

ФОРМУВАННЯ МОДЕЛІ КАТАЛОГУ ГЕНЕРАТОРІВ ІНФОРМАЦІЙНОГО ОБРАЗУ ВНЗ

© Пелешишин А., Корж Р., 2014

У статті уведено поняття коректно визначеного генератора, побудовано формальну теоретико-множинну модель генератора, запропоновано формалізацію структури віртуальної спільноти як середовища інформаційної діяльності ВНЗ.

Ключові слова – вищий навчальний заклад (ВНЗ), генератор інформаційного образу, віртуальна спільнота

In the paper the notion of correctly defined generator is introduced, formal set-theoretic model of generator is built and formalization of structure of virtual community as a platform of university information activity is suggested.

Key words – university, generator of information image, virtual community

Вступ

Перспективні задачі з організації, автоматизації та оптимізації інформаційної діяльності у віртуальних спільнотах охоплюють широкий спектр теоретичного та практичного інструментарію як зі сфери програмно-технічних рішень, так і зі сфери соціальних комунікацій. Тому модель віртуальної спільноти повинна максимально повно відображати характеристики в указаних сферах.

Відповідно у модель віртуальних спільнот введено такі групи додаткових характеристик:

- технологічні показники;
- кількісні показники (показники обсягів та інтенсивності);
- соціально-демографічні характеристики;
- комунікативні характеристики;
- спеціальні показники освітньої спільноти.

Указані характеристики та показники є базовими, оскільки отримуються в результаті безпосереднього аналізу віртуальних спільнот і є фактично первинними даними про спільноти. Проте не в усіх випадках описують спільноти як генератори інформаційного образу, адже генератор може бути як окремою спільнотою, так і лише її фрагментом (набором дискусій або розділом) або «квазіспільнотою» на основі матеріалів онлайн-ЗМІ.

Місце взаємодії та спільноту, що взаємодіє, називатимемо генератором інформаційного образу. Як видно з даного означення генератора, для великих форумів або сайтів, що мають активну аудиторію, яка створює контент («спільноту»), генератором можна вважати певний тематичний сегмент такого сайту.

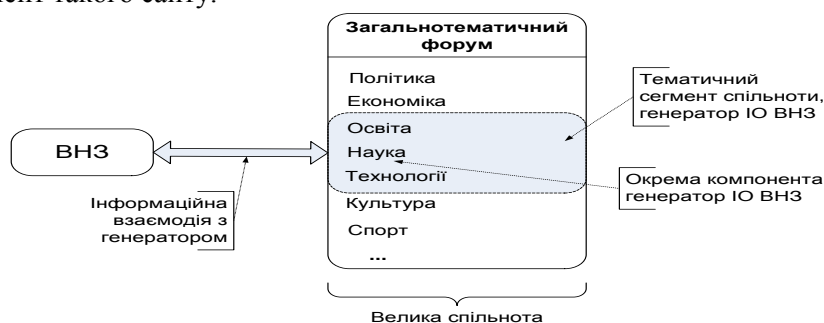


Рис. 1. Інформаційний генератор як частина віртуальної спільноти

Виділення генератора як частини спільноти (рис. 1) є виправданим кроком, який допомагає усунути колізію між необхідністю захисту інтересів ВНЗ у великих спільнотах та нерелевантністю окремих частин наповнення інтересам ВНЗ.

Зазначимо, що на основі однієї окремої спільноти можуть формуватись кілька окремих генераторів, значною мірою незалежних один від одного [1].

Постановка проблеми

Розглянемо питання достатності базової моделі каталогу для процесу формування інформаційного образу ВНЗ. Теоретично, за умови вузького набору спільнот, в який не входять загально-тематичні спільноти, професійні спільноти, проблемні спільноти, «квазіспільноти» і який обмежується лише освітньо-організаційними спільнотами, зіставлення «спільнота-генератор» було би коректним. Проте у цьому разі виникають такі, критичні для великих ВНЗ упущення:

- наукові та проблемно-орієнтовані підрозділи відсторонюються від інформаційної діяльності, оскільки не мають доступу до спільнот, що відповідають їхній сфері компетенції та інформаційному потенціалу;
- більша частина спільнот з частковою релевантністю або релевантністю окремим напрямам ВНЗ лишається поза сферою інформаційної діяльності.

