

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АВТОМАТИЗАЦІЇ АНАЛІЗУ ВІДПОВІДНОСТІ ТЕСТОВИХ ЗАВДАНЬ ЛЕКЦІЙНИМ МАТЕРІАЛАМ НАВЧАЛЬНИХ ДИСЦИПЛІН

У статті досліджено особливості розвитку дистанційної освіти як сучасного способу надання якісних освітніх послуг. Встановлено велике значення оцінки відповідності тестових завдань лекційним матеріалам. Запропоновано інформаційну технологію автоматизації аналізу відповідності тестових завдань лекційним матеріалам навчальних дисциплін. Розглянуто відповідний тестовий додаток та проведено аналіз результатів його роботи. Визначено області й особливості застосування запропонованої інформаційної технології автоматизації аналізу відповідності тестових завдань лекційним матеріалам.

This article studies the development of distance education as a modern way to provide high-quality educational services. It establishes an importance of assessing an accordance between tests items and lectures content, and suggests an information technology for automatizing the analysis system that check an accordance of the lectures contents with the education standards. The article considers the corresponding prototype program and analyzes the results of this program. It defines the uses of suggested information technology for automatic analyzing analysis system that check an accordance of the lectures contents with the education standards.

Дистанційна освіта на сучасному етапі є найбільш розповсюдженим засобом застосування інформаційних технологій для надання освітніх послуг та контролю отриманих знань [1]. Існує значна кількість наукових досліджень, в яких відображені різні підходи до навчання, контролю знань та створенню автоматизованих систем навчання і контролю знань [2, 3]. Одним із найбільш складних завдань визначається об'єктивізація процесу оцінювання знань та умінь [4].

Хоча існує багато можливих засобів надання знань (відео- та аудіо матеріали, лекційний матеріал, онлайн-семінари та ін.), проте загальноприйнятим є підхід [5] по застосуванню в якості інструменту навчання текстових навчальних матеріалів визначеної структури.

Одним із найбільш ефективних методів контролю є тестування, що дає об'єктивну, зіставну і кількісну оцінку якості підготовки [6]. До головних переваг тестового контролю відносять: високу змістовну, прогностичну і критеріальну валідність; високу надійність; можливість стандартизації; диференціацію шкали тестових балів; єдиність вимог до всіх учнів; всебічність та об'єктивність; індивідуальний характер контролю і можливість регулярного систематичного його проведення [7].

Проте, виникає проблема правильного складання тестів, які будуть охоплювати рівномірно весь навчальний матеріал. Лише після складання такого тесту є можливість правильно оцінити рівень знань учнів.

Метою роботи є розробка інформаційної технології автоматизації аналізу відповідності розроблених тестових завдань навчальним матеріалам дисциплін.

Хоча результатом аналізу відповідності тестових завдань навчальним матеріалам може виступати цілий ряд різноманітних параметрів, можна визначити два ключових напрямки такого аналізу:

- визначення повноти покриття тестами ключових положень лекційного матеріалу;
- визначення рівномірності покриття тестами структури лекційного матеріалу.

Припускається, що виявлення спільних рис та характеристик у цих двох напрямках дозволить коректно формалізувати досліджуваний процес та знайти комплексний розв'язок поставленої задачі.

Повнота покриття тестами ключових положень лекційного матеріалу може визначатися через аналіз використання у тестових завданнях ключових термінів із лекційних матеріалів. Рівномірність покриття тестами структури лекційного матеріалу визначається через аналіз прив'язки тестових завдань до ієрархії структурних елементів лекційного матеріалу. Така прив'язка може бути забезпечена аналізом контенту тестових завдань на предмет наявності у них ключових термінів із лекційних матеріалів, за умови прив'язки ключових термінів до ієрархії структурних елементів лекційного матеріалу.

Якщо розглядати ієрархію змістовних блоків лекційного матеріалу як рівні вертикальної онтології відповідної навчальної дисципліни, то, з семантичної точки зору, ключові терміни є найнижчим рівнем онтології лекційних матеріалів навчальної дисципліни. Така формалізація досліджуваного процесу надає інструмент вирішення задачі аналізу відповідності тестових завдань лекційним матеріалам навчальних дисциплін.

