



**POLITECHNIKA WARSZAWSKA**

**WYDZIAŁ ELEKTRONIKI  
I TECHNIK INFORMACYJNYCH**



## **Repozytorium PW**

### **Instrukcja importowania danych**

Opracowanie:  
mgr inż. Waław Struk

Warszawa, lipiec 2013

## Spis treści

1.	Generowanie formatu BibTeX z danych ankiety jednostki (Excel).....	3
1.1.	Opis ogólny.....	3
1.2.	Kolejne kroki przy konwersji danych z formatu ankiety jednostki do formatu BibTeX:.....	4
1.3.	Uwagi.....	5
2.	Pozyskiwanie danych z Internetu za pomocą Zotero .....	6
2.1.	Instalacja i konfiguracja.....	6
2.2.	Korzystanie z zasobów Internetowych na przykładzie google scholar .....	6
2.3.	Korzystanie z zasobów Internetowych na przykładzie BazTech .....	7
3.	Import danych w formacie BibTeX do Repozytorium .....	8

## 1. Generowanie formatu BibTeX z danych ankiety jednostki (Excel)

### 1.1. Opis ogólny

W ramach dostarczonych danych przygotowanych na potrzeby ankiety jednostki występują arkusze programu Excel odpowiadające wyróżnionym typom publikacji:

- Artykuły z czasopism z listy A
- Artykuły z czasopism z listy B
- Książki – monografie
- Książki – tomy (w których opublikowano rozdziały)
- Rozdziały w języku polskim
- Rozdziały w języku angielskim
- Materiały konferencyjne

Dla każdego typu publikacji opisywanych w ankiecie jednostki, przygotowano plik konwersji (w formacie programu Excel) zawierający dwa arkusze:

1. Publikacje                      arkusz wejściowy - służy do wkopiowania danych danego typu z ankiety jednostki (jako dane wejściowe do konwersji)
2. BibTex\_nazwa-typu            arkusz konwersji - służy do uzyskania formatu BibTeX i wygenerowania odpowiedniego pliku tekstowego

Typ publikacji	Nazwa pliku konwersji i jego arkuszy
Artykuły z czasopism z listy A	A_J-BibTeX_Artykuly_A.xls Publikacje BibTeX_Artykuly
Artykuły z czasopism z listy B	A_J-BibTeX_Artykuly_A.xls Publikacje BibTeX_Artykuly
Książki – monografie	A_J-BibTeX_Ksiazki_Mono.xls Publikacje BibTeX_Ksiazki
Książki – tomy (w których opublikowano rozdziały)	A_J-BibTeX_Ksiazki_Tomy.xls Publikacje BibTeX_Tomy
Rozdziały w języku polskim	A_J-BibTeX_Rozdzialy_pl.xls Publikacje BibTeX_Rozdzialy
Rozdziały w języku angielskim	A_J-BibTeX_Rozdzialy_en.xls Publikacje BibTeX_Rozdzialy
Materiały konferencyjne	A_J-BibTeX_MatKonf.xls Publikacje BibTeX_MatKonf

Każdy z plików konwersji zawiera również funkcje napisane w języku VBA, które zasadniczo nie są widoczne dla użytkownika, ale spełniają istotną rolę przy konwersji danych.

W arkuszach konwersji zdefiniowano odpowiednie formuły konwersji zapewniające utworzenie w tych arkuszach, 1-kolumnowego wykazu publikacji w formacie BibTeX na podstawie danych zawartych w bieżącym arkuszu Publikacje.