Аналіз досліджень та публікацій

Тематику формалізації ВНЗ як суб'єкта інформаційної діяльності в соціальних середовищах Інтернету започатковано в попередніх дослідженнях авторів. Так, у статті [2] визначено комплекс показників віртуальної спільноти для вищих навальних закладів, а саме: технологічні, кількісні, соціальної релевантності, комунікативної цінності, комунікативної складності та тематичної релевантності. Розглянуто та проаналізовано соціально-демографічні характеристики, показники обсягів та інтенсивності, технологічні показники та комунікативні характеристики віртуальних спільнот, які є основою для визначення окремих інтегрованих показників.

У статті [3] подана роль веб-сайту кафедри, її функції та інформаційне наповнення, значення веб-сайту кафедри, як ключового структурного підрозділу формування порталу ВНЗ, що дозволяє при формуванні моделі каталогу генератора інформаційного образу врахувати всі підрозділи ВНЗ.

Серед інших досліджень виділимо працю [4], у якій запропоновано методику розробки інформаційної моделі контенту університетського веб-сайту, що функціонує в інтегрованому інформаційно-комунікаційному просторі ВНЗ.

Формулювання мети статті

У статті вирішується задача формування моделі каталогу генераторів інформаційного образу ВНЗ.

Виклад основного матеріалу

1. Компонентне визначення генератора інформаційного образу

На практиці виникають складнощі з формальним описом та подальшим опрацюванням генераторів, які є частиною спільноти (а не цілою спільнотою):

- може бути відсутня єдина фізична адреса генератора;
- визначення показників обсягів аудиторії та авторитетності потребує значних затрат ресурсів, внаслідок великих обсягів ручної роботи.

У разі відсутності єдиної фізичної адреси накопичують адреси, які дають змогу максимально точно ідентифікувати генератор. Такими адресами для визначення генератора, залежно від технічної платформи, є:

- адреси тематичних колонок чи рубрик – у онлайн-ЗМІ («ідентифікація за рубрикою»);
- адреси розділів, що релевантні завданням ВНЗ, – у веб-форумах та групах у соціальних мережах («ідентифікація за розділом»);

- адреси облікових записів авторів, що створюють релевантні дискусії, – у колективних блогах та великих слабо-структурованих спільнотах («**ідентифікація за автором**»);
- адреси окремих релевантних дискусій чи статей з коментарями – у великих слабо-структурованих спільнотах («**атомарна ідентифікація**»);
- сторінки термінів («тегів»), що ідентифікують релевантні дискусії у колаборативних платформах та блогах («**ідентифікація за терміном**»).

Наведені вище адреси вказують на окремі структурні елементи генератора, які надалі називатимемо **компонентами генератора**.

У кожному з перерахованих випадків з високою імовірністю для одного генератора існує більше ніж один компонент.

Компоненти, своєю чергою, містять окремі змістові дискусії, у межах яких і відбувається генерація інформаційного образу. Аналогічно до компонент генератора, дискусії, залежно від платформи, оформляються як:

- дискусії – на форумах, групах у соціальних мережах;
- безпосередні коментарі до публікації – в онлайн-ЗМІ, блогах, колективних блогах;
- коментарі до статті – у колаборативних системах, зокрема у вікі-системах (Вікіпедія та аналогічні).

Проте за суттю про різноманітність форм дискусій можна вести мову лише у площині візуального подання для користувача, сама структура дискусії як послідовності окремих дописів користувачів залишається незмінною [5,6].

Розглянемо детальніше питання співвідношення між компонентами генератора та дискусіями. У всіх випадках, крім одного, між компонентою та дискусіями існує відношення «один до багатьох», а у випадку атомарної ідентифікації є тривіальним, зумовлює відношення «один до одного». Формально його можна звести до загальнішого, який об'єднує решту варіантів, що показано на рис. 2.

