

О.П. Мінцер, М.А. Попова, В.В. Приходнюк, О.Є. Стрижак

ОНТОЛОГІЯ В СИСТЕМНІЙ БІОМЕДИЦИНІ



Київ
«Інтерсервіс»
2021

УДК 57+61]:001.4](02) 0-59

*Рекомендовано до друку вченою радою
Національного центру «Мала академія наук України»
(протокол № 5 від 19 травня 2021)*

Рецензенти:

Зозуля Іван Савович - др мед. наук, професор, професор кафедри медицини невідкладних станів Національного університету охорони здоров'я України імені П. Л. Шупика;

Глоба Лариса Сергіївна - др техн. наук, професор, завідувач кафедри інформаційно-телекомунікаційних мереж Інституту телекомунікаційних систем НТУУ «Київський політехнічний інститут імені Ігоря Сікорського»

Мінцер **О. П.**, Попова М. А., Приходнюк **В. В.**, Стрижак **О. С.**

М62 Онтологія в системній біомедицині: Монографія / О. П. Мінцер, М. А. Попова, В. В. Приходнюк, О. С. Стрижак. - К. : ТОВ «КАЛЕНДАР ТМ», 2021. - 300 с.

ISBN 978-966-999-150-8

В монографії розглядаються теоретичні та практичні аспекти аналітичної обробки і представлення значних обсягів біомедичної мережевої інформації у вигляді онтології біомедицини. Описана модель багаторівневої онтології, здатна забезпечити спільний словник в межах біомедичної спільноти, скорочення помилок через відсутність уніфікованої термінології, інтеграцію та семантичну інтероперабельність інформаційних ресурсів і систем, обмін та багаторазове використання інформації. Описаний вітчизняний та світовий досвід застосування онтології в практичній медицині, медичній освіті, наукових дослідженнях. Розглянуті процеси створення онтології знань із системної біомедицини. Наведений практичний досвід використання онтології «Системна біомедицина» в середовищі когнітивної ІТ-платформи «ПОЛПЕДР» для забезпечення управління біомедичними знаннями в процесах дослідницької діяльності та дистанційної й інтегрованої освіти медичних кадрів.

Жодна частина цієї публікації не може бути відтворена, збережена в пошуковій системі або передана в будь-якій формі і будь-якими засобами, електронними, механічними, шляхом відеозапису або іншим чином, без попереднього дозволу авторів.

ISBN 978-966-999-150-8

О. П. Мінцер, М. А. Попова, В. В. Приходнюк, О. С. Стрижак, 2021.

ЗМІСТ

Перелік скорочень	5
Передмова	8
Частина 1. Сучасні технології представлення біомедичної мережевої інформації.....	35
Розділ 1.1. Технології представлення мережевої інформації, рекомендовані I^3C	36
1.1.1. ХМБ - синтаксична основа Семантичної Павутини	38
1.1.2. Використання підходу KOP до організації даних у вигляді триплетів.....	39
1.1.3. Мова представлення онтології O^B	41
Розділ 1.2. Класифікація онтологій. Онтологічні моделі знань різного рівня	45
1.2.1. Онтології верхнього рівня.....	56
1.2.2. Онтології предметної області (домену)	65
1.2.3. Багаторівнева онтологія - засіб багаторівневої класифікації знань.....	67
Розділ 1.3. Систематизація штеніу. Засоби навігації для користувачів..	72
1.3.1. Онтологія як засіб структуризації медичних знань	81
1.3.2. Оптимальність структури та змісту онтології.....	85
Розділ 1.4. Сучасні уявлення про метрику онтологій	92
Частина 2. Практична реалізація онтології біомедицини. Прикладні задачі (Ладичук О. К., Мінцер О. П., Попова М. А., Приходнюк В. В., Стрижак О. Є.).....	102
Розділ 2.1. Опис створеної онтології	102
2.1.1. Інтеграції та індексації інформації. Інтеграція різнорідної інформації.....	102
2.1.2. Використання стандартів	107
2.1.3. Інтероперабельність онтології	116
2.1.4. Логіка і практика функціонування моста між даними системної біології та системної медицини.....	123
2.1.5. Реалізація багаторівневості онтології	128
2.1.6. Кількісні характеристики створеної онтології.....	160
2.1.7. Використання біомедичної онтології як корпоративної пам'яті. Збір і систематична організація інформації з різних джерел. Забезпечення потрібної інформації за запитом (пасивна форма) і за необхідності (активна форма).....	161
Розділ 2.2. Застосування онтології в практичній медицині, медичній освіті, наукових дослідженнях	171

