

*Розглянуто поняття семантичного репозиторію у сучасних системах менеджменту знань. Запропоновано архітектуру захищеного семантичного репозиторію документів з можливістю контролю версій. Розглянуто його інтеграцію з розподіленим онтологічним порталом*

*Ключові слова: семантичні технології, семантичний репозиторій, онтологічний портал, архітектура*

*Рассмотрено понятие семантического репозитория в современных системах менеджмента знаний. Предложена архитектура защищенного семантического репозитория документов с возможностью контроля версий. Рассмотрена его интеграция с распределенным онтологическим порталом*

*Ключевые слова: семантические технологии, семантический репозиторий, онтологический портал, архитектура*

*The notion of a semantic repository in modern systems of knowledge management is considered. The architecture of secure semantic repository of documents with the ability to version control is proposed. Its integration with distributed ontological portal is considered*

*Keywords: semantic technologies, semantic repository, ontology portal, architecture*

УДК 004.89:[004.75:002.53/55]

# ПРОТОТИП СЕМАНТИЧЕСКОГО РЕПОЗИТОРИЯ ДОКУМЕНТОВ РАСПРЕДЕЛЕННОГО ОНТОЛОГИЧЕСКОГО ПОРТАЛА

**Н.В. Рябова**

Кандидат технических наук, доцент, исполняющий обязанности заведующего кафедры\*

**А.Ю. Шевченко**

Кандидат технических наук, доцент\*

**М.В. Белоиваненко**

Научный сотрудник\*

**А.А. Воскобойникова**

Старший преподаватель\*

**М.В. Головянко**

Кандидат технических наук, старший преподаватель\*

**Н.А. Волошина**

Старший преподаватель\*

\*Кафедра искусственного интеллекта

Харьковский национальный университет радиоэлектроники  
пр. Ленина, 16, г. Харьков, Украина, 61166

E-mail: ai@kture.kharkov.ua

## Введение

Область образования и науки Украины является сложным и многогранным объектом исследований. Поэтому, для поддержки процессов, которые происходят в этой предметной области (ПрО) группой разработчиков Харьковского национального университета радиоэлектроники (кафедра искусственного интеллекта) был выбран онтологический подход к представлению и обработке знаний в распределенных гетерогенных средах, наиболее перспективный на сегодняшний день. Разработанный в рамках научно-исследовательской работы ИТ/543-2009 «Разработка и внедрение распределенной архитектуры онтологического портала МОНУ для надежного, безопасного и эффективного менеджмента и интеграции образовательных ресурсов» онтологический портал должен обеспечивать решение, прежде всего, следующих задач в области образования и науки:

- обеспечение эффективных процедур менеджмента образовательных и научных ресурсов, которые подчинены Министерству образования и науки Украины;
- обеспечение контроля качества высшего образования Украины на основе поддержки процедур аккредитации и лицензирования ВУЗов;

– обеспечение прозрачности формирования показателей деятельности высших учебных заведений и возможности их анализа в определенный период.

База знаний (БЗ), которая является основой онтологического портала менеджмента национальных образовательных ресурсов, хранит в себе все ресурсы ПрО, их характеристики, параметры и механизмы их обработки. Это дает возможность (на основе установленных семантических связей между ресурсами ПрО, их параметрами и значениями этих параметров) в значительной степени автоматизировать процедуры анализа показателей деятельности ВУЗов и качественных характеристик ресурсов ПрО.

Для расширения функциональных возможностей онтологического портала и обеспечения представления в нем актуальной на текущий момент, корректной и правдивой информации в рамках научно-исследовательской работы ИТ/543-2009 была предусмотрена разработка семантического репозитория документов, который будет интегрирован с порталом.

Основной задачей семантического репозитория является обеспечение и регламентация процессов заданной ПрО. В контексте онтологического портала менеджмента образовательных ресурсов такой репозиторий должен содержать документы, которые регламентиру-

ют все виды деятельности высших учебных заведений на всех структурных уровнях заданной ПрО. Документы, хранящиеся в семантическом репозитории, устанавливают допустимые нормы для значений показателей деятельности учебных заведений; подтверждают достоверность всей информации, которая наполняет онтологический портал. Таким образом, семантический репозиторий будет содержать не только знания о происходящих процессах и установленные нормативы для ресурсов в области образования и науки, но и обеспечит необходимый уровень доверия к наполнению онтологического портала. Для достижения этой цели разработчиками созданы механизмы, позволяющие устанавливать связи между значениями параметров ресурсов в онтологическом портале и документами репозитория, содержащими заявленное в портале значение параметра и/или предоставляющими информацию о его происхождении, подписанными или согласованными ответственными должностными лицами.

