

## ПРОБЛЕМИ ОЦІНЮВАННЯ ЯКОСТІ ТЕКСТОВИХ МІТОК В МОДЕЛЯХ БІЗНЕС-ПРОЦЕСІВ

*В.Д. Шевелєв<sup>1</sup>, А.М. Копп<sup>2</sup>*

<sup>1</sup> *магістрант кафедри ICT, НТУ «ХПІ», Харків, Україна*

<sup>2</sup> *доцент кафедри ПІІТУ, Ph.D., НТУ «ХПІ», Харків, Україна*

[vadym.sheveliev@cs.khpi.edu.ua](mailto:vadym.sheveliev@cs.khpi.edu.ua)

Зараз багато компаній документують свої бізнес-процеси за допомоги концептуальних моделей. Ці моделі забезпечують основу для діяльності, пов'язаної з життєвим циклом управління бізнес-процесами, наприклад, аналізу процесу, перепроєктування процесу, оцінювання процесу. Багато ініціатив щодо моделювання цих процесів призвели до сотень моделей, створених проєктувальниками з різним досвідом. Однією з головних перешкод на шляху до більш ефективного використання цих моделей процесів є часто недостатнє забезпечення якості. Це спостереження закладає основу для визначення методів автоматичного аналізу, які можуть забезпечувати таку гарантію якості. В останні роки дослідження забезпечення якості моделей процесу та відповідних методів аналізу запропонували багато нових ідей. Метою цієї роботи є узагальнення деяких істотних внесків у цій галузі.

Дослідження концептуального моделювання часто розрізняють синтаксис, семантику та прагматику моделей бізнес-процесів з посиланням на семіотичну теорію. Ідея цього розрізнення полягає в тому, що базова ідея, кодифікована як концептуальна модель, спочатку має бути зрозумілою з точки зору її синтаксису читачем моделі, перш ніж семантику можна буде інтерпретувати. Тоді розуміння на семантичному рівні забезпечує основу для вжиття відповідних дій у прагматичний спосіб. Ця “семіотична драбина” (explanatory regress) має велике значення для моделювання бізнес-процесу як специфічної області концептуального моделювання та є важливим дослідницьким напрямом. Наслідком семіотичної точки зору на моделювання бізнес-процесів є те, що розуміння моделі бізнес-процесу читачем цієї моделі має розглядатися як центральна основа для розбору її якості. Оскільки відповідна прагматика, тобто дії, вжиті читачем моделі, визначає успішне просування по “семіотичній драбині”, дослідження має визначити глибоке розуміння того, як можна досягти якості на кожній “сходинці” цієї “драбини”. Емпірично вже було доведено, що жодним із трьох ступенів “семіотичної драбини” не можна нехтувати, і що всі кроки є однаково важливими для концептуального моделювання [1]. І хоча більшість досліджень з моделювання бізнес-процесів стосуються поглибленого аналізу синтаксису та семантики виконання моделей, вони нехтують текстовою семантикою та прагматичними аспектами. Через те, важливим напрямком для цього дослідження є доповнення синтаксичного аналізу розумінням семантики та прагматики моделі бізнес-процесу. Отже, пропонується зосередити увагу на текстових характеристиках моделі бізнес-процесу [2].

Емпіричне дослідження виявило, що моделі бізнес-процесів на практиці не завжди відповідають умовам іменування дій, таким як стиль “дієслово-об'єкт”. Існує три загальні класи стилів найменування цих дій. По-перше, стиль “дієслово-об'єкт” визначає мітку діяльності як дієслово, за яким слідує відповідний бізнес-об'єкт. По-друге, існують різні способи визначення діяльності за допомогою міток “дія-іменник”. Для такої мітки дія формулюється не як дієслово, а скоріше як герундій або істотне дієслово. Існує також третя категорія міток діяльності, у яких відсутні посилання на дію [3].

Після визначення цих категорій слід зазначити, що стиль маркування є фактором із характеристиками, які значно відрізняються від структурних показників. У той час, як ті показники можна виміряти за метричною шкалою, стилі маркування можна розрізняти лише номінально. Це означає, що в найпростішому випадку вхідну змінну можна визначити у двійковому вигляді, розрізняючи використання стилю “дієслово-об’єкт” від використання іншого стилю. З точки зору визначення параметрів якості, це полегшує завдання: у той час як метрики вимагають порогу, щоб розрізняти “хороші” та “погані” моделі, стилі маркування можна прямо порівнювати, щоб визначити, чи вони є кращими чи гіршими [3].

Незважаючи на те, що використання стилю маркування добре висвітлено в літературі, все ще існують різні проблеми, власне коли справа доходить термінології. З точки зору якості, терміни повинні мати чітке значення. Це означає, що при моделюванні бізнес-процесу слід уникати синонімів і омонімів. Ця проблема визнається в різних документах, однак належне рішення для автоматичного забезпечення якості відсутнє.

Загальна ідея рефакторингу була сформульована для програмного забезпечення і визначається як “зміни до програмної системи, які не впливають на зовнішню поведінку коду, але покращують його внутрішню структуру”. Для моделей бізнес-процесів під час рефакторингу часто враховується поняття еквівалентності трасування або одне з понять бісимуляції. В подальшому необхідно підсумувати роботу над рефакторингом текстових міток у моделях бізнес-процесів.

Експерименти та передові практики галузі свідчать про надання переваги стилю “дієслово-об’єкт” для текстових міток. Проблема в цьому підході полягає в тому, щоб досягти точного аналізу різних стилів маркування, щоб їх можна було трансформувати в стиль “дієслово-об’єкт”. Точний аналіз англійської мови є складним з двох причин. По-перше, багато іменників в англійській мові утворюються від дієслова таким чином, що іменник морфологічно еквівалентний дієслову. По-друге, текстові мітки дій у моделі бізнес-процесу зазвичай не охоплюють повної граматично правильної структури речення. Тому було виявлено труднощі з використанням стандартних засобів обробки природної мови. У підході, про який йдеться в одному з досліджень, використовується різного рода контекстна інформація для зіставлення кожної мітки з її правильним стилем маркування [4]. Після того, як стиль маркування відомий, такі інструменти, як WordNet, можна використовувати для пошуку дієслова, що відповідає дії, яка була сформульована у вигляді іменника. Було доведено, що цей підхід точно працює для трьох різних колекцій моделювання, що загалом містять понад 10 000 текстових міток [4].

Також одним з напрямків майбутнього дослідження є те, як ці методи рефакторингу можна визначити таким чином, щоб вони не залежали від багатого набору інструментів обробки природної мови, доступних для англійської. Альтернативою може бути пряма робота з анованими корпусами. Крім того, у зв’язку з термінологічною проблемою, визначеною вище, досі не зрозуміло, як проблеми синонімів та омонімів можуть бути автоматично виправлені.

#### **Список літератури:**

1. Lindland, O., Sindre, G., Sølberg, A. Understanding quality in conceptual modeling // IEEE Software. – 1994. – №11(2) – С. 42–49
2. Krogstie, J., Sindre, G., Jørgensen, H. Process Models Representing Knowledge for Action: Revised Quality Framework // European Journal of Inf. Syst. – 2006. – №15(1) – С. 91-102
3. Mendling, J., Reijers, H.A., Recker, J. Activity Labeling in Process Modeling: Empirical Insights and Recommendations // Information Systems. – 2010. – №35(4) – С. 467–482
4. Leopold, H., Smirnov, S., Mendling, J. On the refactoring of activity labels in business process models // Information Systems. – 2012. – №37(5) – С. 443–459