

Wybrane problemy semantyczne we współczesnych systemach informatycznych

1. Wstęp

Systemy komputerowe, obecne we wszystkich dziedzinach współczesnego życia, w rzeczywistości tylko przyspieszyły te czynności intelektualne, które charakteryzowały się automatyzmem działania i dały się sformalizować oraz zalgorytmizować. Należy pamiętać, że system komputerowy operuje na danych zapisanych binarnie i na modelach formalnych, czyli na pewnym przybliżeniu rzeczywistego świata – wynegocjowanym między twórcami oprogramowania a ich zleceniodawcami. Działanie każdego systemu informatycznego jest zatem związane z pewną umową-zgodą na ograniczenie obrazu środowiska, w którym system informatyczny ma funkcjonować, do jego najważniejszych elementów. Uproszczenie złożonej rzeczywistości jest konieczne do jej modelowania. Zadaniem systemu informatycznego jest przetwarzanie danych. W wielu jednak zastosowaniach pojawiła się potrzeba poszukiwania podobieństw (czy wręcz identyczności) nie tylko syntaktycznych, ale również znaczeniowych. Braki jedynie syntaktycznego przetwarzania próbujemy wzbogacić standardami semantycznymi, mającymi poprawić użyteczność np. wyszukiwania w Internecie.

Wzbogacanie modeli metadanymi nie rozwiązuje problemu. Nawet jeśli oznaczamy pewne pojęcia tagami zgodnymi ze standardami RDF czy OWL, to nadal systemy będą operować na znakach tworzących owe tagi – czyli na danych w systemach formalnych. Analiza „poprawności” przetwarzania informatycznego jest możliwa jedynie na poziomie syntaktycznym. Analiza semantyki, czyli znaczenia, pozostaje zatem poza możliwościami tych systemów przetwarzania.

Celem artykułu jest próba wyjaśnienia – z pragmatycznej i socjologicznej perspektywy badawczej – problemów semantycznych, wynikających z tzw. luki semantycznej, czyli rozbieżności między modelem a jego interpretacją w konfrontacji

¹ Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych.

² Uniwersytet Łódzki, Wydział Zarządzania.

z rzeczywistością, oraz różnic występujących często między pojęciami danych i informacji, które są (i nadal mogą być – z milczącym rozróżnieniem) używane zamiennie, a także zwrócenie uwagi na problem fetyszyzacji technologii i informacji. Problemy te dotyczą także wyszukiwarek semantycznych.

2. Informacja a dane

W dyskursie współczesnym pojęcie informacji jest często stosowane, ale czasami można stwierdzić jego nadużywanie, a nawet fetyszyzację. R. Barbrook³ wskazuje, że na przestrzeni (już ponad 100) lat fetyszym towarowy doprowadził do uniwersalnego fetysyzmu technologicznego i że w ślad za tym obserwujemy fetyszyzację informacji. Chociaż wartość informacji w procesach podejmowania decyzji jest niezaprzeczalna, to nadal pojęcie to jest często używane zamiennie z pojęciem danych (a także wiedzy) w kontekście, np.: ICT (służących przede wszystkim do przetwarzania danych), systemów informatycznych (oprogramowanie, aplikacje), baz danych (zgromadzonych, przetwarzanych, udostępnianych itd.). Jak widać z tych przykładów, chodzi przede wszystkim o dane.

W literaturze przedmiotu już od dłuższego czasu funkcjonuje swoista „piramida pojęć”: dane–informacja–wiedza (mądrość), lecz różnica między nimi nie do końca jest wyraźna. W normie PN-ISO/IEC 2382–1:1996⁴ informacja (01.01.01) jest określana jako „wiedza o”; dane (01.01.02) – „jako reprezentacja mająca interpretację” (istniejąca sprzeczność widoczna jest przy porównaniu z definicją „tekstu” poniżej). Nie do końca wiadomo, na czym dokładnie polega różnica między komputerowym przetwarzaniem informacji (01.01.05) a przetwarzaniem danych (01.01.06) – w przypadku systemu formalnego oba zapisy są nierozróżnialne. Natomiast wyraźne jest utożsamianie tekstu (01.01.03) z danymi, przy czym interpretacja taka w istotny sposób opierałaby się na wiedzy (kompetencji) użytkownika w zakresie języka naturalnego lub sztucznego. Zamieszanie wokół używanych pojęć wymaga ich uporządkowania przy określeniu przedmiotu badania i jego charakteru.

