*, Mariola Abkowicz(1), Tomasz Hoffmann(2), Michał Kozak(2)*

(1) Związek Karaimów Polskich, Katedra Studiów Azjatyckich UAM

(2) Poznańskie Centrum Superkomputerowo-Sieciowe

**Rękopisy i starodruki Karaimskiego Archiwum Cyfrowego – od plików MS Word do TEI P5**

Gromadzenie metadanych dotyczących zasobów historycznych takich jak rękopisy i starodruki wymaga określenia ich struktury. Składowanie danych w postaci plików MS Word zawierających opisy tych obiektów wydaje się rozwiązaniem prostym i efektywnym. Problem pojawia się jednak gdy chcemy ujednolicać powtarzające się informacje, poddawać je analizie oraz upublicznić szerokiemu gronu badaczy i społeczeństwu. Wówczas z pomocą przychodzi format TEI P5 (<http://www.tei-c.org/>) - standard reprezentowania tekstów, ich semantyki i metadanych, który zdobywa coraz większą popularność wśród naukowców humanistów różnych dziedzin.

W referacie zaprezentujemy przykład zastosowania standardu TEI P5 do zbudowania systemu informatycznego umożliwiającego tworzenie, przechowywanie, edycję i udostępnianie opisów obiektów historycznych należących do dziedzictwa kulturowego Karaimów. Dane były początkowo opracowywane w formie plików MS Word, a podstawowym formatem ich zapisu była tabela klucz–wartość, przy czym wartość mogła zawierać dane tekstowe jak i hierarchiczne np. listy wielopoziomowe. W ramach wspólnego projektu Poznańskiego Centrum Superkomputerowo-Sieciowego i Związku Karaimów Polskich, pracownicy Działu Bibliotek Cyfrowych i Platform Wiedzy PCSS przygotowali dedykowane narzędzie do transformacji plików MS Word do formatu TEI P5. Następnie zbudowali serwis internetowy operujący na danych w formacie TEI P5 przechowywanych w XML-owej bazie danych. Serwis jest dostępny pod adresem [http://jazyszlar.karaimi.org](http://jazyszlar.karaimi.org/). Obejmuje on obok Katalogu rękopisów i starodruków w zbiorach prywatnych również Karaimską bazę literacko-bibliograficzną, transformowaną z bazy danych MS Access.

Serwis internetowy udostępnia redaktorom dedykowane formularze edycji formatu TEI P5. Dzięki temu nie ma konieczności bezpośredniej edycji pliku XML, którego struktura jest bardzo złożona. Rozwiązanie takie nie tylko ułatwia wprowadzanie danych poprzez zastosowanie odpowiednich komponentów, ale także pozwala na sprawne korygowanie ewentualnych błędów. Dla zaawansowanych użytkowników istnieje możliwość podłączenia webowego edytora plików XML, ze stowarzyszonej z systemem bazy eXistdb (<http://www.exist-db.org/>).

Przejrzysty interfejs graficzny portalu umożliwia sprawne przeglądanie powiązanych ze sobą danych, w tym oferuje takie funkcjonalności jak wyszukiwanie po zdefiniowanych węzłach TEI P5 oraz prezentacje źródłowych obiektów w postaci zdigitalizowanej – galeria skanów.