

УДК 004.633

НЕВИКОРИСТАНИЙ ПОТЕНЦІАЛ ТЕХНОЛОГІЇ XML ДЛЯ ЦІЛЕЙ ЕЛЕКТРОННОГО ДОКУМЕНТООБІГУ В СФЕРІ ЗЕМЛЕУСТРОЮ

Р. Пересоляк

Ужгородський національний університет

Ключові слова: XML, кадастровий файл обміну, мова розмітки.

Мета. Розроблення мови розмітки, виконаної засобами XML технології, та розроблення нового формату кадастрового файла обміну на її основі. **Методика.** Застосування технології XML, зокрема XMLSchema, для визначення мови розмітки, та аналіз файла-схеми, доступного через програмне забезпечення Digitals. **Результати.** Розроблено мову розмітки, котра отримала назву UCML, із урахуванням недоліків, які визначено під час аналізу сучасного стану кадастрового файла обміну. Мова розмітки отримала розширений функціонал для роботи із даними, наприклад, механізм збереження графічної інформації, що також передбачено як основний формат для обмінних файлів інших кадастрів. Визначає централізований доступ до структур даних, котрі використовують для опису інформації в інших кадастрах. **Наукова новизна.** Уперше: розроблено мову розмітки кадастрових файлів обміну для локального застосування, чітко визначено ієрархію та типи даних, які можуть використовуватись у різних сферах обліку даних для передавання їх у вигляді електронних документів. Визначено механізм внесення змін до структури та змісту кадастрового файла обміну. **Практична значущість.** Отримані результати дають змогу покращити якість кадастрових файлів обміну та кадастрового інформаційного наповнення, а також повністю перейти на новий формат кадастрових файлів. Цей підхід дає змогу скласти вичерпний список даних та їх структур, котрі можуть міститись у межах всіх кадастрів, а також використовувати ці структури, визначені в межах одного кадастру, для інших кадастрів.

Вступ

Вже шість років поспіль на території України функціонує та використовується оновлений кадастровий файл обміну, який оснований на технології XML та який, своєю чергою, замінив кадастровий файл обміну попереднього зразка – In4(info). Очікували, що використання XML приведе до стрімкого поліпшення якості документа, його наповненості, простоти використання та функціональної зручності. Розпочинаючи дослідження, автор передбачав запропонувати та внести низку правок, котрі б стосувались

часткового вдосконалення структури кадастрового файла обміну, однак, водночас, не змінювали б його істотно. Однак виконаний аналіз стану кадастрового файла обміну показав низку структурних та модельних проблем, після вирішення яких отримали істотно інший формат кадастрового файла обміну. Низький рівень якості даних, котрі зберігаються засобами наявного кадастрового файла обміну, частково підтверджується й іншими джерелами.

З огляду на особливе місце земельного кадастру серед інших кадастрів, він є основою для ведення таких кадастрів:

- кадастру водних ресурсів;
- кадастру природних рослинних ресурсів;
- кадастру ресурсів тваринного світу;
- кадастру природних лікувальних ресурсів;
- кадастру мінерально-сировинних ресурсів, корисних копалин родовищ, проявів, а також корисних копалин техногенних родовищ;
- кадастр лісових ресурсів;
- кадастр природних територій курортів;
- кадастр територій та об'єктів природно-заповідного фонду;
- містобудівний кадастр;
- кадастр антропогенних викидів та абсорбції парникових газів.

З огляду на це, кадастровий файл обміну земельного кадастру також посідає особливе місце серед кадастрових файлів обміну інших кадастрів. Тому вимоги, котрі стосуються питань формування, наповнення та базової функціональності кадастрового файла обміну, повинні бути високими та потребують додаткового дослідження.

Проведення детального аналізу наявного кадастрового файла обміну виявило такі недоліки:

- Особливості середовища(Digitals)
 - 101 (елемент наявний, коли є хоча б один елемент першого рівня).
- Дані не стосуються особи, що формує файл:
 - 102 (непідвідомча інформація).
 - 103 (метаінформація).
- Показчик на застосування атрибута
 - 104 (назва блока залежить від опції середовища, за допомогою якого сформовано документ).

