

ANDRZEJ BLIKLE  
Instytut Podstaw Informatyki PAN  
ORCID 0000-0002-8534-1108

WITOLD STANISZKIS  
Intelligent Document Engineering Sp. z o.o.

## POLSKO-ANGIELSKI CYFROWY TEZAUZUS TERMINÓW INFORMATYCZNYCH

### 1. O POTRZEBIE DOBREJ POWSZECHNEJ DYDAKTYKI

W dniu 3 lipca 2013 roku odbyła się w Sali Kolumnowej Pałacu Prezydenckiego konferencja inauguracyjna powołanie *Szerokiego Porozumienia na rzecz Umiejętności Cyfrowych w Polsce*, którym stała się nieformalna grupa instytucji, organizacji i firm gotowych wspierać rozwój kompetencji cyfrowych Polaków. Inicjatorami Porozumienia byli Minister Cyfryzacji i Administracji Michał Boni oraz lider cyfryzacji w Polsce Włodzimierz Marciński. Misją porozumienia miało być działanie na rzecz zdobywania przez obywateli naszego kraju cyfrowych umiejętności niezbędnych zarówno na rynku pracy, jak i w życiu publicznym. Porozumienie zostało objęte honorowym patronatem Prezydenta Bronisława Komorowskiego, a w spotkaniu udział wzięła m.in. wiceprzewodnicząca Komisji Europejskiej i jednocześnie komisarz europejski ds. agendy cyfrowej Neelie Kroes.

Uczestnikom spotkania rozdano arcyciekawy dokument: *Manifest w sprawie e-umiejętności* [del Pilar Castillo i in. 2012, 10]<sup>1</sup>, będący dziełem wielu autorów, ze wstępem Antonio Tajaniego, wiceprzewodniczącego Komisji Europejskiej ds. Przemysłu i Przedsiębiorczości. Z tego dokumentu dowiadujemy się, że Europa cierpi na poważny deficyt ludzi z profesjonalnym przygotowaniem informatycznym, a w europejskiej gospodarce cyfrowej marginalizowanych jest 300 milionów osób. Mimo że bezrobocie wśród młodzieży sięga średnio 22%, pracodawcy nie mogą znaleźć pracowników na stanowiska wymagające przygotowania naukowego i technicznego.

W UE tylko niecałe 1,5% osób w wieku 20–29 lat studiuje nauki ścisłe i techniczne na poziomie akademickim. Wpływ na tę sytuację mają złe programy nauczania, a problem zaczyna się już na poziomie szkoły podstawowej. Informatyka jest w większości krajów unijnych przedmiotem fakultatywnym. Chlubny wyjątek stanowią Szwajcaria i Austria, gdzie informatyka jest przedmiotem obowiązkowym

---

<sup>1</sup> International Data Corporation przeprowadziło na ten temat badania przed rokiem 2012.

(w Austrii od pierwszej klasy szkoły podstawowej). A dzieje się tak, mimo że zdaniem IDC [del Pilar Castillo i in. 2012, 10], do roku 2015 aż 90% wszystkich stanowisk pracy będzie wymagało podstawowych e-umiejętności (badania).

Interesujące jest też obalenie mitu o powszechnym informatycznym przygotowaniu młodych Europejczyków. Okazuje się, że zaledwie 25% młodych ludzi w UE określa siebie jako „mających wysoki poziom kompetencji informatycznych”, przez co rozumieją umiejętność korzystania z wyszukiwarek internetowych, umieszczania wiadomości na forach dyskusyjnych, wysyłania e-maili (dumnie podkreślają, że „z załącznikami”), dokonywania zakupu w Internecie plików muzycznych i wymienia się tymi plikami. W tej sytuacji nie dziwi fakt, że aż 41% do 56% firm we wszystkich sektorach bezskutecznie poszukuje informatyków, którzy zajęliby się ich systemami informatycznymi [del Pilar Castillo i in. 2012, 39].

Autorzy *Manifestu* wskazują również na fakt, że kraje takie jak Indie i Chiny, które do tej pory łagodziły deficyt informatyków w Europie (i nie tylko), wkrótce będą potrzebowały tych specjalistów u siebie. A o szybkości, z jaką rośnie na świecie zapotrzebowanie na usługi informatyczne, może świadczyć choćby fakt, że podczas gdy w roku 2003 wysyłano dziennie 12 milionów e-maili, dziesięć lat później krążyło ich dziennie już 247 miliardów. Wykładniczo rośnie też wielkość oprogramowania mierzona liczbą linii kodu. Podczas gdy kod Unixa 1.0 w roku 1971 obejmował 20 tys. linii, kod Windows-7 już 40 mln linii, a amerykański portal zdrowotny healthcare.gov – 500 mln [OfficeObjects® Ontology Manager – Podręcznik Eksperta 2014].