Dane (jak sama nazwa wskazuje) to wszystko, co jest nam dane w postaci utrwalonej na nośniku materialnym, niezależnie od stopnia przetworzenia,

³ R. Barbrook, *Przyszłości wyobrażone. Od myślącej maszyny do globalnej wioski*, Muza, Warszawa 2009.

⁴ <http://www.wsb.edu.pl/download.php?id=1182&source=pr> [dostęp 08.06.2016].

zorganizowania itd. W procesie digitalizacji (konwersji danych analogowych na cyfrowe) obiekty świata zewnętrznego (fakty, zdarzenia, prawidłowości itp. jako znaki, dźwięki, obrazy) są kodowane i utrwalane w postaci cyfrowej – łatwiejszej do przetwarzania i następnego odtwarzania w postaci tekstów (dokumentów, książek), muzyki, głosów, zdjęć, filmów. Wszystkie procesy przebiegają według określonych reguł za pomocą różnych algorytmów w sposób formalny. Dane te są produkowane, gromadzone, przechowywane, ochraniające, przesyłane i przekazywane – początkowo – w stanie „surowym” (nieprzetworzone, nieposegregowane, nieskatalogowane). Następnie potrzebne są: konwersja danych ze źródłowego nośnika na docelowy, ich oczyszczanie (interpretacja, redukcja redundancji, integracja, standaryzacja, agregacja), składowanie (w tym budowanie modelu danych w rozumieniu bazy danych) oraz rejestracja ich kategorii i wymiarów⁵ i – na koniec – utworzenie pliku cyfrowego odpowiednio opisanego (metryczki, tagi, deskryptory lub hasła przedmiotowe). Większość z tych czynności wymaga wykwalifikowanych i kompetentnych zasobów ludzkich⁶.

Należy zwrócić uwagę na fakt, że raz utrwalone na materialnym nośniku dane pozostają danymi. W każdej postaci będzie potrzebne analizowanie „od początku” pod kątem możliwości automatyzacji przetwarzania do celów użytkowych. Należy pamiętać, że ten sam zestaw danych (ciąg sygnałów/znaków) może być źródłem różnych informacji dla różnych użytkowników.

Informacja to każdy czynnik abstrakcyjny, który może być wykorzystany do racjonalnego postępowania/sterowania przez organizmy żywe lub/i urządzenia automatyczne⁷. Jest potrzebna do podejmowania decyzji w konkretnej sytuacji – w zależności od kontekstu i indywidualnej kategoryzacji priorytetów. Informacja⁸ pojawia się dopiero wtedy, gdy nastąpi interpretacja danych oraz nadanie im znaczenia (sensu) przez podmiot w konkretnym kontekście⁹. Jako prerekwizyt takiej interpretacji danych jest konieczna uprzednia wiedza podmiotu (kapitał

⁵ <http://www.manperion.com/zarządzanie-informacją-zi/informacja-w-biznesie/etapy-produkcji-informacji/> [dostęp 10.01.2016].

⁶ A.Z. Wassilew, *Cena a wartość w gospodarce cyfrowej – wybrane problemy*, „Gospodarka XXI Wieku. Ekonomiczne Problemy Usług”, nr 122, Uniwersytet Szczeciński, Szczecin 2016, s. 168.

⁷ *Encyklopedia powszechna PWN*, Wydawnictwo Naukowe PWN, Warszawa 1995.

⁸ Już na wstępie należałoby zauważyć, że w swojej klasycznej pracy *A mathematical theory of communication* C. Shannon zaznacza, że „semantyczne aspekty komunikacji są nieistotne dla problemu inżynierskiego”. C.E. Shannon, *A mathematical theory of communication*, „ACM SIGMOBILE Mobile Computing and Communications Review” 2001, vol. 5(1), s. 3–55. Z tego względu pojęcie znaczenia w dalszej części tego artykułu należałoby rozumieć w sensie adekwatności formalnej.