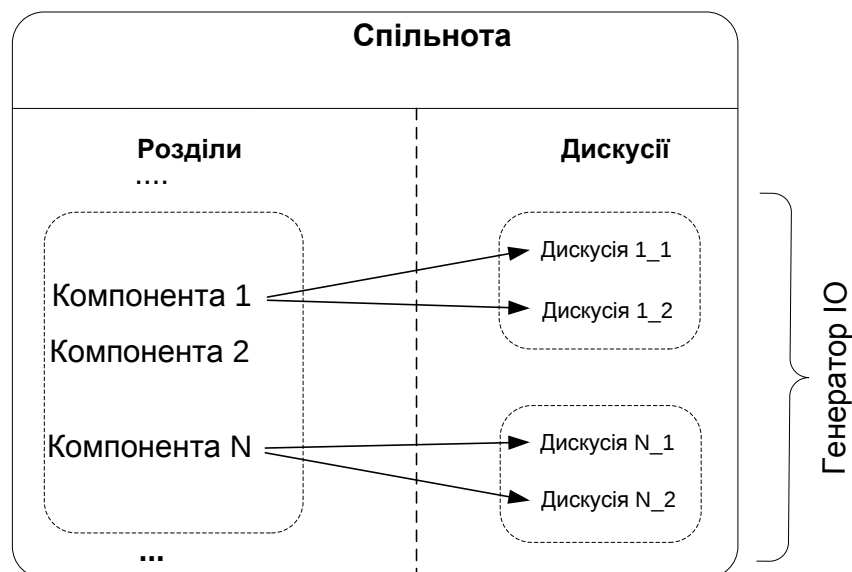


Рис. 2. Співвідношення компонент та дискусій у генераторі інформаційного образу

Лише у випадку, коли генератор визначається через набір термінів (компонента генератора визначається через термін), та сама дискусія може потрапити у різні компоненти генератора. Ця ситуація виникає, якщо дискусія маркується одночасно кількома термінами, які одночасно входять до набору релевантних термінів для визначення компонент.

Відношення компонент до дискусій «багато до багатьох», що існує в таких випадках, значно ускладнює процес планування діяльності в генераторі та моніторингу активності спільноти у межах генератора. Як буде показано далі, під час адресації деякі важливі кількісні характеристики

генератора не можна визначити коректно. Тому до вказаного методу доцільно вдаватися лише за крайньої потреби, коли іншого способу визначення компонент генератора не існує [7].

На практиці, у разі початкової ідентифікації компонент за адресами релевантних термінів, доцільним є подальший аналіз компонент для виділення іншої ознаки для адресації, основні кроки якого наведено на рис. 3.



Рис.3. Поетапне керування процесом формування інформаційного образу

З незначними модифікаціями запропонований алгоритм можна застосувати і для інших перетворень одного способу ідентифікації компонента на другий (рис. 4).

Цей підхід достатньо ефективний, оскільки для багатьох платформ найпростішим способом виявлення тематично релевантної частини сайту є саме ідентифікація за терміном.

