

УДК 004.91

В.В. Карасюк, Н.А. Кошечая

Национальный юридический университет имени Ярослава Мудрого

## АНАЛИЗ СПОСОБА ПРЕДСТАВЛЕНИЯ ПРАВОВОЙ ИНФОРМАЦИИ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

*Рассматривается язык разметки GJXDM для представления правовых данных. Проведен анализ особенностей правовой информации и информационных систем в юриспруденции. Обращено внимание на возможность XML в описании структурированных данных. Сделан обзор основных подходов для описания правовых данных в GJXDM. Сформулирован вывод о перспективности применения GJXDM в украинских правовых информационных системах.*

**Ключевые слова:** правовая информация, формат правовых документов, модель данных, XML, GJXDM.

**Введение.** Юристы в основном работают с документами. Они используют различные документы, относящиеся к определенным юридическим ситуациям: нормативные документы, подзаконные акты, правоприменительные документы, решения, текущие правовые документы и т.п. Объем циркулирующих в правовой сфере документов необычайно велик. Как правило, каждый из таких документов является независимым от других. Однако для описания любой правовой ситуации используется группа документов. Целесообразно, чтобы они имели между собой информационные связи. К тому же большая часть правовых документов имеет четко определенную структуру, которая определена нормативными документами. Эта структура, точнее говоря, стандартные поля документа могут быть связующим звеном для создания системы связанных документов, описывающих конкретную юридическую ситуацию.

Современные информационные технологии давно имеют в своем арсенале средства, позволяющие строить информационные системы со связанным контентом. Из средств организации современных информационных систем с функциями навигации и поиска на основе web-представления документов наиболее перспективным является использование возможностей XML. Опираясь на существующую стандартизацию правовых документов, можно разработать подмножество XML, ориентированное на особенности этих документов. Таким средством является модель GJXDM – Global Justice XML Data Model, которая начинает применяться на практике. Задача разработки и внедрения подобного механизма представления правовых данных является весьма актуальной с учетом существующих и перспективных объемов правовой информации и задач, стоящих перед системой правового обеспечения. С учетом имеющихся планов создания

в стране информационного (электронного) правительства и, в целом, информационного общества, проблема информационно-правовой поддержки этих процессов становится приоритетной.

**Постановка проблемы.** В сфере правовой деятельности наблюдается стойкая тенденция увеличения числа циркулирующих документов. Это множество документов имеет недостаточные связи между собой, что затрудняет практическую юридическую деятельность и поиск релевантных конкретным юридическим ситуациям документов. Также затрудняется процедура обмена текущей правовой информацией между различными правовыми структурами. Возникает задача разработки средств (или адаптации, имеющихся в других сферах) для создания структур связанных правовых документов, доступных множеству пользователей. То есть, в первую очередь необходимо провести анализ возможностей построения информационных систем правовых документов на основе средств стандартной разметки и структуризации документов, что обеспечит реализацию эффективного поиска и доступа к содержанию множества, связанных с конкретной проблемой документов.

**Цель статьи.** Анализ возможностей описания правовой информации средствами модели GJXDM, что обеспечит связывание документов в единую систему. Решение этой задачи предусматривает:

рассмотрение особенностей правовой информации и ее представления в существующих информационных системах;

анализ возможностей стандарта GJXDM по описанию правовой информации.

### **Чем правовая информация отличается от других видов информации?**

Специфика правовой информации заключается в том, что независимо от своего содержания, она

всегда обладает определенной социальной значимостью. Правовая информация регулирует отдельные стороны хозяйственной и социально-культурной деятельности общества, что и определяет ее особый, прагматичный характер. Среди правовых задач стоит в первую очередь отметить задачи, решаемые при осуществлении правотворческой, правоприменительной и правоохранительной деятельности.

В соответствии с Законом Украины «Об информации», информация – это документированные или публично объявленные сведения о событиях и явлениях, происходящих в обществе, государстве и окружающей природной среде. Как один из видов информации этот Закон выделяет правовую информацию – совокупность документированных или публично оглашенных сведений о праве, его системе, источниках, реализации, юридических фактах, правоотношениях, правопорядке, правонарушениях и борьбе с ними и их профилактике и т.п.

