

WITOLD ABRAMOWICZ
PIOTR STOLARSKI
TADEUSZ TOMASZEWSKI

Ontologie jako narzędzie budowy modeli w ubezpieczeniowych systemach informacyjnych – wprowadzenie

W niniejszym tekście autorzy dokonują podsumowania prac prowadzonych w ramach modelowania wiedzy i informacji ubezpieczeniowej. Przedstawiają ontologię w informatyce, czyli nową strukturę służącą do agregowania i porządkowania informacji. Artykuł ponadto wskazuje na dwa kluczowe obszary dotyczące modelowania ubezpieczeń, czyli: reprezentację wiedzy o regulacjach prawnych oraz normalizację pojęcia ryzyka. Obydwa obszary są bardzo istotne dla poprawności procesów obiegu informacji w przedsiębiorstwach rynku ubezpieczeń. Teoretyczny przegląd oparty na badaniach literaturowych pozwala w sposób syntetyczny przedstawić koncepcję oraz rozpocząć dyskusję nad możliwymi zastosowaniami dla modelowania informacji ubezpieczeniowej za pomocą modeli semantycznych – w szczególności w dwóch wskazanych wcześniej obszarach. Zastosowanie ontologii i rozpoczęcie operowania na modelach semantycznych będzie przyczynkiem do uzyskania nowych możliwości związanych z gromadzeniem i przetwarzaniem informacji w systemach ubezpieczeniowych.

1. Wprowadzenie

Obserwujemy ciągły i szybki rozwój technologii informacyjnych. Mimo to w większości przypadków ich zastosowanie w ubezpieczeniach polega w znacznej mierze na wykorzystaniu informacyjnych mechanizmów przetwarzania informacji w działalności zarządczej, marketingowej lub dystrybucyjnej produktów¹. Konceptualizacja produktów ubezpieczeniowych w systemach informatycznych² wynika przede wszyst-

-
1. Handschke J. *Internet w działalności ubezpieczeniowej w Polsce: ze szczególnym uwzględnieniem dystrybucji* / red. nauk. Jerzy Handschke, Poznań: Katedra Ubezpieczeń Akademii Ekonomicznej, 2004.
 2. Poprzez modele ubezpieczenia rozumiemy tutaj zarówno prawną konstrukcję produktu ubezpieczeniowego jak też sposób oceny ryzyka.

kim z tradycyjnych metod modelowania ryzyka, kładąc szczególny nacisk na ilościowe aspekty informacji o przeszłych zdarzeniach, powiązanych z powstaniem wypadku ubezpieczeniowego, przekształcając je z wykorzystaniem metod statystycznych w dane³. Na ich podstawie formułowane są, ograniczone z punktu widzenia zawartości informacji, abstrakcyjne zbiory zagregowanych danych statystycznych. Sposób powstawania informacji z danych oraz ich reprezentacji i wykorzystania uartał się w trakcie ewolucji tych metod. Na ich rozpowszechnienie wpłynęły także ograniczenia w przetwarzaniu informacji. W szczególności do ograniczeń tych należały niska efektywność technologii przetwarzania zasobów informacji w wiedzę oraz zarządzania wiedzą⁴.

Rzówj technologii semantycznych pozwala na zmianę podejścia, umożliwiając automatyczne przetwarzanie nie tylko danych liczbowych, lecz także informacji tekstowych i ich znaczenia.

Postęp technologiczny w dziedzinie zarządzania wiedzą rodzi niepowtarzalną okazję do podjęcia prób wykorzystania nowatorskich rozwiązań wszędzie tam, gdzie istniejące do tej pory ograniczenia wyznaczały standardy działania i rozwiązywania problemów⁵.

Sposób wykonywania działalności ubezpieczeniowej uzależniony jest od dwóch rozłącznych czynników: z jednej strony od konstrukcji instytucji prawnych, z drugiej od podejścia do ryzyka, a w konsekwencji do metod zarządzania nim⁶. Tak się składa, że z semantycznym przetwarzaniem informacji z obu tych dziedzin wiązane są obecnie duże nadzieje.

Artykuł ten ma być wstępem do dyskusji o wykorzystaniu nowych technik usprawniających zarządzanie informacją ubezpieczeniową. Celem autorów jest wprowadzenie do opisów projektowanych zastosowań w ubezpieczeniach.

1. Modelowanie informacji ubezpieczeniowej

OECD definiuje modelowanie ekonomiczne, jako „upraszczającą reprezentację rzeczywistości gospodarczej wskazującą na współzależności pomiędzy wartościami ekonomicznymi (zmiennymi)”⁷. Ciekawe spostrzeżenia nawiązujące również do definio-

3. Definiujemy odrębnie pojęcia: danych, informacji oraz wiedzy. Za: Abramowicz W. *Filtrowanie Informacji*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2008, s. 25-30.

4. Tomaszewski T. *Przykłady zastosowań koncepcji semantycznej reprezentacji ryzyka w ubezpieczeniowych systemach informacyjnych*, w *Studia Ubezpieczeniowe*, red. nauk. Jerzy Handschke, Zeszyty Naukowe Nr 127, s. 232-242, Uniwersytet Ekonomiczny w Poznaniu, Poznań 2009.

5. *Ibidem*.

6. Interesującym zagadnieniem samym w sobie może być tutaj uogólnienie polegające na stwierdzeniu, iż oba składniki są silnie powiązane z wymiarem czasowym, tj. o ile ryzyko jest miarą opisującą zdarzenia w przyszłości, o tyle konstrukcje prawne ze swej natury ulegają zmianom w czasie.

7. Alexander P., Baden S., 2000, *Glossary on macroeconomics from a gender perspective*, Institute of Development Studies, University of Sussex. Odczytane z <http://stats.oecd.org/glossary/detail.asp?ID=6813>, 24 maja 2010.

wania pojęcia modelu ekonomicznego, aczkolwiek sformułowane historycznie wcześniej, pojawiają się w wypowiedzi Lange⁸, w której autor przyrównuje modele do „dedukcyjnych zbiorów teoremów [...] będących generalizacją obserwacji, i mogących być poddanymi testowi empirycznemu”. Autor ten jednocześnie sygnalizuje, w kontekście modeli dostarczanych przez ekonomię teoretyczną, iż „rzeczywistość (doświadczenie) jest znacznie bogatsza niż naukowy język” w związku z czym postuluje dalej: „w celu załatwienia rozłamów pomiędzy (wytworzonymi) w teorii pojęciami a empirią, konieczne jest [...] zorganizowanie reguł określających odnośniki pomiędzy tymi dwoma”. W naukach o systemach informacyjnych modele nie są niczym innym, kładzie się jednak tutaj duży nacisk na sposoby organizacji informacji oraz jej wykorzystania. Stąd np. Guarino badając współczesne metody modelowania informacji, pisze o „wysoc interdyscyplinarnym podejściu, w którym filozofia oraz lingwistyka odgrywają istotną rolę w analizie struktury (zjawisk) rzeczywistości na wysokim poziomie ogólności”⁹. W dalszej części wywodu podaje on także definicję formalną modelu.