- 126 (опція середовища замінює блок, наявний у документі).
- Показчик групування:
 - 121 (блок є елементом списку).
- Показчик на зміну найменування елемента
 - 105 (однакові назви для блоків різного призначення).
 - 108 (найменування блока не відповідає його змісту).
 - 109 (найменування блока занадто спрощене).
 - 110 (найменування не відповідає єдиному стилю найменувань).
 - 111 (найменування містить транслітерацію).
 - 113 (найменування містить назва батьківського блока).
- Невідповідність нормативно-правовому акту:
 - 107 (найменування не відповідає назві у нормативно-правовому документі).
- Показчик на комплексний тип:
 - 114 (блок та його вміст неодноразово повторюються у документі).
 - 106 (вміст блока неодноразово повторюється незалежно від назви блока).
 - 119 (блок містить ознаки типового елемента в незмінній формі).
- Додаткова характеристика елемента:
 - 112 (блок є посиланням на інший об'єкт).
 - 116 (вміст блока залежить від опції середовища, за допомогою якого сформовано документ).
 - 122 (блок містить вичерпний список значень).
 - 123 (значення блока представлене кодом).
 - 124 (значення блока містить пробіли).
 - 125 (блок може бути представлений списком значень).
- Ускладнення документа:
 - 118 (надлишковий елемент).
 - 129 (відсутність альтернатив).

Варто зазначити, що ця класифікація не ставить за мету висвітлити лише негативні моменти формування кадастрового файлу обміну, до класифікації також зараховано особливості формування, котрі дають змогу краще описати особливості внутрішньої будови обмінного файлу за допомогою технології XML.

Як результат, виявлено:

- істотну залежність від середовища формування кадастрового файлу обміну;
- відсутність чіткого механізму внесення змін у структуру кадастрового файлу обміну в разі зміни нормативно-правового регулювання;

- слабка типізація наявних у будові об'єктів;
- відсутність шаблонного підходу до схожих за значенням об'єктів;
- надлишкове ускладнення внутрішньої будови;
- слабке обмеження поля дій із документом.

Також варто згадати про те, що XML являє собою цілу сім'ю технологій, а не лише спосіб опису даних (рис. 1). Тому реалізація будови кадастрового файлу обміну повинна передбачити можливість залучення інструментарію XML у повному обсязі.

Ще однією особливістю кадастрового файлу обміну та нормативно-правової системи, котра його підтримує, є відсутність механізму контролю будови кадастрового файлу обміну як на практичному, так і на нормативному рівні. Це чудово простежується у разі переходу від ліцензій на виконання землепорядних робіт до сертифікатів землепорядників. На нормативному рівні перехід відбувся, як і на практичному, – будову кадастрового файлу змінили, однак нормативно-правові документи, котрі визначають та фіксують зміст кадастрового файлу, залишились незмінними. Ця прогалина унеможливує відстеження змін кадастрового файлу обміну та доступ сторонніх розробників до деталей як архівних, так і актуальних версій кадастрового файлу обміну. А це призводить до неявного домінування одних програмних продуктів над іншими, яке, своєю чергою, підтримує законодавство. Як наслідок, цей підхід, на нашу думку, призводить до погіршення якості кадастрового файлу обміну, його наповненості та чіткості, а також унеможливує опрацювання попередньої версії кадастрового файлу обміну без попереднього їх аналізу.

З огляду на виявлену кількість проблем, незважаючи на початкову мету покращення та вдосконалення структури кадастрового файлу обміну, прийнято рішення розробити абсолютно новий формат кадастрового файлу обміну, перехід на який можна було розділити на декілька етапів. Розпочинаючи роботу, ми поставили низку базових вимог, котрі новий кадастровий файл обміну повинен задовольнити, а саме:

- кадастровий файл обміну є системним, тому позитивні та негативні його прояви впливають та повинні впливати на систему кадастрів України;
- всі структури даних повинні бути перебудовані та чітко систематизовані;
- необхідний чіткий механізм внесення змін у будову кадастрового файлу обміну.

Модель нової мови розмітки кадастрових файлів обміну побудована на декількох базових елементах:

- всі атрибути – це об'єкти;
- кадастрові файли обміну всіх кадастрів побудовані на основі мови розмітки кадастрових файлів обміну.