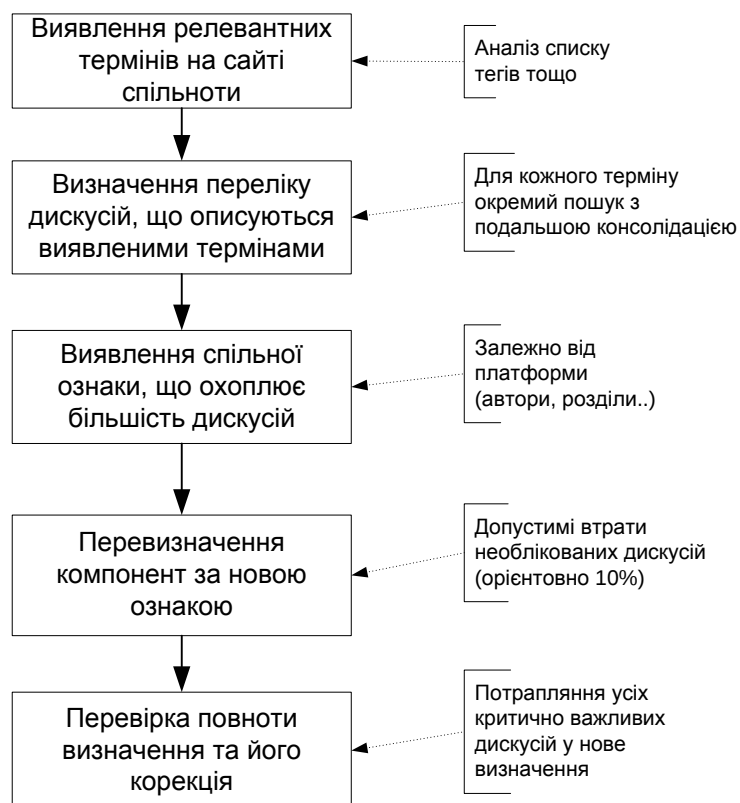


Рис. 4. Зміна способу ідентифікації компоненти

Аналогічно до розглянутої вище ситуації з відношенням між компонентами та дискусіями, яке пропонується зводити до «один до багатьох», виникає питання щодо співвідношення між генераторами в одній спільноті та їхніми компонентами.

Під час попереднього аналізу спільноти та виділення в ній компонент, які можуть бути складовими генераторів, можлива є ситуація, коли окремі компоненти увійдуть до кількох генераторів одночасно. Наприклад, розділ «Дипломне проектування» освітнього форуму може одночасно входити до генератора «наукова діяльність ВНЗ» та до генератора «магістратура ВНЗ». Тобто у загальному випадку між компонентами генератора та генератором може існувати відношення «багато до багатьох».

Проте у разі, коли виділення генераторів здійснюється заради формування та закріплення окремих команд для інформаційної діяльності, виникає ряд ризиків для ВНЗ:

- можливість конфлікту між виконавцями внаслідок неузгоджень;
- розміщення суперечливої інформації;
- «перегрів» окремого компонента через надмірний потік інформації.

Як наслідок, для планування розподілу ресурсів та призначення завдань для інформаційної діяльності ВНЗ необхідно усунути такі колізії, перейшовши до відношення «один до багатьох» із зарахуванням кожної компоненти лише до одного генератора.

Водночас, для деяких завдань аналітичного характеру (пошуку оцінкових суджень, споживацького досвіду тощо) таке спрощення може не знадобитись.

Далі вважатимемо, що, якщо не вказане інше, використовується варіант «один генератор – кілька компонент» [8].

2. Формалізація структури віртуальної спільноти як середовища інформаційної діяльності ВНЗ

Узагальнюючи структуру спільноти як середовища інформаційної діяльності ВНЗ, отримуємо таку схему (рис. 5):