⁹ <http://www.inzynieriawiedzy.pl/wiedza/wiedza-informacje-dane> [dostęp 09.01.2016].

ludzki), jak je rozumieć (model danych), oraz doświadczenie użytkownika dotyczące tego, w jaki sposób może z tak „wydobytej” informacji skorzystać¹⁰.

Wiedza jako synteza doświadczenia może być deklaratywna (propozycyjna) bądź habitualna (proceduralna). Podmiot nie tylko wie „co”, ale i „jak” to wykorzystać. Według I. Nonaki i H. Takeuchiego wiedza jest jego wytworem, zakorzenionym w przekonaniach i oczekiwaniach odbiorcy¹¹. U Platona zaś wiedza to prawdziwe, uzasadnione przekonanie.

Na poziomie danych możemy badać relacje między znakami (syntaktyka), na poziomie informacji – relacje między znakami a ich znaczeniami (semantyka), natomiast na poziomie wiedzy – relacje między znakami a ich użytkownikami oraz ich użyciem w określonym kontekście (pragmatyka).

3. Standardy semantyczne

Gromadzenie i szybkie wyszukiwanie danych przez systemy informatyczne przyczyniły się pierwotnie do ich popularności i rozwoju. Sieć Internet również dawała możliwość szybkiego, ale niestety nie zawsze skutecznego poszukiwania. Z problemem tym nie radziły sobie narzędzia przeszukujące miliony stron WWW: od moderowanych katalogów do działających automatycznie wyszukiwarek i metawyszukiwarek¹². Przyczynia się do tego nieetyczne pozycjonowanie (farmy linków, zaplecze pozycjonerskie) oraz masowe generowanie treści publikowanych przez użytkowników aplikacji Web 2.0, w efekcie czego bardzo często na szczycie SERP (*search engine results page* – strona z wynikami wyszukiwania) pojawiają się linki do witryn, które nie spełniają oczekiwań szukającego.

Wraz z upowszechnianiem się Internetu i cyfryzacji treści zasoby zgromadzone w formie cyfrowej dynamicznie przyrastały, a stosowane metody wyszukiwania nie pozwalały ich w pełni wykorzystać. Pod koniec XXw. twórca Web 1.0, T. Berners-Lee, rozpoczął prace nad projektem Semantic Web (sieci semantycznej), którego celem było stworzenie i rozpowszechnienie standardów opisywania treści w Internecie w sposób umożliwiający robotom wyszukiwarek czy autonomicznym agentom przetwarzanie danych w sposób adekwatny do ich

¹⁰ A.Z. Wassilew, op.cit., s. 171.

¹¹ I. Nonaka, H. Takeuchi, *Kreowanie wiedzy w organizacji*, Poltext, Warszawa 2000, s. 81.

¹² J. Papińska-Kacperek, B. Gontar, *Wyszukiwarki semantyczne*, w: *Wiedza i komunikacja w innowacyjnych organizacjach*, Wydawnictwo UE w Katowicach, Katowice 2011, s. 134–149.

znaczenia. Sieć semantyczna ma korzystać z istniejącego w Web 1.0 protokołu komunikacyjnego, ale przesyłane dane dzięki nowym standardom będą „rozumiane” przez algorytmy aplikacji. Przekazywane dane wymagają także danych o nich samych, czyli tzw. metadanych, które ułatwiają dostrzeganie powiązań (*Linked Data*). Efektem budowania sieci semantycznej jest ustrukturyzowanie zasobów Internetu i połączenie ich w logiczną sieć. Dzięki metadansom algorytmy mogą interpretować dane i łączyć je znaczeniowo, czyli hierarchizować oraz kategoryzować. Umożliwi to nowym wyszukiwarkom, nazywanym semantycznymi, na powiązanie znaczeń między wyrazami, a nie tylko na wyszukiwanie słów kluczowych.

