

METODYKA KORZYSTANIA Z BAZ DANYCH WIELOJĘZYCZNEGO
INTERDYSCYPLINARNEGO TERMINOLOGICZNEGO SŁOWNIKA
I LEKSYKONU GEOINFORMATYCZNEGO KOMISJI GEOINFORMATYKI PAU
ZA POŚREDNICTWEM INTERNETU**

1. Wstęp

Stale wzrastające zapotrzebowanie na międzybranżowe ujednoczenie geoinformatycznej terminologii [1] być może uda się zaspokoić w ramach inicjatywy podjętej przez Komisję Geoinformatyki Polskiej Akademii Umiejętności w Krakowie. Interdyscyplinarny Terminologiczny Słownik Geoinformatyczny, powstający w ramach formuły otwartej [4], umożliwiającej wszystkim zainteresowanym bieżący udział w redakcji słownika na słownikowych stronach www, powinien nie tylko zaspokoić z czasem zapotrzebowanie, ale będzie bieżąco umożliwiał śledzenie powstających haseł i otwartą dyskusję [5]. Będzie też umożliwiał nieograniczone wprowadzanie do słownika terminologii równoległej (synonimów) używanych lokalnie w różnych zespołach, też w zespołach produkcyjnych. Zapewni to słownikowi właściwą witalność, nadążanie za potrzebami rozwijającej się dziedziny wiedzy i techniki.

Wielojęzyczny Interdyscyplinarny Terminologiczny Słownik i Leksykon Geoinformatyczny Komisji Geoinformatyki Polskiej Akademii Umiejętności w Krakowie, zwany dalej Słownikiem, powstaje jako efekt pracy zespołów specjalistów z wielu dyscyplin związanych z geoinformatyką. Słownik dzieli się na kilka autonomicznych działów przedmiotowych:

- o Fotogrametria i Teledetekcja
- o Geologia, Geofizyka i Ochrona Środowiska
- o Geografia
- o Kartografia

* Akademia Górniczo-Hutnicza, Wydział Geodezji Górniczej i Inżynierii Środowiska, Zakład Fotogrametrii i Informatyki Teledetekcyjnej

** Prace związane z opracowaniem systemu słownikowego oprogramowania przystosowanego do pracy w sieci internetowej realizowane są w Zakładzie Fotogrametrii i Informatyki Teledetekcyjnej AGH w ramach tematu badań statutowych nr.11.150.459, oraz w ramach specjalnej dotacji Rektora AGH (badania własne nr 10.150.710), a także przy finansowym wspomóżeniu ze strony Fundacji Fanni i Teodora Blachutów, oraz Fundacji Kościuszkowskiej, Amerykańskie Centrum na rzecz Kultury Polskiej z funduszami zapewnionymi przez fundację im. Alfreda Jurzykowskiego

- o Geodezja i Systemy Informacji Przestrzennej
- o Górnictwo i Szkody Górnicze
- o Inżynieria Środowiska
- o Gleboznawstwo
- o Informatyka

Każdy z działów prowadzi w zasadzie jedną bazę danych, w której znajdują się terminy specjalistyczne, ich synonimy, krótka definicja terminu, przykłady użycia terminu, obszernie omówienie terminu, oraz uwagi. Informacje te mogą być podawane dla każdego hasła w kilku językach.

W razie potrzeby w dziale przedmiotowym Słownika można zakładać dodatkowe, autonomiczne bazy danych. I tak np. dotychczas w dziale Fotogrametria i Teledetekcja są dwie bazy: stara i nowa. Nowa baza służy do bieżącego edytowania haseł słownikowych. Stara baza zawiera cyfrową kopię wydanego drukiem w 1992 roku słownika [2], opracowanego przez zespół specjalistów pod redakcją profesora Zbigniewa Sitka (słownik zawiera ponad 1200 terminów w 5 językach; wpisywanie tego materiału wykonano dotychczas w 80%).

Otwarta, internetowa formuła pracy nad Słownikiem Geoinformatycznym PAU powoduje, że jest on permanentnie aktualizowany [4], a efekty aktualizacji, wraz z całą bazą danych słownikowych, są udostępniane bez ograniczeń, za pośrednictwem sieci internetowej.