Gdzie więc szukać rezerw? Zdaniem autorów *Manifestu* – w prawidłowo kształtowanych programach edukacyjnych od szkoły podstawowej po uczelnie wyższe, a także w podniesieniu prestiżu zawodu informatyka na poziomie firm, gdzie informatycy powinni być współtwórcami projektów innowacji, a nie wyłącznie konserwatorami sprzętu i oprogramowania.

## 2. DLACZEGO DLA EDUKACJI WAŻNA JEST TERMINOLOGIA

Nie ulega wątpliwości, że dla rozwinięcia powszechnej edukacji w obszarze informatyki niezbędne jest opracowanie i stałe aktualizowanie terminologii fachowej z tej dziedziny. Obecnie wielu wykładowców woli posługiwać się terminami angielskimi, obawiając się – zresztą całkiem słusznie – niespójności terminologicznej występującej między różnymi uczelniami, co spowoduje, że absolwenci tych uczelni w przyszłości nie będą mogli się porozumieć. Kwitną więc terminy w rodzaju „apgrejdowanie łebów”. Ta przypadłość dotyczy zresztą nie tylko języka informatyki. W jeszcze chyba większym stopniu anglicyzmami zachwaszczony jest język biznesu. Oto co posłyszeliśmy jakiś czas temu z ust konferencyjnego prelegenta: „nasze sejsforsory dobrze performują, bo są sfokusowane na masmarkecie”.

Używanie terminów anglojęzycznych ma jednak dwie zasadnicze wady:

- czyni język polski mniej zrozumiałym dla osób nieznających języka angielskiego,
- niszczy poczucie, że język jest pewnym dobrem publicznym, który należy chronić, by Polacy mogli się porozumieć między sobą w sposób pełny i zrozumiały.

W tym właśnie celu nasz zespół postanowił podjąć wyzwanie uruchomienia projektu, który zaowocowałby powstaniem powszechnie dostępnego, profesjonalnego i na bieżąco aktualizowanego encyklopedycznego słownika polskich terminów z obszaru szeroko rozumianej informatyki z odniesieniami do terminologii anglojęzycznej.

### 3. BIZNES, PRAWO I PRAWODAWSTWO

Istotne znaczenie informatyki w nowoczesnej gospodarce i administracji publicznej powoduje, iż ta właśnie dziedzina jest podstawowym czynnikiem rozwoju społecznego i gospodarczego, a tym samym istotnym elementem konkurencyjności poszczególnych krajów i związków gospodarczo-politycznych, takich jak Unia Europejska.

Mimo dynamicznego rozwoju informatyki, zarówno w zakresie badań i rozwoju, jak i zastosowań gospodarczych, trudno jest uznać, że ta dziedzina osiągnęła poziom dojrzałości porównywalny z takimi tradycyjnymi obszarami, jak budownictwo czy gospodarka rolna. Jednocześnie pole działalności informatycznej jest bardzo rozległe – obejmuje wiele dyscyplin badawczych od badań podstawowych w dziedzinie matematyki i inżynierskich w dziedzinie elektroniki, przez rozwój metod i technik tworzenia oprogramowania, do zaawansowanych rozwiązań aplikacyjnych.

W konsekwencji język zawodowy informatyki powstaje w sposób spontaniczny w rytmie osiągania kolejnych poziomów rozwoju tej dziedziny, a wprowadzana terminologia często wynika ze szczególnych cech produktów wprowadzanych na rynek przez konkurujących ze sobą producentów. Przykładem prób uporządkowania terminologii w języku angielskim stosowanej w różnych obszarach informatyki są ukazujące się taksonomie i tezaury terminologii informatycznych.

Polska terminologia informatyczna, niezależnie od wszystkich wymienionych już czynników, nie powstaje zazwyczaj w ramach prowadzonych w kraju prac badawczo-rozwojowych czy produkcyjnych, lecz stanowi mniej lub bardziej wierne odbicie znaczeń terminów angielskich powstających w wiodących krajach produkcji i zastosowań informatyki. Dodatkowo chaos terminologiczny wynika z zupełnej dowolności wprowadzania nowych terminów informatycznych przez osoby przekładające na język polski techniczne teksty z języka angielskiego.