2.2.1.	Особливості застосування онтологій у медицині.....	173
2.2.2.	Застосування онтологій у медичній освіті	183
2.2.3.	Управління знаннями та компетенціями на основі онтологій	191
2.2.4.	Метрика онтології «Системна біомедицина»	195
2.2.5.	Приклади завдань використання створеної онтології.....	205
Розділ 2.3.	Використання онтології в пошукових операціях.....	211
2.3.1.	8МАКТ технології	211
2.3.2.	Отримання не заданих явно знань шляхом використання додаткових пошукових можливостей.....	214
2.3.3.	Додаткові пошукові можливості	219
2.3.4.	Пошук інформації в інтегрованих розподілених мережевих джерелах (на прикладі даних вебресурсу відомих та прогнозованих взаємодій білків 8ТКШО).....	222
Розділ 2.4.	Особливості онтології із системної біомедицини	231
Частина 3. Опис роботи з онтологією «Системна біомедицина» в середовищі КІТ «ПОЛЕДР» (Ладичук О. К., Попова М. А., Приходнюк в. В.).....		234
Розділ 3.1.	Доступ та керування онтологією з системної біомедицини	234
Розділ 3.2.	Функціональні можливості та створення онтологічних пошукових призм.....	252
3.2.1.	Використання додаткових пошукових запитів.....	255
3.2.2.	Створення онтологій пошукового запиту	257
Розділ 3.3.	Керування індексацією.....	263
Література	270

Перелік скорочень

ЛІОУБЕ8 - АБіліу агіз ОізаБіліу Опіоіо^у гог Опііпе Беатіп§ акі 8еуісе8
АРІ - Арріісаііоп Про§гаттіп§ Іпіеггасе
Д8СП - Атегісап §іапсІагсІ Сосіе Рог Іпіоітпаііоп ІпіегсЬап^е
АГС - Апаіотісаі ТЬегареііс СЬешісаІ сіаззіісаііоп
МРО - Вазіс Роітпаі ОпІ:оІо§у
ВКО - Віотесіісаі Яезоігсе Опіоіо^у
(ІЗА - СІпсіаІ Оосітесі АгсЬіесШге
(08M0 - Соттоп 8етапІ:іс Мосіеі
СР Г - Сіггепі Проесііггаї Тегшіпоіо^у
ІЗС - ОиБІп Соге
ІЗІСОМ - Оі§іІаІ Іта§іп§ апсі СОТтіпсіаІіопз іп Месіісіпе 1310 -
Ош^Іпсісаііоп ОаІаБазе [ЗО - Оізеазе Опіоіо^у
ІЗОБСЕ - Оезсгірііуе ОпІоІо§у Рог Еіп§іІ8Іс апсі Со§міІІуе
Еп§іпеегін§ ІЗЯ - Оосітесі Яеігіеуаі
[38М-[У - Оіа§по8Іс апсі §іаІІ8іісаІ МапіаІ оРМенІаІ Оізогсіегз,
РоігіЬ
ССШІОП
[ЗТО - Оосітесі Туре Оейпіііоп
І ТІК8 - ЕІесІгопіс НеаІЬ Яесогсіз
сМЕЯСЕ - ЕІесІгопіс Месіісаі Яесогсіз апсі Сепотісз
Б М А - Роіпсіаііопаі Мосіеі оР Апаіоту
РОАР - Ргіепсі оР а Ргіепсі
РЯАРО - Ріпсііп§, ЯезеагсЬ Адтіпізігаііоп апсі РојесІ8 ОпІоІо§у
СІРО - Оепеггаї Роггаї Опіоіо^у ОО - Оепе ОпІоІо§у
НЕЯО - Ні§Ьег Есіісаііоп ЯеРегепсе Опіоіо^у
Н[РАА - НеаІЬ Іпзігапсе РогаІЬШіу апсі АссоіпІаЬШіу Асі
ИЕ7 - НеаІЬ ЕЕУЕІ 8еуеп Іпіегпаііопаі
НЕА - Ні§Ь ЕЕУЕІ АгсЬіесІіге
НМО - НіегагсЬіаІ Ме88а§е Оезсгіріог
НРО - Нітап РЬепоІуре ОпІоІо§у
р[ТМЕ - НурегТехІ Магкір Еап§іа§е
ІСО - Іпіегпаііопаі Сіаззіісаііоп оРОізеазе
Ю- ІсіепіШсаІіоп
[ЕС - Іпіегпаііопаі ЕІесІгоІесЬпісаІ Соттізіоп
[НТ800 - Іпіегпаііопаі НеаІЬ Тегтіпоіо^у §іапсІагсІ8 ОеуеіортепІ
Ог^апігаііоп
[08 - іРЬопе ОрегаІіп§ 8узІет