Użytkownicy Słownika mogą również pośrednio uczestniczyć w wypełnianiu go treścią poprzez uwagi merytoryczne wpisywane w specjalnej rubryce programu słownikowego, albo w dowiązaniu do wybranego hasła (można np. proponować synonimy), albo w ogólnej rubryce uwag i komentarzy czytelników, do której można też wpisywać propozycje nowych haseł, nie zawartych jeszcze w Słowniku.

W wykazie redaktorów i współautorów dostępne są adresy, również poczty elektronicznej, które ułatwiają nawiązanie bezpośrednich kontaktów i wpływanie, poprzez dyskusję, na zawartość Słownika.

Słownik Geoinformatyczny PAU jest pomyślany w ten sposób, że można go w każdej chwili poszerzyć o dowolny dział tematyczny, lub dodać do każdego działu nową wersję językową. Językiem wiążącym opracowania w różnych wersjach językowych jest język angielski.

Przedmiotowe bazy danych będą mogły w każdej chwili, w razie potrzeby, zostać scalone w ogólnej bazie danych Słownika Geoinformatycznego PAU. Po kilku latach, gdy etap opracowania zostanie zamknięty, można będzie wydać drukiem nasz Słownik. Prace redakcyjne w zespołach przedmiotowych będą jednak trwały permanentnie, ponieważ rozwój nauki i techniki kreuje nowe słownictwo, które powinno być uwzględniane w Słowniku.

2. Rozpoczynanie pracy ze Słownikiem

Pracę ze Słownikiem, czyli wyszukiwanie haseł, zaczynamy od wpisania adresu strony www w dowolnej przeglądarce internetowej. Adres strony internetowej, pod którym znajduje się Słownik jest następujący:

<http://sownik.fotogrametria.agh.edu.pl>

Po wpisaniu adresu pojawia się strona główna Słownika (Rys.1). Znajdują się na niej informacje ogólne o Słowniku oraz pola wyboru. W pierwszej kolejności możemy wybrać wiodącą opcję językową: polską albo angielską. Polska jest ustawiona automatycznie. Aby wybrać angielską należy kliknąć na flagę angielską. Powrót do wersji polskiej odbywa się analogicznie.

Kolejnym polem wyboru jest rodzaj działu Słownika, w którym chcemy pracować. Wybór działu jest bardzo ważnym elementem, ponieważ wiąże się z linkiem do konkretnej bazy danych. W Słowniku mamy 9 działów tematycznych (9 osobnych baz danych, połączonych ze sobą numerem hasła). Podczas pracy ze Słownikiem, możemy w każdej chwili zidentyfikować nazwę działu w którym pracujemy, ponieważ pojawia się ona zawsze na górnym, zielonym pasku strony. W dowolnym momencie możemy zmienić dział tematyczny powracając do pierwszej strony Słownika, poprzez kliknięcie na wspomnianym, górnym zielonym pasku na napisie *Słownik Geoinformatyczny PAU*.

Kolejnym polem wyboru są pola językowe. W pierwszej kolejności wybieramy język wiodący opracowania (język nr 1). Język wiodący jest to język w którym podawane powinny być nagłówki i objaśnienia. W tym języku będą prowadzone wyszukiwania i jednocześnie będzie to język w którym będzie można wprowadzać nowe terminy. Docelowo w języku tym, po jego wyborze, będą podawane również pola menu pozostałych podstron Słownika.

Język drugi jest językiem, w którym będzie można wprowadzać nowe terminy niejako równoległe z pierwszym, tzn. wprowadzając nowy termin , należy wprowadzić go w dwóch pierwszych językach – a przynajmniej pola „termin” i „definicję”

Wybór pozostałych języków (trzeci, czwarty i piąty) powoduje, że będziemy mieli do nich dostęp w momencie wyszukiwania haseł, tj. w tych językach będzie można znaleźć szukane hasła i je porównać. Powrót do wyboru języków może być dokonany w dowolnym momencie pracy ze Słownikiem; przez kliknięcie na górnym zielonym pasku na napisie *Słownik Geoinformatyczny PAU* nastąpi powrót do pierwszej strony Słownika.