W pracach badawczych i rozwojowych, a także do pewnego stopnia w edukacji, można uzyskać porozumienie bazując na terminologii angielskiej. Takie rozwiązanie jest jednak niemożliwe w biznesie i w prawie. Dodatkowym czynnikiem ryzyka w tym obszarze jest polskie prawo zamówień publicznych, które dając absolutną władzę instytucjom zamawiającym, nie dopuszcza żadnej interpretacji poza nie zawsze uzasadnioną intuicją osób tworzących opis warunków zamówienia publicznego.

Ponieważ w ciągu ostatnich dziesięciu lat, między innymi dzięki otrzymanym funduszom spójności z Unii Europejskiej, osiągnęliśmy miliardowy poziom wydatków administracji publicznej na zastosowania informatyki, taki stan niejednokrotnie prowadzi do powstawania ogromnych strat finansowych zarówno po stronie wykonawców, jak i zamawiających.

Często rozstrzygnięcia sądów powszechnych, opierające się na opiniach nie zawsze w pełni merytorycznie przygotowanych biegłych, sięgają milionowych kwot, stanowiąc

istotny czynnik podnoszący ryzyko działalności gospodarczej w dziedzinie informatyki. Oczywiście taki stan rzeczy, niezależnie od jego korupcyjności, powoduje wzrost cen za dostawy i usługi informatyczne.

Proponowane rozwiązanie, wprowadzające tezaurus polskiej terminologii informatycznej, w naszej opinii – w odpowiednio długim okresie – doprowadzi do profesjonalizacji wymiany handlowej na rynku informatycznym i przyczyni się do podniesienia merytorycznego poziomu wprowadzanych rozwiązań informatycznych w sferze administracji publicznej oraz prawnej, a także w ogólnym obrocie gospodarczym.

#### 4. PRZEDMIOT PROJEKTU

Projekt jest poświęcony realizacji i publikacji w sieci Internet ontologicznego Polsko-Angielskiego Tezaurusu Terminów Informatycznych (PATTI). Pilotażowa wersja tezaurusu PATTI obejmie podstawowe terminy z obszaru szeroko rozumianej informatyki z odniesieniami do terminów anglojęzycznych. Tezaurus powinien obejmować również definicje terminów z odniesieniami do literatury.

Celem przedsięwzięcia jest wdrożenie tezaurusu dostępnego powszechnie w sieci Internet i rozwijanego przez kwalifikowanych wolontariuszy pod nadzorem zespołu redaktorów – członków Zespołu Terminologii Informatycznej RJP.

Proponowany model tezaurusu, pokazany na rysunku 1, może zostać zrealizowany jako sieć pojęć umożliwiającą swobodną nawigację pomiędzy terminami zgodnie z modelem Topic Maps (ISO 13250). Oczywiście nic nie stoi na przeszkodzie, aby rozwinąć przedstawiony model klas o dalsze powiązania i pojęcia. Planujemy, że ostateczna weryfikacja modelu pojęciowego tezaurusu nastąpi w trakcie wdrożenia pilotażowego rozwiązania.

Interfejs użytkownika może być generowany automatycznie dla każdej klasy pojęć zgodnie z jej definicją. Alternatywnym pojęciem implementacyjnym jest stworzenie szablonów HTML5 dla każdej klasy pojęcia lub wybranych klas, aby uzyskać pożądane walory estetyki i ergonomii.

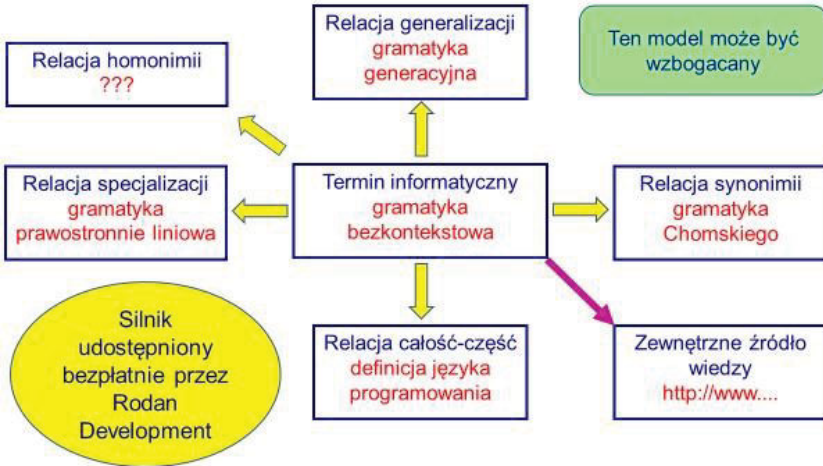
Zastosowanie modelu sieci pojęciowej obejmującej dobrze określone powiązania występujące pomiędzy terminami informatycznymi pozwoli na zachowanie odpowiedniej dyscypliny w zakresie definiowania znaczenia i kontekstu użycia wprowadzanej terminologii.