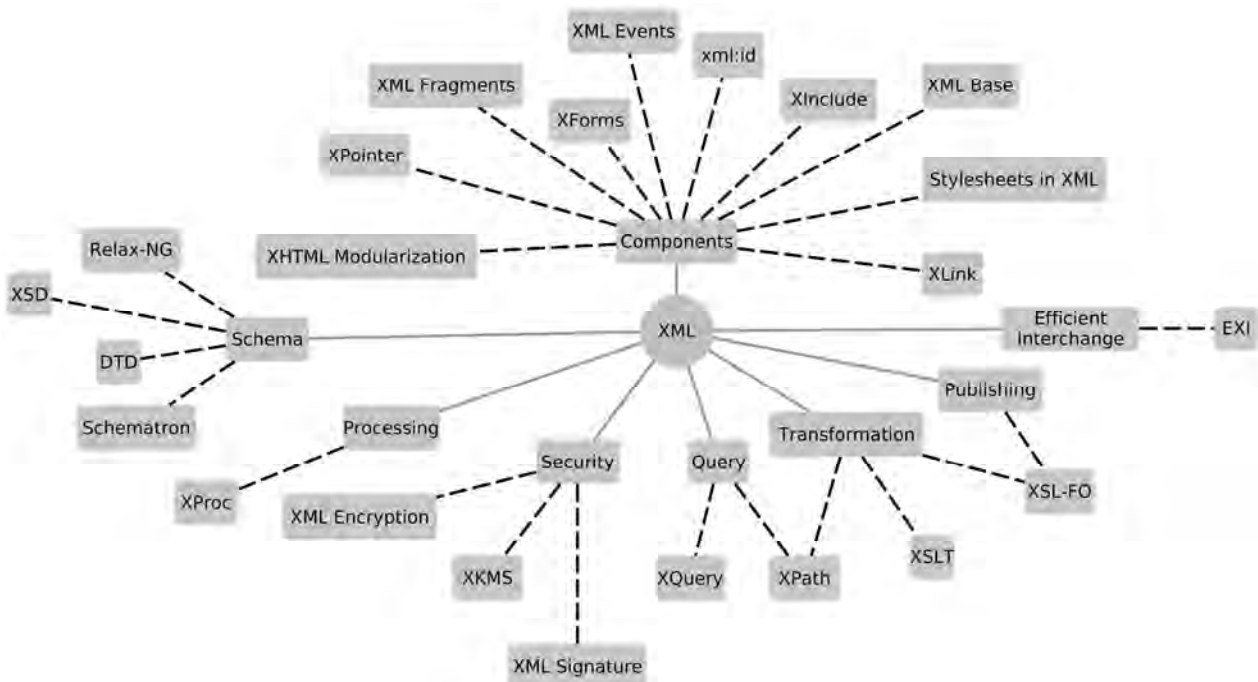


Рис. 1. Інструментарій XM

Основна частина

Важливою особливістю нового підходу є використання XMLSchema 1.1, на відміну від XMLSchema 1.0, яку використовують сьогодні під час формування кадастрового файлу обміну. До цих груп належать:

- об'єкти;
- шаблони файлів.

Об'єктна група (рис. 2) складається із двох файлів "data_types.xsd" та "geometry.xsd". Перший відповідає за визначення усіх об'єктів мови розмітки, котрі можуть бути розподілені між вісьмома кадастрами системи кадастрів України, а другий файл відповідає

за визначення типів геометрії, які підтримує мова кадастрових файлів.

Група шаблонів файлів відповідає за збирання усіх типових об'єктів та геометрії у новий композитний об'єкт. З огляду на різноманіття кількостей можливих композицій та на різноманіття кадастрів та їх обмінних файлів, цей тип об'єктів виділений окремою групою файлів формату XSD. Кількість цих файлів визначається кількістю необхідних кадастрових файлів обміну. На період написання роботи визначено лише один шаблонний файл, котрий відповідає за форму та зміст кадастрового файлу обміну земельного кадастру.

