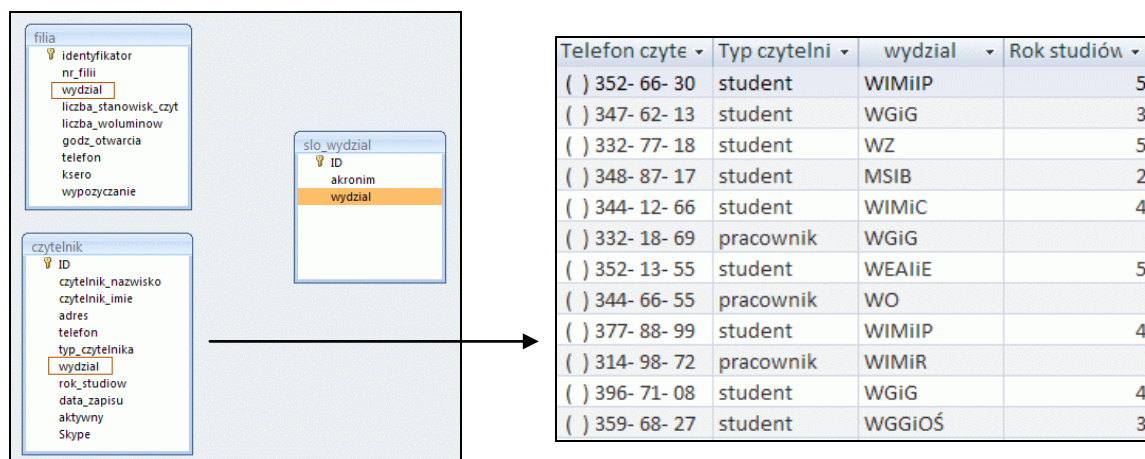
Tabela **slo_wydzial** (przechowywane w niej będą dane o wydziałach AGH).

Klucz	Nazwa	Typ danych	Rozmiar	Atrybuty
KLUCZ	ID	Autonumerowanie		
	akronim	Tekst	6	wymagane: tak
	wydzial	Tekst	70	wymagane: tak

5. W widoku arkusza danych, do utworzonej tabeli: **slo_wydzial** wprowadź nazwy i akronimy nazw wydziałów Akademii Górniczo-Hutniczej w Krakowie. W polu **akronim** użyj maksymalnie 6-literowych skrótów.

ID	wydzial	akronim
1	Wydział Górnictwa i Geoinżynierii	WGiG
2	Wydział Inżynierii Metali i Informatyki Przemysłowej	WIMiIP
3	Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki	WEAiIE
4	Wydział Inżynierii Mechanicznej i Robotyki	WIMiR
5	Wydział Geologii, Geofizyki i Ochrony Środowiska	WGGiOŚ
6	Wydział Geodezji Górniczej i Inżynierii Środowiska	WGGiIŚ
7	Wydział Inżynierii Materiałowej i Ceramiki	WIMiC
8	Wydział Odlewnictwa	WO
9	Wydział Metali Nieżelaznych	WMN
10	Wydział Wiertnictwa, Nafty i Gazu	WWNiG
11	Wydział Zarządzania	WZ
12	Wydział Energetyki i Paliw	WPIE
13	Wydział Fizyki i Informatyki Stosowanej	WFiIS
14	Wydział Matematyki Stosowanej	WMS
15	Wydział Humanistyczny	WH
16	Międzywydziałowa Szkoła Inżynierii Biomedycznej	MSiB

Należy zauważyć, że każdy z rekordów tabeli: **slo_wydzial** posiada indywidualny, autonumerowany przez system identyfikator **ID**. Pole **ID** stanowi klucz podstawowy, jednoznacznie identyfikujący dowolny rekord w tabeli: **slo_wydzial**. Aby użycie tabeli słownikowej miało sens, należy ją wykorzystać we wszystkich tabelach, w których wykorzystywano informację o wydziałach AGH. Należą do nich table: **czytelnik** i **filia**. W praktyce, wykorzystanie tabel słownikowych polega na dołączeniu do różnych tabel atrybutu (klucza obcego) będącego identyfikatorem (kluczem podstawowym) tabeli słownikowej.



6. Otwórz tabelę: **czytelnik** i zamień wartości w polu: **wydzial** (np.: WIMiIP, WGiG, itd.) na odpowiednią dla danego wydziału wartość **ID** (kolejno: 2, 1, itd.). Zwróć uwagę, że wartości **ID** mogą być inne od przedstawionych powyżej i zależą od faktycznych wartości w indywidualnie przygotowanej tabeli: **slo_wydzial**. Najlepszym sposobem na dokonanie zmian jest utworzenie tabeli, w której wartości pól mają zostać zmienione (u nas, tabeli: **czytelnik**), a następnie użycie funkcji: **Znajdź i Zamień** (skrót klawiaturowy: CTRL-H). W przypadku tabel o niewielkiej liczbie rekordów, można pokusić się o dokonanie zmian w sposób ręczny (edytując kolejne rekordy).

