

**Tomasz Boiński, Anna Jaworska, Radosław Kleczkowski, Piotr Kunowski,
Julian Szymański**

**Politechnika Gdańska, Wydział Elektroniki, Telekomunikacji i Informatyki, Katedra
Architektur Systemów Komputerowych**

ZESPOŁOWA BUDOWA ONTOLOGII Z WYKORZYSTANIEM SYSTEMU OCS ORAZ EDYTORA PROTÉGÉ*

Streszczenie

Konstruowanie ontologii wymaga współpracy wielu osób. W idealnej sytuacji nad pojedynczą ontologią pracować będzie duża, rozproszona społeczność tworząca tym samym wspólną reprezentację wiedzy z danej dziedziny. W publikacji zaprezentowano propozycję modelu pracy grupowej nad ontologią. Zdefiniowano model zarządzania wersjami ontologii. Przedstawiono system Ontology Creation System (OCS) oraz architekturę i implementację rozszerzenia edytora Protégé umożliwiającego zespołowe wytwarzanie ontologii na podstawie zaproponowanego modelu pracy grupowej. Omówiono również alternatywy i modyfikacje zaprezentowanego modelu.

1. WSTĘP

Sieć Internet stanowi wszechstronną składnicę informacji. Wiedza ta jest jednak w znacznym stopniu ukierunkowana na człowieka – składowane w sieci dokumenty prezentowane są zazwyczaj w postaci graficznej oraz nie posiadają jawnie zdefiniowanego kontekstu. Czynniki te powodują, że przetwarzanie tych informacji w sposób komputerowy jest niezwykle trudne. Wychodząc naprzeciw temu problemowi Berners Lee [1] zaproponował ideę sieci semantycznej [2], w której zasoby są powiązane ze sobą znaczeniowo. Dzięki takiemu podejściu składowana w sieci informacja mogłaby być przetwarzana oraz interpretowana zarówno przez człowieka jak i maszyny, gdyż opatrzona by była kontekstem.

Budowa takiej sieci wymaga jednak formalizacji zapisu informacji. W3Cache Consortium [3] zaproponowało ontologie, czyli reprezentację wiedzy poprzez zdefiniowanie zbioru pojęć i relacji między nimi, jako podstawę światowej sieci semantycznej. Jako język budowy ontologii zaproponowano język OWL [4].

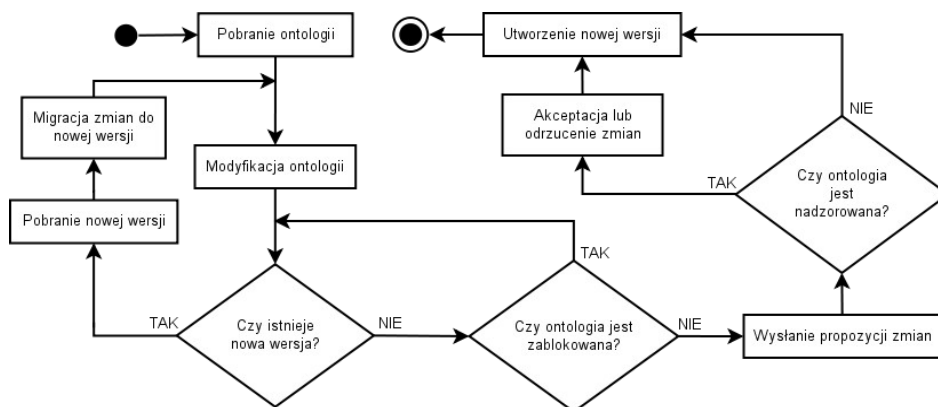
* Publikacja zrealizowana w ramach grantu MNiSW N519 432338

2. PRACA GRUPOWA NAD ONTOLOGIĄ

Aby możliwe było sprawne przetwarzanie wiedzy przez wszystkich potencjalnych odbiorców, konieczne jest by reprezentowała ona wspólne spojrzenie na dane zagadnienie. Człowiek charakteryzuje się jednak mocno zindywidualizowanym poglądem na otaczającą go rzeczywistość. Stąd konieczne jest by praca nad ontologią przebiegała w jak najszerzym gronie, pozwalając tym samym na ujednoczenie wizji opisu wybranego fragmentu świata. Drugim ważnym aspektem ontologii jest dostępność odpowiednich narzędzi umożliwiających prace nad ontologiami również specjalistom nie będących informatykami.