Stworzony zasób wiedzy terminologicznej zorganizowanej na podstawie standardu Topic Maps będzie zawierał terminy informatyczne w języku polskim i powiązaną z nimi terminologię w języku angielskim.

Wprowadzenie klasy pojęcia „dziedzina informatyki” pozwoli na zbudowanie taksonomicznej struktury terminów informatycznych, która może być wykorzystywana jako narzędzie klasyfikacji zasobów wiedzy informatycznej, takich jak między innymi publikacje naukowe, oprogramowanie i programy dydaktyczne. Przykładem takiej taksonomii jest opracowana przez wolontariuszy z organizacji Association for Computing Machinery (ACM) w USA klasyfikacja dziedzin informatyki (ang. Computing Classification System) [<http://taxonomies.labs.crossref.org/?p=109>].

Przykład opisu terminu informatycznego *proces biznesowy*, który przedstawia rysunek 2, jest elementem prototypu tezaurusu zrealizowanego w oparciu o produkt

### Model sieci pojęciowej tezaurusa

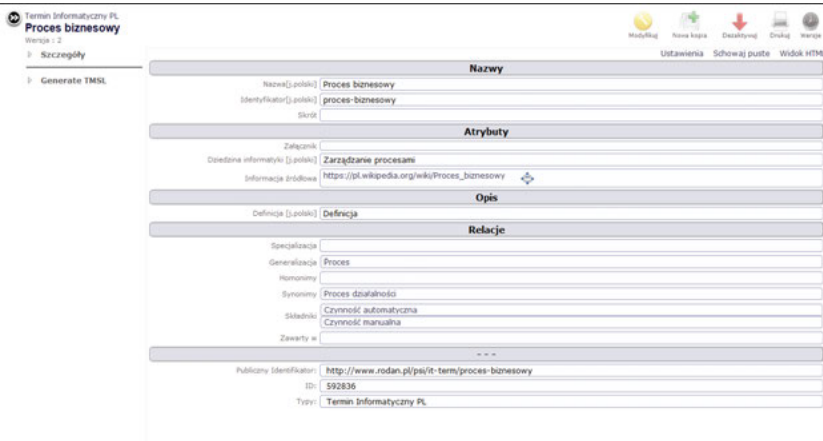


18 paź 2019

A. Blikle, W. Staniszkis - PL-GB Tezaurus Terminów Informatycznych

5

Rysunek 1  
Przykład modelu pojęciowego tezaurusa PATTI



Rysunek 2  
Przykład terminu informatycznego w prototypie tezaurusa PATTI

OfficeObjects® Ontology Manager [OfficeObjects® Ontology Manager – Podręcznik Eksperta, 2014]. Przykład obejmuje wszystkie relacje ontologiczne przewidziane w pojęciowym modelu tezaurusa.

Opracowana w ramach projektu pilotażowa baza polskiej terminologii informatycznej, zachowująca przyjęte reguły pojęciowego modelu tezaurusa, będzie stanowiła

przykład dla kwalifikowanych wolontariuszy podejmujących wysiłek uzupełniania zakresu publikowanej terminologii informatycznej. W pilotażowej fazie wdrożenia tezaurya przewidujemy opracowanie 500 pozycji z zakresu powszechnie stosowanej polskiej terminologii informatycznej.

Graficzny model powiązań opisywanych terminów informatycznych będzie wspomagany przez menu kontekstowe automatycznie tworzone w interfejsie prezentacyjnym, jak również przez możliwości nawigacji pomiędzy pojęciami prezentowanymi w grafie powiązań.