Następnie w zależności od celu pracy wybieramy przycisk: *Szukaj*, *Edytuj*, *Weryfikuj*.

Opcja *Szukaj* pozwala dowolnym użytkownikom na pracę ze Słownikiem (wyszukiwanie haseł i czytanie opisów). Po naciśnięciu tego przycisku, przechodzimy do kolejnej strony.

Dostęp do części programu aktywowanych po naciśnięciu opcji *Edytuj* i *Weryfikuj* jest ograniczony. Mogą dokonywać tego tylko osoby uprawnione. *Edycja* pozwala na dodawanie nowych rekordów i edytowanie istniejących. Opcja *Weryfikuj* służy recenzentom do korekty zaproponowanych haseł .

W każdej chwili możemy skorzystać z pomocy Słownika, poprzez naciśnięcie pola *Pomoc*.



Rys.1 Strona główna Słownika

3. Wyszukiwanie haseł w systemie wielojęzycznym

Po wybraniu odpowiednich pól możemy przystąpić do wyszukiwania haseł. Jak już wspomniano w poprzednim rozdziale, aby tego dokonać należy wybrać opcję *Szukaj*. Dostęp do tej części Słownika jest nieograniczony.

Po dokonaniu odpowiednich ustawień na pierwszej stronie Słownika (patrz poprzedni punkt) i po naciśnięciu przycisku *Szukaj*, przechodzimy do strony wyszukiwania. Charakteryzuje się ona tym, że na górnym zielonym pasku pojawia się nazwa działu

tematycznego, który wybraliśmy na pierwszej stronie oraz pojawiają się nowe pola wyboru. Ponieważ większość działów może się opierać częściowo na bazach już istniejących (jak np. "Słownik pięcioletni z zakresu fotogrametrii i teledetekcji" wydany pod redakcją prof. Zbigniewa Sitka w roku 1990 przez AGH [2]), istnieje możliwość wyboru bazy Słownika starej i nowej dla każdego działu.

Przykładowo wspomniany „Słownik pięcioletni z zakresu fotogrametrii i teledetekcji” w 2004 roku został wprowadzony do bazy danych Słownika Geoinformatycznego PAU pod nazwą Stara Baza Danych Działu Fotogrametria i Teledetekcja. Natomiast nowa baza tego działu powstaje aktualnie i może być permanentnie przedawana i uzupełniana.

Kolejne dwa pola wyboru pozwalają na wybór języka wyszukiwań hasła oraz formę prezentacji wyników. Przykład przedstawiono na Rys.2.

Wybierz	język	wyszukiwania	hasel:	
<input type="checkbox"/> <i>Polski</i>	<input type="checkbox"/> <i>Angielski</i>	<input type="checkbox"/> <i>Francuski</i>	<input type="checkbox"/> <i>Niemiecki</i>	<input type="checkbox"/> <i>Rosyjski</i>
Wybierz	prezentację	wyników	przeszukiwania	bazy:
<input type="checkbox"/> <i>Podstawowa</i>	<input type="checkbox"/> <i>użytkownika</i>	<input type="checkbox"/> <i>kompletna</i>		

Rys.2.

Po wyborze odpowiednich ustawień można wprowadzić wyszukiwany wyraz i nacisnąć przycisk *Szukaj*. Słownik przeszukuje wszystkie hasła znajdujące się w tym momencie w konkretnej bazie (wybranim działem tematycznym) i szuka w języku wybranym w polu *Wybierz język wyszukiwania hasel* (przeszukiwane są pola termin i termin robocze w wybranej bazie).

W Słowniku jest wprowadzona dodatkowa możliwość wyszukiwania dowolnego ciągu znaków występującego w jakimkolwiek polu Słownika w całej bazie, poprzez zaznaczenie opcji *Szukaj wszędzie w bazie*.

Sposób prezentacji wyników wybierany jest opcją: *Wybierz prezentację wyników przeszukiwania bazy*. Pozwala ona na wybór pól bazy jakie zostaną docelowo zaprezentowane na ekranie. Po wybraniu opcji: *Podstawowa* pojawiają się pola: *Termin*, *Synonimy*, *Definicja*, *Termin Angielski*, *Synonimy Angielskie*. W opcji *Kompletna* pojawia się pełny zakres pól z bazy danych dla danego działu.