Правовая информация входит, как правило, в состав государственных информационных ресурсов. Государственные информационные ресурсы должны содержаться в форме и в состоянии, позволяющем свободно осуществлять доступ к находящейся в них информации. Так должно быть. Однако на практике мы видим достаточно проблем, которые препятствуют простому доступу к правовой информации.

Правовой информации присущи некоторые особенности [1]:

большой объем, например, действующая нормативная база Украины (указы Президента, законы и постановления Верховной Рады, распоряжения Кабинета Министров и другие) объединяет около 500 тысяч документов. Объем информации, которую необходимо использовать в процессе эффективной юридической деятельности, настолько велик и разнообразен, что даже опытные юристы испытывают затруднения при поиске необходимой информации для принятия правильных решений и тратят на это значительное время;

эта информация в основном представлена в виде неструктурированной текстовой информации, которую трудно формализовать, а в качестве элементов структуризации этих данных используется всего лишь механизм гиперссылок для связывания различных версий документов и их частей;

эта информация должна быть достоверной, то есть строго соответствовать действующему законодательству и правоприменительным документам;

должна быть своевременной – вовремя поступать

субъекту юридической деятельности, иначе она потеряет свое практическое значение;

имеет временные ограничения своей легальности; расположена в отдельных базах, данных, хранилищах, локальных сетях и для каждого ее вида применяется, к сожалению, свой формат представления.

Правовая информация является основой функционирования правовых систем. В целом юридическая деятельность реализуется через совокупность систем государственно-правового характера. Такие системы имеют сложную организацию, состоят из комплекса подсистем, каждая из которых выполняет определенную функцию. Часть подсистем выполняет функции управления (функциональные системы управления); часть – функции информационного обеспечения (подсистемы информационного обеспечения). Основной по объему в этих системах обращается правовая информация, которая циркулирует в них согласно своим функциям. При этом следует иметь в виду, что правовая информация используется не только в существующих структурах юстиции, она должна быть доступной для всех субъектов права, любой формы собственности и всех граждан страны.

В отличие от многих других видов информации (Закон Украины «Об информации» [2]) правовая информация нуждается в постоянном соотношении с точными формулировками законодательных норм, то есть функции поиска при работе с нею стоят на первом месте. Сегодня безопасный, надежный и своевременный обмен информацией разной юрисдикции признается как наиболее важный и фундаментальный вопрос для функционирования всей правовой системы, в том числе безопасности страны. Многие интерфейсы были разработаны исключительно для функций обмена информацией внутри системы, существуют в собственном формате, и не в состоянии быть повторно использованными в какой-либо другой системе.

Юридическая обработка на  $\frac{3}{4}$  – это выявление взаимосвязей между документами и фиксирование выявленных связей с помощью определенных форм (ссылок, примечаний, справочных сведений), а также создание новых редакций документов при их изменении. Юридическая обработка документов включает следующие этапы:

- классификацию (рубрикацию) документа, ввод ключевых слов,
- выявление взаимосвязей документов,
- формирование перекрестных ссылок между документами,
- составление примечаний, справочных сведений к документу,
- подготовку новой редакции документа при

издании официальных изменений.

Взаимосвязи документов оформляются в виде ссылок, которые можно разделить на два основных вида: смысловые и формальные. Оба вида ссылок являются гипертекстовыми. Это означает, что в тексте документа они выделены цветом и при необходимости можно перейти по этой ссылке в текст соответствующего документа.

Информационно-поисковые языки, получившие название дескрипторных, основаны на применении принципов координатного индексирования, при котором смысловое содержание документа может быть с определенной степенью точности и полноты задано списком ключевых слов, содержащихся в тексте. Дескрипторные языки привязаны к лексике текстов. Ключевые слова из текстов выбираются исходя из разных целей. Для построения дескрипторного языка критерием отбора ключевых слов, как правило, служат информативность слова и частота его встречаемости в тексте.