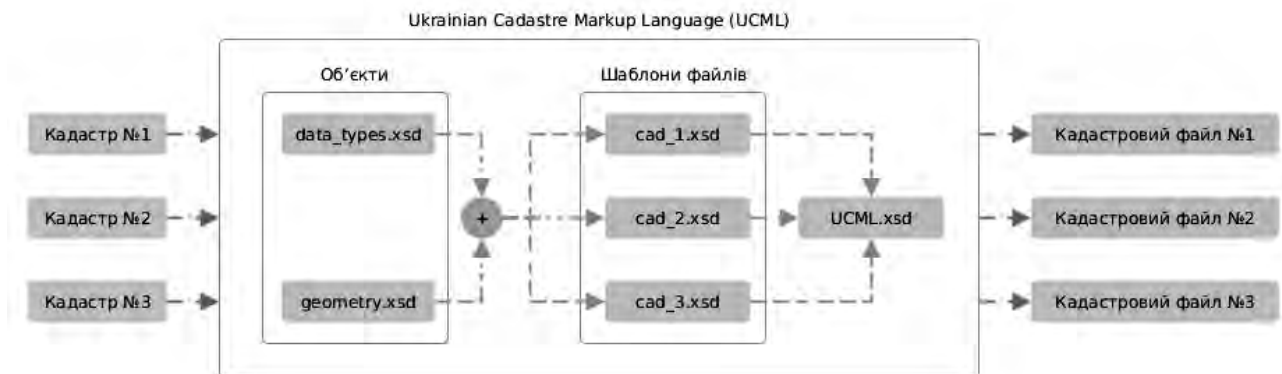


Рис. 2. Схема UCML

Також на рис. 2 можна побачити, що, хоч UCML складається з великої кількості файлів, визначено єдину точку доступу до всіх об'єктів мови – файл "UCML.xsd". Це дає змогу, використовуючи мову, вказати тільки на один файл, котрий і реалізує всі можливості мови розмітки.

Структури даних, котрі трапляються більше ніж один раз, виділено в окремі об'єкти. UCML визначає такі види об'єктів:

- статичні об'єкти (об'єкти з чітко визначеною будовою, однак вони можуть містити в собі інші типи об'єктів, зокрема змінні).

- Address;
- CadastralNumber;
- CertificatedPerson;
- CheckingAction;
- Comment;
- Contacts;
- Country;
- Duration;
- Image;
- LandParcel;
- ProjectExecutor;
- Registration;
- RegulatoryAct;
- змінні об'єкти (об'єкти, котрі змінюють свою будову та наповненість залежно від критеріїв опису об'єкта, визначених декількома XML атрибутами):
 - Document;
 - LandProject;
 - LandUnit;
 - Name;
 - Payment;
 - Subject;
- об'єкти-списки (об'єкти, котрі організують у списки декілька однотипних об'єктів):
 - ApprovalList;
 - CertificatedPersonsList;
 - ChackingActionsList;
 - ImagesList;
 - DocumentsList;
 - LandParcelsList;
 - LandUnitsList;
 - PrivilegesList;
 - SubjectsList;
- об'єкти-списки посилань (об'єкти, котрі визначають списки посилань на інші об'єкти документа):
 - DlinksList;
 - GlinksList.

Яскравою демонстрацією відмінностей у внутрішній організації є ознака того, що в нинішній реалізації декілька різних типів структур представляють один, однак модифікований на різних рівнях об'єкт. До таких представників можна зарахувати:

- Autentification (інформація про власника земельної ділянки);
- Beneficiary (особа, на користь якої встановлено обмеження);
- Chief (дані про відповідальну особу);
- DeliveryPerson (відомості про особу, яка отримала державний акт);
- Leasees (блок опису усіх орендарів);
- LegalEntityInfo (дані про юридичну особу);
- NaturalPersonInfo (дані про фізичну особу);
- Proprietor (власник або розпорядник);
- Subleases (блок опису усіх суборендарів).