Po wyborze opcji *Użytkownika*, program daje możliwość użytkownikowi dowolnego doboru pól do wyświetlania. Jest to opcja wygodna szczególnie dla tych, którzy często korzystają z własnych udogodnień systemu.

Słownik dodatkowo podświetla na niebiesko ciąg znaków, który został zdefiniowany do poszukiwania, co znacznie ułatwia pracę.

Każde hasło posiada również dodatkowo pole *Dyskusja*, które umożliwia każdemu użytkownikowi przeczytanie i wprowadzenie aktualnych opinii i uwag na temat danego terminu.

Pełną informację o danym hasle można uzyskać klikając na nim. Wchodzimy wówczas do trybu przeglądaj i wyświetlane są wszystkie zapisy znajdujące się w danym hasle.

4. Gwarancja trwałości słownikowej bazy danych

Praktyczne zrealizowanie idei słownika wymagało połączenia kilku elementów odpowiadających za poszczególne fazy przechowywania i dostępu do danych. Rozwiązanie musiały cechować m.in. uniwersalność (różne rodzaje danych – tekst, liczby, obraz), możliwość dostępu za pośrednictwem sieci komputerowej oraz niezawodność, stabilność i szybkość działania. Należało zdecydować o tym jakiego użyć systemu operacyjnego, którą bazę danych wybrać oraz jaki zastosować serwer realizujący usługę dostępu do danych przy pomocy protokołu http. Należało również zadbać o odpowiednie zabezpieczenie danych.

Działający obecnie system charakteryzują wszystkie podstawowe cechy założone w projekcie i możliwy jest jego dalszy rozwój. Pomimo stosowania jednych z najlepszych dostępnych aktualnie technik pracy w sieci i odpowiednich zabezpieczeń przed niepożądanym dostępem, w dalszym ciągu pozostaje jednak problem błędów czynnika ludzkiego i awarii samego sprzętu, których nie można wyeliminować. Mamy do czynienia z ważnymi danymi, które gromadzi się wysiłkiem specjalistów szeregu dyscyplin, a problem bezpiecznego przechowywania i udostępniania zawartości baz słownikowych jest sprawą priorytetową. Dlatego w trakcie opracowania i testów jest system zapasowych kopii danych. Podstawowym założeniem tego systemu jest automatyczne tworzenie kopii bezpieczeństwa w określonym interwale czasowym oraz przechowywanie tych danych na fizycznie innym serwerze. Automatyzacja zapisu kopii baz słownikowych jest konieczna ze względu na potrzebę eliminacji czynnika ludzkiego z tego procesu. Przechowywanie kopii bezpieczeństwa na odrębnym fizycznie serwerze praktycznie całkowicie eliminuje ryzyko utraty danych w przypadku awarii sprzętu. Dla realizacji tych założeń wykonano już szereg odpowiednich kroków gwarantujących bezpieczeństwo danych po wprowadzeniu słownika do użytku publicznego. Został zakupiony dodatkowy komputer i utworzony serwer pracujący w charakterze właśnie serwera przechowującego zapasowe pliki bazy danych. W trakcie

tworzenia i jednocześnie testów jest system odpowiednich skryptów zarządzających automatycznym tworzeniem, transportem i selekcją kopii zapasowych, zupełnie niezależnych od warstwy roboczej dostępnej na serwerze poprzez sieć internetu. Innymi słowy: przypadkowe skasowanie nawet całkowicie roboczej bazy roboczej nie powoduje kompletnej utraty danych.