IK - Інформаційні Яєієієуаі
180 - Інформаційні Організації тог Віапіяєсііаііоп
Кпоу\^Е - Кпо\^V1e(1şe \^ікі Епуігоптепі
ЕОШС - Ео\^саі ОЪзегуаііоп Ісіепіієгз Натез апсі Сосіез
МесіОКА - Месісаі Оісііопару &г Яе\^іаіаіогу Асйуіііез
Ме8Н - Месісаі 8иЪ]есІ Неайіп8
МО\^Е - Микітесіа \^еЪ Опіоіо\^у Еап8иаşe
М80 - Микіріе 8СІСГo8I8 опіоіо\^у
NAN^A - НогіЪ Атегісап Нигзіп8 Оіа\^позіз Аззосіаііоп N060 - Найопаі
Сепіег тог Віотесісаі Оп1о1о8у N01 - Na\^іопа1 Оапсер Іпзікііе ОВО -
Ореп Віоіо\^ісаі апсі Віотедісаі Опіоіо\^іез ОЕ AP - Опііпе апауіісаі
прозезіп^
ОрепОА^EN - Оепегатізей АгсБіесШге тог Еап8иаşeз, Епсусіораесііаз,
апсі Нотепс1а\^и\^е8 іп Медісіне
0КЕ)0 - Оп1о1о8у ої Яаге Оізеазез ОгрЪапел
0\^Е - \^еЪ Оп1о1о8у Еап\^иа\^е
0\^Е ОЕ - 0\^Е Оезсгірйоп Ео8іс
РАРІ - Рибліс апсі Ргіуаіе Іпгогтаііоптог Ееагпегз
РЕІВ - Рагі ЕіЪгагу
Р8Е - Проесезз 8ресШсаі іоп Еап\^иа\^е
РА - \^иа1і1у-Аззі\^апсе
ЯОР - Яезоигсе Оезсгірйоп Ргате\уогк
ЯІМ - Яегегепсе Іпгогтаііоп Мосіеі
8ЕО - 8еагсЪ Еп8іне Оріітіхаііоп
8К08 - 8ітр1е Кполуіед\^е бг8ап1заііоп 8узлет
8МАЯТ - 8а1 іоп'з Ма\^іс Аиіотаііс Яешевег оРТехІ
8НОМЕ^-ОТ - 8узлета12ед Нотепс1а\^и\^е оРМесісііне - Оііпсаі Тегтз
80А - 8егуісе-Огіеп;ед АгсБкесШге
80АР - 8ітр1е ОЪ]есІ Ассезз Проіосіі
8ТЕР - 8іапйагс1 Рог ібе ЕхсЪап8е ог Проіосііс
8ТЯШО - 8еагсЪ Тооі Рог ібе Яеієієуаі оР Іпегасйп\^ Оепез/Проіепз
81]МО - 8и88ез1ес1 Уррег Мер\^есі Опіоіо\^у 8\^Я0 - Зетатііс \^еЪРог
ЯезеагсЪ Оопшшпкієз ТЯ - Техі Яеієієуаі
иООІ - Іпіуегзаі Оезсгірііоп, Оізсоуегу та ІпІе8гаііоп ІМЕ8 - ОпШесІ
Месісаі Еап8иаşe 8узлет ЕПЧІОООЕ - Іпіуегзаі ОЪагагс1ег Епсоеліп8
ІЯІ - ІпіРогт Яезоигсе ІепіШег ІЯЕ - ОпіРогт Яезоигсе Еосаіог