### 1. Понятие семантического репозитория в современных системах менеджмента знаний

В современных системах менеджмента знаний и поддержки принятия решений значительная часть информации, которую необходимо хранить и обрабатывать, представлена в виде текстовых документов. К таким системам относятся и современные корпоративные системы, которые возможно будут функционировать в Интернет-пространстве, а также порталы знаний. В последнем случае ядром системы является онтологическая база знаний, в которой эксплицитно представлена концептуальная модель ПрО, основные понятия (концепты) и термины, релевантные ПрО, а также система отношений между ними. Корпоративный портал должен обеспечивать единый способ доступа к внутренней информации организации, который позволяет сотрудникам устанавливать взаимосвязь информации с коллективным пониманием, системой ценностей и опытом принятия оптимальных решений. Корпоративный портал можно рассматривать как интегрированное хранилище информации (репозиторий), доступное для оперативного обобщения и анализа. Особую актуальность при этом приобретают методы построения репозитория текстовых документов [1].

Для повышения эффективности работы с репозиторием документов в данной работе используется семантический подход, который предусматривает выделение составляющей каждого документа в виде семантической аннотации, которая может быть представлена в виде RDF-файла [2]. Кроме стандартных тегов, которые характеризуют документ (автор, название, тип документа, год, источник и т.п.), семантическая аннотация включает также и семантические теги, в которых определены основные концепты и термины (предварительно полученные из документов, например, методами Text Mining). Представление данных в виде RDF-графов позволяет использовать онтологический подход как для концептуального индексирования семантических аннотаций, так и для их сравнения при поиске похожих документов или документов из одного смыслового поля. Вопросы предварительной

обработки текстовых документов, формирования семантических аннотаций и работы с ними рассмотрены в [3,4].

Благодаря стремительному развитию и распространению семантических технологий возникло новое видение Semantic Web как среды менеджмента знаний, которая предъявляет новые требования к хранению, поиску и обработке больших объемов информации в Web-ориентированных системах. За последние годы наиболее эффективным подходом для этого стали так называемые семантические репозитории. Следует отметить, что до сих пор отсутствует общепринятое определение «семантического репозитория». К неполному списку приблизительных синонимов, которые используются параллельно с термином «семантический репозиторий» относятся, например, такие понятия:

- блок логических рассуждений, «ризонер» (reasoner);
- онтологический сервер (ontology server);
- семантическое хранилище (semantic store);
- метахранилище (metastore);
- RDF база данных (RDF database).

Такая разница в формулировках относительно схожих понятий является отображением базовых отличий в реализации, выполнении, внедрении, продуктивности, области и задач применения и т.п. В определении «семантический репозиторий» отображена попытка охватить основную функциональность и целевое назначение конечного продукта, создаваемого с помощью разнообразного инструментария и объединенного общим названием.

Семантические репозитории комбинируют в себе свойства СУБД и машин логического вывода (inference engine). Таким образом, семантический репозиторий рассматривается как машина логического вывода, по функциональности близкая к СУБД потому, что тоже предназначена для хранения, опроса и управления структурированными данными. Семантические репозитории играют роль Web-серверов потому, что способны хранить, интерпретировать и обслуживать запросы различных категорий пользователей к большим массивам данных. Следует отметить, что они все еще находятся на начальной стадии своего стремительного развития. С 2004 года каждую пару лет машины логического вывода становятся на порядок быстрее и более масштабируемыми. Однако с каждым новым достижением в создании семантических репозиториев, помимо новых возможностей, возникают новые проблемы. Такие системы метафорично сравнивают с машинами-путеукладчиками (track-laying machines), которые продолжают «железнодорожный путь» области действия данных, шаг за шагом меняя организацию и структуру данных целых предметных областей, предоставляя возможности для организации все более и более сложных данных, которые можно обрабатывать с меньшими временными затратами.