Большая часть государственных структур для опубликования на сайте правовых документов использовала формат графических баз. Объясняется это достаточно просто. Формат графических образов оказался наиболее приемлемым для создания баз на ведомственных сайтах. Ведущие компании-разработчики правовых систем (во многих странах) содействовали в разработке и поддержании в основном именно графических баз правовых документов, так как именно этот формат позволял оперативно, без ущерба достоверности, размещать правовые акты в базе данных. Эволюционно сложилась ситуация, когда на ведомственных сайтах правовая информация по большей части представлена в графике. В то же время некоторые ведомства по-прежнему используют иные, не графические, форматы баз данных. Однако ни один из таких форматов не обеспечивает все свойства правовой информации. Так, текстовый формат не может обеспечить 100% достоверности, аутентичности документа. Выходом из данной ситуации могло бы стать использование электронно-цифровой подписи (ЭЦП), но на данный момент ЭЦП не получила широкого распространения. В первую очередь это связано с отсутствием полноценной технической базы, позволяющей повсеместно, в любом регионе страны осуществлять использование и идентификацию ЭЦП; существенными финансовыми затратами для использования ЭЦП; проблемами с использованием в правоприменительной практике документов с ЭЦП. Все это ограничивает использование электронно-цифровой подписи, и потому размещение правовой информации на сайтах с ее применением не дает возможности говорить об

обеспечении доступности правовой информации для всех граждан и организаций.

В графических базах, данных документы размещаются в виде копий, получаемых путем сканирования с оригиналов документов. Такие графические копии являются точной копией бумажного аналога документа, что позволяет их широко использовать в правоприменительной практике. Кроме того, обеспечение доступа к документам в графическом формате не требует каких-либо дополнительных программных или аппаратных средств, как при использовании документов с ЭЦП. Использование графического формата для размещения правовой информации на сайтах ведомств позволяет организовать полноценный свободный доступ для любых заинтересованных организаций и граждан к правовым актам. В 2010 г. одним из реализованных в Соединенных Штатах Америки проектов в области развития правовой информатизации стало создание федеральной цифровой системы (FDsys) (<http://www.gpo.gov/fdsys/>), доступ к которой осуществляется на бесплатной основе. Данная усовершенствованная цифровая система заменила действовавшую правовую систему Правительственной службы печати GPO Access, в которой до 2011 года размещались правовые издания и тексты актов федерального правительства. Система FDsys содержит правовые издания и документы федерального правительства, включая текстовые, графические, видео-, аудио-, цифровые и другие формы информационного содержания.

Однако в таких документах невозможно осуществлять контекстный поиск и связь между документами на основе их ключевых понятий. Возникает проблема структурирования и описания правовой информации таким образом, чтобы обеспечить эффективное исполнение функций создания, обработки, поиска в соответствии с этапами жизненного цикла правовых документов

#### **Возможности XML в описании структурированных данных**

XML начал вызывать огромный интерес и поддержку с того момента, как о нем было впервые заявлено в 1996 году. Он обеспечивает стандартный способ кодирования содержания, обеспечивая гибкий способ создания структур данных. В XML для разметки содержания на основе правил, составляемых разработчиком документа, используются теги. Этот набор правил называется описанием типа документа (Document Type Definition, DTD), и он позволяет разработчикам, применяющим XML, размечать различные документы. Например: стандартные документы, которые могут содержать текст и ссылки на графические изображения и внешние ресурсы;

более структурированные документы или записи, такие как формы HTML, со структурой, реально закодированной непосредственно в документе (заказ покупки, медицинское назначение или адресная книга); объект, содержащий данные и методы, например, объект Java или элемент ActiveX; записи баз данных, которые могут быть представлены на Web-странице, на основе запросов пользователей к поисковым средствам баз данных; что-либо, обеспечивающее обмен данными между компьютерами, между компьютерами и людьми, либо другими типами машин, применяемых для обработки данных. XML позволяет создавать документы, содержащие термины и определения, в соответствии с конкретными пожеланиями пользователей. Спецификации, которые применяются для описания документов, можно помещать в описание типа документа (DTD). DTD можно поместить непосредственно в сам документ [3].

Поскольку сами данные отделены от информации об их представлении, одни и те же данные XML могут быть представлены различным образом на экранах компьютеров пользователей. Сам по себе XML-документ не указывает, будет ли, и каким образом, информация отображена на экране. XML-документ содержит лишь данные. С помощью механизма таблиц стилей HTML отображает данные. Как Web-сервер, так и браузер управляют преобразованием XML-данных в формат HTML. К тому же данные XML могут обновляться автоматически, без обновления всей страницы в целом.