Godny zaakcentowania jest fakt, że wszystkie stosowane w projekcie rozwiązania mają charakter otwartego źródła (*open source*), co pociąga za sobą wysoką adaptacyjność do różnych systemów operacyjnych. Stosunkowo niewielkim nakładem pracy można by, w razie potrzeby, uruchomić podobny zestaw na wszystkich systemach typu UNIX, w tym prawie wszystkich dystrybucjach Linuxa, FreeBSD, Solaris, QNX itp. Możliwe jest to nawet w systemie Windows, niestety kosztem mniejszej stabilności i wydajności. Inną zaletą jest minimalne prawdopodobieństwo stagnacji oprogramowania, ze względu na brak czynników komercyjnych, dostępny kod źródłowy i charakter licencji którą objęte jest oprogramowanie. W przypadku produktów komercyjnych czasem ma to miejsce, np. ze względu na zakończenie działalności danej firmy. W przypadku *open source* rozwój trwa nieprzerwanie, a każdy może sam udoskonalać istniejące źródło i dostosowywać je do swoich potrzeb, nawet bez udziału autorów.

5. Zakończenie

Rozwój geoinformatyki jest coraz szybszy i obejmuje ona coraz szerszy zakres tematyczny. Wiąże się to z dynamicznym rozwojem terminologii geoinformatycznej, często niezależnie w różnych dyscyplinach naukowych i technicznych. Daje się zauważać niejednoznaczność interpretacji tekstów, wynikająca z baraku uzgodnień terminologicznych między dyscyplinami. Problem ten dostrzegany jest przez wielu specjalistów, a znalazł wyraz na forum Komisji Geoinformatyki PAU w dyskusji zorganizowanej przez jej przewodniczącego, Profesora Janusza Kotlarczyka [1]. Wyrazem zrozumienia dla potrzeby ujednoczenia terminologii jest słownik fotogrametryczno-teledetekcyjny opracowany pod redakcją Profesora Zbigniewa Sitka [2] i włączony do bazy danych naszego Słownika internetowego, oraz, wydany ostatnio „Leksykon geomatyczny” autorstwa Profesora Jerzego Gaździckiego [3], dotyczący terminologii typowej dla geomatyki, geodezyjnego zakresu zastosowań geoinformatyki.

Stale wzrastające zapotrzebowanie na międzybranżowe ujednoczenie geoinformatycznej terminologii być może uda się zaspokoić w ramach inicjatywy podjętej przez Komisję Geoinformatyki Polskiej Akademii Umiejętności w Krakowie.

Interdyscyplinarny Terminologiczny Słownik Geoinformatyczny, powstający w ramach formuły otwartej [4], umożliwiającej wszystkim zainteresowanym bieżący udział w redakcji słownika na słownikowych stronach www, powinien nie tylko zaspokoić z czasem zapotrzebowanie, ale będzie bieżąco umożliwiał śledzenie powstających haseł i otwartą dyskusję [5]. Będzie też umożliwiał nieograniczone wprowadzanie do słownika terminologii równoległej (synonimów) używanych lokalnie w różnych zespołach, też w zespołach produkcyjnych. Zapewni to słownikowi właściwą witalność, nadszanie za potrzebami rozwijającej się dziedziny wiedzy i techniki.

Mamy nadzieję, że możliwość wypowiedzania się w otwartej dyskusji na tematy terminologiczne wpłynie też dodatnio na staranność języka specjalistycznego, który używamy.

Bibliografia

- [1] Kotlarczyk J.: „Dyskusja panelowa poświęcona zagadnieniom terminologicznym geoinformatyki”. Edycja: Janusz Kotlarczyk, Geoinformatica Polonica, 3:2001, PAU, Kraków
- [2] „Słownik pięcioletni z zakresu fotogrametrii i teledetekcji z objaśnieniami w języku polskim” pod redakcją Z.Sitka, Zakład Fotogrametrii AGH, Kraków 1990...
- [3] Gaździcki J.: Leksykon geomatyczny, Polskie Towarzystwo Informacji Przestrzennej, Warszawa 2001
- [4] Jachimski J.: „Design of an open formulae for the interdisciplinary multilingual terminological dictionary for geoinformatics”. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences Vol.XXXV, part B6, Comm.VI, Istanbul 2004.
- [5] Jachimski J., Mikrut S., Majewski M. „Struktura bazy danych Wielojęzycznego, Interdyscyplinarnego Terminologicznego Słownika i Leksykonu Geoinformatycznego Komisji Geoinformatyki PAU”. Półrocznik Akademii Górniczo-Hutniczej „Geodezja”. Wydawnictwo AGH, Kraków (w druku)