Як відомо, будь-який XML документ, з огляду на сегментацію та внутрішню організацію інформації, можна розглядати як представника об'єктно-орієнтованих баз даних. Однак реалізація доступу до необхідних даних, а також зручність та простота пошуку, хоч й реалізовані самостійно XML технологією, однак повністю ґрунтуються на внутрішній розмітці даних. Тобто, інакше кажучи, внутрішня будова будь-якого XML документа визначає коло можливих запитів до структурованої інформації. Всі зазначені вище елементи є елементами одного типу (особи або суб'єкта), однак для кращого виділення ролі їх помістили у різні вузли та ще й, користуючись термінологією XML, частину визначили як тип, а іншу частину як іменованний екземпляр типу. Це призвело до того, що їх усіх неможливо ідентифікувати у межах одного простого запиту такого вигляду: “Знайти всіх осіб, як юридичних, так і фізичних, котрі тією чи іншою мірою були залучені до формування певних робіт із землеустрою”.

З огляду на це дизайн кадастрового файлу, що реалізований на XML, потребує додаткового дослідження та модифікації. Яскравим представником, котрий абсолютно відрізняє наявну схему кадастрового файлу обміну від пропонованого нового дизайну кадастрового файлу обміну, є елемент електронного документа, котрий містить інформацію про суб'єктів стосовно долученості тією чи іншою мірою до формування земельпорядної документації.

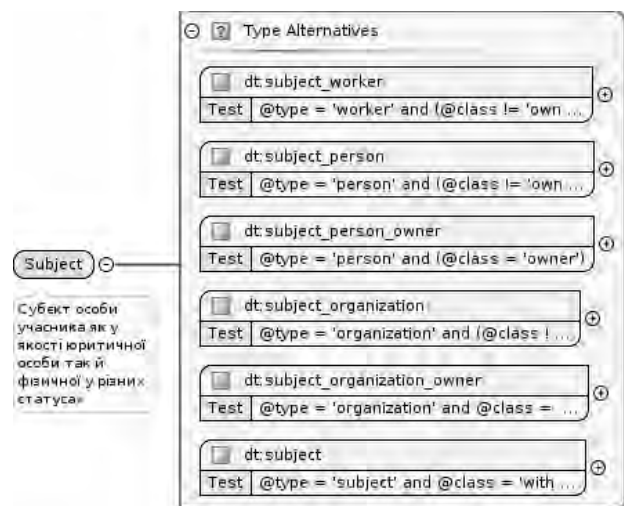


Рис. 3. Клас суб'єкта

Тому UCML пропонує новий підхід до вирішення цього питання. Об'єкт типу “Subject” відповідає за збереження інформації щодо будь-яких осіб, причетних до створення, формування чи коригування того чи іншого виду документації. Незважаючи на аналоги, які пропонує наявна версія кадастрового файлу обміну, цей об'єкт об'єднує різні типи структур даних, які описують відомості щодо осіб чи організацій. Також варто зауважити, що цей об'єкт разом із наявними реалізовує нові типи суб'єктів,

відсутні в наявній версії кадастрового файлу, зокрема беручи до уваги те, що відомості щодо осіб, котрі погоджують документацію, також повинні зберігатись засобами кадастрового файлу обміну.

“Subject” оснований на базовому типі, який у цьому випадку представлений об’єктом “subject” (рис. 4).

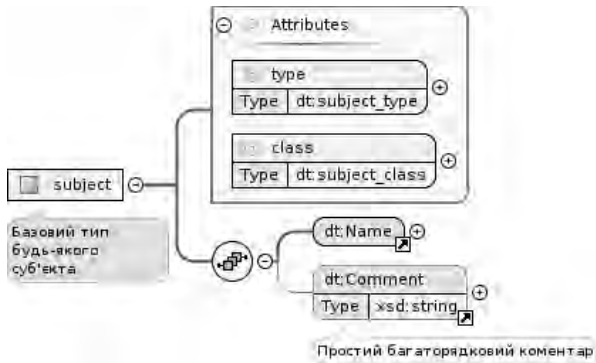


Рис. 4. Базовий тип суб’єкта

Як бачимо на рис. 3, базовий об’єкт складається із двох XML-елементів, що зберігають відомості про ім’я суб’єкта, незалежно від того, фізична це особа чи юридична, та необов’язковий коментар. Також, з огляду на базовість, цей об’єкт визначає обов’язкові XML-атрибути “клас” та “тип”, які відповідають за керування об’єктом “Subject”. Як бачимо на схематичному зображенні об’єкта “subject”, “клас” та “тип” визначені спеціальними типами даних, котрі містять вичерпний перелік можливих значень, та значення яких обмежені у всіх похідних від базового класу класів.