При помощи XML можно определять собственные наборы тегов и необходимые имена атрибутов. Структуры документа могут быть вложены друг в друга, обеспечивая любой уровень сложности, если при этом соблюдаются соответствующие правила вложения документов XML. Если "разобрать" XML-документ, то можно обнаружить, что он состоит из двух типов информации: из содержания и разметки. Содержание определяет, какой будет разметка.

Важно отметить, что в отличие от языка HTML язык разметки XML определяет не форматную разметку XML-документов, а разметку и структурирование их содержимого (контента). Для форматной разметки служат языки XSL и CSS. Предполагается, что XML-документ состоит из элементов документа различных типов. Элементы документов могут иметь свое содержимое, а также параметры, называемые атрибутами элемента.

Однако в технологиях XML наряду с автономно существующими метаданными, отчужденными от самих данных, существуют также метаданные, встроенные в XML-документы. Встроенные метаданные представляются разметкой документов и обеспечивают их самоописание. Для

извлечения значений таких метаданных необходимо сканировать содержание документа с помощью синтаксического анализатора (парсера). Таким образом, в технологиях XML сосуществуют принципы самоописания данных и их автономного описания. Разметка XML-документа определяет при этом структурные его компоненты и описывает некоторые простейшие свойства конкретного документа.

Интенсивное развитие технологической платформы Web, быстрое расширение сферы ее практического использования привели к интенсивному наращиванию объема разнообразных информационных ресурсов XML, поддерживаемых не только в среде Web, но и вне ее: XML - документов, составляющих их данных и описывающих их метаданных. Для эффективного управления репозиториями таких информационных ресурсов, поддерживаемых вне Web, естественно использовать технологии баз данных. Стали разрабатываться специализированные СУБД, основанные на стандартах платформы XML и ее окружения, которые непосредственно поддерживают XML-представление данных, а также соответствующие интерфейсы конечных пользователей и интерфейсы прикладного программирования для различных языков программирования. Базы данных, содержащие информационные ресурсы XML, стали называться базами данных XML, а СУБД, обеспечивающие возможности работы с такими базами данных - XML-ориентированными СУБД. Корпорация IBM представила серию комплектов программного обеспечения для сервера баз данных DB2 9, которые позволяют заказчикам быстро адаптировать свои среды управления информацией к новым стандартам обработки XML-данных. Одна из важнейших особенностей продукта DB2 9 - способность работать как с реляционными данными, так и с XML-данными.

#### Стандарт GJXDM для описания правовых данных

Обратимся к практическому опыту наиболее развитых государств. В США, в течение последних 30 лет, отсутствие стандартов увязки работы информационных систем правосудия привело к значительным затратам, связанным с обменом информацией и принесло серьезные проблемы в обмене информацией между агентствами. Было принято решение о необходимости разработки единых стандартов обмена данными и создания общего языка для обмена информацией между правовыми системами. После публикации отчета комиссии "9/11" президент Буш издал несколько распоряжений, обязывающих уполномоченные силовые ведомства руководствоваться политикой

“делиться друг с другом известной им информацией”, т. е. снять межведомственные информационные барьеры. В результате была создана государственная модель обмена информацией (National Information Exchange Model, NIEM), ставшая одной из лучших мировых практик построения информационного взаимодействия с большим количеством объектов обмена и его участников. Фактически NIEM — это стандарт межведомственного обмена информацией в США, созданный на базе референсной (эталонной) модели GJXDM (Global Justice XML Data Model) и распространенный на множество других предметных областей. Специализированные отраслевые форматы используются также в федеральных государственных организациях (та же модель GJXDM) и здравоохранении (модель HL7). Обращаем ваше внимание, что модель GJXDM имеет в своей основе концепцию XML. Еще одним важным условием к использованию платформы

XML как стандарта в системе правосудия является то, что партнеры договорились об общем словарном запасе XML, чтобы более четко представить информацию для совместного использования [4]. Сегодня этот словарь уже существует.

National Information Exchange Model (Национальная модель обмена информацией, NIEM) — это, как уже упоминалось, открытый язык для межведомственного обмена информацией. Эта модель: имеет словарь общих понятий и определений; является схемой описания XML для обмена компьютер-компьютер; поддерживает расширяемые стандарты для обмена информацией. Используя GJXDM, субъекты правоотношений могут описать структуры документов и обмениваться информацией по вопросам правосудия и правонарушителей в пределах населенного пункта, государства, между государствами или с федеральными структурами.