Znaczenia zasobów informacyjnych określa się za pomocą ontologii. Dzięki nim komputery (czyli algorytmy) są w stanie „zrozumieć” zawartość dokumentów w sieci. Ontologia jest obecnie integralną składową współczesnej nauki. Pierwotnie pojawiła się w filozofii jako dziedzina teorii bytu¹³. W informatyce ontologia to formalny sposób reprezentacji wiedzy przez zdefiniowanie pojęć, ich właściwości oraz relacji między nimi. Stanowi zbiór twierdzeń sformułowanych przy pomocy istniejących standardów. Praktyczna realizacja ontologii polega na zapisaniu wiedzy w postaci drzewa lub grafu, w którym w wierzchołkach znajdują się pojęcia, a krawędzie opisują typ relacji między nimi. W jego korzeniu zapisuje się pojęcie najbardziej ogólne. Zapis metadanych składa się ze zbioru atrybutów niezbędnych do opisu danych. Znany już z innych zastosowań XML pozwala na wprowadzenie dowolnej struktury znaczników, ale nie określa ich znaczenia. To jest możliwe w strukturze RDF (*Resource Description Framework*), opartej na graficznym modelu danych, który składa się z trójelementowych obiektów (*RDF triples*¹⁴). RDF daje możliwość łączenia ich i tworzenia różnych struktur. Bardziej wyrafinowanymi narzędziami do tworzenia ontologii są specjalizowane języki, jak: OWL (*Ontology Web Language*), OIL (*Ontology Inference Layer lub Ontology Interchange Language*), DAML+OIL (*DARPA Agent Markup Language*).

Stosowanie standardów sieci semantycznej w założeniach miało pozwolić na precyzyjniejsze przeszukiwanie sieci. Dzięki nim semantyczna wyszukiwarka lub inna aplikacja wybierze tylko te strony, na których występuje dane pojęcie,

¹³ B. Paż, *Ontologia versus metafizyka? Geneza, rozwój i różne postaci nowożytnej teorii bytu*, „*Filo-Sofija*” 2011, nr 15, s. 817–847.

¹⁴ Zbudowanych według schematu: podmiot, obiekt i predykat; podmiot i obiekt są węzłami, a predykat reprezentuje relację, jaka tu zachodzi (strzałka między nimi).

ale w ściśle zdefiniowanym znaczeniu, a nie tylko dane słowo kluczowe, które jest przecież często wieloznaczne.

Przykładem aplikacji korzystających ze standardów semantycznych są wyszukiwarki semantyczne: analizujące „znaczenie” indeksowanych dokumentów (Hakia, Bing – dawniej Powerset czy zarzucony już projekt Google Squared) oraz przeszukujące zasoby sieci semantycznej (Swoogle, Sindice, Falcons, Watson). Wyszukiwarki analizujące znaczenie przeszukują zawartość witryn WWW i starają się same zinterpretować ich treść poprzez semantyczną i gramatyczną analizę języka naturalnego. Przekształcenie treści wyrażonej w języku naturalnym w zrozumiałą dla algorytmu jest niezmiernie trudne i wymaga zastosowania metod sztucznej inteligencji. Wyszukiwarki analizujące „znaczenie”, przeszukując strony WWW, tworzą własną bazę ontologii. Uczą się zatem nowych pojęć i relacji między nimi, czyli budują „wiedzę”, dzięki czemu na kolejne zapytania dostarczają bardziej relewantne i odpowiednie do zapytań wyniki¹⁵.