Структура лекційних матеріалів як електронних документів регламентується мовами розмітки документів (наприклад, WordprocessingML для XML) [8] й реалізується через систему заголовків. Зважаючи на вимоги стандартів освіти до структури навчальних дисциплін (Назва дисципліни, Модуль, Тема) [9, 10], можна зробити висновок про відповідність системи заголовків лекційних матеріалів як електронних

документів верхнім рівням онтології навчальної дисципліни (наприклад, назви дисциплін відповідатимуть «Заголовку 1», назви модулів – «Заголовку 2», назви тем – «Заголовку 3, тощо).

За такого підходу ключові терміни (поняття, визначення) є найнижчим рівнем онтології лекційних матеріалів навчальної дисципліни. Причому особливістю цього рівня онтології є наявність властивості рівня прив'язки: адже поряд із термінами, що використовуються суто в рамках певних підтем, інші терміни можуть бути актуальними в масштабах тем, модулів, і навіть всієї навчальної дисципліни. Хоча структура лекційного матеріалу визначає місцезнаходження відповідних термінів, але вона не включає їх безпосередньо. Для аналізу відповідних змістовних блоків лекційного матеріалу є доцільним використання алгоритмів аналізу інформаційного вмісту текстових матеріалів [11]. За результатами попередніх досліджень [12] було встановлено доцільним використання дисперсійної оцінки [13] для аналізу інформаційного вмісту навчальних матеріалів. Оскільки терміни бувають не тільки словами, а й словосполученнями, для пошуку останніх довів [12] свою ефективність метод пошуку безперервних послідовностей ключових слів із відбором за порогом частоти вживання.

Таблиця 1 – Приклад таблиці лем та типів набору тестових питань

	Тест 1	Тест 2	Тест 3	Тест 4	...	Тест N
Термін 1	1	1				
Термін 2		1				
Термін 3			1	1		
Термін 4				1		
...						
Термін N						1
Тип тесту	2	2	1	2		1

На основі побудованої описаним способом повної онтології лекційного матеріалу, можна проводити аналіз повноти та рівномірності покриття тестами обсягу лекційного матеріалу. Аналіз проводиться шляхом пошуку співпадінь елементів таблиці лем та типів набору тестових питань із нижнім рівнем повної онтології лекційного матеріалу. Таблиця лем та типів набору тестових питань подається двовимірною системою координат, у якій одна з осей включає номери тестових завдань у наборі, а друга – список ключових термінів лекційного матеріалу у лематизованому вигляді.