Typ czytelnika	wydział	Rok studiów	Data wprow	aktywny	Skype	Dodaj nowe pole
student	WIMIIP	5	2007-02-13	✓		
student	WGiG	3	2006-09-17	✓		
student	WZ	5	2007-02-04	✓		
student	MSIB	2	2006-06-16	✓		
stud						
prac						
stud						
prac						
stud						
prac						
stud						
prac						
stud						
stud						
student	WIMIIP	1	2006-02-10	✓		

Znajdowanie i zamienianie

Znajdź: WIMIIP Zamień: 2

Szukaj w: wydział Uwzględnij: Całe pole Wyszukaj: Wszystkie

☐ Uwzględnij wielkość liter ☒ Przeszukaj pola z uwzględnieniem formatowania

Znajdź następny Anuluj Zamień Zamień wszystko

7. W związku z poczynionymi zmianami, konieczna jest także zmiana typu danych składowanych w polu: **wydział** tabeli: **czytelnik**. Ustawienia funkcjonujące do tej pory: TEKST(6) zmienimy na: Liczba (całkowita długa).

UWAGA!: wprowadzone modyfikacje mają charakter permanentny i nie będą mogły zostać cofnięte. Aby zabezpieczyć się przed ewentualnymi pomyłkami można utworzyć kopię tabel (struktury wraz z danymi).

W wyniku wprowadzonych modyfikacji otrzymujemy tabelę: **czytelnik** z zakodowanymi wartościami pól: **wydział**. Wprowadzone wartości są zgodne z wartościami pól: **ID** tabeli: **slo_wydzial**. Są więc wartościami kluczy obcych.

Typ czytelnika	wydział	Rok studiów	Data wprow	aktywny	Skype
student	2	5	2007-02-13	✓	
student	1	3	2006-09-17	✓	
student	11	5	2007-02-04	✓	
student	16	2	2006-06-16	✓	
student	7	4	2007-04-17	✓	
pracownik	1		2007-02-18	✓	
student	3	5	2006-03-24	✓	
pracownik	8		2006-06-05	✓	
student	2	4	2007-01-06	✓	
pracownik	4		2007-01-18	✓	
student	1	4	2006-03-06	✓	
student	5	3	2006-02-18	✓	
pracownik	3		2007-02-17	✓	
student	2	1	2006-09-04	✓	
student	5	2	2007-01-08	✓	
student	2	1	2006-02-10	✓	

Zastanów się: Co by się stało, gdyby operacje konwersji danych oraz zmiany typu danych zostały przeprowadzone w odwrotnej kolejności – tzn. najpierw konwersja danych tekst->liczba, a potem edycja danych w arkuszu?

8. Utwórz tabelę o nazwie: **slo_osoba_status** według wytycznych:

Tabela słownikowa: **slo_osoba_status** (przechowywane w niej będą dane o rodzaju czytelnika: pracownik/student)

Klucz	Nazwa	Typ danych	Rozmiar	Atrybuty
KLUCZ	ID	Autonumerowanie		
	status	Tekst	10	wymagane: tak

9. Wprowadź w widoku arkusza danych tej tabeli dwa rekordy: **pracownik**, **student**.

czytelnik	slo_osoba_status
ID	Status
1	pracownik
2	student
*	(Nowy)