Na tym etapie rozważań podkreślić należy, iż modele służą do redukcji złożoności rozpatrywanych zjawisk od stopnia, w którym analiza przedmiotu badania jest z takich lub innych powodów bezcelowa, do stopnia umożliwiającego ich dokładne poznanie. Ułatwiają zrozumienie zjawisk przeszłych i mogą umożliwiać ich przewidywanie w przyszłości¹⁰. Niewątpliwym zaś problemem przy ich konstrukcji jest właściwe określenie ich poziomu złożoności.

Chcąc lepiej to uzmysłowić, wskażemy tutaj na jeszcze jedną inspirację, płynącą ze strony informatyki. Na Ilustracji 1 przedstawione są relacje (aproxymacja i denotacja) pomiędzy światem rzeczywistym, jego przyjętym modelem oraz ekwiwalentnym opisem teoretycznym. Jak można zaobserwować, wraz z zastosowaniem kolejnych odwzorowań, następuje redukcja możliwych wartości charakteryzujących konkretne przekształcenie¹¹.

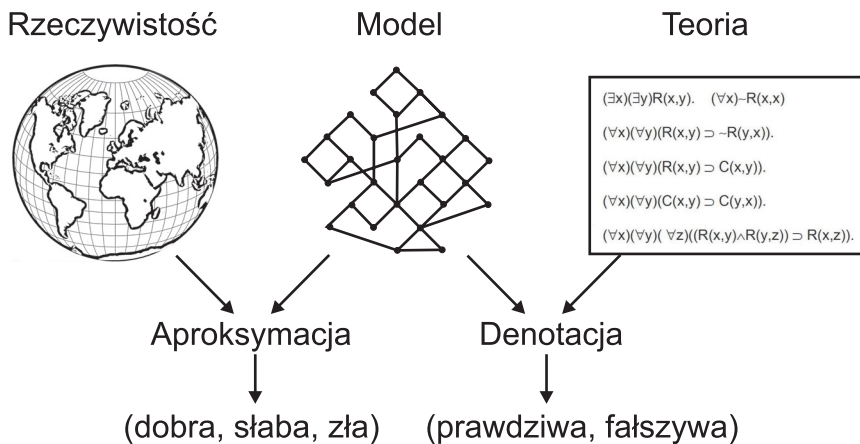
8. Lange O., *The Scope and Method of Economics*, The Review of Economic Studies, Vol. 13, No. 1 (1945 - 1946), s. 19-32.

9. Guarino N., *Formal Ontology and Information Systems*, Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, s. 3-15.

10. Poprawska E., Ronka-Chmielowiec W., *Metody aktuarialne*, PWN, Warszawa 2006.

11. O ile odwzorowaniu aproxymacyjnemu przypisać można cechę charakteryzującą jego jakość, o tyle opis za pomocą teorii jest albo prawdziwy, albo fałszywy.

Ilustracja 1. Świat rzeczywisty, model i teoria – porównanie



Źródło: Sowa F. J., *Guided Tour of Ontology*; <http://www.jfsowa.com/ontology/guided.htm>, odczytano: 26 marca 2010 r.

Ilustracja ta dobrze odzwierciedla prawidłowość: stosowanie modeli abstrakcyjnych, jakimi są np. statystyki lub ich przekształcenia, może dawać dobre rezultaty w wielu dziedzinach zastosowań, jednak związane jest także z utratą (świadomym pominięciem) informacji o poszczególnych obiektach. W kontekście nauk o ubezpieczeniach strata taka może być o tyle wyzwaniem, że przedmiotem modelowania są po części zdarzenia¹² – byty trudne do odzwierciedlenia.

W naukach ekonomicznych opieramy się przede wszystkim na modelach abstrakcyjnych. Abstrakcja to wyodrębnienie pewnych cech danego pojęcia, przedmiotu, zjawiska i zapomnienie o innych¹³. Jest to forma aktywności umysłowej, dzięki której dana treść świadomie jest uwalniana od swych skojarzeń z elementami z nią niezwiązanymi¹⁴. W modelowaniu ubezpieczeń wyróżnić można trzy nie zawsze rozłączne typy przedmiotów modeli¹⁵: modele danych i informacji, modele ryzyka, pozostałe modele^{16,17}.

Modele ryzyka mają przede wszystkim charakter ilościowy i wykorzystują w znacznej mierze metody matematyczne¹⁸ oraz miary statystyczne. Takie podejście jest przed-

12. Henderson, Glenn V., Jr. *Problems and Solutions in Conducting Event Studies*. Journal of Risk and Insurance, 1990 57(2), s. 282.

13. Szurek M. *Matematyka dla humanistów*, Wydawnictwo RTW, Warszawa 2000.

14. Sharp D., *Leksykon pojęć i idei C.G. Junga*, Wydawnictwo Wrocławskie, 1999.

15. Dużo bardziej złożona typologia badań nad ryzykiem i ubezpieczeniami przedstawiona jest w: Bar-Niv R., Bickelhaupt D. L., *Research in International Risk and Insurance: Summary, Synthesis, and Prospects*. Journal of Risk and Insurance, 1986, 53(1), s. 119.