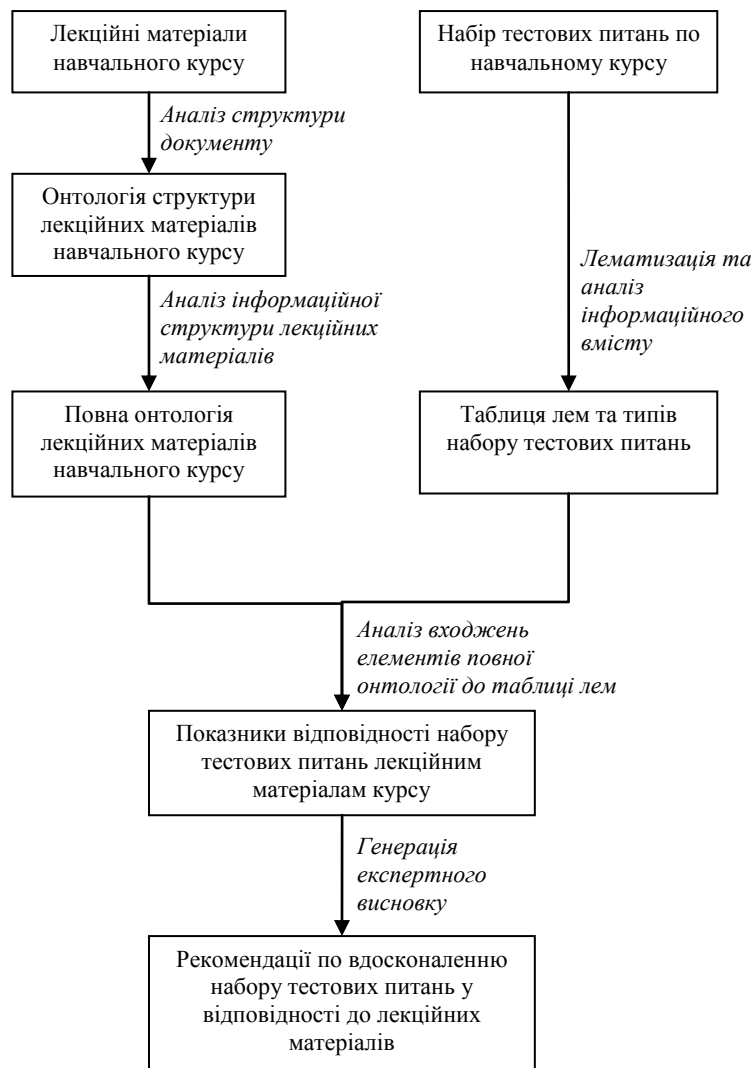


Рисунок 1 – Узагальнена схема роботи експертної системи аналізу відповідності тестових завдань лекційним матеріалам навчальних дисциплін

На перерині значень осей виставляється позитивне значення, якщо даний термін присутній у даному тестовому завданні (табл. 1). Такий вигляд таблиці зручний як для візуального спостереження, так і для програмної обробки, і є формою вхідних даних в аналітичну систему для генерації експертного висновку (рис. 1).

Повнота покриття тестами ключових положень лекційного матеріалу може визначатися через аналіз включення у тестові завдання ключових термінів із лекційних матеріалів. Очевидно, що смислова відповідність тестових завдань лекційним матеріалам має встановлюватись через їх прив'язку до певних елементів лекційних матеріалів. Це доцільно реалізовувати не шляхом прямого пошуку спільних фрагментів тексту, а знаходженням смислової відповідності. У текстових комунікаціях ключем до смислу є ключові слова та словосполучення (терміни), в той час як решта контенту призначена для опису смислових відтінків та зв'язування слів. Розглядаючи терміни як смислові осередки контенту лекційних та тестових матеріалів, очевидно, що інструментом для їх смислового співставлення є пошук відповідності і зв'язування термінів, а характеристикою відповідності тестових завдань перелікам термінів є згадування. Отже, облік згадувань, пошук відповідностей та зв'язування тестових завдань із блоками лекційних матеріалів є початковими задачами в аналізі відповідності тестів лекційним матеріалам.

Одиницею обліку термінів є згадування терміна, а не власне тестове завдання. Якщо деякий термін має кілька згадувань в одному тестовому завданні, то вважається, що цей термін має в завданні лише одне згадування. В одному тестовому завданні можуть міститись одночасно згадування кількох різних термінів.

Як зазначалося вище, однією з ключових вимог до тестів є повне та рівномірне охоплення матеріалу, по якому проводиться тестування. У відповідності з цим, змістовно якісними (тими, що максимально відповідають вимогам повноти та рівномірності охоплення матеріалу) є тестові матеріали, у яких:

- всі основні терміни лекцій задіяні в тестуванні;
- кількість згадувань термінів у тестах є максимально однаковою (концентрація уваги до одних термінів автоматично веде до зменшення ролі й уваги до інших термінів);

- обсяг тестового навантаження рівномірно розподілений серед елементів простої онтології лекційних матеріалів (максимально однакова кількість завдань для кожного з елементів одного рівня);
- кількість завдань, у яких не використано жодного з термінів лекційних матеріалів (і які відповідно не мають зв'язку з лекційними матеріалами), є мінімальною;
- розподіл тестових завдань відносно вузлів онтології лекційних матеріалів за іншими критеріями (наприклад, типом чи рівнем складності тестових завдань) є максимально рівномірним.

Інші вимоги, хоча й є важливими для оцінки тестів, є такими, що виходять за рамки завдання визначення відповідності тестів лекційним матеріалам. Наприклад, кількість тестових завдань (можна визначити мінімальну кількість, але не максимальну, оскільки збільшення кількості тестових завдань екстенсивно підвищує якість тесту). Адже метою дослідження не загальна оцінка якості тестів, а оцінка відповідності тестів відповідним лекційним матеріалам навчальних дисциплін.