Prace nad standardami semantycznymi rozpoczęły się jeszcze przed próbami uporządkowania bałaganu w Web 2.0. Pod koniec lat 90. XXw. powstał standard Dublin Core (1995, przyjęty przez ISO w 2003 r.), podstawowy standard metadanych dedykowany opisywaniu zasobów, a także CBIR-Content Based Image Retrieval (1992), do wyszukiwania obrazów na podstawie ich zawartości (przez podanie hasła QBSE Query By Semantic Example lub próbki obrazu QBVE Query By Visual Example). Również w latach 90. powstał Topic Maps (od 2003 r. standard ISO) do usprawnienia tworzenia indeksów, glosariuszy oraz tezaurusów dla zbiorów dokumentów elektronicznych. W trakcie prac nad różnymi zastosowaniami zauważono istnienie luki semantycznej, czyli rozbieżności między informacją opracowaną na podstawie np. analizy obrazu a jej interpretacją przez użytkownika, innymi słowy – różnicy pomiędzy prostotą opisów, które można obecnie automatycznie generować, a niejednoznacznością semantyczną w zapytaniach użytkowników¹⁶.

Każdy opisywany metadanymi obiekt może być różnie interpretowany, w zależności od kontekstu. Słowny opis interpretacji może mieć wiele synonimów w danym języku naturalnym, a ponadto twórcy używają różnych języków. Zaufanie jest ważnym elementem wizji sieci semantycznej T. Bernersa-Lee¹⁷. Często odnosi się ono do mechanizmów walidacji, że źródłem informacji jest

¹⁵ J. Papińska-Kacperek, B. Gontar, op.cit.

¹⁶ C. Dorai, S. Venkatesh, *Bridging the semantic gap with computational media aesthetics*, „IEEE Multimedia” 2003, vol. 10, no. 2, s. 15–17.

¹⁷ D. Artz, Y. Gil, *A survey of trust in computer science and the semantic web*, „Web Semantics: Science, Services and Agents on the World Wide Web” 2007, vol. 5(2), s. 58–71.

naprawdę ten, kto się za źródło podaje. Powinniśmy mieć także zaufanie do poprawności metadanych – że zostały stworzone obiektywnie, bez intencji wprowadzenia w błąd czy reklamowania swoich poglądów. W przypadku globalnego zastosowania należałoby zatem stosować ten sam język do opisu metadanych. Powinniśmy się posługiwać tymi samymi pojęciami, by algorytmy korzystające z metadanych interpretowały je tak samo. Współcześnie nie chodzi już tylko o poprawniejsze wyszukiwanie odpowiedzi dla użytkownika-człowieka, ale także dla użytkownika-algorytmu lub usługi – wobec rozwijających się usług webowych (czyli komponentów programowych udostępnianych innym komponentom), mashupów czy aplikacji *internet of things* (IoT). Konsumentem informacji zatem jest już nie tylko człowiek.

4. Systemy formalne

W sieci występuje obieg (wyłącznie) wszelakich form digitalnych w postaci bitów (danych) – tworzonych, wymienianych i „konsumowanych”. Systemy informatyczne operują na bitach interpretowanych przez konkretne systemy jako znaki i symbole. Wszelkie tego typu operacje na danych podlegają regułom systemów (czysto logicznych) formalnych.

System formalny jest konstruktem abstrakcyjnym. Do opisu jego przestrzeni potrzebne są:

- **alfabet** – skończony lub przeliczalny zbiór symboli, znaków, liter (abstrakcyjne, niemające więc *implicite* żadnego znaczenia), z których są konstruowane dowolnej długości ciągi – **formuły**;
- **gramatyka** (składnia, syntaktyka) – skończony zbiór **reguł** pozwalających określić, czy dana formuła jest poprawna, dopuszczalna (razem z alfabetem tworzą język formalny);
- skończony zbiór „początkowych” **aksjomatów** *a priori* (bez uzasadnienia) uznanych za prawdziwe;
- zbiór reguł **inferencji** – (skończona liczba) reguł wyprowadzania, następowania, przekształceń (według reguł poprawności) kolejnych formuł (zaczynając od aksjomatów);
- zbiór **twierdzeń** – wszystkich aksjomatów oraz skończony ciąg formuł wyprowadzonych (dowodzone) zgodnie z regułami inferencji.

Należy zwrócić uwagę na fakt, że poprawność danej formuły nie jest równoznaczna z istnieniem rozstrzygalności (algorytmu), czy jest ona twierdzeniem.