Jednym z najbardziej znanych takich narzędzi jest Protégé [5][6][7]. Edytor ten nie był jednak tworzony z myślą o pracy grupowej nad ontologiami. Co prawda równolegle do prezentowanego w tej publikacji systemu powstało rozszerzenie do Protégé zwane Collaborative Protégé [8], jednak nie przewiduje ono funkcjonalności składowania ontologii w ogólnodostępnym repozytorium czy też kontroli nad kształtem ontologii przez jej właściciela – przyjęto model głosowania nad zmianami. Dzięki takiemu podejściu możliwe jest szybkie rozwiązanie konfliktów, jednak twórca ontologii może stracić kontrolę nad jej rozwojem.

W proponowanym rozwiązaniu, jakim jest system OCS (Ontology Creation System) [9], przyjęto odmienny model pracy grupowej. W opinii autorów, jeżeli ontologia ma być stosowana przez szerokie gremium odbiorców, proces jej powstawania również powinien obejmować szeroką grupę pomysłodawców. Zakładając system głosowania przyjęty w Collaborative Protégé autor ontologii może łatwo stracić kontrolę nad kierunkiem jej tworzenia a ostateczny efekt może całkowicie odbiegać od przyjętych założeń. Stąd w proponowanym rozwiązaniu uprzywilejowany użytkownik, będący twórcą ontologii, lub ich grupa nominowana przez twórcę ontologii ma prawo akceptować zmiany w ontologii. Każdy inny użytkownik ma prawo zgłaszać propozycje zmiany do dowolnej z ontologii, jednak nie zostaną one uwzględnione aż do etapu akceptacji zmian.



Rys.1. Proces zgłaszania propozycji zmian oraz tworzenia nowych wersji ontologii w systemie OCS

W przypadku ontologii publicznych, których rozwój nie jest nadzorowany, proces zgłaszania propozycji może być również połączony z tworzeniem nowej wersji realizując algorytm znany z repozytoriów kodu, np. Subversion [10] ale w ujęciu semantycznym. W

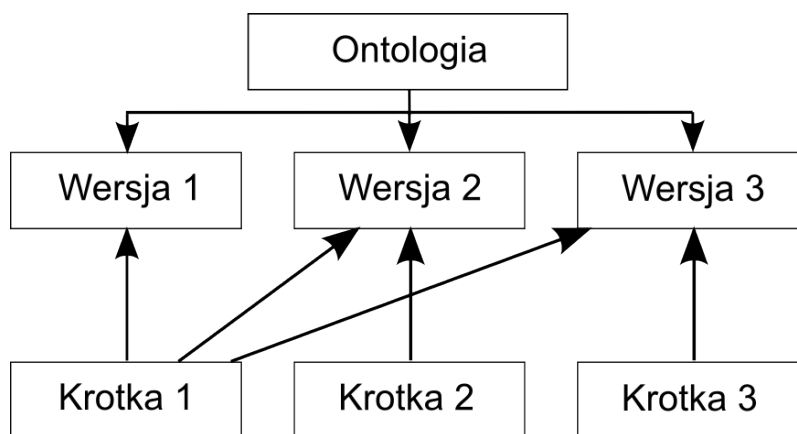
tym przypadku każdy użytkownik ma prawo tworzenia nowej wersji i zgłoszone przez niego zmiany są automatycznie zatwierdzane.

Oba powyższe procesy zostały przedstawione na Rys. 1. Należy pamiętać, że podobnie jak to ma miejsce w przypadku repozytoriów kodu, to użytkownik tworzący nową wersję odpowiedzialny jest za zachowanie spójności ontologii.