Так як забезпечення ідентифікації виконання усіх встановлених вимог вимагає визначення певної кількості параметрів, є доцільним формування єдиних вимог до цих параметрів. Вони ідентичні вимогам до параметрів оцінки відповідності лекційних матеріалів навчальних дисциплін вимогам, і є наступними:

- діапазон допустимих значень складає $[0;1]$;
- збільшення значення параметра виражає позитивну тенденцію щодо відповідності лекційних матеріалів вимогам;
- внаслідок фрактальної згортки простої онтології лекційних матеріалів, кожен параметр певного рівня онтології має враховувати також вплив аналогічних параметрів наслідуваного нижчого рівня.

Таким чином, параметри відповідності тестів лекційним матеріалам навчальних дисциплін визначено наступними моделями.

Модель номінального обсягу використання термінів формує параметр b_1 й виражає покриття переліку термінів – елементів нижнього рівня повної онтології лекційних матеріалів (незалежно від кількості згадувань кожного з термінів):

$$b_1 = \frac{k_1 b_1^0 + k_0 \bar{b}_1^{-1}}{2}, \quad (1)$$

де b_1^0 – параметр номінального обсягу використання термінів визначеного рівня онтології;

\bar{b}_1^{-1} – середнє значення параметрів номінального обсягу використання термінів нижчого рівня онтології;

k_1 та k_2 – значення відповідних вагових коефіцієнтів.

Параметр b_1^0 визначається:

$$b_1^0 = \frac{T_\Phi}{T_\Pi}, \quad (2)$$

де T_Π – кількість ключових термінів у переліку елементів визначеного рівня повної онтології лекційних матеріалів;

T_Φ – кількість слів (словосполучень) у текстах тестових завдань, що відповідні елементам переліку T_Π . Випадок $T_\Phi = T_\Pi$ означає, що всі терміни переліку мають хоча б одне згадування в тестових завданнях.

Модель рівномірності уваги до термінів формує параметр b_2 й виражає рівень рівномірності згадувань термінів у тестах:

$$b_2 = \frac{k_1 b_2^0 + k_0 \bar{b}_2^{-1}}{2}, \quad (3)$$

де b_2^0 – параметр рівномірності уваги до термінів визначеного рівня;

\bar{b}_2^{-1} – середнє значення параметрів рівномірності уваги до термінів нижчого рівня;

k_1 та k_2 – значення відповідних вагових коефіцієнтів.

Параметр b_2^0 визначається:

$$b_2^0 = \frac{\sum_{j=1}^m b_{2j}^0}{m}, \quad (4)$$

$$b_{2j}^0 = \frac{1}{|\bar{U} - U_j| + 1}, \quad (5)$$

$$\bar{U} = \frac{\sum_{i=1}^n U_i}{n}, \quad (6)$$

де \bar{U} – середня кількість згадувань термінів у всьому наборі тестових завдань;
 U_i – кількість згадувань i -го терміна у всьому наборі тестових завдань;
 n – кількість термінів у всьому наборі тестових завдань;
 m – кількість термінів на заданому рівні простої онтології лекційних матеріалів.

Модель достатності уваги до термінів формує параметр b_3 й виражає кількість термінів, кількість згадувань яких вища або рівна за вказану користувачем:

$$b_3^0 = \frac{T_D}{T_{\Pi}}, \quad (7)$$

де T_{Π} – кількість ключових термінів у переліку елементів визначеного рівня повної онтології лекційних матеріалів;

T_D – кількість слів (словосполучень) у текстах тестових завдань, що відповідні елементам переліку T_{Π} і мають кількість згадувань не менше ніж задано умовами.

Модель номінального обсягу актуальних завдань формує параметр b_4 й виражає обсяг завдань, у яких присутнє хоча б одне згадування хоча б одного із термінів лекційних матеріалів (визначається один параметр для всього досліджуваного обсягу матеріалів):

$$b_4 = \frac{L_{\Pi} - L_H}{L_{\Pi}}, \quad (8)$$

де L_{Π} – загальна кількість тестових завдань;

L_H – загальна кількість тестових завдань, у яких не використано жодного з термінів лекційних матеріалів (і які відповідно не мають зв'язку з лекційними матеріалами).