### 2. Взаимосвязь параметров ресурсов в онтологическом портале и документов, которые их подтверждают

Процесс формирования семантических аннотаций для различных документов является достаточно сложным, поскольку должен включать формализацию есте-

ственно-языковой информации. В связи с этим в данной разработке (онтологическом портале) реализован автоматический процесс формирования «неполных семантических аннотаций» документов. Онтологический портал создан таким образом, что содержит большое количество параметров, каждый из которых может быть подтвержден тем или другим документом из семантического репозитория. В строке, где редактируется параметр, предусмотрена кнопка, позволяющая подтвердить введенные данные отсканированной копией и/или электронной версией документа (рис. 1).

Дисципліна та її забезпечення	
Цикл, до якого належить дисципліна	Дисципліни вільного вибору ВНЗ
<i>Кількість примірників основної літератури згідно з програмою</i>	
Підручники та навчальні посібники	2
Конспекти лекцій	1
Методичні вказівки	2
<i>Навчально-методичне забезпечення</i>	
Забезпеч. навчальною літ-рою, %	100
Наявність фахових періодичних видань	<input checked="" type="checkbox"/>
Підручників на одного студента	1

Рис. 1. Внесение параметров в онтологический портал МОНМС Украины

Такая связь может быть использована для автоматического формирования семантической аннотации документа. Это очень удобно: если компьютер сам не в состоянии провести анализ отсканированного документа со 100% соответствием, то все же определенный анализ можно сделать благодаря учету связей документа с параметрами, которые он подтверждает.

В онтологическом портале структура БЗ содержит описание всех типов объектов, их параметров и механизмов их вычисления, а также семантическое описание страниц онтологического портала. Портал комбинирует эти описания и, при необходимости, выполняет действия по вычислению параметров, на основании заданных методов. На основании этой информации динамично создаются Web-страницы портала. Онтология хранит различные элементы знаний, организованные в графовую структуру. Однако при формировании страниц онтологического портала используется только та информация, которая связана с описанием вкладок портала, при этом логические связи между объектами и классами портала сохраняются, но не используются для дальнейших вычислений. Для формирования рабочих вкладок онтологического портала на данный момент используется упрощенная версия онтологического дерева, которая содержит ссылки только на объекты классов. Онтология ресурсов портала может расширяться для наращивания необходимой функциональности. Кроме того, ее объекты имеют большое количество параметров, которые также можно задать в рамках онтологического портала. Просмотр набора параметров любого из классов онтологии образовательного процесса осуществляется с использованием заданного пользователем набора параметров фильтрации. Благодаря полному онтологическому описанию параметров появляется возможность формирования достаточно детальной RDF-аннотации для каждого из

них. При ссылке любого параметра на документ, его описание и значение добавляется к аннотации этого документа. Таким образом, для каждого документа репозитория накапливается достаточно подробная семантическая аннотация. Такой подход позволяет эффективно решать задачи семантического поиска, сравнения документов, предоставляет возможность узнать, чем предыдущая версия документа отличается от последующей, провести сравнение версий документа и его аннотации для выявления расхождений.

### 3. Описание документа семантического репозитория и разработка его формального представления

Архитектура онтологического портала менеджмента образовательных и научных ресурсов МОНМС Украины основана на принципах разработки распределенных систем,

ориентированных на работу в Интернет-пространстве и рассчитанных на большое количество пользователей. Портал представляет собой специализированную информационную систему, снабженную эргономичным Web-интерфейсом пользователя. Информационную основу портала составляет онтология и связанные с ней описания соответствующих ресурсов. Одной из основных задач выполнения данного этапа работы являлась реализация прототипа семантического репозитория документов с возможностью контроля версий, а также его интеграция с распределенным онтологическим порталом. В связи с этим рассмотрим подробнее понятие документа семантического репозитория, введем его основные характеристики и покажем его связи с другими ресурсами системы.

В семантическом репозитории документы хранятся в виде двух основных составляющих: непосредственно самого документа и семантической аннотации, сформированной на основе связи между онтологическим порталом и семантическим хранилищем. Введем основные характеристики документа, которые учитываются на данном этапе разработки, а также набор связей между документами и ресурсами онтологии.

