

Janusz Kaczmarek
Katedra Logiki i Metodologii Nauk UŁ
janusz.kaczmarek@filozof.uni.lodz.pl

**RECENZJA pracy doktorskiej
p. mgr Rafała Trójezaka**

**Reprezentacja, przetwarzanie i udostępnianie
wyników badań naukowych
w systemach informatycznych opartych na ontologii**

Temat, który podjął doktorant, obejmuje kwestie filozoficzne (w szczególności ontologiczne i metodologiczne, a po części z filozofii nauki) oraz informatyczne (bazy wiedzy, ontologia inżynierska), na których się mało znam. Skupię się zatem na analizie problemów filozoficzno – ontologicznych w nadziei, że pozostałe kwestie omówi specjalista w dziedzinie informatyki.

We wstępie, na stronie 15, Autor, p. mgr R. Trójezak, przedstawia główny cel swojej pracy: rozwiązanie powyżej przedstawionego problemu przez stworzenie ontologii praw nauki (praw naukowych); owym „powyżej przedstawionym problemem” jest wg mnie – o ile dobrze odczytuję – zbudowanie bazy wiedzy dotyczącej określonej gałęzi naukowej w taki sposób, by a) łatwiejsze stało się wyszukiwanie informacji naukowej oraz by baza umożliwiała automatyczną „produkcję” nowej wiedzy (na podstawie już istniejącej).

Dla realizacji tego celu (tych celów) Autor proponuje wspomnianą ontologię praw nauki, na którą składają się a) taksonomia praw (z opisem wewnętrznej struktury praw) oraz b) zbiorów praw. Ontologia tak zaproponowana jest nazwana SCIENCE. Według słów samego Autora głównym osiągnięciem filozoficznym rozprawy ma być „stworzenie zunifikowanej klasyfikacji praw nauki, zapis ich wewnętrznej struktury oraz ich formalna reprezentacja w języku logiki pierwszego rzędu”. Ta kwestia zostanie niżej oceniona.

Obok filozoficznego celu Autor wskazuje również na cel aplikacyjny: opracowanie algorytmów i stworzenie aplikacji, która przetwarza i udostępnia prawa, a także generuje problemy badawcze. Ten cel ma charakter informatyczny i należy do inżynierii ontologicznej.

Praca składa się z pięciu rozdziałów dotyczących kolejno 1) ontologii i tzw. ontologii w informatyce, 2) dotychczasowych prób związanych z reprezentacją wyników badań naukowych, 3) analizy praw naukowych i ich różnych klasyfikacji, 4) problemów baz wiedzy i przetwarzania praw naukowych oraz 5) zagadnień związanych udostępnianiem praw naukowych w oparciu o aplikację OntoBeef. Ponadto Autor zawarł w pracy a) Zakończenie, w którym podsumowuje swoje dociekania i wskazuje na kolejne możliwe etapy badań, b) Dodatek, w którym przedstawione są wyniki kwerendy wykrywającej tranzytywność praw funkcyjnych, c) indeks rzeczowy oraz d) indeks osobowy.

Nie jest moim zamiarem streszczanie poszczególnych rozdziałów. Odniosę się jednak do następujących spraw przedstawionych przez Doktoranta.

1. Praca jest przedstawiona do oceny przez Wydział Filozofii Katolickiego Uniwersytetu Lubelskiego, rozumiem więc, że doktorant stara się o doktorat z pewnej specjalności filozoficznej (nota bene informacja o tym powinna być zawarta w liście do recenzenta – to nie uwaga, ale podpowiedź). Jako taką zatem oceniam niniejszą pracę.

2. Prezentacja problemów ontologii oraz ontologii inżynierskiej została przedstawiona jak najbardziej poprawnie, ale – co zauważa sam Doktorant – b. skrótowo. Zdaję sobie sprawę, że nie było celem tej pracy szczegółowe analizowanie poszczególnych koncepcji ontologicznych, a jedynie przedstawienie zasadniczych rysów, pojęć i sposobów badania z podkreśleniem takich zagadnień, które będą lub mogą być wykorzystane przez informatyków (m. in. problem ujęcia istoty rzeczy i zjawisk (s. 21), czy badania Czeżowskiego dotyczące przekonań (s. 23)). Pojawia się jednak pytanie: dlaczego pomija się – jako ważne dla roboty informatycznej – ontologie o charakterze formalnym, np. ontologia kombinacyjna Perzanowskiego, ontologia sytuacji Wolniewicza (która jest konsekwencją głównych tez Traktatu Wittgensteina), ontologia sytuacji Zalty, czy logika istoty Fine'a. Być może szczegółowe tezy wypracowane w tych ontologiach nie znajdują jeszcze zainteresowania w ontologii inżynierskiej

3. Podstawowy dla tej pracy rozdział 3 b. dobry i twórczy; Autor rzeczywiście podjął tu problemy o charakterze ontologicznym (bo w rozdziale 3 mamy do czynienia z przeglądem stanowisk, co należy raczej do historii filozofii lub ma wymiar sprawozdawczy ale nie filozoficzny) i metodologicznym a te należą do filozofii; widać z pewnością pełne panowanie ze strony Autora nad terminologią, pojęciami i twierdzeniami ontologicznymi i metodologicznymi. Doktorant b. zgrabnie przedstawił stanowiska Ajdukiewicza i Krajewskiego dotyczące klasyfikacji praw naukowych, a następnie (punkt 3.1.5) zaproponował własną klasyfikację, która oparta jest na robocie Ajdukiewicza, Krajewskiego oraz własnych doświadczeń uzyskanych przy opracowywaniu publikacji naukowych z dziedziny badań rolno-spożywczych. Schemat tej klasyfikacji (taksonomii) przedstawiony jest na str. 83. Podkreślmy też, że przedstawiona klasyfikacja posłużyła do:

