Використання онтологічного аналізу для семантичної розмітки стандартів з інформаційної безпеки

© Рогушина Ю.В.
Інститут програмних систем НАН України, Київ, Україна
ladamandraka2010@gmail.com
© Гладун А.Я.
Інститут спеціального зв'язку та захисту інформації, НТУ-КПІ ім. Сікорського,
Міжнародний науковий центр інформаційних технологій і систем НАНУ,
Київ, Україна
glanat@yahoo.com

Анотація

Аналіз публікацій вказує на велику кількість національних стандартів, безпосередньо пов'язаних з інформаційною безпекою, які вже розроблені або знаходяться на стадії узгодження. Між цими стандартами та їх елементами існує складна та слабо структурована система ієрархічних відношень, взаємозв'язків, посилань та відповідностей, що потребує відповідної формалізації. Тому розробка нормативно-правової бази України для підтримки інформаційної безпеки потребує створення методів та засобів обробки і аналізу цієї інформації на семантичному рівні.

Через недостатньо глибоко стандартизовану та усталену україномовну термінологію в стандартах використовуються різні варіанти перекладу основних понять та назв. Це призводить до проблем із пошуком потрібної користувачам інформації у неструктурованих природномовних документах. Тому пропонується використовувати Wiki-технології для подання знань, пов'язаних з проблемами інформаційної безпеки. Wiki забезпечує динамічне поновлення інформації, пошук та навігацію у мультилінгвістичних ресурсах, а семантизація Wiki-ресурсу (приміром, за допомогою Semantic MediaWiki) дозволяє виконувати розмітки та здійснювати запити на рівні знань.

На жаль, вбудовані механізми менеджменту знань у Semantic MediaWiki не достатньо функціональні, і тому ми пропонуємо інтегрувати їх із засобами онтологічного аналізу: онтологія предметної області "Інформаційна безпека" використовується як базис для семантичної розмітки відповідного Wiki-ресурсу.

В роботі запропоновано приклад використання такого підходу, що демонструє семантичну розмітку Wiki-сторінок, побудованих на основі аналізу стандартів з інформаційної безпеки, та виконання запитів до них.

Ключові слова: Wiki-ресурс, онтологія предметної області, інформаційна безпека, стандарт

1 Вступ

Останнім часом посилюються вимоги до використання нормативно-правового забезпечення України щодо оцінки відповідності систем управління інформаційною безпекою правилам та процедурам, які відповідають міжнародним стандартам та кращим міжнародним практикам. Разом з цим відбуваються якісні зміни у процесах керування знаннями для підтримки пошуку знань в системах нормативно-правової документації та забезпечення інтероперабельності термінологічної бази у сфері інформаційної безпеки, зумовлені інтенсивним впровадженням сучасних інтелектуальних інформаційних технологій.

У зв'язку з важливістю розробки національного нормативно-правового забезпечення інформаційної безпеки, великим обсягом міжнародних стандартів в цій сфері та термінологічними проблемами гармонізації національних стандартів України стає актуальною організація ефективного пошуку та навігації у різноманітних документах, пов'язаних з цією предметною областю. Важливість цієї проблеми обумовлена тим, що стандарти мають великий обсяг та складну структуру і семантичні методи дозволяють підвищення рівня пошуку інформації в галузі інформаційної безпеки. Це викликає потребу в створенні таких методів та засобів забезпечення доступу до необхідних відомостей, які легко використовувати, але які мають достатню виразну потужність для вирішення такої задачі.

2 Захист мереж. Основні концепції

Швидкий розвиток загальнодоступних мережних технологій (дротових, бездротових, Інтернету) пропонує значні можливості для діяльності різних організацій, як для транспортування інформації, так для створення і надання онлайнових сервісів. Однак, хоча це середовище забезпечує значні переваги, з'являються нові ризики безпеки, які потребують ефективного керування. Організаціям, які довіряють використанню великого об'єму інформації та відповідних мереж для своєї діяльності, втрата конфіденційності, цілісності і доступності інформації та сервісів може створити істотний несприятливий вплив на робочий процес. Отже, на перший план виступає належна захищеність мереж і пов'язаних з ними інформаційних систем і інформації.

Іншими словами, реалізація і підтримка адекватної мережної безпеки є абсолютно критичними для успіху діяльності будь-якої організації.