Об’єкт “subject_organization” є похідним об’єктом “subject” та розширює його структуру додатковими XML-елементами для збереження адреси “Address” і списку посилань на документи “DLinksList”. Список допустимих значень для “тип” додатково обмежений лише одним значенням – “організація”. Варто зауважити, що кінцевий варіант значень XML-атрибутів для усіх альтернативних значень визначено в “Subject”, однак для жорсткішої структуризації, а також можливості чітко визначати список можливих документів додано таке обмеження.

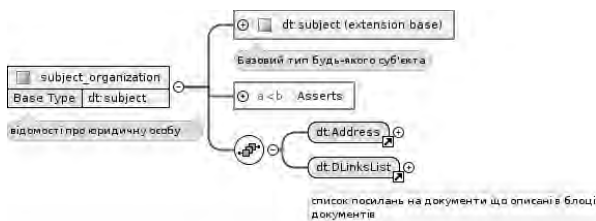


Рис. 5. Суб’єкт організації

Об’єкт “subject_organization_owner” є розширенням класу “subject_organization” та додає два нові XML-елементи “Part” та “PrivilegesList”.

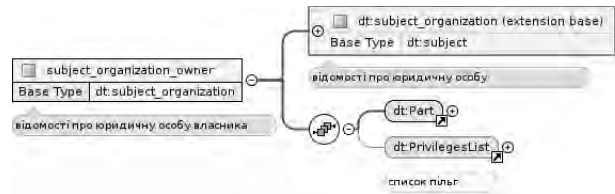


Рис. 6. Суб’єкт організації власника

Об’єкт на рис. 7 є розширенням базового об’єкта “subject” та аналогом “subject_organization” для фізичної особи, відповідно містить такий самий набір додаткових елементів.

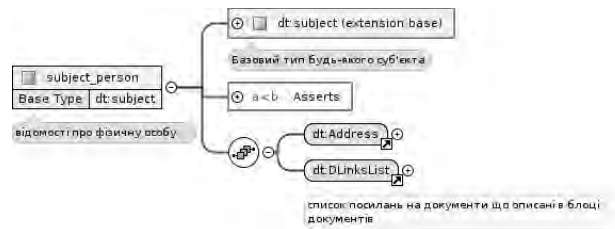


Рис. 7. Суб’єкт фізичної особи

Об’єкт на рис. 8 є розширенням базового об’єкта “subject_person” та аналогом “subject_organization_owner” для фізичної особи.

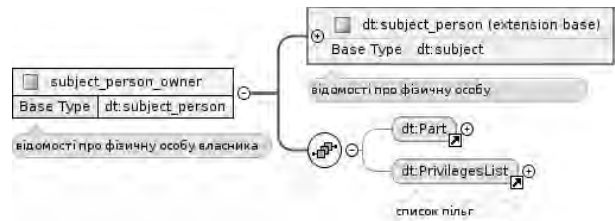


Рис. 8. Суб’єкт фізичної особи – власника

Об’єкт типу “subject_worker” є розширенням базового класу та доповнює його XML-елементом “Position” для збереження відомостей про посаду особи. Варто зазначити, що цей об’єкт має обмеження на застосування значень атрибутів базового типу та передбачає використання лише значення “person” для “type”. Це забезпечує захист від використання у цьому об’єкті назв для юридичних осіб організацій.

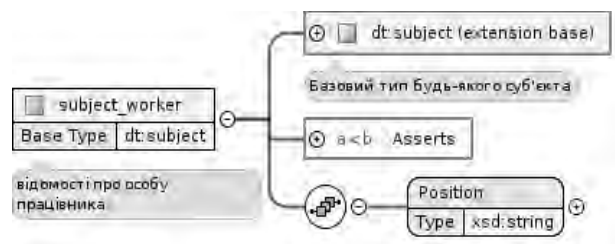


Рис. 9. Суб’єкт працівника

Висновки

Використання технології XML як для земельного кадастру, так і для інших кадастрів – перший крок до

ефективного та якісного використання даних. Однак для реалізації можливостей на якісно іншому рівні підхід, взятий за основу для електронного документообігу в землеустрої, потребує змін. Зокрема, запропоновано використання нового підходу для формування кадастрових файлів обміну, перехід на нову мову розмітки UCML та обов'язкове визначення у нормативному полі механізму публікації змін структури кадастрового файла обміну, за допомогою фіксації процедури публікування шаблону кадастрового файла обміну (в разі використання UCML – схеми мови розмітки).