Загальна оцінка b відповідності тестів лекційним матеріалам навчальної дисципліни вираховується наступним чином:

$$b = \frac{\delta_1 b_1 + \delta_2 b_2 + \delta_3 b_3 + \delta_4 b_4}{4}, \quad (9)$$

де $\delta_1, \delta_2, \delta_3, \delta_4$ – вагові коефіцієнти, що виражають вплив окремих параметрів відповідності тестів лекційним матеріалам навчальної дисципліни на загальну оцінку.

Вагові коефіцієнти $k_1, k_2, \delta_1, \delta_2, \delta_3, \delta_4$ визначаються методом експертних оцінок та залежать від вимог до деталізації аналізу (k_1, k_2) та впливу окремих параметрів відповідності тестів лекційним матеріалам на загальну оцінку ($\delta_1, \delta_2, \delta_3, \delta_4$), лежать у межах $[0;1]$ й по замовчуванню можуть приймати значення «1».

Характер впливу змін наведених вагових коефіцієнтів є наступним.

Вагові коефіцієнти деталізації аналізу (k_1, k_2) дозволяють визначити, наскільки на оцінку критерію якості на обраному вузлі онтології впливають оцінки критерію якості на підлеглих вузлах онтології.

Ваговий коефіцієнт деталізації аналізу поточного рівня k_1 впливає на значення відповідного критерію якості, що обраховується для даного вузла. При зменшенні значення k_1 вплив критерію якості поточного рівня на загальний критерій якості буде зменшуватись, а при збільшенні значення вплив критерію якості поточного рівня на загальний критерій якості буде зростати. При $k_1=0$ вплив критерію якості поточного рівня на загальний критерій якості буде ігноруватись, а при $k_1=1$ вплив критерію якості поточного рівня на загальний критерій якості буде враховуватись без змін. Діапазон допустимих значень k_1 лежить у межах $[0;1]$.

Ваговий коефіцієнт деталізації аналізу нижнього рівня k_2 впливає на значення критерію якості на підлеглих вузлах онтології, що обраховується для даного вузла. При зменшенні значення k_2 вплив критерію якості на підлеглих вузлах онтології на загальний критерій якості буде зменшуватись, а при збільшенні значення вплив критерію якості на підлеглих вузлах онтології на загальний критерій якості буде зростати. При $k_2=0$ вплив критерію якості на підлеглих вузлах онтології на загальний критерій якості буде ігноруватись, а при $k_2=1$ вплив критерію якості на підлеглих вузлах онтології на загальний критерій якості буде враховуватись без змін. Діапазон допустимих значень k_2 лежить у межах $[0;1]$.

Вагові коефіцієнти впливу окремих параметрів відповідності тестів лекційним матеріалам на загальну оцінку ($\delta_1, \delta_2, \delta_3, \delta_4$) дозволяють визначити, наскільки на загальну оцінку відповідності тестів лекційним матеріалам впливають значення окремих критеріїв відповідності тестів лекційним матеріалам.

Ваговий коефіцієнт номінального обсягу використання термінів δ_1 впливає на значення загальної оцінки відповідності тестів лекційним матеріалам. При зменшенні значення δ_1 вплив оцінки номінального обсягу використання термінів на значення загальної оцінки відповідності тестів лекційним матеріалам буде

зменшуватись, а при збільшенні значення δ_l вплив оцінки номінального обсягу використання термінів на значення загальної оцінки відповідності тестів лекційним матеріалам буде зростати. При $\delta_l=0$ вплив оцінки номінального обсягу використання термінів на значення загальної оцінки відповідності тестів лекційним матеріалам буде ігноруватись, а при $\delta_l=1$ вплив оцінки номінального обсягу використання термінів на значення загальної оцінки відповідності тестів лекційним матеріалам буде враховуватись без змін. Діапазон допустимих значень δ_l лежить у межах $[0;1]$.