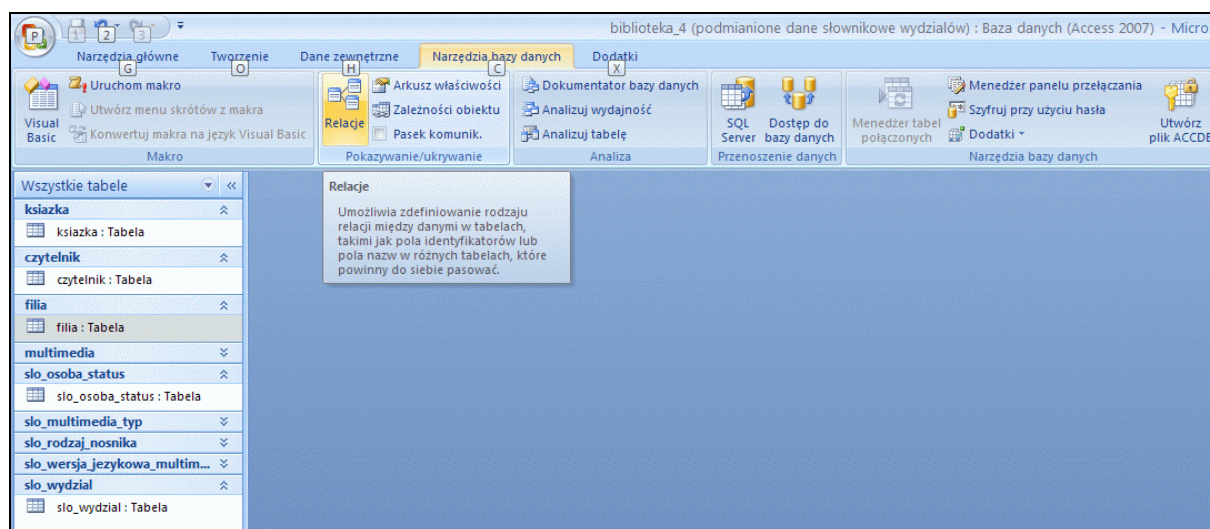
10. W sposób analogiczny jak poprzednio, zamień w tabeli: *czytelnik* wartości pól: *student* i *pracownik* na odpowiednie wartości *ID* z tabeli słownikowej: *slo_osoba_status*.

Typ czytelnika	wydział	Rok studiów	Data wprow	aktywny
2	2	5	2007-02-13	<input checked="" type="checkbox"/>
2	1	3	2006-09-17	<input checked="" type="checkbox"/>
2	11	5	2007-02-04	<input checked="" type="checkbox"/>
2	16	2	2006-06-16	<input checked="" type="checkbox"/>
2	7	4	2007-04-17	<input checked="" type="checkbox"/>
1	1		2007-02-18	<input checked="" type="checkbox"/>
2	3	5	2006-03-24	<input checked="" type="checkbox"/>
1	8		2006-06-05	<input checked="" type="checkbox"/>
2	2	4	2007-01-06	<input type="checkbox"/>
1	4		2007-01-18	<input checked="" type="checkbox"/>
2	1	4	2006-03-06	<input checked="" type="checkbox"/>
2	5	3	2006-02-18	<input checked="" type="checkbox"/>
1	3		2007-02-17	<input checked="" type="checkbox"/>
2	2	1	2006-09-04	<input checked="" type="checkbox"/>
2	5	2	2007-01-08	<input checked="" type="checkbox"/>
2	2	1	2006-02-10	<input checked="" type="checkbox"/>

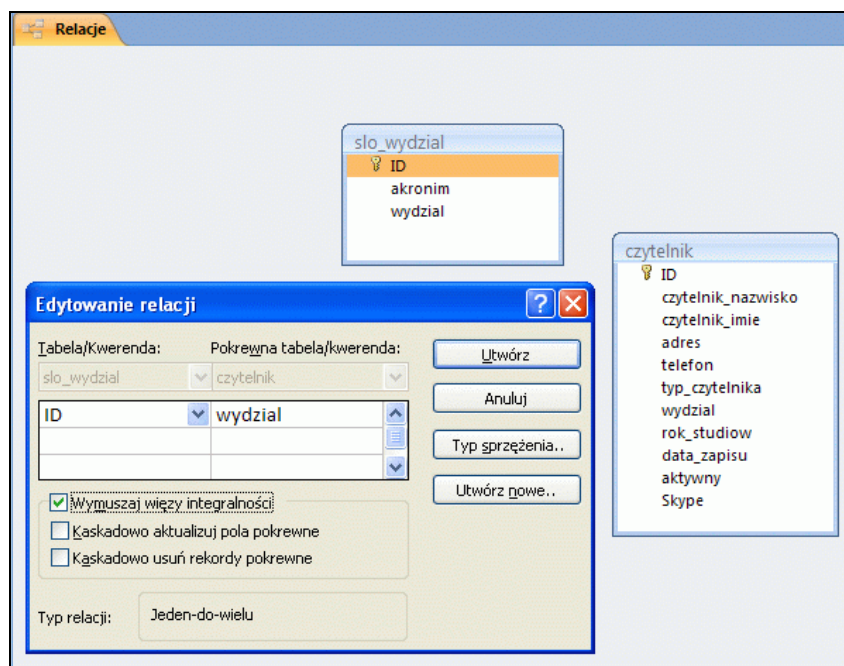
11. Zmień typ danych pola: *typ_czytelnika* tabeli: *czytelnik* z tekstowego na Liczbowy (całkowita długa).
 12. Otwórz tabelę: *filia* i przeanalizuj jej zawartość pod kątem normalizacji.
 13. Zrób odpowiednie poprawki. Wskazówka: jest jedno pole, które musi zostać zmienione.