16. Mamy tutaj na myśli modele uwzględniające lub opisujące czynniki otoczenia, jak np. uwarunkowania prawne.

17. Craig H., *New directions in risk management*. Canadian Insurance, 1996, October, 101(11), s. 33-34.

18. Djehiche B., *Review of Non-Life Insurance Mathematics – An Introduction with Stochastic Processes*. Scandinavian Actuarial Journal, 2004, (6), s. 475.

miotem intensywnych badań. Frees i Wang¹⁹, na przykładzie ubezpieczeń samochodowych, rozwijają model predyktorów zagregowanej straty opartych na ramie danych obserwacyjnych, w których napotykane są dane dla różnych klas ryzyka z historią liczby pozwów. Uogólnione modele regresji liniowej są wykorzystane do szacowania rozkładu krańcowego pozwów. Proces liczby pozwów jest generowany za pomocą modelu regresji Poisson'a uzależnionego od sekwencji ukrytych zmiennych. W innych pracach wskazuje się na praktyczne możliwości wykorzystania połowicznych procesów Markov'a, na przykładzie zastosowania dla ubezpieczeń społecznych²⁰. Autorzy argumentują, że klasyczne podejście do modeli ryzyka może okazać się zawodne, ponieważ wiele rzeczywistych zjawisk ma charakter niehomogeniczny czasowo. Z perspektywy podziału na przedmiot modelowania, jaki wprowadzono w pierwszym akapicie obecnego podrozdziału, znacznie trudniejszą pracą do skategoryzowania jest tekst²¹. Publikację tę umiejscowić można w przyjętym systemie klasyfikacyjnym na pograniczu tematyki modelowania ryzyk oraz pozostałych modeli. Autorzy prezentują ramę wykorzystującą logitowe lub probitowe modele probabilistyczne na potrzeby predykcji nadużyć ubezpieczeniowych²². Opisano także implementację dla ubezpieczeń rolniczych. Kolejne podejście jest równie trudne do klasyfikowania²³. Autor na podstawie modelu ICAPM^{24,25} analizuje przychód z przypisanej składki kilku rodzajów ubezpieczeń. Na podstawie zbudowanego modelu próbuje następnie przewidywać przyszłe kształtowanie się modelowanych wielkości jednocześnie zwracając uwagę na konsekwencje modelu – ryzyko systematyczne jest większe w środowiskach rozmytych niż w środowiskach dobrze określonych. Jednocześnie też zmianie środowiska towarzyszą zmiany w parametrach finansowych ubezpieczyciela. Walorem tego podejścia jest wskazanie na problemy gromadzenia informacji. W pracy Yin, Liu, Yang²⁶ poruszona jest tematyka modelowania ryzyk w warunkach zmiany zasad wpływających na te modele. Efekt uzyskany jest poprzez zastosowanie łańcuchów Markov'a o ciągłym czasie. Rezultat badań sprowadza się do wniosku argumentującego za ograniczaniem danych decyzyjnych. W pracy autorstwa Gori i Padrone²⁷ próbuje się wykorzystać mechanizmy teorii gier do modelowania ubezpieczenia na wypadek aktu terrorystycz-

19. Frees E. W., Wang P., *Copula credibility for aggregate loss models*. *Insurance, Mathematics & Economics*, 2006, 38(2), s. 360-373,
20. Janssen J., De Dominicis R. *Finite Non-Homogeneous Semi-Markov Processes: Theoretical and Computational Aspects*. *Insurance, Mathematics & Economics*, 1984, 3(3), s. 157.
21. Jin Y., Rejesus R. M., Little B. B., *Binary choice models for rare events data: a crop insurance fraud application*. *Applied Economics*, 2005, 37(7), s. 841-848.
22. A więc także specyficznego ryzyka jednakże pochodnego dla podstawowego przedmiotu ubezpieczenia i obciążającego tym razem bezpośrednio firmę ubezpieczeniową nie zaś na nią przenoszonego.
23. Lai L., *Underwriting profit margin of P/L insurance in the fuzzy-ICAPM*. *Geneva Risk and Insurance Review*, 2006, 31(1), s. 23-34.
24. Intertemporal Capital Asset Pricing Model.
25. Fama E. F., *Multifactor Portfolio Efficiency and Multifactor Asset Pricing*. (The Journal of Financial and Quantitative Analysis), Vol. 31, 1996, No. 4, Dec.
26. Yin G., Liu Y. J., Yang H., *Bounds of ruin probability for regime-switching models using time scale separation*. *Scandinavian Actuarial Journal*, 2006, (2), s. 111-127.
27. Gori, S., Padrone, M. *Is game theory a useful tool for terrorism insurance?* *International Journal of Risk Assessment and Management*, 2007, 7(8), s. 1176.

nego. Zastosowano ewolucyjne podejście do teorii gier. Podsumowuje się też problemy w uzyskaniu dostatecznie satysfakcjonujących wyników.

Przechodząc do modeli informacji i danych, należy omówić tutaj szereg prac, choć zauważyć trzeba, że jest ich zdecydowanie mniej niż w wyżej wymienionym podejściu do przedmiotu modeli ryzyka. Gobble i Windeler²⁸ wskazują na znaczący postęp w systemach oceny ryzyka katastrof ze względu na napływ danych z nowej generacji systemów informacji geograficznej, cechujących się dużo wyższą jakością, w tym ich dokładnością. Spowodowało to, zdaniem autorów, znaczący wzrost możliwości zróżnicowania ocen ryzyka dla celów ubezpieczeniowych. Lane i Dennis²⁹ badają wpływ użycia różnych danych socjoekonomicznych oraz ich estymacji dla celów tworzenia profili klientów oraz ich oceny ze względu na możliwości prognozowania samych wypadków ubezpieczeniowych (bez względu na szkodowość). Dyskutowany jest także wpływ współzależności pomiędzy poszczególnymi elementami profilu. Na marginesie rozważań nad modelowaniem danych dla ubezpieczeń, warto zwrócić uwagę na dwa dodatkowe konteksty: możliwości i zastosowanie komercyjnych systemów informatycznych³⁰, a także próby budowy pierwszych systemów semantycznych dla firm ubezpieczeniowych³¹.

Trzeci z wyróżnionych na początku podrozdziału typów modeli, to pozostałe modele. Zaliczyliśmy tutaj prace niepodlegające prostemu przyporządkowaniu lub też będące w pewnej mierze meta-modelami, których przedmiotem są systemy ubezpieczeniowe jako całość. Jednym z podejść jest propozycja całościowego modelu przedsiębiorstwa ubezpieczeniowego³². Środowisko symulacyjne skupia się na aktuariacie oraz operacjach finansowych i podatkowych. W rezultacie model jest w stanie wskazać wartość firmy. Do przykładowych zastosowań autorzy zaliczają: testowanie różnego rodzaju portfeli ubezpieczeń albo badanie zależności pomiędzy wielkościami ekonomicznymi ubezpieczycieli. Autorzy kolejnego opracowania³³ zastanawiają się nad rolą ubezpieczeń oraz poziomami ich nasycenia w gospodarce rynkowej. W tym celu konstruują dynamiczny model, w którym agenty programowe mogą wybrać pomiędzy różnymi wariantami ubezpieczenia lub jego braku, jako odpowiedź na wystawienie na dane ryzyka. Model został przystosowany do uwarunkowań gospodarki amerykańskiej. W przedostatnim z omówionych w tej części artykułów³⁴, przedstawione jest podejście z zakresu teorii gier do modelowania efektów

28. Gobble E., Windeler D., *Catastrophe modeling: Shifting perceptions*. Canadian Underwriter, 2003, June, 70(6), s. 38.