3. WERSJONOWANIE ONTOLOGII

Kolejnym istotnym aspektem pracy grupowej nad ontologią jest możliwość śledzenia zachodzących w niej zmian oraz dostęp do dowolnej wcześniejszej wersji ontologii. Obie te operacje są możliwe dzięki zastosowaniu języka OWL, a tym samym oparciu konstrukcji ontologii o trójki RDF [11][12].

W systemie ontologie składowane są jako trójki pogrupowane w wersje (Rys. 2). W tym modelu raz zapisana krotka pozostaje w systemie aż do usunięcia ontologii.



Rys.2. Sposób składowania ontologii w systemie OCS

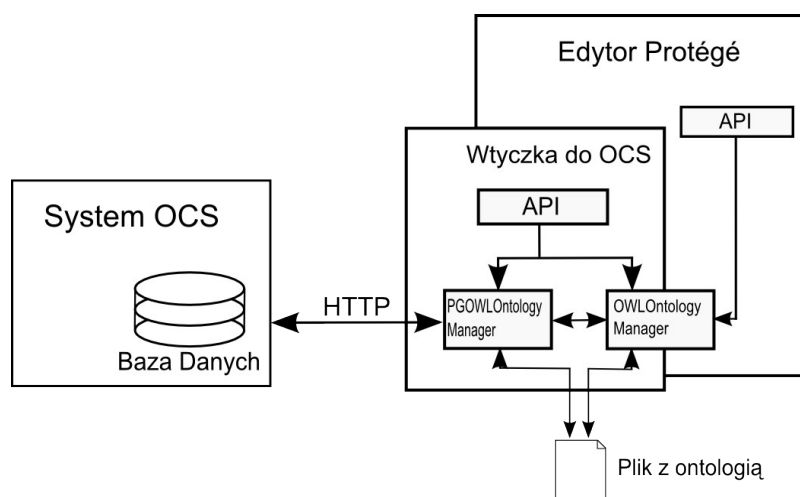
Umożliwia to bezpośrednie uzyskanie dostępu do dowolnej z wersji ontologii bez konieczności przetwarzania danych historycznych. Również porównanie dwu wersji ontologii na poziomie semantycznym wymaga jedynie porównania dwóch zbiorów trójek i określeniu różnic pomiędzy nimi. Wynik takiego porównania może zostać następnie przekształcony w aksjomaty języka OWL.

4. INTEGRACJA SYSTEMU OCS Z EDYTOREM PROTÉGÉ

Bogata funkcjonalność edytora Protégé oraz jego popularność skusiły autorów do rozszerzenia go o możliwość pracy grupowej opartej o przepływ zadań opracowany dla systemu OCS. Jako cel postawiono sobie napisanie modułu pozwalającego na pobieranie dowolnej wersji ontologii z serwera i wysyłanie propozycji zmian, czyli dostarczenie rozwiązania dla zwykłego użytkownika modyfikującego ontologię. Właściciel ontologii nadal zatwierdzałby zmiany i tworzył nowe wersje korzystając z edytora OCS. W

późniejszym etapie prac rozwiązanie zostało rozszerzone o możliwość dodawania nowych ontologii i tworzenia ich kolejnych wersji.

Programiści Protégé umożliwili dodawanie nowych funkcji do edytora poprzez moduł obsługi „wtyczek” (ang. plugin) [13]. Ponadto, zarówno system OCS jak i edytor Protégé, do operowania na ontologiach wykorzystują bibliotekę OWL API [14]. Biblioteka ta dostarcza manager ontologii, pozwalający m.in. na wczytywanie i zapis plików ontologii. Pozwala również na edycję składowych wczytanych ontologii. Klasa managera w systemie OCS (PGOWLOntologyManager) jest rozszerzeniem klasy z biblioteki OWL (OWLOntologyManger) o funkcje pozwalające na komunikację z serwerem składającym i wersjonującym ontologie. Zezwala również na zarządzanie i uwierzytelnianie użytkowników, pobierania ontologii o zadanym URI, porównywania ontologii, czy też wysyłania propozycji zmian. Przewidzianą architekturę wtyczki zaprezentowano na Rys. 3.