З метою дослідження ефективності наведених вище моделей аналізу інформаційного вмісту навчальних матеріалів було розроблено тестове програмне забезпечення (рис. 2), що реалізує запропоновану інформаційну технологію.

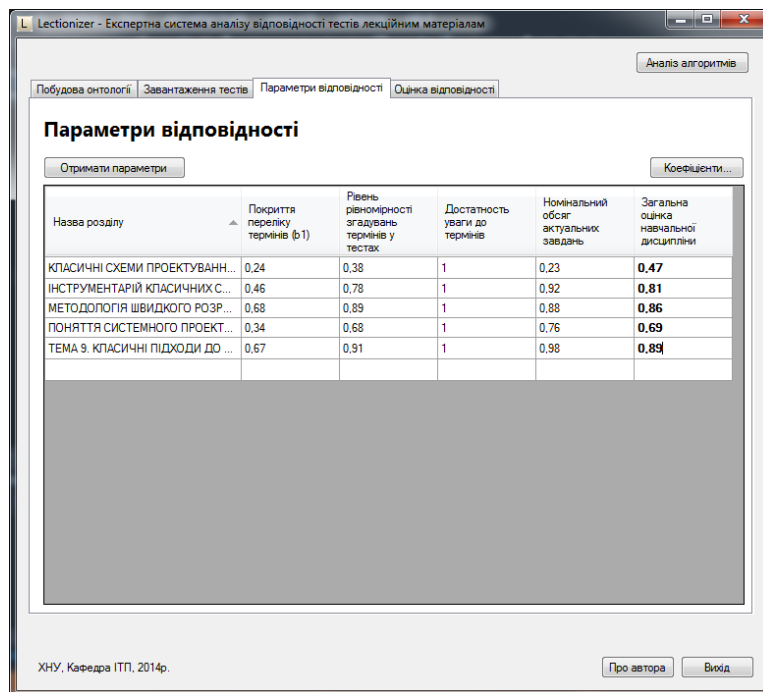


Рисунок 2 – Застосування тестового програмного забезпечення для автоматизованого визначення відповідності тестів навчальним дисциплінам

Наприклад, аналіз навчального курсу «Системний аналіз» [14] із модульної системи Хмельницького Національного університету виявив наступні характеристики відповідності даного навчального курсу вимогам:

- номінальний обсяг використання термінів $b_1=95\%$;
- рівномірність уваги до термінів $b_2=73\%$;
- достатність уваги до термінів $b_3=76\%$;
- номінальний обсяг актуальних завдань $b_4=89\%$;
- загальна оцінка $b=83\%$.

За результатами аналізу вибірки з 10 навчальних курсів, для яких проводився аналіз, виявив характеристики відповідності навчальних курсів вимогам, представлені в таблиці 2 (вагові коефіцієнти, що виражають вплив окремих параметрів відповідності лекційних матеріалів вимогами до навчальної дисципліни на загальну оцінку, взято рівними 1).

Таблиця 2 – Результати аналізу на відповідність вибірки навчальних курсів

Модель відповідності	Мінімальне значення для вибірки	Максимальне значення для вибірки	Середнє значення для вибірки
Номінальний обсяг використання термінів	87%	100%	93%
Рівномірність уваги до термінів	38%	91%	74%
Достатність уваги до термінів	62%	94%	81%
Номінальний обсяг актуальних завдань	71%	98%	86%
Загальна оцінка	67%	94%	84%

Показники відповідності виражають оцінку повноти та рівномірності покриття тестами ключових положень лекційного матеріалу в рамках всіх рівнів структури лекційного матеріалу, а також оцінку балансу використання тестових завдань різних типів і складності.

Загалом, застосування запропонованої інформаційної технології аналізу відповідності набору тестових питань лекційним матеріалам курсу у рамках відповідної експертної системи дозволяє досягти наступних основних результатів:

- визначення показників відповідності набору тестових питань лекційним матеріалам навчальної дисципліни як результат аналізу входжень елементів повної онтології до таблиці лем та типів;
- рекомендації по вдосконаленню набору тестових питань у відповідності до лекційних матеріалів у вигляді згенерованого експертного висновку за показниками відповідності тестових питань лекційним матеріалам навчальної дисципліни.