Необходимым условием формального представления документа в рамках решения задачи менеджмента ресурсов является выделение его значимых элементов, отображающих специфику процессов интеграции семантического репозитория и онтологического портала МОНМС Украины. Иерархия таких элементов охватывает, прежде всего, реквизиты, которые характеризуют данный информационный источник. Если рассматривать документооборот в высших учебных заведениях, то можно выделить, например, следующие основные структурные характеристики документа:

1) організація, відділ, випускаючий документ (*DocAuth*):

- відділи і підрозділи університету, к которым можуть бути віднесені наступні (назви окремих підрозділів в різних ВУЗах можуть нескількох відрізнятися): ректорат, відділ організації навчального процесу, факультет, навчально-методический відділ, приймальна комісія, центр довузовської підготовки, науково-дослідницька частина, відділ кадрів (студентів, професорсько-викладацького складу, навчально-вспомогательного персоналу), бухгалтерія, планово-фінансовий відділ, відділ охорони праці і др.;

- структурні підрозділи кафедр: філіал кафедри, навчальна лабораторія, науково-дослідницька лабораторія і т.п.;

2) тип документа (*DocType*), який задається списком, наприклад: наказ, розпорядження, виписка, отчет, акт, план, договір, протокол, журнал учета, графік роботи, паспорт, нормативний документ, заявка і т.п.;

3) назва документа *DocName*;

4) реєстраційний номер документа *RegNumber*;

5) дата випуску документа *ReleaseDate*;

6) дата підписання документа *ConsummatedDate*;

7) строки дієвості документа *ValidityPeriod*;

8) рахунок документа *AccountNumber*;

9) кількість сторінок документа *PageQuantity*;

10) особи, з якими документ погоджений (обов'язково, ФІО) *AgreementPerson*.

В зв'язі з інтеграцією онтологічного порталу МОНМС України і семантичного репозиторія, виділяють деяке число ознак самого файлу, який завантажено в систему, і для якого з часом буде сформована семантична анотація з урахуванням інформації, яка зберігається в системі. Крім того, передбачено велика кількість посилань *Pr* між документом (завантаженим файлом) і онтологічним порталом, які також враховують контроль версій і захищають інформацію цифровою підписом. Їх можна виділити окремо, а саме: прив'язка до певного ресурсу порталу *PortalResource*, посилання на першу попередню версію *PreviousVersion*, кількість попередніх версій *PreviousVersionNumber*, URI для доступу до файлу *DocFileURI*, розмір в байтах *FileSize*, дата завантаження файлу *FileLoadingDate*, тип контенту *ContentType*, користувач, який завантажив документ *FileLoadingUser*, користувач, який поставив підпис *FileSignUser*, текст підпису *Signature*, статус валідації підпису *SignValidation*. Таким чином формуються багато, що характеризує інформацію об'єкта, який зберігається в семантичному репозиторії,  $Pr = \{PortalResource, PreviousVersion, PreviousVersionNumber, DocFileURI, FileSize, FileLoadingDate, ContentType, FileLoadingUser, FileSignUser, Signature, SignValidation\}$ . Це описання дозволяє отримати зв'язок між ресурсами онтологічного порталу і завантаженим в систему документом, який використовується далі для формування семантичних анотацій.

Формально узагальнену структуру документа семантичного репозиторія можна представити в наступному вигляді:

$$M^D = \{DocAuth, DocType, DocName, RegNumber, ReleaseDate, ConsummatedDate, ValidityPeriod, AccountNumber, PageQuantity, Pr\}.$$

Приведене структурне описання є зображенням інформації про документ, завантаженою в систему, узагальнена модель якого представлено на рис. 2. Всі характеристики, які були наведені, використовуються для надання максимальної інформативності даним, які зберігаються в семантичному репозиторії. Слід зауважити, що запропонована модель  $M^D$  є актуальною лише для завдань, які розглядаються в даному дослідженні. Будь-яке змінення інтеграції семантичного репозиторія і онтологічного порталу призведе до певних змін в описанні вимог і основних характеристик, які вводяться для документів.