29. Lane, J., Glennon, D., *The Estimation of Age/Earnings Profiles in Wrongful Death and Injury Cases*. Journal of Risk and Insurance, 1985, 52(4), s. 686.

30. *New Version of DFA's Insurance Modeling Technology Adds Multi-Currency Capabilities and Modeling for Businesses with Complex Structures*. (16 October). Business Wire lub Eckler Ltd. *Signs Multi-Year Deal for ADVISE(R) and GEMS(R) Solutions from DFA*. (15 December). Business Wire,

31. Po szersze omówienie tej problematyki odsyłamy, do: Tomaszewski T. *op. cit.*

32. Goldstein, A. B., Markowitz, B. G. *SOFASIM: A Dynamic Insurance Model with Investment Structure, Policy Benefits and Taxes*. The Journal of Finance, 1982, 37(2), s. 595.

33. Braun H., Koeniger W., *On the role of market insurance in a dynamic model*. Geneva Risk and Insurance Review, 2007, 32(1), s. 61-90.

34. Kihlstrom R. E., Roth A. E., *Risk Aversion and the Negotiation of Insurance Contracts*. Journal of Risk and Insurance, 1982, 49(3), s. 372.

awersji do ryzyka w trakcie negocjowania kontraktów ubezpieczeniowych. W rezultacie wykazano, że podejście uczestników zmienia się w przypadku ubezpieczeń negocjowanych w stosunku do zdecydowanie częstszej sytuacji ubezpieczeń masowych. Z kolei badania Leblanc'a³⁵ są zorientowane na współczesne problemy ryzyk związanych z inwestycjami w niekonwencjonalne źródła energii. Zaproponowany model przeznaczony jest do wyboru portfela ubezpieczeń dotyczących tego rodzaju inwestycji. Autor stwierdza, że ryzyka w niektórych tego rodzaju przedsięwzięciach, modeluje się lepiej za pomocą parametrów finansowych.

Na zakończenie tej części tekstu wskazujemy jeszcze skrótowo³⁶ dwa rozwijane od dłuższego czasu rozwiązania, które przyporządkować można do wstępnych prób ontologizowania domen ubezpieczeń (ACORD³⁷) oraz szerzej – sprawozdawczości finansowej (XBRL³⁸).

ACORD (Association for Cooperative Research and Development) jest niezależną organizacją zrzeszającą interesariuszy rynku ubezpieczeń, która od lat 70. zajmuje się i rozwija standardy gromadzenia i obiegu informacji w ubezpieczeniach. W roku 1999 zasoby w tym zakresie zostały przeniesione na ramę definiującą słownictwo do opisu procesów i operacji w przedsiębiorstwach rynku ubezpieczeniowego oraz pomiędzy nimi. W tym celu wykorzystano język XML. Tak powstał promowany przez tę organizację ACORD XML.

Z kolei XBRL (*eXtensible Business Reporting Language*) to język oparty na XML mający na celu dostarczenie słownictwa do reprezentowania dokumentów finansowych w zakresie wewnętrznej i zewnętrznej sprawozdawczości. Dzięki zastosowaniu jednoznacznych reguł użycia, racjonalnemu podziałowi na poddomeny zdefiniowane taksonomie w obecnej wersji standardu dostarczają bardzo wyczerpujące i elastyczne narzędzie do reprezentacji sprawozdawczości finansowej w formie elektronicznej przyczyniając się istotnie do wzrostu interoperacyjności. Do potencjalnych korzyści z wykorzystania XBRL w ubezpieczeniach zaliczyć można: skonsolidowany zarząd i raportowanie, analizę aktuarialną, ocenę sytuacji partnerów i reasekuranatów, analizę portfeli inwestycyjnych, określenie ryzyk korporacyjnych, wsparcie narzędzi regulacyjnych oraz właścicielskich.

2. Ontologie

W poprzednim podrozdziale przedstawiliśmy różne perspektywy badań nad rozwojem modeli dla potrzeb systemów informacyjnych w dziedzinie ubezpieczeń. Wspomnieliśmy także o możliwości pojawienia się nowego podejścia wykorzystującego technologie semantyczne. Ich podstawą są ontologie. W tej i w następnych sekcjach postaramy się szczegółowo przedstawić, czym są ontologie i na czym to nowe podejście ma polegać.

35. Leblanc E., *Challenges of the Renewable Energy Industry Generate New Demands for Risk Advisory: How to Value an Insurance Package from a Financing Perspective?* Geneva Papers on Risk & Insurance, 2008, 33(1), s. 147.

36. Szersze opracowanie przedstawimy w innym artykule.

37. <http://www.acord.org>, odczytano 25 lipca 2010 r.

38. <http://www.xbrl.org>, odczytano 25 lipca 2010 r.

Pojęcie ontologii wywodzi się z filozofii. Sama nazwa pojawiła się stosunkowo późno. Użyta została po raz pierwszy jako synonim słowa „abstrakcja” w słowniku filozoficznym zatytułowanym *Lexicon philosophicum quo tamquam clavae philosophiae fors aperiuntur* wydany w Frankfurtu w 1613 roku przez Rudolfa Gocleniusa. Jednak sens, jaki jej współcześnie nadajemy, konstruowany był przez filozofów starożytnej Grecji.

Ontologie w naukach związanych z przetwarzaniem informacji upowszechniają się głównie w związku z gwałtownym rozwojem nauk kognitywnych (w szczególności sztucznej inteligencji). Co prawda pierwsze pojawienie się pojęcia „ontologia” w informatyce pochodzi z roku 1967, jednak wówczas było to tylko pojęcie pomocnicze techniki modelowania danych. Dopiero gwałtowny rozwój informatyki, a w szczególności Internetu, a co za tym idzie porządkowania pojęć dla późniejszego wykorzystania ich w reprezentacji wiedzy i automatycznym wnioskowaniu, spowodowało nadanie „ontologii” nowego sensu.