## **1.2. Kolejne kroki przy konwersji danych z formatu ankiety jednostki do formatu BibTeX:**

1. Otworzenie pliku Excela odpowiedniego dla przetwarzanego typu danych, np. przy przetwarzaniu artykułów z czasopism (lista A), należy otworzyć plik o nazwie A\_J-BibTeX\_Artykuly\_A.xls
2. O ile arkusz Publikacje zawiera nieaktualne dane (np. z poprzednich przetwarzań wskazane wyczyszczenie tych danych)
3. Wkopiowanie do arkusza Publikacje danych z odpowiedniego arkusza ankiety jednostki, np. przy przetwarzaniu artykułów z czasopism (lista A), należy do arkusza Publikacje wkopować opisy z arkusza, który zawiera dane artykułów z listy A
4. Przejście na arkusz konwersji, np. przy przetwarzaniu artykułów, przechodzimy na arkusz BibTeX\_Artykuly
5. Arkusz konwersji zawiera w pierwszej kolumnie formuły, których zadaniem jest wygenerowanie odpowiednich danych na podstawie danych z arkusza Publikacje. Ustawiamy się na komórce Pierwszego wiersza pierwszej kolumny i po ustawieniu kursora na prawym dolnym rogu komórki przeciągamy myszką w dół, aby zdefiniowana formuła zadziałała na tylu wierszach, ile zawiera arkusz Publikacje. Może zająć potrzeba wyczyszczenia nadmiarowych wierszy pozostałych po wykorzystaniu arkusza w poprzednich konwersjach. Trzeba jednak uważać, aby nie usunąć formuły z pierwszego wiersza, aby nie utracić treści formuły.  
W ten sposób powinniśmy uzyskać w arkuszu konwersji dane z arkusza Publikacje w formacie BibTeX.
6. Po wygenerowaniu w arkuszu konwersji danych w formacie BibTeX należy zapisać je do pliku tekstowego, który będzie służył do importowania danych do Repozytorium PW. W tym celu musimy być ustawieni na arkuszu konwersji i wybieramy opcję "Plik" / "Zapisz jako" a jako typ zapisu danych "CSV (rozdzielany przecinkiem)". Następnie akceptujemy lub zmieniamy nazwę/położenie pliku i kończymy proces tworzenia pliku wyjściowego.
7. Utworzony plik wyjściowy może jeszcze wymagać konwersji kodowania znaków polskich do UTF8.  
W tym celu można posłużyć się dostępnymi konwerterami albo edytorami tekstowymi, które umożliwiają taką konwersję. Jednym z takich edytorów jest dostępny w internecie edytor Notepad++

### 1.3. Uwagi

1. Przed załadowaniem opisów rozdziałów należy koniecznie załadować dane książek i/lub tomów, aby podczas ładowania danych rozdziałów mogły być na bieżąco rozwiązywane odesłania do opisów książek/tomów.
2. W opisach materiałów konferencyjnych brak informacji dotyczących czasopisma i z tego powodu standardowo wszystkie takie materiały będą ładowane do Repozytorium jako rozdziały z książki.  
Jeśli wiadomo, że dana publikacja ukazała się jako artykuł w czasopiśmie, to dodatkowo należy w kolumnach J oraz K arkusza Publikacje (plik A\_J-BibTeX\_MatKonf.xls) wpisać odpowiednio ISSN oraz tytuł tego czasopisma. W takim przypadku publikacja będzie przenoszona do Repozytorium jako artykuł z czasopisma, a nie jako rozdział z książki.
3. W ramach testów zauważono, że wielu autorów posiada w danych jednostki drugie imię, natomiast w bazie Repozytorium PW występują oni tylko z jednym imieniem, gdyż takie dane zostały przekazane z systemu kadrowego. Ze względu na te różnice, w trakcie importowania danych z ankiety jednostki zachodzi potrzeba uzgadniania tożsamości tych autorów. Gdyby przed importowaniem tych danych uzupełnić drugie imiona, to w fazie importowania następowaloby automatyczne dopasowanie autorów.

## 2. Pozyskiwanie danych z Internetu za pomocą Zotero

### 2.1. Instalacja i konfiguracja

#### 1. Instalacja zotero

Należy wejść na stronę: <http://www.zotero.org/download/>

Wybrać zakładkę "Zotero for Firefox" i uruchomić pobieranie/instalowanie Zotero

Akceptujemy ostrzeżenia i inne polecenia procesu instalowania

#### 2. Przygotowanie wyszukiwarek internetowych

W przypadku scholar, wchodzimy na stronę: <http://scholar.google.pl/>

W ustawieniach wybieramy maksymalną liczbę rekordów na stronie (100)

#### 3. W przeglądarce FireFox, otwieramy okno Zotero - na dole z prawej strony napis zotero

#### 4. W oknie Zotero tworzymy struktury katalogowe (kolekcje i podkolekcje) w celu zapewnienia dobrej organizacji pracy z pobranymi danymi, np. tworzymy kolekcję: PW

WIS (Faculty of Environmental Engineering)

Juda-Rezler\_K

WMech

Cieslicki\_K

### 2.2. Korzystanie z zasobów Internetowych na przykładzie google scholar

#### 1. Wykonujemy wyszukiwanie na stronie scholar

Stosujemy raczej język angielski, zapytanie o autora/naukowca (z ewentualnym dodaniem Warsaw University of Technology)