Сьогодні актуальною є проблема інформаційного убезпечення. Досвід експлуатації інформаційних систем і ресурсів у різних сферах життєдіяльності показує, що існують різні й досить реальні загрози втрати інформації, що приводять до значних матеріальних та інших збитків [1-4]. Розробка міжнародної та національної нормативно-правової бази документів для підтримки інформаційної безпеки, потребує застосування інтелектуальних методів та засобів обробки і аналізу релевантної інформації на семантичному рівні. Це створює передумови для усунення технічних бар'єрів між країнами в сфері технічного регулювання, в тому числі з питань безпеки.

Як відомо, процес розроблення стандартів у сфері інформаційних технологій (ІТ) і, зокрема, в напрямі інформаційної безпеки відбувається безупинно. Розробляють стандарти для відкритих систем, зокрема у сфері безпеки ІТ, спеціалізовані міжнародні організації і консорціуми. Критерії оцінювання інформаційної безпеки є методологічною базою для визначення вимог захисту комп'ютерних систем від несанкціонованого доступу, створення захисних систем та оцінювання ступеня захищеності. Послідовно публікуються проекти та версії стандартів на різних стадіях узгодження та затвердження. Значну кількість стандартів поетапно поглиблюють і деталізують у вигляді сукупності взаємозалежних груп щодо концепцій та визначених структур [5,6].

В процесі використання результатів гармонізованих міжнародних та національних стандартів у користувачів дуже часто виникає проблема зіставлення стандартів, яка на нашу думку може бути ефективно вирішена на основі семантичного підходу. Проблема зіставлення стандартів виникає при наявності в нормативно-правовій бази документів, розроблених, гармонізованих, модернізованих різними міжнародними та національними органами зі стандартизації, наприклад, ISO/IEC, ITU-T, IEEE, CEN, CENELEC, IAB, WOS, ANSI, W3C, ECMA, X/Open, OSF, OMG тощо. В якості узагальнених показників, що характеризують стандарти інформаційної безпеки, пропонується використовувати універсальність, гнучкість, гарантованість, реалізованість і актуальність.

3 Постановка задачі

Для забезпечення користувачам ефективного використання наявного національного та міжнародного нормативно-правового забезпечення з інформаційної безпеки, що включає динамічне поновлення інформації, пошук та навігацію у взаємопов'язаних мультилінгвістичних ресурсах, пропонується використовувати семантичні Wiki-технології для подання таких знань та підвищення ефективності розв'язання поставлених задач. Щоб інтегрувати знання з різних джерел, потрібно розробити онтологію предметної області "Інформаційна безпека", елементи якої будуть використовуватися для семантичної розмітки Wiki-сторінок. Використання онтологічного підходу забезпечить інтероперабельність цих знань та можливість їх застосування в різних інформаційних системах. Це вимагає створення методів співставлення елементів Wiki-сторінок з елементами онтології та засобів автоматизації семантичної розмітки природномовних текстів.

4 Онтології як засіб подання розподілених знань

Для представлення та спільного використання розподілених знань різних предметних областей (ПрО) в сучасних інформаційних системах зараз широко застосовуються онтології [7]. Онтології базуються на фундаментальному теоретичному базисі дескриптивних логік, для них вже існують загальноприйняті стандарти опису, мови і програмні засоби.

Онтологічний аналіз дозволяє перетворювати представлення певної особи або групи осіб про зовнішній світ у формалізований набір термінів і правил їхнього використання, придатний для автоматизованої обробки. Онтологію можна розглядати як базу знань (БЗ) спеціального виду із семантичною інформацією про визначену ПрО [8]. Компоненти, які використовують в різних формалізаціях онтологічних описів ПрО, залежать від як парадигми представлення, так і від цілей побудови такої онтології. Більшість формальних моделей онтологій містять класи та їх екземпляри, властивості класів, відношення між класами та екземплярами класів і обмеження їх використання. В рамках проекту Semantic Web створено формальну мову подання онтологій OWL [9] та мову опису метаданих RDF, які фактично стали стандартом для подання знань у розподілених Web-застосунках. Для них створено велику кількість інструментальних засобів обробки, приміром, Protégé.

5 Semantic MediaWiki

Одним за найбільш поширених сьогодні Wiki-двигунів є MediaWiki– вільне програмне забезпечення з відкритим кодом для гіпертекстового Wiki-середовища, що є платформою для створення довідників, енциклопедій і каталогів. Ця програма дозволяє створювати та редагувати різноманітні Wiki-ресурси, надає користувачам зручний та інтуїтивно зрозумілий інтерфейс та підтримує більшість функцій, необхідних для зручної колективної роботи.