Poprzez termin „ontologia” rozumie się „generalnie sformalizowane definicje pojęć i ich konceptualizację oraz formalne określenie powiązań zachodzących pomiędzy nimi”³⁹.

Z definicji tej wynika, że ontologia zajmuje się opisem i zdefiniowaniem „tego, co jest”, występującego pod różnymi postaciami – bytów napotkanych rzeczywiście, stworzonych i wyobrażonych, jako pojęcia i myśli, czy wreszcie pod postacią symboli, które można zapisać. Ontologia z zasady zajmuje się pewnym wycinkiem rzeczywistości – ten wycinek (domena) musi być określony. Jednoznaczność przekazu opisywanej wiedzy wymaga wykorzystania różnorodnych technik opisu.

Ontologia jest abstrakcyjnym modelem rzeczywistości, przy czym w zależności od podejścia może ona zawierać także elementy formalnej teorii⁴⁰. Do podstawowych rodzajów ontologii będących jednocześnie wczesnymi i najbardziej prymitywnymi formami prostych ontologii są kategoryzacja oraz hierarchizacja. Definiowane są one następująco:

- Kategoryzacja – zdolność przyporządkowania symbolu występującego w komunikacji do określonej grupy obiektów, które to obiekty posiadają określone cechy, np. „ubezpieczenie” – klasa ubezpieczeń, pojęcie ogólne ubezpieczenie. Zestaw tych grup można określić jako zewnętrzny model pojmowania świata
- Hierarchizacja – umiejscowienie określonej klasy w hierarchicznej strukturze. Instancja klasy, poza oczywistymi charakterystykami wynikającymi z przynależności do klasy, posiada także cechy dziedziczone z klas nadrzędnych⁴¹.

39. Gruber T.R., *A Translation Approach to Portable Ontology Specifications*, Knowledge Acquisition, Vol. 5, 1993.

40. Formalizmu użytego do tworzenia ontologii. Np. języki takie OWL DL czy WSML oferują do wykorzystania aparat logiczny – logikę deskryptywną lub formalizm zbliżony do logik pierwszego rzędu.

41. Maedche, A., Staab, S., *Measuring Similarity between Ontologies*. In: Proc. of the European Conference on Knowledge Acquisition and Management – EKAW-2002, Madrid, Spain, LNCS/LNAI 2473, Springer, 2002.

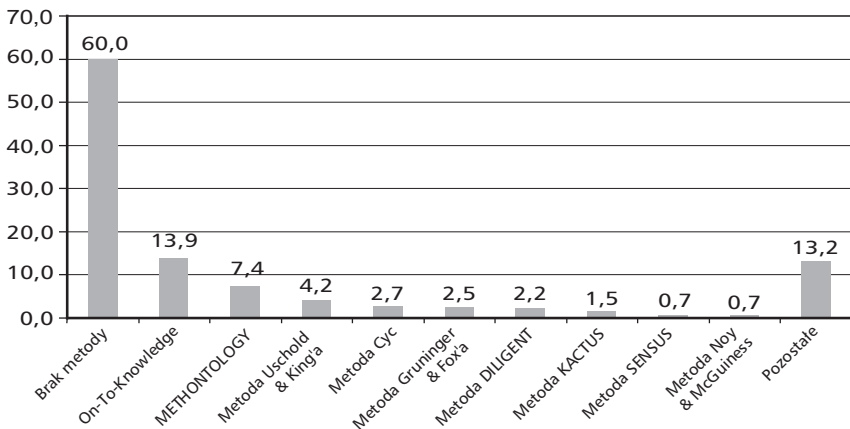
Zastosowanie zarówno kategoryzacji, jak i hierarchizacji, pozwala budować proste abstrakcyjne modele rzeczywistości. W przypadku zastosowania do zapisu takiego modelu wyodrębnionego wycinka świata formalnego aparatu (języka opisu), daje w rezultacie prostą ontologię.

Warto wyraźniej podkreślić cechy charakterystyczne ontologii. W literaturze często pojawiają się następujące postulaty dotyczące nie tyle samej ich konstrukcji, co ontologii jako takich:

- Ontologia nie stanowi listy, katalogu czy taksonomii obiektów, stwarza natomiast formalne przesłanki, wedle których takowe mogą być budowane
- Ontologia jest oderwana od epistemologii (teorii poznania), powiązana jest z obiektem, a nie jego subiektywnym odbiorem
- Ontologia musi uchwycić rzeczywistość na różnych poziomach atomizacji, jak również relacje pomiędzy tak wyodrębnionymi warstwami
- Naturalne jest istnienie wielu ontologii – uznanie braku możliwości stworzenia jednej ogólnej ontologii, przy czym każda ontologia może być poprawna
- W przeciwieństwie do typowego podejścia analitycznego relacje między obiektami nie są ujęte w postaci funkcji – zależności nie muszą być ilościowe
- Nauka rozpoczyna proces od mierzenia i predykcji, ontologia zaś od stworzenia taksonomii⁴².

Tworzenie oraz wykorzystanie ontologii jest już stosunkowo dobrze rozpoznane. Istnieje określona liczba opisanych i przebadanych metod oraz narzędzi te metody wspierających. Wykres 1 zawiera wyszczególnienie przykładowych metod.

Wykres 1. Zestawienie metod tworzenia ontologii wraz z częstością ich zastosowania (w proc.)

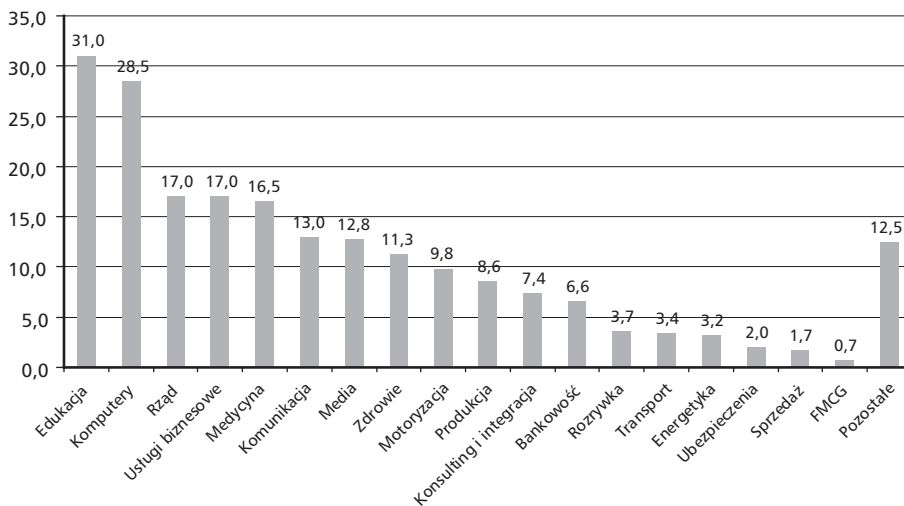


Źródło: Cardoso J., „The Semantic Web Vision: Where are We?” *IEEE Intelligent Systems*, September/October 2007, s. 22-26, 2007.