Очевидно, що найвищий показник відповідності теоретично може бути 100% проте недосяжний практично через специфіку використання одних і тих же ключових термінів в різних тестах а також через використання суб'єктивні особливості лексичної і семантичної побудови у тестових завданнях.

Отже, у статті було досліджено особливості розвитку дистанційної освіти як сучасного способу надання якісних освітніх послуг. Встановлено велике значення оцінки відповідності тестових завдань лекціям навчальних дисциплін. Запропоновано інформаційну технологію автоматизації аналізу відповідності тестових завдань лекційним матеріалам навчальних дисциплін. Розглянуто відповідний тестовий додаток та результати його роботи. Визначено області й особливості застосування запропонованої інформаційної технології автоматизації аналізу відповідності тестових завдань лекціям навчальних дисциплін.

Література

1. Мазур М.П. Розвиток дистанційного навчання в Україні як складової інформатизації сучасного суспільства. - Інформатика та інформаційні технології в навчальних закладах. - №1, 2007. - с. 71-75
2. Беляєв С.В., Воливач А.П., Поляничко А.Л., Самойлова О.М. Компьютерная обучающая тестовая система контроля и самоконтроля знаний // Вестник Восточноукраинского национального университета им. В. Даля. - 2011.- № 7 (161). - Ч. 1. - С. 206-208.
3. Нові інформаційні технології в освіті [Електронний ресурс]. - Режим доступу: <http://it-tehnolog.com/statti/novi-informatsiyini-tehnologiyi-navchannya/>
4. Снитюк В.Е., Юрченко К.Н. Интеллектуальное управление оценением знаний. - Черкассы, 2013. - 262 с.
5. Мазур М.П., Яновський М.Л. Нова модель цифрової дистрибуції на прикладі навчального процесу // Вимірювальна та обчислювальна техніка в технологічних процесах. - Хмельницький, 2009. - №1. - с.159-166.
6. Павлыш В.Н., Зайцева М.Н. Теоретические основы контроля в учебном процессе высшей школы // Научные труды ДНТУ. - 2011. - № 9. - С. 130-136.
7. Глибовец Н.Н., Крись А.А. Реализация подсистемы тестирования в системах дистанционного обучения // Управляющие системы и машины. - 2001. - №3. - С. 70-78.
8. Сайт Microsoft для розробників OpenXMLDeveloper [Електронний ресурс]. - Режим доступу: <http://openxmldeveloper.org/>
9. Луговий, В. І. Стандартизація стандартів вищої освіти / В. І. Луговий. //Вища освіта України : теорет. та наук.-метод. часоп. Дод. 2. Педагогіка вищої школи: методологія, теорія, технології. Т.1 / Ін-т вищ. освіти АПН України. - Київ: Гнозис, 2008. С.33-44.
10. Ковалюк Т.В. Освітній стандарт з напрямку "Комп'ютерні науки" як інструмент підготовки кваліфікованих фахівців для ІТ-галузі економіки / Проблеми освіти: наук. зб. / Київ: Ін-т інновац. дослідж. і змісту освіти. 2009. - Вип. 60. - С.24-29.
11. Ortuño M., Carpena P., Bernalda P., Muñoz E., Somoza A.M. Keyword detection in natural languages and DNA // Europhys. Lett., 2002, 57. - P. 759-764.
12. Поліщук А.О., Мазурець О.В. Експертна система аналізу відповідності тестових завдань лекційним матеріалам навчальних дисциплін // Збірник наукових праць за матеріалами восьмої міжнародної науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2014». Хмельницький - 2014. - С.263-272.
13. Д.В. Ландэ, А.А. Снарский, Е.В. Ягунова Использование графов горизонтальной видимости для выявления слов, определяющих информационную структуру текста // Труды 15-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции — RCDL-2013» - Ярославль: 2013. - с. 158-164.
14. Системний аналіз - Модульне середовище для навчання ХНУ. [Електронний ресурс]. - Режим доступу: <https://msn.khnu.km.ua/enrol/index.php?id=1903>