W teorii czysto formalnej – nieopierającej się na intuicji (odczuwania faktu czy przekonania) i niezależnej od świata zewnętrznego – trudno jest sformalizować pojęcie prawdy niezależne od znaczenia¹⁸. Mogą więc istnieć formuły, w których przypadku nie można jednoznacznie orzec, czy są prawdziwe, czy nieprawdziwe (bo nic nie znaczą).

Systemy formalne znajdują zastosowanie przy tworzeniu modeli formalnych – opisów abstrakcyjnych wybranych cech badanych obiektów ze świata realnego¹⁹. Możliwe jest stworzenie modelu (odnoszącego się do świata zewnętrznego) poprzez przyporządkowanie obiektom rzeczywistym symboli i formuł w przypadku, gdy możliwa jest interpretacja oraz orzeczenie o prawdziwości twierdzenia formalnego. Możliwe to jest tylko przez werbalizację – w języku naturalnym!

Język naturalny jest również trudny do sformalizowania. Nawet spójna pod względem logicznym gramatyka generatywno-transformacyjna N. Chomsky'ego ogranicza się tylko do poziomu syntaktycznego. Dopuszcza bowiem istnienie poprawnych gramatycznie zdań, lecz bezsensownych. Reguły semantyczne odnoszą się do świata realnego i nie jest możliwe ich pełne sformalizowanie. Nie sposób jednoznacznie sprecyzować pojęć aksjomatu, prawdy, fałszu i dowodu. Relacje semantyczne w systemach całkowicie sformalizowanych są pozbawione znaczenia. Każde zwerbalizowane zdanie jest tylko strukturą powierzchniową myśli głębokiej (którą można rozbudowywać w nieskończoność).

5. Komputery jako silniki syntaktyczne

Działanie komputerów jest realizowane na algorytmach – czyli pewnego rodzaju procesach formalnych. „Każdy program komputerowy to algorytm”. Kluczowe (trudne do zdefiniowania) cechy²⁰ algorytmów to:

- 1) neutralność substratu (niezależność od wykorzystywanych środków oraz symboli) – skutek swojej struktury logicznej;

¹⁸ S. Wereński, *Co intelektualista wiedzieć powinien o... twierdzeniu Gödla*, „Problemy” 1984, nr 451, s. 18–19.

¹⁹ R. Tadeusiewicz, *Biocybernetyka. Metodologiczne podstawy dla inżynierii biomedycznej*, Wydawnictwo Naukowe PWN, Warszawa 2013, s. 23.

²⁰ D. Dennett, *Dźwignie wyobraźni i inne narzędzia do myślenia*, Copernicus Center Press, Kraków 2015, s. 195.

- 2) bezmyślność u podstaw – banalnie proste kroki każdego elementu wykonywanej procedury, niewymagające „od wykonującego mądrych decyzji, subtelnych sądów lub intuicji”²¹;
- 3) gwarantowane rezultaty – niezawodność algorytmu.

Komputery elektroniczne po prostu wykonują wszystkie kroki szybciej.

Ze względu na to, że systemy komputerowe jako systemy fizyczne są zasilane własnościami fizykalnymi („przetłumaczonymi” na kod binarny), reagują na różnice w wartościach szczegółów fizykalnych, a wyniki wszelkich procesów są prezentowane również jako własności fizykalne (dane na nośnikach materialnych), więc za D. Dennettem możemy je określić jako **silniki syntaktyczne**²². Generalizując, możemy stwierdzić, że każdy program działania komputera jest oparty na algorytmach. Wykonując mechanicznie krok za krokiem instrukcje działania, potrzebuje wyłącznie zestawu reguł formalnych bez potrzeby „rozumienia”.

6. Podsumowanie

Wykorzystanie współczesnych narzędzi ICT zależy od możliwości zautomatyzowania – sformalizowania i zalgorytmizowania – zadania, które chcielibyśmy przekazać do rozwiązania przez systemy informatyczne. Ich działanie polega na przetwarzaniu danych w postaci cyfrowej na zasadach systemów formalnych. „Prawidłowość” ich działania można sprawdzać wyłącznie na poziomie składni – stąd komputery możemy określić jako silniki syntaktyczne, co najwyżej inferencyjne. Poziom semantyczny wymagałby interpretacji danych, „rozumienia” oraz nadania im „sensu” (znaczenia) – werbalizacji w języku naturalnym, który cechują niejednoznaczność i nieprecyzyjność.