1) stworzenia ontologii Science wyrażonej w języku logiki predykatów (ontologia ta definiuje kilkadziesiąt różnych relacji, które – warto zauważyć – łączą prawa naukowe z przedmiotami, których te prawa dotyczą; chodzi więc o relacje między prawami a endurantami, perdurantami i/lub cechami),

2) opracowania ontologii praw naukowych w logice opisowej.

Wskażmy drobne problemy:

1) w ontologii DOLCE, na której oparta jest Science wymienia się także abstrakty. Dlaczego relacja *concerns* nie dotyczy abstraktów?

2) (ad. str. 67) czy Autor posiada definicję (kryterium) rodzaju naturalnego? dlaczego metale stanowią rodzaj naturalny a prezydenci Stanów Zjednoczonych nie (rozumiem, że ludzie stanowią również rodzaj naturalny);

3) (ad. str. 82) prawa metodyczne stanowią podklasę praw inkluzywnych, ale zauważmy, że prawa inkluzywne dotyczą przedmiotów pewnej dziedziny, a prawa metodyczne nie przedmiotów ale samych metod; czy nie ma tu nieporozumienia?

Tę robotę oceniam b. pozytywnie. Jest zaprezentowana czytelnie a wprowadzone pojęcia (i relacje) i zaproponowane aksjomaty są dobrze uzasadnione. Co więcej, Autor pokazuje w kolejnych rozdziałach możliwość praktycznego wykorzystania ontologii Science i ontologii praw naukowych.

4. Interesująco wygląda rozdział 4, w którym Doktorant przedstawia m. in. w jaki sposób odpowiedni algorytm wykorzystujący opracowane ontologie praw pozwala wyprowadzać nowe twierdzenie naukowe na podstawie innych (już zaakceptowanych). Przykład ze str. 102/103 nie jest może wyjątkowo odkrywczy dla umysłu ludzkiego, ale pokazuje, że maszyna istotnie może produkować prawa (dotyczące określonej dziedziny), których jeszcze nie było. Autor nadmienia również, że uzyskiwanych w ten sposób twierdzeń czy wykrywanych niespójności w bazach nie należy traktować jako definitywnych, ale takich, które mogą być źródłem nowych analiz.

5. Ponieważ p. Trójczak publikował artykuły wieloautorskie (m. in. z prof. Trypuzem i prof. Kulickim) stąd może być zastanawiające, ile w tych artykułach zasługi Doktoranta a ile pozostałych autorów. Tego nie potrafię jednoznacznie rozstrzygnąć. Niemniej fakt, że pp. prof. Kulicki i Trypuz zaprosili Doktoranta do współautorstwa musi świadczyć o tym, że miał on twórczy wkład w prowadzonych badaniach, a to – obok wspomnianych wyżej autorskich dokonań w rozprawie – dodatkowo przesądza, że jest dojrzałym badaczem. Miałem zresztą przyjemność słuchać p. Trójczaka na prezentowanym przez niego wykładzie konferencyjnym w Kazimierzu Dolnym (kilka lat temu), gdzie wyrobiłem sobie opinię o niemałych kompetencjach Autora, zarówno gdy chodzi o sam wykład jak i Jego dyskusje z uczestnikami i udzielane odpowiedzi.

6. Praca jest zredagowana b. porządnie. Autor posługuje się zrozumiałym językiem, nie popełnia błędów semantycznych i nawet trudne kwestie przedstawia w sposób czytelny, uporządkowany i metodycznie poprawny. Nie ma niepotrzebnych dygresji, myśl jest przedstawiona liniowo. Zauważyłem jedynie kilka (może kilkanaście) literówek, braku słowa czy niezręczności. M. in. na stronie 89 w linii: prawa korelacyjne – powinno być: *is first component in (y, x)*, a nie: *is first component in (x, y)*. Inne tego typu braki pomijam. Duże braki zawiera indeks rzeczowy; np. termin SPARQL występuje na wielu stronach a nie tylko na 18 i 106 – 108. Podobnie jest z innymi terminami. Praca zawiera Dodatek A, ale B nie ma, więc wystarczyłoby napisać: Dodatek. Na stronie 37, gdzie omawia się OWL, Autor używa terminu *Encja*, który nie jest terminem polskim. Skoro pozostałe pojęcia standardu OWL to aksjomat i wyrażenie, to może warto znaleźć lepszy odpowiednik dla *entity* (może obiekt, jednostka, indywiduum, albo byt podstawowy?) Oczywiście nie umniejsza to wartości merytorycznej pracy.

Reasumując, stwierdzam, że Pan Rafał Trójczak przedłożył rozprawę doktorską, która spełnia wymogi ustawowe określone dla prac doktorskich i wnioskuję o dopuszczenie Go do dalszych etapów przewodu doktorskiego.

Janusz Kaczmarek