## 5. KONCEPCJA ZAWARTOŚCI TEZAUURUSA

Tezaurus powinien zawierać terminy z zakresu szeroko rozumianej informatyki wraz z ich odpowiednikami angielskimi i przykładami użycia. Definicje powinny pojawiać się tylko wówczas, gdy można wskazać ich źródło o charakterze autorytatywnym (tzn. nie encyklopedycznym). Źródłami mogą być normy, słowniki z aktów prawnych (szczególnie unijnych, zawierających odpowiednik angielski), metodyki (Prince2), dobre praktyki (ITIL, Cobit) i firmowa dokumentacja produktowa.

Praca nad stworzeniem tezaurya powinna przebiegać analogicznie do prac nad Wikipedią – ten model przy wszystkich swoich wadach okazał się zdecydowanie najbardziej skuteczny: duża społeczność autorów, niewielka grupa moderatorów/administratorów. Na pewno można liczyć na subskrybentów grupy Facebookowej PTI-Polskie-Słownictwo-Informatyczne. Promowaniem swojej terminologii mogą też być zainteresowane firmy (na przykład Microsoft czy Oracle) oraz organizacje takie jak ISACA.

## 6. PRZEWIDYWANE EFEKTY REALIZACJI PROJEKTU

Ontologiczny tezaurus polskich terminów informatycznych powinien zostać opublikowany w otwartej sieci Internet jako aplikacja dostępna w przeglądarkach internetowych z zapewnionym dostępem do zasobów tezaurya zarówno z komputerów stacjonarnych jak i z urządzeń mobilnych (smartphone, tablet).

Podstawowym celem projektu jest dostarczenie referencyjnego źródła terminologii informatycznej opracowanej i udokumentowanej na podstawie ontologicznego modelu prezentacji wiedzy. Wstępna analiza potrzeb, skrótowo omówiona w rozdziałach 2 i 3, wskazuje, że istnieje duże zapotrzebowanie na takie źródło nie tylko w informatycznym środowisku akademickim i biznesowym, lecz również, ze względu na wysokie koszty i zakres oddziaływania inwestycji informatycznych, w szeroko pojętych środowiskach użytkowników rozwiązań informatycznych.

W odpowiedniej fazie wdrożenia projektu przewidywane jest podjęcie działań promocyjnych zarówno w środowiskach potencjalnych użytkowników systemu, jak również w środowisku specjalistów informatyki. W tym drugim przypadku istotne jest zainteresowanie potencjalnych autorów haseł tezaurya, których ciągła współpraca powinna doprowadzić do dynamicznego wzrostu liczby opracowanych terminów informatycznych.

Platformą działań promocyjnych będą przede wszystkim systemy sieci społecznościowych, jak również strony internetowe stowarzyszeń zawodowych, takich jak Polskie Towarzystwo Informatyczne, Izba Telekomunikacji i Informatyki, Naczelna Organizacja Techniczna, Naczelna Rada Adwokacka, Izba Radców Prawnych.

## ŹRÓDŁA

(del) Pilar Castillo V., Curley M., Fabry E., Gottiz M., Hagedorn P., Herczog E., Higgins J., Joyce A., Korte W., Lanvin B., Parola A., Straub R., Tapscott D., Vassallo J., 2012, Manifest w sprawie e-umiejętności, European Schoolnet (EUN Partnership AISBL)<sup>2</sup>, <http://docplayer.pl/58305308-Manifest-w-spr-awie-e-umiejtnosci-wstep-don-tapscott-autor-wi-kinomii.html> (dostęp: 15.10.2020).

OfficeObjects® Ontology Manager – Podręcznik Eksperta, 2014, Warszawa.

## PODZIĘKOWANIE

Autorzy wyrażają podziękowanie niżej wymienionym osobom (kolejność alfabetyczna z pominięciem stopni i tytułów), które wniosły uwagi i korekty do początkowej wersji naszego artykułu: Krzysztof Goczyła, Wacław Iszkowski, Paweł Sawicki, Ewa Teleżyńska.

---

<sup>2</sup> Na końcu publikacji umieszczono następujący komentarz: *Opublikowano w czerwcu 2012 roku. Poglądy wyrażone w niniejszej publikacji są poglądami autorów i niekoniecznie są to poglądy European Schoolnet, DIGITALEUROPE, Europejskiego Stowarzyszenia e-umiejętności czy Komisji Europejskiej. Publikacja ta jest dostępna zgodnie z warunkami licencji Creative Commons Uznanie Autorstwa 3.0 (<http://creativecommons.org/licenses/by/3.0>)*