\\^3С - А\^\\^\\^^-консорціум
ОБ - \^еБ 8егuісе Оезсгірііоп
Бап\$иа^е ^V^V^V - ^Vo^Ш \¥i(1е АЅеБ
ХМЕ - еХіепзіБІе Магкир Еап^иа^е БЗ -
база знань
ВООЗ - Всесвітня організація охорони здоров'я
ЕМК - електронна медична карта
ЕОМ - електронно-обчислювальна машина
ЕС - експертна система
ЗМІ - Засоби масової інформації
ІКТ - інформаційно-комунікаційні технології
ІС - інформаційна система
КІТ - когнітивна інформаційна технологія
ККО - канонічна концептуальна онтологія
ЛО - лінгвістична онтологія
Л-середовище - лексикографічне середовище
Л-система - лексикографічна система
МКХ - Міжнародна класифікація хвороб
ММЕ - мікро- і макроелементи
11 ККО - неканонічна концептуальна онтологія
ОПР - особа, що приймає рішення
ПМКЗ - персональні медичні карти здоров'я
ПНЗ - портал наукових знань
СЗЗ - система, заснована на знаннях
СППР - система підтримки прийняття рішень
УЗ - управління знаннями

столітті завдяки працям Христіана фон Вольфа, який відокремив семантику термінів «онтологія» і «метафізика».

Подальший розвиток парадигми онтологій характеризувався співфункціонуванням міждисциплінарних підходів. В дослідженні «Онтологічні основи структурних концептуальних моделей»[^] узагальнено сім тлумачень терміну «онтологія», доступних в літературних джерелах: 1) філософська дисципліна; 2) неформальна концептуальна система; 3) офіційна семантична структура; 4) специфікація концептуалізацій; 5) подання концептуальної системи за допомогою логічної теорії; 6) теорія логічного відображення словникового запасу; 7) специфікація (мета-рівень).

З того часу, як термін «онтологія» вперше було згадано в роботі, присвяченій основам моделювання даних[^], онтології почали застосовуватися в різних галузях інформатики: базах даних та інформаційних системах, програмній інженерії (зокрема, інженерії предметної області) та штучному інтелекті.

Перше помітне зростання інтересу до онтологій в середині 1990-х було спричинене необхідністю створення концептуального представлення знань предметної області в співтоваристві з обміну і повторного використання знань в галузі штучного інтелекту. Велике значення в цьому напрямку мала робота Вільяма Дж.Кленсі^{*}, до оприлюднення якої в традиціях штучного інтелекту знання в інформаційних системах визначалися строго функціональним чином з метою включення до бази знань кроків, які зазвичай

^{*} Сміггарсії О. Ontological Foundations for Knowledge Representation Models. STIT '90 тезис Series, No. 05-74. Тейтатіса Іншішіл Рипсіатеніаі ЯезеарсЬ Series, No. 015 (TI/PK5/015). ЕлзсБейе, ТКе Не[^]Бе[^]1ап(із, 2005.

^{*} Меайу О. Н. АпоіБег Бок аі Оаіа. Процесіін§5 оі Оіе Раїї Іоїні Сотрїег ('опіегепсе, МоуетБег 14-16, АнаБегіт, Саїїогніа (АРІР8 СогГегепсе Процесі- іпцз, Уоїте 31), Уа8Ып§Іоп, ОС: ТЬотрзоп Вокз, Ыогкіоп: Асасіешіс Презз, 525-534, 1967.

^{*} Сіансу \У. 3. ТБе Кпоу1е<1§е Бевеі Кеїніеггреїед: Мо(1е11п§ 8ocio-ТесЬпісal .Чузїетз. Іпїетайїопаї Зоїшаї оРІпІе11і§еп4 8узїетз, 8: 33-49,1993.