42 *Ibidem*.

Rozpoczynając od prostych modeli wiedzy stanowiących punkt odniesienia dla stron przekazu, jako swoisty rodzaj meta-języka, pojawiają się coraz to nowe pomysły zastosowań ontologii. One same zaś przechodzą przeobrażenia do bardzo zaawansowanych baz wiedzy.

Wykres 2. Procentowe zestawienie rozwijanych ontologii (wg danych dostarczonych przez ich twórców) w rozbiciu na dziedziny działalności gospodarczej (w proc.).



Źródło: Cardoso J., „The Semantic Web Vision: Where are We?” *IEEE Intelligent Systems*, September/October 2007, s. 22-26, 2007.

Wykres 2 przedstawia zestawienie dziedzin działalności gospodarczej, dla których tworzone są ontologie⁴³. Wynik osiągnięty dla sektora ubezpieczeń zdecydowanie nie odzwierciedla skali potrzeb i możliwości, jakie byłyby dane przez wykorzystanie mechanizmów, o których mówimy. Jest on naszym zdaniem wynikiem niskiego poziomu dyfuzji wiedzy o nowoczesnych rozwiązaniach w zakresie zarządzania wiedzą do podmiotów tego sektora.

3. Budowa ontologii

W literaturze można spotkać się z twierdzeniem, iż jednym z podstawowych zastosowań tworzonych ontologii jest pełnienie funkcji „(...) systemu pojęć na potrzeby konceptualizacji określonej dziedziny (...)”. A jednocześnie, co za tym idzie, „(...) narzucania sposobu rozumienia tej dziedziny poprzez wskazanie pojęć podstawowych, do których odnoszone są wszystkie inne⁴⁴(...)”. Można zaryzykować twierdzenie, że jednym z ważniejszych celów tworzenia ontologii jest kreacja wąskiego, specjalistycznego aparatu pojęciowego. Aparat ten w zamyśle ma wspomagać i odpowiadać za sprawną

43. Badania zostały oparte na danych ze Stanów Zjednoczonych Am. Półn. z 2007 roku.

44. Vetulani Z., *Systemy ontologiczne wobec rozwoju eDziedzin*, w: *Sztuczna Inteligencja, organizacje wirtualne*, nr 22, WAP, Siedlce 2003.

oraz jednoznaczną wymianę informacji pomiędzy jego użytkownikami. Równocześnie wiedza, która jest zawarta w stworzonym aparacie pojęciowym, może wspomagać rozumowania wykonywane zarówno przez ludzi, jak i systemy informatyczne.

Warto zwrócić uwagę na wyszczególnienie potencjalnych odbiorców ontologii. Kto może, w bezpośredni (lub pośredni) sposób, wykorzystywać tworzone modele rzeczywistości? Przede wszystkim agenty⁴⁵ działające w systemach informacyjnych, w szczególności te o dużym stopniu zautomatyzowania oraz heterogeniczności architektury (systemy eksperckie, w tym w szczególności systemy wspomaganie decyzji), ale także możliwe jest szerokie zastosowanie w ramach koncepcji Internetu Semantycznego⁴⁶. Niemniej warto zwrócić uwagę, że wykorzystywać ontologię mogą także ludzie – w sposób bezpośredni, np. jako źródła referencyjne.

Tak jak stwierdziliśmy powyżej podstawowym celem tworzenia ontologii jest budowa semantycznych elementów języka, którym mogą się posługiwać ludzie i systemy informacyjne. Oprócz celu podstawowego, można także wyróżnić inne cele poboczne:

- Ścisłe (precyzyjne) określenie zakresu dziedziny, dla której ontologia jest tworzona
- Określenie pojęć kluczowych dla danej dziedziny, a także wagi pojęć pomocniczych (mniej istotnych)
- Precyzyjne zdefiniowanie pojęć tworzących ontologię
- Likwidacja wieloznaczności oraz redundancji w wyrażeniach zaczerpniętych ze słownika potocznego
- Utworzenie właściwej hierarchii pojęć oraz określenie zachodzących pomiędzy nimi relacji⁴⁷.

Jak więc wynika wyraźnie z powyższej listy, tworzenie ontologii ma w konsekwencji charakter porządkujący oraz dookreślający. Z kolei wykorzystanie języków formalnych może służyć istotnemu wzrostowi użyteczności zgromadzonej wiedzy poprzez zaprężnięte do tego celu zaawansowane narzędzia informatyczne (wykorzystujące tak potraktowaną ontologię, jako bazę wiedzy). Dodatkowo nie można zapomnieć o kolejnej zalecie właściwie skonstruowanej ontologii, jaką jest możliwość wielokrotnego wykorzystania zamodelowanej wiedzy.

Wreszcie, mając daną określoną ontologiczną strukturę wiedzy, możliwe jest zwiększenie precyzji modelowania złożonych relacji zachodzących w świecie rzeczywistym. Taki wzrost dokładności będzie miał wpływ na wierniejsze respektowanie standardów w danej dziedzinie zastosowań – o ile zostaną one uwzględnione w samej ontologii.

Jak już wcześniej wspomniano, początkowe formy prostych ontologii wykorzystywały takie instrumenty jak kategoryzacja i hierarchizacja. Pewnymi formami ontologii, powstającymi dużo wcześniej, przed epoką informatyzacji, były słowniki. Różniły się

45. Przez termin „agenty” rozumiemy tutaj szeroki zakres bytów wchodzących w interakcje z systemem (mogą to być agenty programowe, użytkownicy – ludzie lub wydzielone podsystemy oraz komponenty).

46. W obu przypadkach warunkiem jest posiadanie odpowiednich narzędzi i infrastruktury – przede wszystkim sformalizowanych języków – do reprezentacji wiedzy zawartej w zamierzonej ontologii.