USTALENIE RELACJI TABEL

14. Ustal wszystkie możliwe relacje między tabelami w swojej bazie danych. W tym celu na karcie: *Narzędzia bazy danych*, w grupie: *Pokazywanie/ukrywanie* wybierz polecenie *Relacje*.

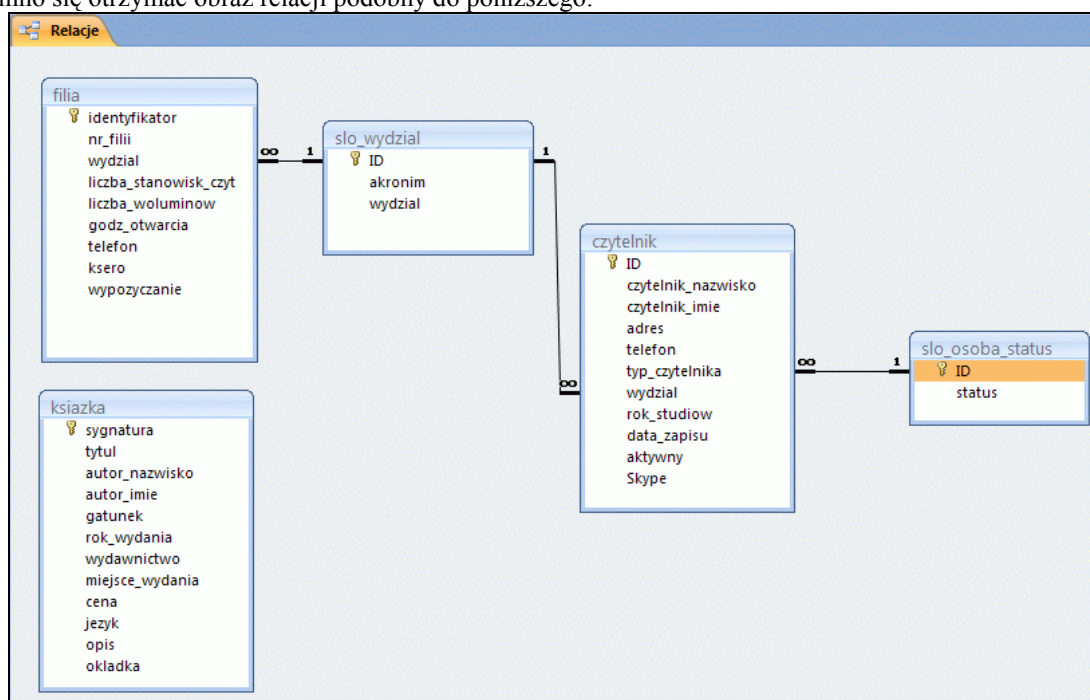


15. Dodaj na pulpit narzędzia *Relacje* table: *filia*, *czytelnik*, *ksiazka*, *slo_wydzial* i *slo_osoba_status*.
 16. Ustal wszystkie możliwe relacje między odpowiednimi polami tabel. W tym celu przeciągaj kursor od odpowiedniego pola jednej tabeli do odpowiadającego mu pola drugiej tabeli. Przy przeciąganiu trzymaj wciśnięty lewy klawisz myszy. MS Access każdorazowo będzie wyświetlać następujące okno:



Włącz zawsze opcję: „Wymuszaj więzy integralności”. Gdy wszystko jest w porządku, naciśnij przycisk: **Utwórz**.

Powinno się otrzymać obraz relacji podobny do poniższego:



Zwróć uwagę, że we wszystkich przypadkach MS Access utworzył relacje typu: **jeden-do-wielu**. Dlaczego? Jak to zinterpretować?