використовуються експертами в цій галузі для вирішення даної проблеми. Кленсі запропонував змінити таку точку зору, стверджуючи, що «основне завдання інженерії знань - моделювання систем в світі, а не копіювання того, як думають люди». Згідно його моделі отримання знань база знань не є сховищем витягнутих з розуму експерта знань (як в так званій концепції передачі), а відноситься до об'єктивної реальності, яка більше пов'язана з класичним поняттям істини, задуманим як відповідність реальному світу. Тобто, моделювання має встановлювати відповідність між базою знань і двома окремими підсистемами: поведінкою інтелектуальної системи (тобто досвідом вирішення проблем) і її середовищем (проблемною областю)^

Варто зауважити, що значення терміну «онтологія» в спільнотах інформаційних систем та інженерії предметної області різняться. Для перших онтологія, як система категорій, не залежить від мови: онтологія Аристотеля одна і та сама, незалежно від того, представлена вона англійською, грецькою мовою або в логіці першого порядку. Навпаки, в більшості інших областей інформатики (включаючи інженерію предметних областей і штучний інтелект) термін онтологія, як правило, використовується як конкретний інженерний артефакт, призначений для певної мети: онтологія має бути побудована в систематичному процесі, аналогічному до процесів традиційної програмної інженерії. Модель процесу онтологічної інженерії зазвичай включає такі дії, як ідентифікація мети і специфікація вимог, охоплення онтології, формалізація, повторне використання та інтеграція, оцінка та документування*.

* Сиагіно N. Роїтпаї ОпІоІо§у, Сопсеріаї Анаїузіз аркі Кпо\уле<1§е КергезепШіоп. Піераііопаї ^оитаІ оГНитап апї Сотриер 8(исІіе5,43(5/6), 625-640, 1995.

РаБо К. А., ОиІ22а<іі О., Оиагіе К. С. Ап оп(оІо§ісаІ арргоасЬ Іо сіотаїп еп§іпеегїн§. Іп ПроceeЛп§5 оР ІБе 14їБ іпіегпаїіопаї сопРегепсе оп 8ой\уаре еп§іпеегїн§ апсі кпо\уле<18е еп§іпеегїн§ (8ЕКЕ '02). Аззосіаїіоп Гог Сотрїлїп§ Масїінегу, Келу Уогк, КУ, В8А, 351-358, 2002.

За короткий час свого існування в інформатиці термін «онтологія» пройшов шлях від повного неприйняття до повсюдного, ме завжди виправданого використання. Ставлення до нього досі лишається неоднозначним. Дослідники, що займаються розробкою інформаційних систем і баз даних, стверджують, що онтології являють собою схеми баз даних. Інженери знань висловлюють думку про те, що поняття моделі предметної області і бази знань використовувалося ще задовго до розповсюдження терміну «онтологія». Залишаються відкритими питання доповнення та верифікації інформації, що зберігається в каркасній системі семантичної бази знань, зокрема щодо її уточнення і деталізації при визначенні понять та існуючих між ними зв'язків. Ключовим же питанням залишається генерація запитів, що строго залежать від задалегідь заданої структури збережених знань.

Найбільш відчутний практичний ефект спричинило застосування онтологій в системах, заснованих на знаннях (СЗЗ), з а в стратегії управління знаннями (УЗ), зокрема в експертних системах (ЕС), системах підтримки прийняття рішень (СППР) з а порталах наукових знань (ПНЗ).

Під *управлінням знаннями* розумітимемо сукупність технологічних практик, метою яких є структуризація корпоративних (або загальнолюдських) знань, спрямованих на вирішення певних задач. Внесок онтологій до УЗ в кінцевому підсумку є стратегією вдосконалення процесів корпоративного спілкування.

Щойно отримаємо позитивні результати від аналізу змісту ін- (юрмаційного ресурсу та виконаємо процедуру оцінювання знань, відповідна онтологія працюватиме як тип організаційної мови, з им самим посилюючи комунікацію та сприяючи ініціативам у с(!)срі управління ресурсами. При цьому оцінка змісту онтології є процесом, орієнтованим на користувача, що полягає у перевірці відповідності отриманих знань тим, що наявні в середовищі виконання процесу набуття знань.