47. Gruber T.R., *A Translation Approach to Portable Ontology Specifications*, Knowledge Acquisition, Vol. 5, 1993.

one jednak przede wszystkim brakiem lub niskim stopniem formalizacji opisywanych pojęć.

Struktury ontologiczne można pogrupować według różnych kryteriów. Przykładowy podział ontologii może wyglądać następująco:

- Taksonomie – wyczerpujące klasyfikacje obejmujące wiele dziedzin, np. kategorie w systemach wyszukiwawczych www
- Klasyfikacje – podziały obejmujące najwyżej kilka dziedzin
- Katalogi – np. katalogi produktów w sklepach *on-line*
- Specjalistyczne słowniki – ontologie definiujące standardową terminologię używaną w określonym kontekście, np. UMLS (Zunifikowany Język Medyczny), UNSPSC (Zunifikowany Język Produktów i Usług)⁴⁸.

4. Podstawowe elementy ontologii oraz meta-ontologie

Aby tworzona ontologia mogła sprostać wszystkim wymienionym wcześniej założeniom i aby możliwe było osiągnięcie stawianych przed nią celów, niezbędna jest odpowiednia konstrukcja. Konstrukcja ontologii złożona jest oczywiście z różnorodnych elementów. Zazwyczaj elementami takimi są:

- Koncepty (ożywione i nieożywione) ściśle związane z daną dziedziną (w tym także osoby)
- Cechy (właściwości) i atrybuty odpowiadające tym konceptom
- Ograniczenia cech i atrybutów
- Konkretni przedstawiciele określonych konceptów
- Relacje.

Wymienione elementy nazywają typy bytów, z których składają się określone rodzaje ontologii. Ontologie mogą różnić się pod względem typów elementów je tworzących. Jeżeli ze zbioru tychże typów utworzymy pojęcia w pewnej ontologii, otrzymamy meta-ontologię.

Stworzenie meta-ontologii definiującej wszystkie elementy, z których mogą być budowane inne ontologie, byłoby trudne. Przy znacznej większości tradycyjnych sposobów użycia, niecelowe. Dlatego zazwyczaj poprzez meta-ontologię rozumieć będziemy ontologię (wyższego rzędu) opisującą elementy (byty) wykorzystywane w związku z projektowaniem i budowaniem danej ontologii niższego rzędu. Inaczej mówiąc meta-ontologia (przy bardziej pragmatycznym podejściu), to ontologia opisująca szereg pojęć i ich relacji, opisujących byty wykorzystywane w procesie kreowania czy opisywania struktury projektowanej ontologii – jest to pewien dobrze skonstruowany podzbiór pełnej meta-ontologii.

Najważniejszymi składnikami tworzącymi meta-ontologie są:

- Podontologie (ontologie podrzędne, mikroteorie) – zbiory porządkujące zawierające powiązane ze sobą koncepty

48. Noy N. F., *Ontology Engineering for the Semantic Web and Beyond*, prezentacja, www.co-ode.org/resources/tutorials/intro/slides/OntologyEngineering.ppt, odczytano: 16 września 2006 r.

- Koncepty – typy, reprezentujące zbiory (grupy) obiektów o wspólnych cechach
- Nadkoncepty – pojęcia nadrzędne – taksonomiczni przodkowie podkonceptów
- Podkoncepty – pojęcia dziedziczące cechy jednego lub kilku nadkonceptów
- Instancje – konkretyzacje konceptów lub relacji, mogące odpowiadać bytom rzeczywistym lub abstrakcyjnym
- Anotacje – działania przeprowadzane przez konkretne obiekty
- Właściwości – cechy obiektów i związki mogące między nimi zachodzić
- Ograniczenia – wartości brzegowe charakterystyczne dla danych właściwości
- Relacje – związki znaczeniowe, specyficzny typ właściwości nie określającej jednego bytu, lecz wzajemne stosunki zachodzące między co najmniej dwoma
- Cechy – pola opisujące dany koncept i mogące przyjąć wartości o określonym typie⁴⁹.

Z częścią tych bytów czytelnik miał okazję już się spotkać przy okazji omówienia składowych samych ontologii. Należy jednak pamiętać, że w przypadku meta-ontologii pojęcia te mają odmienną konotację.

Przykłady przedstawione w jednej z cytowanych już prac⁵⁰ bardzo dobrze obrazują potencjał zastosowań technologii semantycznych oraz ontologii dla rozwoju informatyki ubezpieczeniowej. Jak wynika jednak z badań (Wykres 3), zaledwie 2 proc. tworzonych ontologii przeznaczonych jest dla ubezpieczeń.

Zakończenie

W rozprawie wskazano na ontologie, jako sposób reprezentacji stosowany w zarządzaniu wiedzą. Przygotowano w ten sposób podstawę konceptualną dla zademonstrowania możliwości zastosowania ontologii w ubezpieczeniach poprzez wsparcie dwóch kluczowych dla tej domeny obszarów: reprezentacji wiedzy o regulacjach prawnych oraz normalizacji pojęcia ryzyka. Ponadto wskazano na kierunki rozwoju tworzenia modeli dla potrzeb ubezpieczeń argumentując, że rozpoczęcie operowania na modelach semantycznych może otwierać nowe możliwości związane z gromadzeniem i przetwarzaniem informacji.

49. Noy N. F., McGuinness D. L. *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

50. Tomaszewski T. *op. cit.*

Wykaz źródeł:

Abramowicz W., *Filtrowanie Informacji*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2008.

Alexander P., Baden S., *Glossary on macroeconomics from a gender perspective*, Institute of Development Studies, University of Sussex, <http://stats.oecd.org/glossary/detail.asp?ID=6813>, (24.05.2010).

Bar-Niv R., Bickelhaupt D. L., *Research in International Risk and Insurance: Summary, Synthesis, and Prospects* „Journal of Risk and Insurance”, 1986 nr 53(1).

Braun H., Koeniger W., *On the role of market insurance in a dynamic model*, „Geneva Risk and Insurance Review”, 2007 nr 32(1).

Cardoso J., *The Semantic Web Vision: Where are We?* „IEEE Intelligent Systems”, 2007 nr 9/10.

Craig H., *New directions in risk management*, „Canadian Insurance”, 1996 nr 101(11).

Djehiche B., *Review of Non-Life Insurance Mathematics – An Introduction with Stochastic Processes*, „Scandinavian Actuarial Journal”, 2004 nr 6.