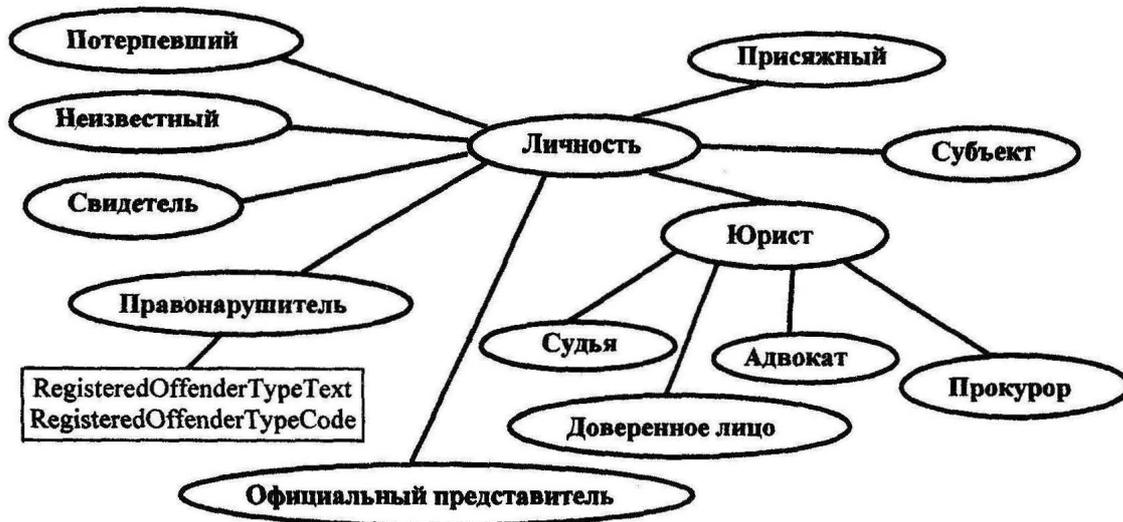


Рис. 1. Описание восьми подклассов концепта «Личность».

При проектировании GJXDM министерство юстиции столкнулось с проблемой создания набора спецификаций, которые бы четко определили структуру данных правосудия, связанных с содержащейся в ней информацией. Описание архитектуры юстиции в терминах XML состоит в том, что в словаре данных (GJXDD) использована четко определенная лексика имен, данных и структур, собранных в объектно-ориентированной модели данных (GJXDM), из которых генерируются последовательные XML-схемы. Были исследованы 33 исходных спецификации, в них исследовано порядка 16000 элементов, на основании которых в иерархической модели представлено более 2700 компонентов, что и составляет словарь GJXDM [5].

Рассмотрим пример описания концепта

«Личность» в рамках словаря GJXDM.

По результатам анализа значительного количества документов разработчики решили, что только восемь подклассов описания концепта «Личность» целесообразно включить в описание PersonType: Потерпевший, Незвестный, Свидетель, Правонарушитель, Официальный представитель, Юрист (Судья, Адвокат, Доверенное лицо, Прокурор), Субъект, Присяжный (рис. 1).

**Направления дальнейших исследований.** Выполнить углубленный анализ системы документооборота в правовой сфере и разработать предложения для стандартизации словаря юридических компонентов для Украины, взяв за основу словарь GJXDM. Разработать описания типовых юридических документов с разметкой

GJXDM. Проаналізувати можливості підключення системи UML для оптимізації процесу обміну інформацією між суб'єктами правоотношений.

**Выводы.** Объем используемых в правовой сфере данных чрезвычайно велик, что требует применения средств их автоматизированной обработки и наличия механизмов объединения в единую информационную структуру.

Правовая информация, как отдельный вид информации, обладает рядом особенностей, на основании которых ее можно формализовать. В США на платформе XML разработана модель данных для глобального описания правовых ситуаций GJXDM, для которой разработан свой словарь профессиональной области. Модель оказалась достаточно удачной и в настоящее время распространяется на другие проблемные области. Проанализировав способ представления данных в этой модели, можно сделать вывод, что данная модель, ввиду ее универсальности, может быть применена к описанию юридических документов в Украине. Министерство юстиции Украины своим приказом № 2563/5/645 от 05.12.2013 [6] устанавливает в качестве стандарта для описания правовых документов модель на основании языка разметки XML с целью гармонизации стандартов с Европейским Союзом и США. Это создаст предпосылки для унифицированного описания правовых ситуаций и автоматизированного обмена стандартизированной правовой информацией как внутри государства, так и с межнациональными правовыми структурами.