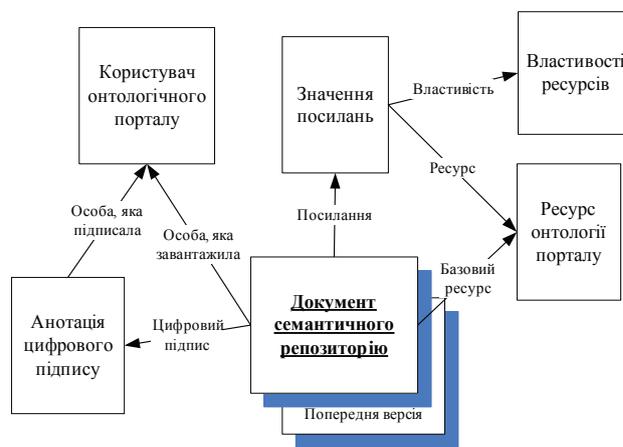


Рис. 2. Узагальнена модель документа семантичного репозиторія

#### 4. Програмна архітектура семантичного репозиторія документів

Репозиторій як система, орієнтована на конкретну Про, зберігає не тільки дані, але і процедури їх аналітичної обробки. Тому, використовуючи технології онтологічного інжинірингу для представлення і обробки знань Про, для якої власне і розробляється репозиторій, справедливим є твердження, що репозиторій стає семантичним, оскільки зберігає в собі «повну картину світу» (предметної області) – знання, правила їх обробки і застосування в межах системи.

Функціональне призначення репозиторія документів відокремлює його від будь-яких інформаційних систем, які забезпечують обробку даних і знань в заданій Про. Виходячи з цього, до семантичного репозиторія документів можна встановити наступні вимоги:

1) семантичний репозиторій повинен бути автономною системою, яка реалізує процедури управління документами, які підтримують і регламентують завдання Про;

2) документи, які зберігаються в семантичному репозиторії, повинні бути структуровані відповідно до областей їх застосування; таким чином

они являются полностью отделенными от структуры (модели) средств их обработки;

3) семантический репозиторий должен содержать исчерпывающую информацию о процессах, которые происходят в заданной ПрО;

4) информация, которая содержится в репозитории, создается за его пределами, не может быть измененной или удаленной из репозитория; то есть данные и знания, которые будут храниться в семантическом репозитории, должны быть доступны только для просмотра;

5) еще одной особенностью репозитория является то, что информация в нем корректна только тогда, когда она связана с конкретным промежутком времени или конкретной датой; таким образом репозиторий, кроме хранения актуальной текущей информации, должен содержать также все предыдущие версии документов;

6) семантический репозиторий должен реализовывать механизм интеграции с той системой, работу которой он поддерживает.

Реализация семантического репозитория, который содержит документы, связанные с задачами менеджмента национальных ресурсов, требует также реализации полностью защищенной системы.

Возможность надежной защиты может быть осуществлена благодаря поддержке следующих функций репозитория:

- распределение ролей пользователей, что обеспечит возможность создания и просмотра документов репозитория лишь теми лицами, которые имеют соответствующие их должности права доступа к информации; реализация распределения ролей должна осуществляться на основе знаний, которые содержатся в онтологическом портале менеджмента национальных образовательных ресурсов;

- реализация возможности цифровой подписи документов, которые будут храниться в репозитории, что даст возможность устанавливать ответственных за любую предоставленную информацию, а также обеспечить высокий уровень доверия ко всем документам репозитория;

- также семантический репозиторий должен содержать защищенное хранилище данных.

На основе обозначенных выше требований к созданию семантического репозитория документов для поддержки задач онтологического портала менеджмента национальных образовательных ресурсов, была разработана программная архитектура репозитория, которая приведена на рис. 3. Разработанный семантический репозиторий реализован как автономная Web-система, которая интегрирована в онтологический портал.

Архитектура семантического репозитория предусматривает, прежде всего, возможность реализации самых важных функций – механизмов загрузки документов в репозиторий с возможностью контроля версий, механизмов просмотра всех существующих документов репозитория согласно заданных параметров просмотра и прав доступа к информации. Эти функции реализованы отдельными модулями системы, которые связаны с моделью распределения прав доступа и защищенным файловым хранилищем. Также в рамках предложенного семантического репозитория предусмотрена возможность реализации модуля поддержки цифровой подписи документов, который предусматривает организацию хранения открытых ключей цифровой подписи и возможность верификации подписей на основе этих ключей.

Документы в семантическом репозитории имеют метаданные в виде схемы документа, которая содержит необходимые реквизиты, необходимые для его

корректного хранения и дальнейшего использования. В схеме документа репозитория предусмотрена возможность назначения подомласти (назначение категории документа) предметной области, к которой он относится. Основной особенностью разработанного семантического репозитория является механизм его интеграции с онтологическим порталом, который реализован благодаря созданию семантической аннотации документа, которая содержит ссылки на значения параметров ресурсов онтологического портала или на сами ресурсы.