Література

1. Шоломицкий А. А. Использование формальных грамматик для анализа кадастровой информации / А. А. Шоломицкий, М. Л. Лепеш // In: 2013. – С. 10–18.
2. David Mertz. XML Matters: Putting XML in context with hierarchical, relational, and object-oriented models. 2001. С. 7. URL: <https://www.ibm.com/developerworks/library/x-matters8/x-matters8-pdf.pdf>.
3. Ronald Bourret. XML and Databases [Електронний ресурс]. 2005. URL: <http://www.rpbouret.com/xml/XMLAndDatabases.htm>
4. Tim Bray et al. Extensible Markup Language (XML). 1.1 (Second Edition). [Електронний ресурс]. 2006. С. 45. URL: <http://www.w3pdf.com/W3cSpec/XML/2/REC-xml11-20060816.pdf>.
5. XML technology. [Електронний ресурс]. URL: <https://www.w3.org/standards/xml/>.

Невикористаний потенціал технології XML для цілей електронного документообігу в сфері землеустрою Р. Пересоляк

Основна мета статті – розроблення мови розмітки та формату кадастрового файла обміну на її основі. В результаті було розроблено мову UCML. Отримані результати покращують якість кадастрових файлів обміну.

Нереалізований потенціал технологій XML для цілей електронних документів в області управління земельними ресурсами Р. Пересоляк

Цель. Разработка языка разметки, выполненной средствами XML технологии, и нового формата кадастрового файла обмена на ее основе. Методика. Применение технологии XML, в частности XMLSchema, для определения языка разметки и анализ существующего файла-схемы, доступного через программное обеспечение Digitals. Результаты. Разработан язык разметки, которая получила название UCML, с учетом недостатков,

выявленных при анализе современного состояния кадастрового файла обмена. Язык разметки получил расширенный функционал по работе с данными, к примеру, механизм сохранения графической информации, а также предполагается как основной формат для обменных файлов других кадастров. Определяет централизованный доступ к структурам данных, который используют для описания информации другими кадастрами. Научная новизна. Впервые разработан язык разметки кадастровых файлов обмена для локального применения. Впервые четко определены типы данных и их иерархия, которые могут использоваться в различных сферах учета данных для передачи их в виде электронных документов. Определен механизм внесения изменений в структуру и содержание кадастрового файла обмена. Практическая значимость. Полученные результаты позволяют улучшить качество кадастровых файлов обмена и кадастрового информационного наполнения, а также полностью перейти на новый формат кадастровых файлов. Этот подход позволяет определить исчерпывающий список данных и их структур, которые могут содержаться в пределах всех кадастров, а также использовать эти структуры, определенные в пределах одного кадастра, другими кадастрами.

Lack of the XML technology potential for the purpose of the electronic documentary cooperation in the field of the earthquake

R. Peresoliak

Goal. Development of markup language implemented by XML technology and development of a new cadastral file format on the basis of it. Method. The use of XML technology, in particular XMLSchema, to determine the markup language and to analyze the embedded file schema available through the Digitals software. Results The markup language, called UCML, was developed, taking into account the disadvantages of analyzing the current state of the cadastral file of exchange. The markup language has an advanced data functionality, for example, the mechanism for storing graphical information, and also provides as the main format for the exchange files of other cadastres. Defines centralized access to the data structures of the court used to describe the information of other cadastres. Scientific novelty. For the first time a cadastral file sharing markup language for local application was developed. For the first time, the types of data and their hierarchy that can be used in different areas of data accounting for transmission in the form of electronic documents are clearly defined. The mechanism of making changes to the structure and content of the cadastral file of exchange is determined. Practical significance. The obtained results allow to improve the quality of cadastral file sharing and cadastral information content, as well as completely switch to the new cadastral file format. This approach allows you to define an exhaustive list of data and their structures that may be contained within all inventories, and also use these structures defined within the same inventory by other cadastres.