Semantic MediaWiki (SMW) [10] забезпечує семантизацію інформації у Wiki-ресурсі. Це дозволяє користувачам додавати семантичні анотації до Wiki-сторінок, перетворюючи MediaWiki у семантичний ресурс. SMW перетворює MediaWiki на семантичний ресурс, дозволяючи автоматично інтегрувати інформацію, що

знаходиться у різних Wiki-сторінках, генерувати відповіді на складні семантичні запити, будувати бази знань та візуалізувати їх результати, що відображають знання щодо семантичних властивостей та категорій інформаційних об'єктів, виконувати логічне виведення тощо, тобто обробляти інформацію на рівні знань.

У Semantic MediaWiki використовуються такі додаткові елементи розмітки, як *семантичні властивості* (для створення даних) та *семантичні запити* (для використання даних). Semantic MediaWiki має досить високу виразну потужність, надійну реалізацію і зручний користувальницький інтерфейс, сформовані на її основі IP часто обновляються і швидко збільшуються. Використання таких джерел обумовлюється тим, що Wiki-ресурси динамічно обновляються всім співтовариством користувачів, мають чітко визначену і просту для розуміння структуру і забезпечують обробку інформації на семантичному рівні, і, таким чином, забезпечують технологічну платформу для групового керування знаннями. Слід відмітити, що, на Semantic MediaWiki базується все більша кількість розробок сайтів, порталів та енциклопедичних видань, що викликає розвиток засобів для автоматизованої обробки таких баз знань. Тому використання саме цього програмного забезпечення значно збільшує можливості використання розроблених на його базі довідників та енциклопедій.

6 Онтологія "Інформаційна безпека"

Ця онтологія описує відповідну ПрО. Вона базується на наборі національних та міжнародних стандартів, пов'язаних з питаннями інформаційної безпеки. На сьогодні побудовано наступні базові класи – "Предметна область", "Стандарт" та "Термін стандарту", для яких побудовано відповідні екземпляри та підкласи. Екземпляри цих основних класів пов'язані об'єктними властивостями "Описано в стандарті", "Посилається на стандарт", "Відноситься до ПрО" тощо. Щоб формалізувати семантику відношень між елементами онтології, що згенеровані за термінологічним базисом стандартів, побудовано специфічні для ПрО об'єктні властивості.