Rys.3. Architektura wtyczki integrującej edytor Protégé z systemem OCS

W celu zachowania kompatybilności wtyczki z systemem OCS interfejs użytkownika został przeniesiony z systemu OCS. Zachowano również zgodność plików projektu, dzięki temu możliwe jest edytowanie tego samego projektu zarówno w Protégé jak i w edytorze OCS.

6. ZAKOŃCZENIE

Zaprezentowane rozwiązanie umożliwia grupową pracę nad ontologią z zachowaniem równowagi pomiędzy swobodą użytkowników a kontrolą twórcy ontologii nad jej kształtem. Pozwala ono na wytwarzanie ontologii zarówno w trybie otwartym, gdzie każdy z użytkowników wprowadzając swoje zmiany tworzy nową wersję ontologii, oraz w trybie bardziej zamkniętym, gdzie to właściciel bądź ekspert ontologii decyduje o fakcie akceptacji bądź odrzucenia danej propozycji zmiany. Integracja z edytorem Protégé, będącym najpopularniejszym i najbardziej elastycznym obecnie dostępnym narzędziem do wytwarzania ontologii, rodzi nadzieje na popularyzację proponowanego rozwiązania.

BIBLIOGRAFIA

- [1] Berners-Lee T., Hendler J., Lassila O.: *The Semantic Web*, Scientific American, May 2001
- [2] The Semantic Web Community Portal, <http://semanticweb.org>, 2008
- [3] W3C, *Semantic Web*, <http://www.w3.org/2001/sw/>, 2001
- [4] W3C, *OWL Web Ontology Language Overview*, 2004, <http://www.w3.org/TR/owl-features/>
- [5] Stanford Center for Biomedical Informatics Research, 2009, <http://protege.stanford.edu/>
- [6] Gennari J. H., Musen M. A., Fergerson R. W., Grosso W. E., Crubézy M., Eriksson H., Noy N. F., Tu S. W.: *The evolution of Protégé: An environment for knowledge-based systems development*, Technical Report SMI-2002-0943, Stanford Medical Institute, 2002
- [7] Noy N. F., Fergerson R. W., Musen M. A.: *The knowledge model of Protégé-2000: Combining interoperability and flexibility*, Lecture Notes in Computer Science, Springer-Verlag, 2000
- [8] Stanford Center for Biomedical Informatics Research, http://protegewiki.stanford.edu/index.php/Collaborative_Protege, 2009
- [9] Boiński T., Budnik Ł., Jakowski A., Mroziński J., Mazurkiewicz K., *OCS - domain oriented Ontology Creation System*, Sejmik Młodych Informatyków, Polish Journal of Environmental Studies, Vol 18, No. 3B, 2009, s. 35-38, 2009
- [10] CollabNet, <http://subversion.tigris.org/>, 2008
- [11] W3C, *RDF/XML Syntax Specification*, <http://www.w3.org/TR/rdf-syntax-grammar/>, 2004
- [12] W3C, *RDF Vocabulary Description Language 1.0: RDF Schema*, <http://www.w3.org/TR/rdf-schema/>, 2004
- [13] Knublauch H., *Protégé-owl api programmer's guide*, <http://protege.stanford.edu/plugins/owl/api/guide.html>, 2009
- [14] Horridge M., Bechhofer S., *The OWL API: A Java API for Working with OWL 2 Ontologies*, 6th OWL Experienced and Directions Workshop, 2009

ONTOLOGY CREATION AS A GROUP PROCESS USING OCS SYSTEM AND PROTÉGÉ EDITOR**Summary**

Ontology creation requires cooperation within a group of people. In perfect situation a single ontology is being created by numerous and distributed community thus gathering common knowledge of that community within a given domain. In the paper a proposition of group work model was described. Ontology versioning management model was also proposed. Ontology Creation System (OCS) and architecture and implementation of Protege editor extension utilizing proposed group work model was described. Possible modifications of proposed model were also presented.