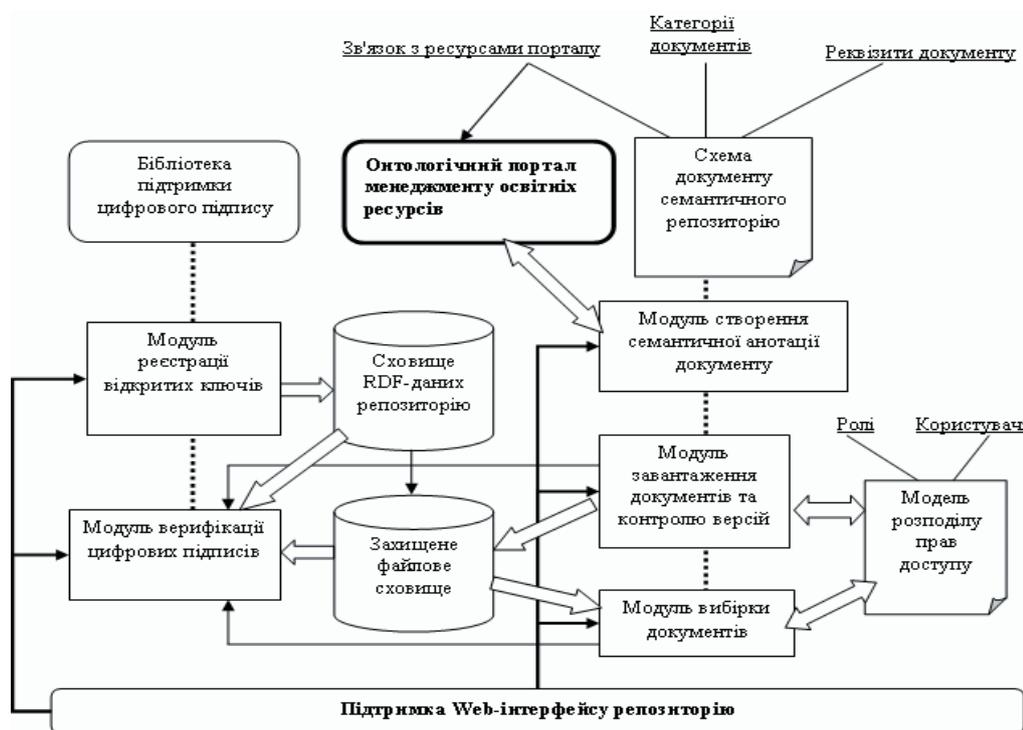


Рис. 3. Архитектура семантического репозитория

## 5. Контроль версий документов семантического репозитория и определение концепции защищенного хранилища

Документооборот организации, которая поддерживается средствами разработанного онтологического портала, является сложным процессом движения документации от ее создания до регистрации в разработанном семантическом репозитории. Важным аспектом является версияльность документов, что является неотъемлемой особенностью процесса накопления и обработки актуальной и достоверной документации любой организации, особенно таких организаций со сложной структурой и многоуровневыми иерархиями подчинения, как МОНМС Украины и высшие учебные заведения.

Определение версии документа связано с временной отметкой, назначаемой каждому документу, который размещается в репозитории. Порядок накопления документов в репозитории выстраивается соответственно с их временной отметкой. Наличие нескольких версий одного и того же документа указывает на то, что произошли изменения текущей версии документа в соответствующих отрезках времени, то есть текущая версия является актуальной на данный момент. Это может быть документ, созданный МОНМС или подразделением университета. При этом различные версии документа содержат информацию, которая является релевантной для определенного временного интервала.

Добавление новой версии конкретного документа требует создания и заполнения актуальной семантической аннотации, которая загружается в репозиторий вместе с самим документом и сохраняет описание его основных характеристик. В случае размещения новой версии документа весь набор версий остается доступным для просмотра пользователями репозитория. Для возможности отслеживания изменений в содержании документа аннотация содержит ссылку на предыдущую версию документа в репозитории, а также информацию о количестве доступных для просмотра версий. Таким образом, можно возобновить всю цепочку версий документа – от первичной, загруженной в репозиторий, до последней версии, которая по умолчанию считается такой, что содержит актуальную и достоверную информацию.