Uświadomienie tych problemów jest istotne w kontekście sprecyzowania naszych oczekiwań w stosunku do wyszukiwarek semantycznych. Użytkownik spodziewa się sensownej odpowiedzi na swoje zapytanie – szuka zgodności semantycznej, a system korzystający z dostępnych w sieci danych może dostarczyć tylko to, co potrafi przetworzyć strukturalnie na poziomie syntaktycznym.

²¹ Ibidem, s. 196.

²² Ibidem, s. 245, 263, 266.

Bibliografia

- Artz D., Gil Y., *A survey of trust in computer science and the semantic web*, „Web Semantics: Science, Services and Agents on the World Wide Web” 2007, vol. 5(2), s. 58–71.
- Dennett D., *Dźwignie wyobraźni i inne narzędzia do myślenia*, Copernicus Center Press, Kraków 2015.
- Dorai C., Venkatesh S., *Bridging the semantic gap with computational media aesthetics*, „IEEE Multimedia” 2003, vol. 10, no. 2, s. 15–17.
- Hein A.M., *Identification and Bridging of Semantic Gaps in the Context of Multi-Domain Engineering*, Abstracts of the 2010 Forum on Philosophy, Engineering & Technology, Colorado 2010.
- Nonaka I., Takeuchi H., *Kreowanie wiedzy w organizacji*, Poltext, Warszawa 2000.
- Papińska-Kacperek J., Gontar B., *Wyszukiwarki semantyczne*, w: *Wiedza i komunikacja w innowacyjnych organizacjach*, Wydawnictwo UE w Katowicach, Katowice 2011, s. 134–149.
- Paź B., *Ontologia versus metafizyka? Geneza, rozwój i różne postaci nowożytnej teorii bytu*, „Filo–Sofija” 2011, nr 15, s. 817–847.
- Shannon C.E., *A mathematical theory of communication*, „ACM SIGMOBILE Mobile Computing and Communications Review” 2001, vol. 5(1), s. 3–55.
- Tadeusiewicz R., *Biocybernetyka. Metodologiczne podstawy dla inżynierii biomedycznej*, Wydawnictwo Naukowe PWN, Warszawa 2013.
- Wassilew A.Z., *Cena a wartość w gospodarce cyfrowej – wybrane problemy*, „Gospodarka XXI Wieku. Ekonomiczne Problemy Usług”, nr 122, Uniwersytet Szczeciński, Szczecin 2016, s. 165–173.
- Wereński S., *Co intelektualista wiedzieć powinien o... twierdzeniu Gödla*, „Problemy” 1984, nr 451, s. 18–19.
- Zhou Y., De S., Wang W., Moessner K., *Search Techniques for the Web of Things: A Taxonomy and Survey*, „Sensors” 2016, vol. 16(5), s. 600.

Źródła sieciowe

- <http://www.inzynieriawiedzy.pl/wiedza/wiedza-informacje-dane> [dostęp 09.01.2016].
- <http://www.manperion.com/zarzadzanie-informacja-zi/informacja-w-biznesie/etapy-produkcji-informacji/> [dostęp 10.01.2016].
- <http://www.wsb.edu.pl/download.php?id=1182&source=pr> [dostęp 08.06.2016].

* * *

Some Semantic Problems in Contemporary Information Systems

Abstract

Information systems operate on bits – data – interpreted by specific systems as signs and symbols. Any of this type of operations of data are subject to ruling in formal systems. Analyses of the “correctness” are possible only on the syntactic level. The analysis of the semantics or meaning thus remains beyond the capabilities of these processing systems. Even if some of the concepts are denoted by tags compliant with the standards of RDF and OWL, the systems still operate on signs constituting those tags. Therefore determination of the research subject and its nature requires ordering the used concepts.

Keywords: data, information, syntax, semantics