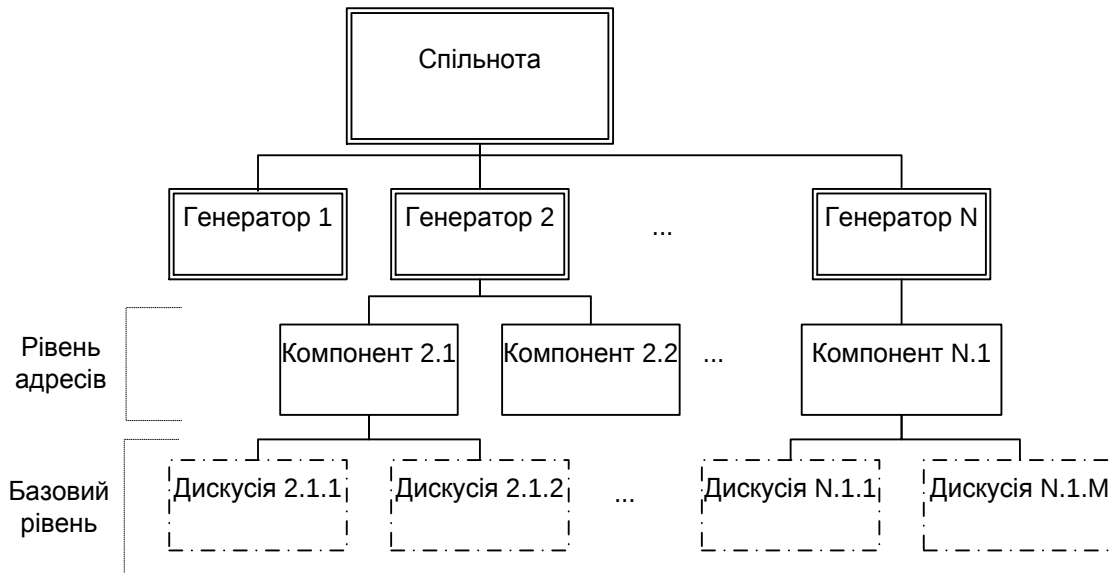


Рис. 5. Багаторівнева структура спільноти
Формалізуємо наведену вище схему віртуальної спільноти. Нехай

$$Diss(VC^*) = \{Diss_k\}_{k=1}^{N^{(Diss)}}, \quad (1)$$

де VC^* – певна віртуальна спільнота; $N^{(Diss)}$ – кількість повідомлень у спільноті VC^* ; $Diss(VC^*)$ – множина усіх дискусій спільноти.

Водночас віртуальна спільнота визначається також набором компонент:

$$Diss(VC^*) = \{Comp_j\}_{j=1}^{N^{(Comp)}}. \quad (2)$$

Компоненти також мають власну множину дискусій $Diss(Comp_j)$, причому:

$$Diss(Comp_j) \subset Diss(VC^*). \quad (3)$$

Множину компонент вважатимемо **коректно визначеною**, якщо:

$$Diss(Comp_j) \cap Diss(Comp_l) = \emptyset, \quad \forall j, l = 1 \dots N^{(Comp)} \quad (4)$$

Алгоритм отримання коректно визначеної множини компонент наведено на рис. 4.

Кожен генератор у спільноті формується з множини компонент генератора $Comp(Gen_i)$.

Тоді множина дискусій генератора Gen_i :

$$Diss(Gen_i) = \bigcup_{Comp_j \in Comp(Gen_i)} Diss(Comp_j). \quad (5)$$

Множину генераторів вважатимемо **коректно визначеною**, якщо:

$$Comp(Gen_i) \cap Comp(Gen_m) = \emptyset, \quad \forall i, m = 1 \dots N^{(Gen)}, \quad (6)$$

де $N^{(Gen)}$ – кількість виділених генераторів інформаційного образу в спільноті VC^* [9].

Очевидно, що:

$$Diss(Gen_i) \subset Diss(VC^*). \quad (7)$$

3. Використання характеристик віртуальних спільнот для опису генераторів інформаційного образу

Розглянемо детальніше питання ужитковості та застосовності пропонованих характеристик віртуальних спільнот до генераторів інформаційного образу.

Як зазначено вище, генератор часто є лише певною частиною спільноти. Це стосується як інформаційного наповнення сайту спільноти, так і її обсягів (аудиторії). Як наслідок, технічні, комунікативні характеристики спільнот прямо переносяться на генератори, утворені на базі спільноти. Окреме, незалежне від спільнот визначення соціально-демографічних характеристик для генератора на практиці також неможливе і недоцільне через безперешкодну участь будь-якого учасника спільноти у дискусіях у межах генератора.

Проте, кількісні характеристики безпосередньо до генераторів застосувати не можна. Для великих спільнот частка генератора інформаційного образу ВНЗ та наповнення, релевантного ВНЗ, достатньо незначна (у деяких випадках навіть менша за 1 %). Як наслідок, враховувати загальні обсяги спільноти як такі, що властиві генератору, некоректно [10].

Враховуючи характер генераторів як частини спільноти, найприроднішим варіантом є визначення кількісних характеристик генератора як частки від показників усієї спільноти. Визначити частку від кожного показника можливо, проте обчислення надто трудомістке і потребуватиме постійного оновлення, зважаючи на високу динамічність процесів, що відбуваються у спільноті між окремими її частинами.