Представленный механизм управления версиями документов позволяет решать задачи поиска информации на разных этапах ее документирования, сравнения количественных и качественных показателей, отраженных в отчетной документации, сопоставления показателей с нормативной документацией и выявления несоответствий, проверку достоверности предоставленной информации и мониторинг деятельности учебного заведения, организации или структурного подразделения с возможностью проверки его развития.

Достоверность, совместимость и, избыточность представленной информации обеспечивается защищенностью репозитория. Возможность пополнять репозиторий предоставляется ответственным за документацию сотрудникам структурных подразделений согласно с распределением прав доступа к информации портала и репозитория, что позволяет влиять на процесс наполнения репозитория документами. Для защиты репозитория используются технологии

цифровой подписи. Каждый пользователь портала, который имеет право пополнять семантический репозиторий, получает цифровую подпись. Это пара ключей для шифрования, дешифрования и валидации документов, выданная специально для работы в репозитории. Также может использоваться внешний ключ, который использует технологию цифровых подписей на основе стандарта OpenPGP (Pretty Good Privacy) [5]. Закрытый ключ, используемый для шифрования документов, хранится в защищенном месте и имеет пароль для использования. При помощи закрытого ключа происходит подписание документов.

Для быстрого выполнения функций генерации ключей, подписания и верификации документов, предложено два независимых подхода:

- применение программы;
- применение online-сервисов, которые реализованы на базе открытого Java-проекта поддержки цифровой подписи Bouncy Castle [6].

С другой стороны, семантические технологии позволяют осуществлять проверку уже полученной информации путем применения методов сравнения семантических аннотаций документов и метода верификации информации, который состоит в доказательстве достоверности и пояснении гибридного вывода знаний, которые хранятся в БЗ интеллектуальной распределенной системы (онтологического портала) на основе данных, которые извлекаются из распределенных разнородных источников с использованием подходов Semantic Web.

## 6. Интеграция семантического репозитория документов и онтологического портала менеджмента национальных образовательных ресурсов

Назначение семантического репозитория состоит в хранении и обеспечении механизмов менеджмента документов, которые поддерживают решение задач в области образования и науки Украины. Поэтому важным аспектом функционирования репозитория является его интеграция с онтологическим порталом менеджмента национальных образовательных ресурсов. Главным требованием к реализации интеграции семантического репозитория и онтологического портала является необходимость создания такого типа связи систем, который бы раскрывал не только механизмы передачи информации и методы ее обработки, а и смысловую нагрузку такой связи. Разработанный механизм интеграции продемонстрирован на рис. 4.

Во время разработки процедур интеграции был создан и реализован двухсторонний механизм связи между интегрированными системами, который обеспечивает возможность наиболее оптимального использования репозитория при решении задач онтологического портала. Между системами существует два типа связи:

- 1) прямая связь (на основе использования семантической аннотации документа) от семантического репозитория к portalу;
- 2) обратная связь (на основе ссылки на документ репозитория) – от портала к репозиторию.

В свою очередь, прямая связь реализуется одним из двух способов:

- 1) неявная связь – в семантической аннотации документа репозитория формируется ссылка на ресурс онтологического портала, который соответствует заданному документу;
- 2) явная связь – в семантическую аннотацию включается триплет, который содержит ссылку на значение параметра ресурса портала, который соответствует содержанию документа.