Fama E. F., *Multifactor Portfolio Efficiency and Multifactor Asset Pricing*, „The Journal of Financial and Quantitative Analysis”, 1996 nr 4.

Frees E. W., Wang P., *Copula credibility for aggregate loss models*, „Insurance, Mathematics & Economics”, 2006, nr 38(2).

Janssen J., De Dominicis R., *Finite Non-Homogeneous Semi-Markov Processes: Theoretical and Computational Aspects*, „Insurance, Mathematics & Economics”, 1984 nr 3.

Gobble E., Windeler D., *Catastrophe modeling: Shifting perceptions*, Canadian Underwriter, 2003 nr 70(6).

Goldstein A. B., Markowitz B. G., *SOFASIM: A Dynamic Insurance Model with Investment Structure, Policy Benefits and Taxes*, „The Journal of Finance”, 1982 nr 37(2).

Gori S., Padrone M., *Is game theory a useful tool for terrorism insurance?*, „International Journal of Risk Assessment and Management”, 2007 nr 7(8).

Gruber T. R., *A Translation Approach to Portable Ontology Specifications*, Technical Report KSL 92-71, KNOWLEDGE SYSTEMS LABORATORY, Computer Science Department, Stanford University <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.3273&rep=rep1&type=pdf> (24.05.2010).

Guarino N., *Formal Ontology and Information Systems*, w „Proceedings of FOIS'98”, Trento, Italy, 6-8 June 1998, IOS Press Amsterdam 1998.

Handschke J., *Internet w działalności ubezpieczeniowej w Polsce: ze szczególnym uwzględnieniem dystrybucji*, red. Jerzy Handschke, Katedra Ubezpieczeń Akademii Ekonomicznej, Poznań 2004.

Hasło – Ontologia, <http://pl.wikipedia.org/wiki/Ontologia>.

Helwig Z., *Rozważania nad istotą modelu ekonometrycznego*, „Economista”, 1974 nr 4.

Henderson, G. V. Jr., *Problems and Solutions in Conducting Event Studies*. „Journal of Risk and Insurance”, 1990 nr 57(2).

Jin Y., Rejesus R. M., Little B. B., *Binary choice models for rare events data: a crop insurance fraud application*, „Applied Economics”, 2005 nr 37(7).

- Kihlstrom R. E., Roth A. E., *Risk Aversion and the Negotiation of Insurance Contracts.*, „Journal of Risk and Insurance”, 1982 nr 49(3).
- Kowaczyk P., Poprawska E., Ronka-Chmielowiec W., *Metody aktuarialne*, PWN, Warszawa 2006.
- Lai L., *Underwriting profit margin of P/L insurance in the fuzzy-ICAPM*, „Geneva Risk and Insurance Review”, 2006 nr 31(1).
- Lane J., Glennon D., *The Estimation of Age/Earnings Profiles in Wrongful Death and Injury Cases*, „Journal of Risk and Insurance”, 1985 nr 52(4).
- Lange O., *The Scope and Method of Economics*, „The Review of Economic Studies”, 1945 - 1946, nr 1.
- Leblanc E., *Challenges of the Renewable Energy Industry Generate New Demands for Risk Advisory: How to Value an Insurance Package from a Financing Perspective?*, „Geneva Papers on Risk & Insurance”, 2008 nr 33(1).
- Maedche, A., Staab, S., *Measuring Similarity between Ontologies*, w: „Proc. of the European Conference on Knowledge Acquisition and Management – EKAW-2002”, Madrid, Spain, Springer-Verlag, Heilderberg, Berlin 2002, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.5761&rep=rep1&type=pdf> (24.05.2010).
- Noy N. F., McGuinness D. L., *Ontology Development 101: A Guide to Creating Your First Ontology*, Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.
- Noy N. F., *Ontology Engineering for the Semantic Web and Beyond*, <http://www.coode.org/resources/tutorials/intro/slides/OntologyEngineering.ppt> (24.05.2010).
- Sharp D., *Leksykon pojęć i idei C.G. Junga*, Wydawnictwo Wrocławskie, 1999.
- Szurek M., *Matematyka dla humanistów*, Wydawnictwo RTW, Warszawa 2000.
- Tomaszewski T., *Przykłady zastosowań koncepcji semantycznej reprezentacji ryzyka w ubezpieczeniowych systemach informacyjnych*, w „Studia Ubezpieczeniowe”, red. Jerzy Handschke, Zeszyty Naukowe Nr 127, Uniwersytet Ekonomiczny w Poznaniu, Poznań 2009.
- Vetulani Z., *Systemy ontologiczne wobec rozwoju eDziedzin*, w: „Sztuczna Inteligencja, organizacje wirtualne”, WAP, Siedlce 2003, nr 22.
- Yin G., Liu Y. J., Yang H., *Bounds of ruin probability for regime-switching models using time scale separation*, „Scandinavian Actuarial Journal”, 2006 nr 2.

Ontology as a tool used for building models in insurance information systems – Summary

In this text, the authors summarise the work carried out as part of modelling knowledge and insurance information. The authors present ontology in IT, that is a new structure used to gather and sort out information. Additionally, the article also indicates two crucial areas related to insurance modelling, namely representation of knowledge of legal regulations and the standardisation of the concept of risk. Both areas are very important for the correctness of information flow processes in insurance market companies. The theoretical review based on literature research helps present, in a synthetic way, the concept and start

discussion on its possible applications for modelling insurance information by means of semantic models – in particular in the two areas mentioned above. Using ontology and starting operations on semantic models will contribute to gaining new possibilities related to gathering, processing and operations on information in insurance systems.

Prof. dr hab. WITOLD ABRAMOWICZ kieruje Katedrą Informatyki Ekonomicznej na Uniwersytecie Ekonomicznym w Poznaniu, prowadzi cykl konferencji *Business Information Systems* odbywających się co roku.

Dr PIOTR STOLARSKI jest pracownikiem Uniwersytetu Ekonomicznego w Poznaniu, w Katedrze Informatyki Ekonomicznej. Absolwent Wydziału Ekonomii Akademii Ekonomicznej w Poznaniu oraz Wydziału Prawa i Administracji UAM.

Dr TADEUSZ TOMASZEWSKI jest niezależnym konsultantem, wcześniej adiunktem i starszym wykładowcą na Uniwersytecie Ekonomicznym w Poznaniu, w Katedrze Informatyki Ekonomicznej.

Recenzenci: dr hab. Grzegorz Bartoszewicz, dr Stefan Szyszko.