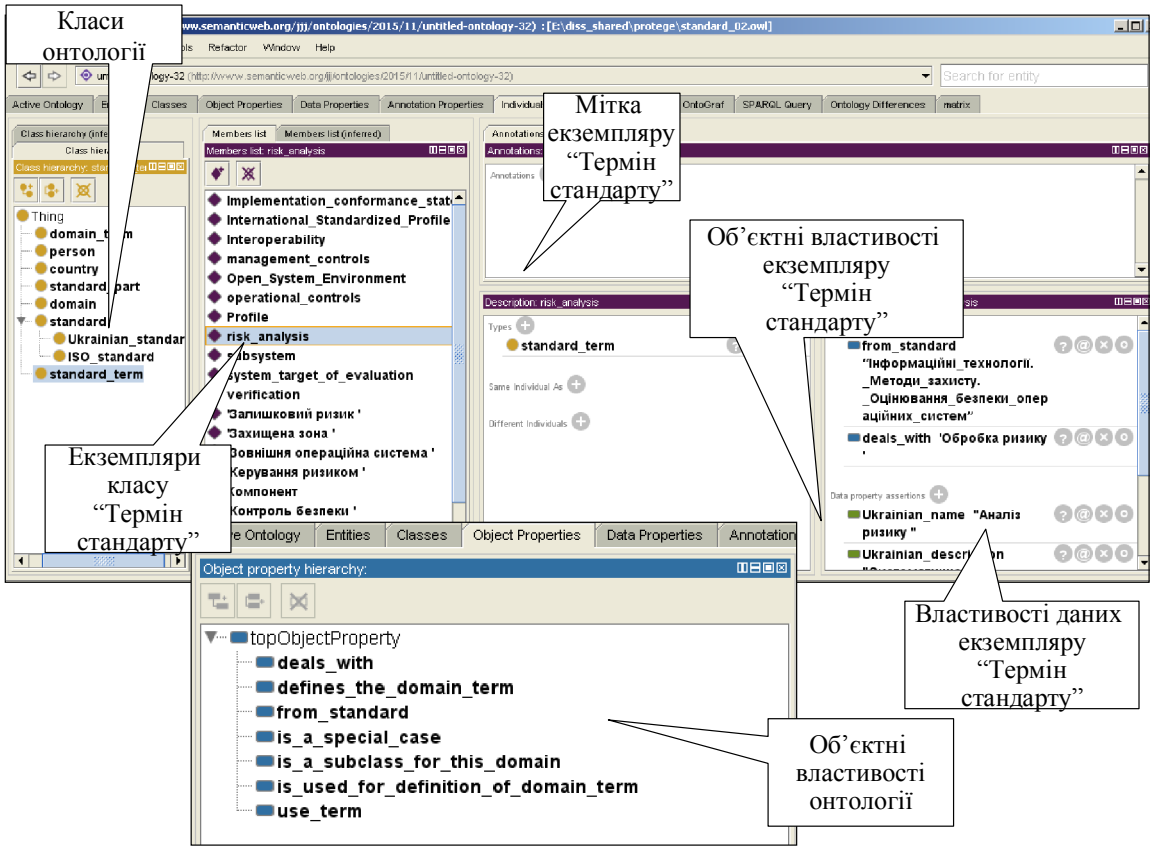

Рис.1. Об'єктні властивості та екземпляри класів в онтології "Інформаційна безпека"

7 Семантична розмітка Wiki-ресурсу термінами онтології "Інформаційна безпека"

У семантичній розмітці Wiki-ресурсу класам онтології відповідають категорії, елементам – сторінки, а об'єктним відношенням – семантичні властивості. Більш детально правила побудови цих відповідностей розглянуто в [11,12,13].

Окремі Wiki-сторінки доцільно створити як для кожного стандарту в цілому (перелік його підрозділів), так і для окремих підрозділів та визначень. Крім того, доцільно використовувати гіперпосилання на інші Wiki-ресурси, приміром, на Вікіпедію або на електронну версію Великої української енциклопедії (е-ВУЕ).