Przy zapytaniu o publikacje danego autora należy poprzedzić nazwisko i imię ujęte w znaki cudzysłowu prefiksem author:. Przykładowe zapytania o publikacje autora Kulpa Krzysztof:

**author:"Kulpa K"**

daje 199 rekordów

**author:"Kulpa Krzysztof"**

daje 94 rekordy

**author:"Kulpa Krzysztof" "Warsaw University of Technology"**

daje 81 rekordów

#### 2. Przy zadawalającym wyniku przeszukiwania możemy ściągnąć odszukane dane do naszego zotero:

W oknie zotero ustawiamy się w odpowiedniej kolekcji

Uruchamiamy ściągnięcie danych do okna Zotero – należy kliknąć ikonkę Zotero, która znajduje się z prawej strony okienka adresu przeglądarki

#### 3. Dane są importowane do okna zotero, przy czym na początku mamy możliwość wskazania, które dane powinny być pobrane

4. Po pobraniu danych do okna Zotero, zaznaczamy rekordy, które chcemy importować do Repozytorium i klikamy prawym przyciskiem, aby wybrać funkcję "Eksportuj zaznaczone elementy", wybieramy format BibTeX i generujemy plik eksportu zapamiętując go pod użyteczną nazwą w katalogu roboczym na dysku
5. Utworzony plik wyjściowy może jeszcze wymagać konwersji kodowania znaków polskich do UTF8.  
W tym celu można posłużyć się dostępnymi konwerterami albo edytorami tekstowymi, które umożliwiają taką konwersję. Jednym z takich edytorów jest dostępny w internecie edytor Notepad++
6. Wygenerowany plik jest gotowy do zaimportowania do Repozytorium

### **2.3. Korzystanie z zasobów Internetowych na przykładzie BazTech**

1. Wykonujemy wyszukiwanie na stronie BazTech (<http://baztech.icm.edu.pl/>)

Na wykazie odszukanych publikacji nie ma możliwości pobrania danych do Zotero, jednak po wejściu do szczegółowego opisu publikacji można taki opis pobrać do Zotero.

Uruchamiamy ściąganie danych do okna Zotero – należy kliknąć ikonkę Zotero, która znajduje się z prawej strony okienka adresu przeglądarki

2. Wykonujemy pobranie danych do Zotero dla poszczególnych publikacji, które zamierzamy importować do Repozytorium
3. Wykonujemy pobranie danych do Zotero dla poszczególnych publikacji, które zamierzamy importować do Repozytorium
4. Po pobraniu danych do okna Zotero, zaznaczamy rekordy, które chcemy importować do Repozytorium i klikamy prawym przyciskiem, aby wybrać funkcję "Eksportuj zaznaczone elementy", wybieramy format BibTeX i generujemy plik eksportu zapamiętując go pod użyteczną nazwą w katalogu roboczym na dysku
5. Wygenerowany plik jest gotowy do zaimportowania do Repozytorium

### 3. Import danych w formacie BibTeX do Repozytorium

1. W interfejsie edytora Repozytorium wybieramy funkcję "Import"
2. W oknie importowania wybieramy "ImportBibtex"
3. Wskazujemy plik, który ma być zaimportowany (musi być w kodowaniu (UTF8))
4. Po załadowaniu pliku na serwer, zostanie zaprezentowana informacja dotycząca procesu importowania oraz przycisk "Kontynuuj":

#### Import rekordów

id: 1358238281240  
status: Rozpoczęty  
Zaimportowanych rekordów: 0/10

5. Kliknięcie przycisku "Kontynuuj" uruchamia proces importowania danych z pliku do Repozytorium
6. W czasie importowania danych mogą pojawiać się zapytania dotyczące sposobu dołączania danych, co do których system wskaże podobieństwo z danymi występującymi w Repozytorium.
7. Z jednej strony, należy dołożyć starań, aby wskazać właściwe działanie systemu w celu zapobieżenia powstawaniu zbyt wielu duplikatów i nieścisłości. Z drugiej strony powinniśmy zadbać, aby importowane dane były właściwie przyporządkowane do danych istniejących w Repozytorium.
8. Identyfikator importu (w przykładzie 1358238281240) warto zapisać w notatkach, gdyż można go potem wykorzystać do wyszukania rekordów zaimportowanych podczas danego importowania