Застосування онтологій в системах, заснованих на знаннях.

Зазвичай онтології застосовуються для проектування і розробки С'ЗЗ, проте можуть бути використані і в якості функціонального компонента системи для:

- формування і фіксації загальних знань, які поділяються спільною предметною областю;
- визначення комплексу аксіом, що обмежують визначення деяких елементів знань для кращого представлення їх передбачуваного значення (релевантність, транзитивність знань);
- концептуалізації предметної області, що дозволяє описувати семантику даних;
- забезпечення можливості використання знань в різних напрямках предметної області;
- опису функціоналу СЗЗ (типів вирішуваних завдань);
- розробки окремих компонентів системи в якості високорівневої специфікації;
- розробки сховищ даних як високорівневого інтерфейсу.
- спільного використання різнорідних даних і знань в рамках однієї системи;
- визначення зв'язків між різними типами семантичних знань;
- здатності визначати семантичну модель даних у поєднанні з відповідними знаннями предметної області;
- формулюванні деяких стратегій пошуку даних;
- процесу вирішення завдань, що становлять функціонал системи;
- кращого розуміння предметної області користувачами системи'.

Застосування онтологій в системах підтримки прийняття рішень та порталах наукових знань. Під системою підтримки прийняття рішень* * мається на увазі інтерактивна автоматизована інформаційно-аналітична система, яка допомагає особі, що приймає рі-

' Загорулько Ю.А., Загорулько Г.Б. Онтологии и их практическое применение в системах, основанных на знаниях. Всероссийская конференция с международным участием «Знания - Онтологии - Теории» (ЗОНТ-2011), Новосибирск, 3-5 октября 2011 г., С. 132-141.

* Ларичев О. И., Петровский А. В. Системы поддержки принятия решений. Современное состояние и перспективы их развития. Итоги науки и техники. Серия; Техническая кибернетика. Т.21. М.: ВИНТИ, 1987, С. 131-164.

ііснення (ОПР), використовувати дані і моделі для вирішення про- (І)ссійних слабкоформалізованих задач. СППР та ЕС' - це системи практично одного класу, часто СППР в своєму складі має кілька різних ЕС, проте є аспекти використання онтологій, специфічні ліиіе для СППР. Так, приймаючи до уваги слабку формалізова- мість розв'язуваних СППР завдань, дуже важливо мати детальний іісуперечливий опис проблемної області, в рамках якої СППР здійснює підтримку вирішення завдань ОПР. Онтологія є незамінним інструментом для створення такого опису. В більшості типів СІ ППР використовуються великі масиви різнорідних даних і знань. Завдяки тому, що онтологія дозволяє явно описувати семантику даних і знань, вона забезпечує базис для їх інтеграції та спільного використання при вирішенні завдань.

Портал наукових знань*® - це спеціалізований інтернет-портал, що забезпечує систематизацію знань та інформаційних ресурсів заданої області знань, їх інтеграцію в єдиний інформаційний простір і змістовний доступ до них. Онтологія використовується як ядро ін(1)ормаційної моделі порталу знань. Вводячи формальний опис системи понять, існуючих в області знань порталу, у вигляді класів об'єктів і відношень між ними, онтологія також задає структуру для представлення реальних об'єктів і зв'язків між ними. Відповідно до цього дані в контенті порталу представлені як множина різнотипних інформаційних об'єктів - екземплярів класів онтології, іюв'язаних між собою відношеннями, також заданими в онтології. Це створює базис для зручної навігації контентом порталу та змістовного пошуку в ньому.

При розробці СЗЗ можливе використання онтологій, раніше розроблених для області знань цих систем. Це дозволяє перевикористовувати вже перевірені на практиці знання та забезпечувати

" І Попов З.В Экспертные системы: Решение неформализованных задач и диалог с ЗВМ. М. : Наука. 1987. 288 с.

Чагорулько Ю.А., Боровикова О.И. Подход к построению порталов научных знаний. Автметрия. № 1, 2008, Т. 44, С. 100-110.