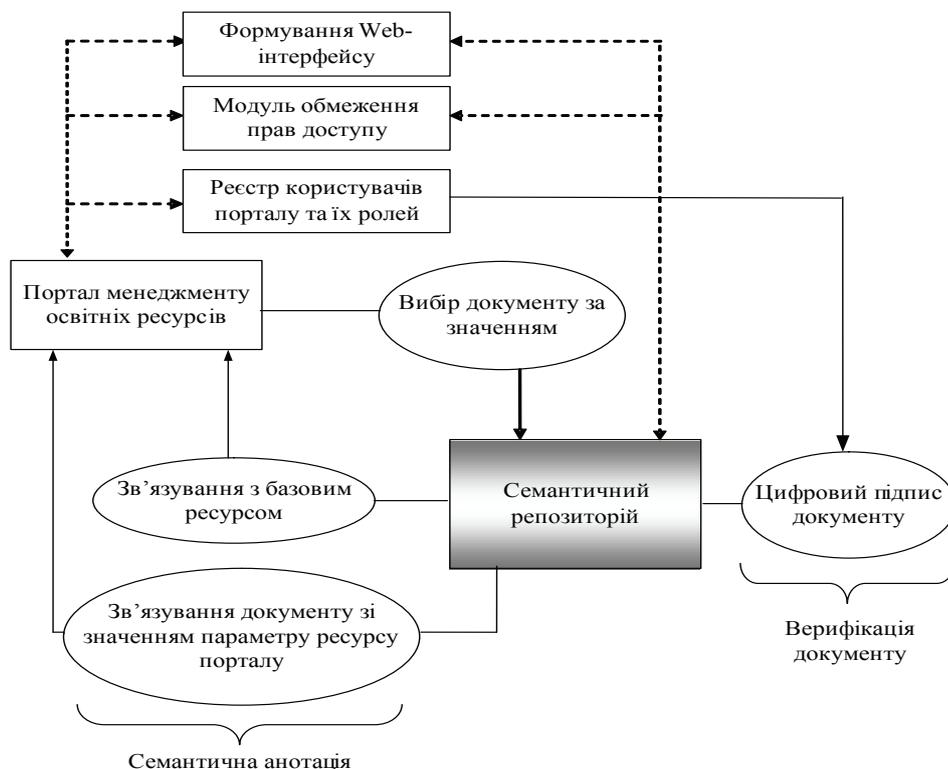


Рис. 4. Схема интеграції семантичного репозиторія документів і онтологічного порталу менеджменту освітніх ресурсів

Таким образом, семантическая аннотация содержит в себе исчерпывающее знание о назначении документа семантического репозитория в рамках заданной предметной области.

Обратная связь обеспечивает реализацию механизмов установления соответствия ресурсов портала к документам репозитория: описание ресурсов портала дополняется ссылкой на соответствующий им документ.

Реализация такого типа интеграции обеспечивает адекватное наполнение репозитория актуальной, необходимой и исчерпывающей информацией на основе онтологической модели предметной области.

Также следует отметить, что к интеграционным процессам между репозиторием и порталом относится процесс верификации документа на основе установления достоверности его цифровой подписи. Это реализовано при помощи связи реестра пользователей портала, который содержит открытый ключ цифровой подписи, с реквизитами документов, среди которых – уникальная цифровая подпись соответствующего лица.

### Выводы

В работе рассмотрено понятие семантического репозитория в современных системах менеджмента знаний, приведено описание прототипа семантического репозитория документов с возможностью контроля версий. Предложен прототип распределенной версии «Онтологического портала менеджмента образовательных и научных ресурсов МОН Украины» с интегрированной системой хранения образовательных документов в защищенном семантическом репозитории, при помощи которого реализованы механизмы загрузки документов в репозиторий с возможностью контроля версий, механизмами просмотра всех существующих документов семантического репозитория согласно заданным параметрам просмотра и прав доступа к информации. Эти функции реализованы отдельными модулями системы, которые связаны с моделями распределения прав доступа к защищенным файловым хранилищам.

### Литература

1. Sullivan D. Document Warehousing and Text Mining. Techniques for Improving Business Operations, Marketing and Sales / D. Sullivan. – Canada: John Wiley & Sons, Inc., 2001. – 542 p.
2. Resource Description Framework: Overview. [Электронный ресурс] // Режим доступа: <http://www.w3.org/RDF/>.
3. Рябова Н. Застосування онтологічної семантики у зіставленні документів / Н. Рябова, Г. Козопольська, К. Дяденко // Матеріали III Міжн. конф. молодих вчених “Комп’ютерні науки та інженерія” CSE-2009, 14-16 травня, 2009, Львів. – Львів: НУ „Львівська політехніка”, 2009. – с. 107-108.
4. Волошина Н.А. Онтологический подход к построению хранилищ текстовых документов в системах поддержки принятия решений / Н.А. Волошина, А.А. Козопольская, Н.В. Рябова, О.В. Шубкина. // Праці V Міжн. школи-семінару «Теорія прийняття рішень». – Ужгород, УжНУ, 2010. – с. 46-47.
5. Secure E-Mail and File Encryption using GnuPG. [Электронный ресурс] // Режим доступа: <http://www.gpg4win.org/>.
6. The Legion of the Bouncy Castle. [Электронный ресурс] // Режим доступа: <http://www.bouncycastle.org/>.