Тому пропонується для кожного генератора Gen_j встановити його узагальнену частку в спільноті $Part(Gen_j)$, що визначає рівень впливу генератора та його значущість у функціонуванні спільноти. Визначимо, що $0 \leq Part(Gen_j) \leq 1$. Якщо ж спільнота загалом є генератором, приймемо $Part(Gen_j) = 1$.

Повноцінно визначити такий коефіцієнт сьогодні можна лише експертним способом на підставі доступних даних про кількісні характеристики компонент генератора та дискусій.

Початковим наближенням для визначення частки генератора $Part(Gen_j)$ є відношення між кількістю повідомлень у спільноті взагалі, оскільки саме кількість повідомлень визначається активністю та кількістю користувачів спільноти та кількістю обговорюваних тем.

$$Part^{(Post)}(Gen_j) \approx \frac{\sum_{Diss_i \in Diss(Gen_j)} PostCount(Diss_i)}{\sum_{Diss_i \in Diss(VC^*)} PostCount(Diss_i)}, \quad (8)$$

де $PostCount(Diss_i)$ – кількість повідомлень у дискусії.

Проте отримане наближення не відображає факту, що значною мірою кількість повідомлень утворюється в темах, де існує «флейм», розмішених відповідно в розділах, що не входять у тематичні для ВНЗ (на зразок розділів «Кнайпа», «Курилка», «Балачка» на популярних тематичних форумах). Для зменшення такої похибки доцільно використати частку дискусій, що належать генератору, від усіх дискусій спільноти

$$Part^{(Diss)}(Gen_j) \approx \frac{\sum_{Diss_i \in Diss(Gen_j)} Count(Diss_i)}{\sum_{Diss_i \in Diss(VC^*)} Count(Diss_i)}. \quad (9)$$

Проте і цей показник не може бути універсальним способом для визначення $Part(Gen_j)$, оскільки велика кількість «коротких» дискусій з малою кількістю повідомлень характерна для розділів, у яких низька масова активність. Дискусії та повідомлення створюють лише окремі автори або дискусії започатковуються в автоматичному режимі на основі запозичення чи реплікації контенту з інших (переважно новинних) джерел. У такому разі показник $Part(Gen_j)$ буде відчутно завищеним.

Як наслідок, можемо встановити лише можливі межі для $Part(Gen_j)$. Нехай:

$$\begin{aligned} \overline{Part}(Gen_j) &= \max(Part^{(Diss)}(Gen_j), Part^{(Post)}(Gen_j)) \\ \underline{Part}(Gen_j) &= \min(Part^{(Diss)}(Gen_j), Part^{(Post)}(Gen_j)). \end{aligned} \quad (10)$$

Тоді для визначення узагальненої частки генератора одержуємо:

$$\underline{Part}(Gen_j) \leq Part(Gen_j) \leq \overline{Part}(Gen_j). \quad (11)$$

Так визначаються допустимі межі, у яких повинна міститись експертна оцінка коефіцієнта $Part(Gen_j)$.

Кожен з показників груп VC та VA для генератора Gen_j отримують помноживши відповідний показник для спільноти на $Part(Gen_j)$. А система базових характеристик віртуальної спільноти VC_i описується кортежем:

$$VC_Ch(VC_i) = \langle VT(VC_i), VA(VC_i), VI(VC_i), VD(VC_i), VG(VC_i) \rangle, \quad (12)$$

де $VT(VC_i)$ - технологічні показники, $VA(VC_i)$ - показники обсягів, $VI(VC_i)$ - показники інтенсивності, $VD(VC_i)$ - соціально-демографічні показники, $VG(VC_i)$ - комунікативні показники [11].

Висновки

У статті досліджено задачу формалізації онлайн-спільнот як генераторів контенту, який є складовою частиною інформаційного образу ВНЗ. Визначено, що у загальному випадку генератором є не уся спільнота, але лише її певна, релевантна тематиці ВНЗ, частина. У статті побудовано формальну теоретико-множинну модель генератора, уведено поняття коректно визначеного генератора та запропоновано підходи до ідентифікації генератора як підмножини спільноти за різними ознаками. Вибір ідентифікаційної ознаки визначається технологічними характеристиками платформи, на якій функціонує спільнота.