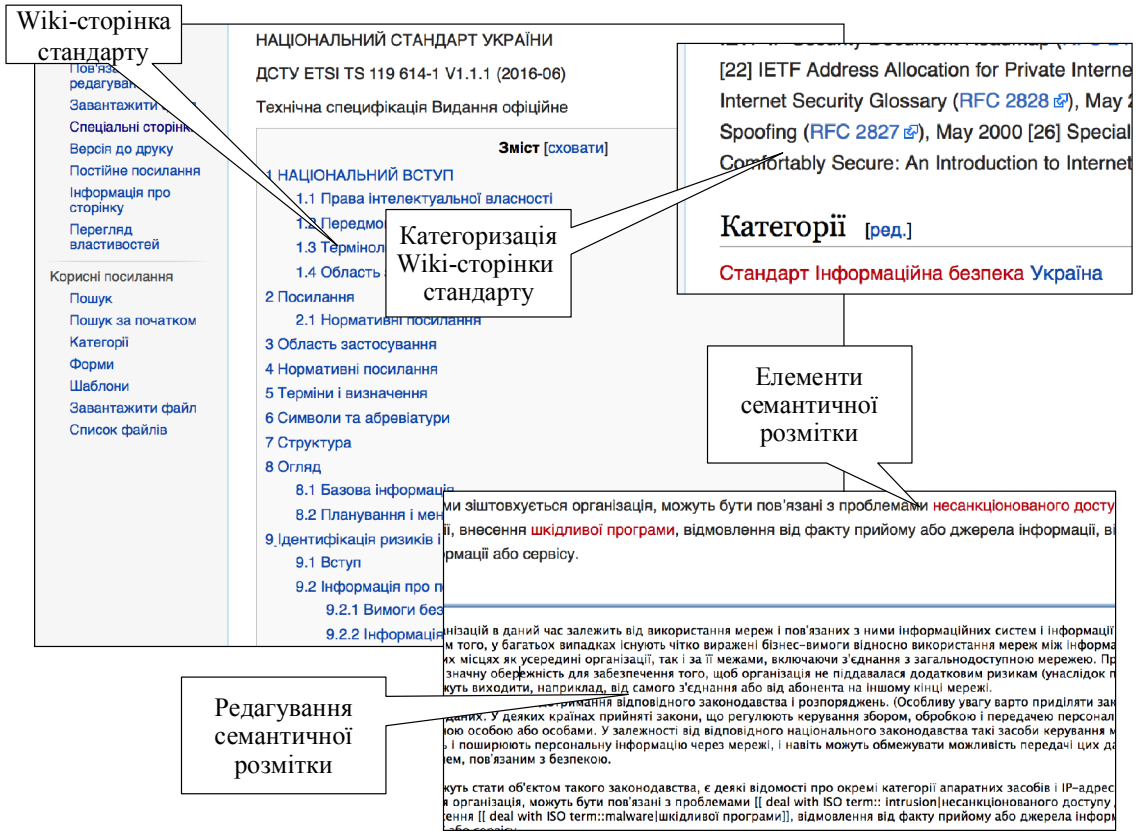

Рис.2. Семантична розмітка Wiki-ресурсу за допомогою онтології “Інформаційна безпека”

Таким чином, якщо вже побудована онтологія мовою OWL, тоді її досить просто використовувати в Semantic MediaWiki. Але зворотний процес не може бути цілком автоматизований. Більш того, при автоматичній генерації онтології за Semantic MediaWiki будуть загублені відомості, що містяться в OWL-онтології та стосуються характеристик класів і властивостей, що не мають аналогів у Wiki (зокрема, про еквівалентність класів і властивостей, їхній неперетин, про їхню область значення і визначення). У той же час, частина контента Semantic MediaWiki не може бути безпосередньо трансформована в онтології [14]. Наприклад, той факт, що в сторінках використаний одні і той же шаблон, говорить про те, що на цих сторінках описані інформаційні об'єкти одного типу, але для відображення цього в онтології треба створити специфічний клас і зв'язати його з елементом сторінки. Але потім за такою онтології неможливо зрозуміти, що треба створити в Wiki – шаблон чи категорію: вибір залежить від користувача, тому що для ІО зі специфічної, але формалізованою структурою доцільно створювати шаблони, а у всіх інших випадках – сторінки, що будуть віднесені до якої-небудь категорії. Крім того, не можливо зв'язати з класом онтології не всю сторінку, а її конкретний фрагмент (крім тих випадків, коли в нього є підзаголовок).

8 Висновки

Використання технології Wiki та її семантизації для подання предметної області інформаційної безпеки є актуальним завданням у сучасну епоху зростання великих обсягів нормативно-правової мультилінгвістичної документації. Онтологічний підхід для вирішення поставленого завдання забезпечує інтероперабельне представлення знань та застосування відповідних технологій і форматів і дозволяє інтегрувати ці елементи знань із знаннями, представленими в Semantic Web. Онтологія національних та міжнародних стандартів (та інших документів) з інформаційної безпеки використовує семантичну розмітку природномовних текстів та пов’язує описи реальних або розроблюваних систем в яких реалізовані ці стандарти. Такий підхід значно спрощує пошук та аналіз таких документів та забезпечує можливість їх автоматичної обробки. Прикладом застосування такого підходу може бути пошук рекомендації щодо ідентифікації та аналізу ризиків мережної безпеки для визначення вимог до безпеки мереж. При цьому користувачеві не потрібно буде самостійно відслідковувати зміни в останній редакції обраного стандарту або передивлятися опис кожної потенційно придатної системи – співставлення має виконуватися автоматизовано. Технологічною основою для такої семантичної розмітки та пошуку може стати середовище Wiki із семантичним розширенням.

Перспективи розвитку у майбутньому результатів нашої роботи передбачають створення *глобальної семантичної мережі стандартів*, яка пов’яже окремі національні та міжнародні стандарти; об’єкти, що використовують ці стандарти та посилаються на них (як матеріальні, так і інформаційні об’єкти); фахівців, що є

експертами в сфері розробки стандартів, та організації різного рівня, що підтримують різні види діяльності, пов'язаної із розробкою та використанням стандартів.