17. Utwórz nową tabelę o nazwie: **wypozyczenia** według wytycznych:

Tabela: **wypozyczenia** (przechowywane w niej będą dane o wypożyczonych książkach)

Klucz	Nazwa	Typ danych	Rozmiar	Atrybuty
KLUCZ	ksiazka_sygnatura	Liczba	całkowita długa	
KLUCZ	czytelnik ID	Liczba	całkowita długa	
KLUCZ	data_wypozyczenia	Data/godzina	data krótka	wartość domyślna: Date ()

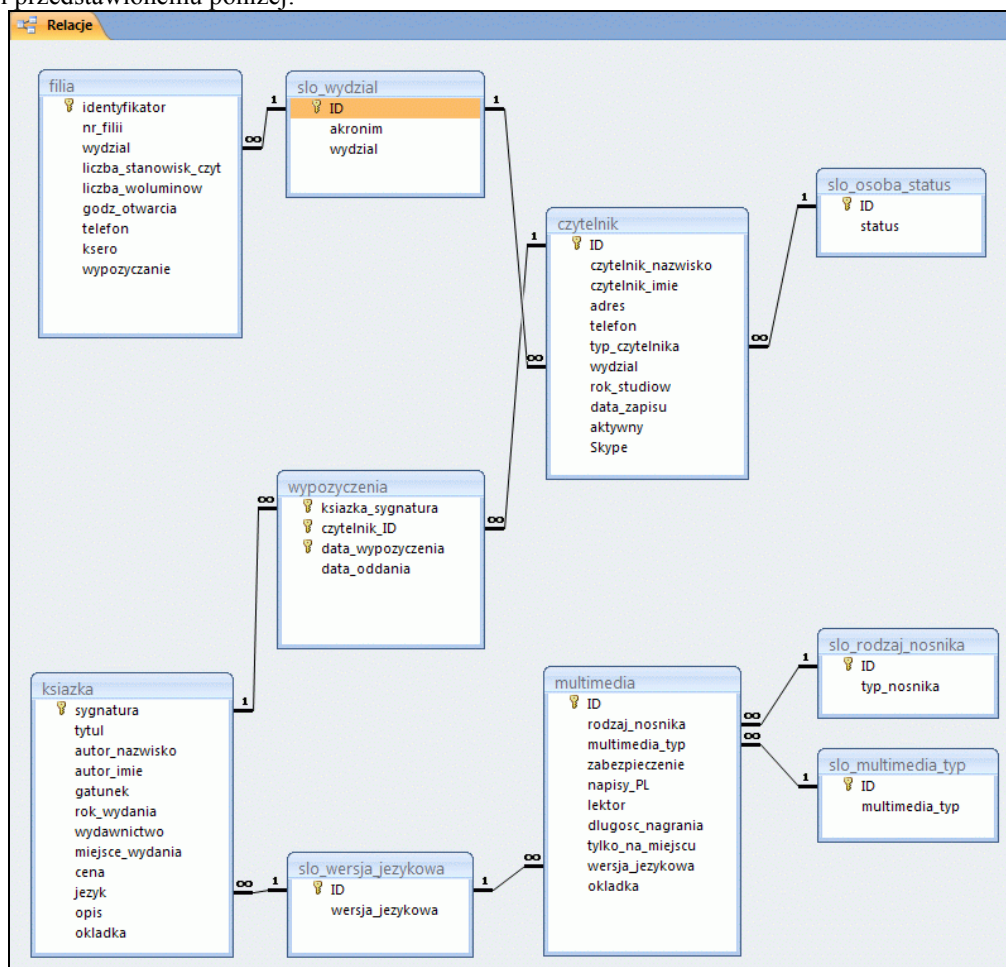
	data zwrotu	Data/godzina	data krótka	
--	-------------	--------------	-------------	--

18. Otwórz ponownie widok relacji, dodaj tabelę: **wypożyczenia** i ustal odpowiednie relacje. Wskazówka: dla tej tabeli trzeba ustalić relacje z dwiema tabelami.

ĆWICZENIE 3.2

Na poprzednim zajęciach, ćwiczenie: 2.2 polegało na samodzielnym rozwiązaniu problemu zasobów multimedialnych biblioteki. Przyjrzyj się utworzonej tabeli bądź tabelom. Czy jest/są one wykonane zgodnie z zasadami normalizacji? Jeśli nie, określ potrzebne zmiany. Czy wystarczy do tego jedna tabela? Jeśli nie, utwórz odpowiednią liczbę niezbędnych tabel słownikowych i ustal odpowiednie relacje.

W zależności od sposobu rozwiązania problemu multimediiów oraz wersji językowej książek i multimediiów, otrzymany obraz relacji w bazie danych: **biblioteka**, w mniejszym bądź większym stopniu powinien odpowiadać obrazowi przedstawionemu poniżej:



Przydatne linki:

<http://office.microsoft.com/pl-pl/access/HP051884441045.aspx>

<http://office.microsoft.com/pl-pl/access/HP051875301045.aspx>

<http://office.microsoft.com/pl-pl/access/HP051875241045.aspx>

Wykorzystywano dane i materiały internetowe autorstwa: mgr inż. Adam Czarnecki, mgr inż. Tomasz Sitek, mgr inż. Artur Ziółkowski