Отримані результати є основою для проектування каталогів та баз даних генераторів інформаційного образу, математичних та імітаційних методів оптимізації інформаційної діяльності ВНЗ у онлайн-спільнотах.

1. Корж Р. О. Формалізація процесу формування інформаційного образу ВНЗ в соціальних середовищах Інтернету / Р. О. Корж, А. М. Пелецишин // *Східно-Європейський журнал передових технологій*. – Харків, 2013. - №5/3 (65) – С. 4-8. Бібліогр.: 10 назв.
2. Пелецишин А. М. Визначення комплексу показників віртуальної спільноти для вищих навчальних закладів / А. М. Пелецишин, Р. О. Корж, О. Р. Трач // *Східно-Європейський журнал передових технологій*. – Харків, 2014. - №2/2 (68) – С. 16-23. Бібліогр.: 12 назв.
3. Сугак Д. Б. Роль веб-сайта в научно-образовательной деятельности вуза / Д. Б. Сугак // *Вестник СПбГУКИ*. - №3(12), сентябрь. – 2012. — С.77–82.
4. Філіпова Л. Інформаційне моделювання контенту університетського веб-сайту: загальні принципи та методики / Л. Філіпова, А. Шелестова // *Вісник Книжкової палати*, №3, 2014.
5. Тимовчак-Максимець О. Ю. Розроблення формалізованих запитів для виявлення веб-форумів та дискусій // О. Ю. Тимовчак-Максимець, А. М. Пелецишин // *Вісник Національного авіаційного університету* : – Київ, 2012. – № 1 (50). – С. 106–111.
6. Пелецишин А. М. Аналіз комунікативної взаємодії учасників спільнот веб-форумів / А. М. Пелецишин, О. Ю. Тимовчак-Максимець // *Східно-Європейський журнал передових технологій*. – Харків, 2010. – № 6/8 (48). – С. 44-49.
7. Корж Р. О. Етапи формування цілісного інформаційного образу ВНЗ та їхня характеристика / Р. О. Корж, А. М. Пелецишин // *Інформація, комунікація, суспільство 2014 : матеріали 3-ої Міжнар. наук. конф. ІКС – 2014, 21-24 трав. 2014р., Україна, Львів, Славське / М-во освіти і науки України, Нац. ун-т «Львів. політехніка»*. – Л., 2014. – Бібліогр.: 2 назви.
8. Lisa Harris (2010) *Community in the Electronic Classroom: Virtual Social Networks and Student Learning*. VDM Verlag Dr. Müller. P. 304.
9. Korz, R., Peleschyshyn, A., Syerov, Y., Fedushko, S. (2014). "The cataloging of virtual communities of education thematic." *Webology*, 11(1), Article 117. Available at: <http://www.webology.org/2014/v11n1/a117.pdf>
10. Кудін А.П., Кархут В.Я., Франчук В.М. Інформаційно-комунікаційні технології та управління діяльністю вищого навчального закладу: освітній портал. Збірник наукових праць Кам'янець-Подільського національного університету імені Івана Огієнка. Серія педагогічна / [редкол.: П.С.Атаманчук (голова, наук. ред.) та ін.]. – Кам'янець-Подільський: Кам'янець-Подільський національний університет імені Івана Огієнка, 2010. – Вип. 16: Формування професійних компетентностей майбутніх вчителів фізико-технологічного профілю в умовах євроінтеграції. С. 26-29.
11. Verification of the Socio-demographic Characteristics of Virtual Community Members / S. Fedushko, A. Peleshchyshyn, R. Korzh, Yu. Syerov // *Сучасні проблеми радіоелектроніки, телекомунікацій, комп'ютерної інженерії : матеріали XII Міжнар. конф., присвяч. 170-річчю заснування Національного університету «Львівська політехніка», 25.02–01.03 2014 р., Львів, Славське, Україна / Нац. ун-т "Львів. політехніка"*. – Л. : Вид-во Львів. політехніки, 2014. – С. 632.