Література

1. Інформаційні технології. Методи захисту. Захист мережі. Частина 1. Огляд і поняття (ISO/IEC 27033-1:2015, IDT): ДСТУ ISO/IEC 27033-1:2018. — Вид. офіц. — Вперше введ. 2016-10- 01. — К. : Держспоживстандарт України, 2017. — 89 с. (Національний стандарт України).
2. ISO/IEC 27033-2:2015 «Information technology – Security techniques – Network security – Part 2: Guidelines for the design and implementation of network security (Інформаційні технології. Методи захисту. Захист мережі. Частина 2. Настанови щодо розроблення та впровадження безпеки мережі).- www.iso.org.- 28p.
3. ISO/IEC 27033-3:2015 «Information technology – Security techniques – Network security – Part 3: Reference networking scenarios - Threats, design techniques and control issues (Інформаційні технології. Методи захисту. Захист мережі. Частина 3. Еталонні сценарії побудови мереж. Загрози, методи розроблення та проблеми керування) .- www.iso.org.- 30p.
4. ISO/IEC 27033-3:2015 «Information technology – Security techniques – Network security – Part 4: Securing communications between networks using security gateways (Інформаційні технології. Методи захисту. Захист мережі. Частина 4. Захист обміну даними між мережами за допомогою шлюзів безпеки) .- www.iso.org.- 22p.
5. Гладун А.Я., Хала К.О. Таксономія стандартів з інформаційної безпеки // Наука, технології, інновації, сер. Інформаційні технології, 2017, №2 .-С.53-67.
6. ДСТУ ISO/IEC 15408-1:2017 Інформаційні технології. Методи захисту. Критерії оцінки. Частина 1. Вступ та загальна модель (ISO/IEC 15408-1:2009, IDT).-223с.
7. Gruber, T.R. The role of common ontology in achieving sharable, reusable knowledge bases. – http://www.cin.ufpe.br/~mtcfa/files/ 10.1.1.35.1743.pdf.
8. Guarino N. Formal Ontology in Information Systems. Formal Ontology in Information Systems. Proceedings of FOIS'98, 3-15, 1998.
9. OWL 2 Web Ontology Language Document Overview. W3C. 2009 (Електронний ресурс). – http://www.w3.org/TR/owl2-overview/.
10. Krotzsch M., Vrandecic D., Volkel M.: Semantic MediaWiki. – http://c.unik.no/images/6/6d/SemanticMW.pdf
11. Рогушина Ю.В. Сопоставление семантических информационных ресурсов web на основе онтологического анализа // International Journal "Information Technologies & Knowledge" Volume 11, Number 1, 2017. – С.49-71. – http://www.foibg.com/ijitk/ijitk-vol11/ijitk11-01-p03.pdf.
12. Гладун А.Я., Рогушина Ю.В. Онтологічній підхід до проблем підвищення якості розроблення національних стандартів України // Стандартизація, сертифікація, якість, № 2, 2016. – С. 19–28.
13. Гладун А.Я., Рогушина Ю.В. Семантичні технології: принципи та практики: монографія; ред. С. Кузнецов. — К. : ТОВ "ВД "АДЕФ- Україна", 2016. — 387 с.
14. Рогушина Ю.В., Гладун А.Я., Снігир Г.В. Онтологічний підхід до розробки національних стандартів України з оцінювання безпеки інформаційних технологій// Информационные технологи и безопасность. ИПРИ НАНУ. Киев, 2017.-С.58-72.

Usage of Ontological Analysis for Semantic Marking of Information Security Standards

© Julia V. Rogushina
Institute of Software Systems of National Academy of Sciences of Ukraine, Kyiv, Ukraine
ladamandraka2010@gmail.com
© Anatoly Y. Gladun
Institute of Special Communication and Information Protection of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute",
International Research Center of Information Technologies and Systems of National Academy of Sciences of Ukraine, Kyiv, Ukraine
glanat@yahoo.com

Abstract

Analysis of publications indicates a large number of national standards directly related to information security that were developed or are processed now. The complex and semistructured system of hierarchical relationships,

interconnections, references and correspondences between these standards and their elements needs in relevant formalization. Therefore the development of the normative legal framework of Ukraine oriented on support of the information security requires the development of methods and means for processing and analyzing of this information on the semantic level. Due to the lack of deeply standardized and well-established Ukrainian terminology, these standards use various translation variants of basic concepts and names. This fact causes to problems with retrieval of relevant information into the unstructured natural language documents. Therefore, we propose to use Wiki technology to represent knowledge related to information security domain. Wiki provides a dynamic updating of information, search and navigation into multi-linguistic resources. The semantization of Wiki resource (for example, with the help of Semantic MediaWiki) provides knowledge level of their mark up and execution of queries.

Unfortunately, the built-in knowledge management tools of Semantic MediaWiki are not purposeful enough, and therefore we propose to integrate them with means of ontological analysis where domain ontology "Information Security" is used as the basis for semantic markup of the corresponding Wiki-resource.

The paper proposes an example of such approach use that demonstrates the semantic markup of Wiki pages built on the basis of an analysis of information security standards and the execution of requests for them.

Keywords: Wiki-resource, domain ontology, information security, standard.