#### Список литературы

1. Tatsyi V. *Semantic network of knowledge in science of law* / V. Tatsyi, A. Getman, S. Ivanov, V. Karasiuk, O. Lugoviy, O. Sokolov // *Proceedings of the IASTED International Conference on Automation, Control, and Information Technology, held June*

15 – 18 2010 in Novosibirsk, Russia / *The International Association of Science and Technology for Development. – Anaheim, USA, Calgary, Canada, Zurich, Switzerland: ACTA Press. – 2010. – p. 218 – 222.*

2. Закон Украины «Об информации» № 2567-ХІІ от 02.10.1992.- ВВР, 1992. -№ 48. -ст.650.

3. Козаловский М.Р. *Технологии xml и xml-данные // Материалы международной научной конференции Базы данных и информационные технологии XXI века. Москва, 29-30 сентября 2003 г. Институт проблем рынка РАН. – М.: РГГУ, 2004. [Электронный ресурс]. – Режим доступа: <http://www.cemi.rssi.ru/mei/articles/koga03-2.pdf>*

4. Miller F.P., Vandome A.F., McBrewster J. *GJXDM. – Dusseldorf: Alphascript Publishing. – 2010. – 64p.*

5. *Building Exchange Content Using the Global Justice XML Data Model: A User Guide for Practitioners and Developers.* U.S. Department of Justice. SEARCH: – 2005 – 163 p. [Электронный ресурс]. – Режим доступа: <http://it.ojp.gov/documents/GJXDMUserGuide.pdf>

6. *Наказ Міністерства Юстиції України та Адміністрації Державної служби спеціального зв'язку та захисту інформації України № 2563/5/645 від 05.12.2013 «Про затвердження переліків стандартів у сфері електронного цифрового підпису, перспективних для перегляду та гармонізації з європейськими та міжнародними стандартами відповідно до встановлених законодавством процедур»* [Електронний ресурс]. – Режим доступа: [zakon3.rada.gov.ua/laws/show/v2563323-13](http://zakon3.rada.gov.ua/laws/show/v2563323-13)

**Рецензент:** проф. кафедри ІС І.А. Золотарева, Харківський національний економічний університет, Харків.

**Автор:** Карасюк Володимир Васильевич

*Національний юридичний університет імені Ярослава Мудрого, Харків, кандидат технічних наук, доцент. Раб. тел. – 704-92-02, дом. тел. – 370-49-48, E-mail – vl\_karasuk@ukr.net*

**Автор:** Кошева Наталья Анатольевна

*Національний юридичний університет імені Ярослава Мудрого, Харків, кандидат технічних наук, доцент. Раб. тел. – 704-92-02, дом. тел. – 704-00-41, E-mail – [inform@mulaw.edu.ua](mailto:inform@mulaw.edu.ua)*

#### АНАЛІЗ СПОСОБУ ПОДАННЯ ПРАВОВОЇ ІНФОРМАЦІЇ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

В.В. Карасюк, Н.А. Кошева

*Розглядається мова розмітки GJXDM для представлення правових даних. Проведено аналіз особливостей правової інформації та інформаційних систем в юриспруденції. Звернуто увагу на можливості XML в описі структурованих даних. Зроблено огляд основних підходів для опису правових даних в GJXDM. Сформульовано висновок про перспективність застосування GJXDM в українських правових інформаційних системах.*

**Ключові слова:** правова інформація, формат правових документів, модель даних, XML, GJXDM.

#### ANALYSIS OF THE MODE OF REPRESENTATION OF THE LEGAL INFORMATION IN INFORMATION SYSTEMS

V.V. Karasiuk, N.A. Kosheva

*GJXDM considered markup language to represent the legal data. The analysis of features legal information and information systems in the jurisprudence. Attention is drawn to the power of XML to describe structured data. A review of the main approaches to describe the legal data GJXDM. Conclusions on the application prospects GJXDM in Ukrainian legal information systems.*

**Keywords:** legal information, legal documents format, the data model, XML, GJXDM.