

## ПРОГРАММНЫЕ ИНФОРМАЦИОННЫЕ ПРОДУКТЫ КАК ИНСТРУМЕНТ УНИФИКАЦИИ И ОПТИМИЗАЦИИ ЮРИДИЧЕСКИХ ТЕКСТОВ

### SOFTWARE INFORMATION PRODUCTS AS A TOOL FOR UNIFICATION AND OPTIMIZATION OF LEGAL TEXTS

**Макович Г. В.**

Профессор кафедры иностранных языков Уральского юридического института МВД России, доктор филологических наук, профессор

**Makovich G. V.**

Professor of the Department of Foreign Languages of the Ural Law Institute of the Ministry of Internal Affairs of Russia, Doctor of Philology, Professor

**Аннотация:** В статье изучаются возможности программных информационных продуктов для анализа юридической текстовой информации. Задачей автоматизированной обработки текстов с помощью информационных продуктов является поиск и интерпретация правовой информации. Также в статье показано, что автоматизированная обработка информации позволяет выявлять и закреплять для последующего употребления в юридической практике стандартные модели выражения правовых смыслов. В статье показано, как происходит маркировка текста, его обработка и как формируется семантико-синтаксическая инфраструктура юридических текстов.

**Abstract:** The article examines the possibilities of software information products for the analysis of legal textual information. The task of automated text processing using information products is the search and interpretation of legal information. The article also shows that automated information processing makes it possible to identify and consolidate standard models of expression of legal meanings for subsequent use in legal practice. The article shows how the text is marked, processed, and how the semantic and syntactic infrastructure of legal texts is formed.

**Ключевые слова:** Лексико-семантические шаблоны, правовые Big Data, семантико-синтаксическая инфраструктура юридических текстов, синтаксические шаблоны.

**Keywords:** Lexical and semantic templates, legal Big Data, semantic and syntactic infrastructure of legal texts, syntactic templates.

В юридическом дискурсе проявляются особенности использования языка права в профессионально значимых ситуациях общения. В основе дискурса лежит ситуационная модель общения, она актуализирует юридические тексты применительно к ситуации. Исходный текст трансформируется и приобретает веер реализаций, включая реализации в различных жанровых и поджанровых формах. Юридический дискурс представляет собой текст права в динамике, в процессе толкования и разъяснения.

Дискурс реализует профессиональный тезаурус в речевой деятельности. Базовый слагаемый дискурса являются терминология Ядро юридического дискурса составляют тематические фреймы и лингвокогнитивные модели, которые проявляют правовые концепты.

Задачей является создание семантического инструментария для поиска и интерпретации правовой информации и стандартных моделей выражения юридических смыслов. Обработка текстов одной предметной области позволит сформировать семантико-синтаксическую инфраструктуру юридических текстов.

Эта задача может быть решена путем применения технологии «больших данных», Big Data. Эта технология активно внедряется в разные сферы профессиональной деятельности, основанные на знаниях, и используется для сбора информации о сложных объектах и всестороннего анализа этой информации.

Возможности программных информационных продуктов и, в частности, Big Data, для анализа текстовой информации изучаются в работах А.А.Хорошилова<sup>240</sup>, О.Н.Пошатаева<sup>241</sup>, Ю.М.Богданова<sup>242</sup>, А.В.Старовойтова, Р.А.Субхангулова<sup>243</sup>, А.А.Филиппова. Потенциал Big Data в сфере права изучается в работах В.Д.Чуракова<sup>244</sup>, И.И. Черных<sup>245</sup>, Е.В. Булгаковой<sup>246</sup>, И.И.Черных считает, что возможности информационных технологий, в ряду которых и Big Data, способны существенно повысить объективность оценки существующего уровня правового регулирования и на следующем этапе проектировать модели правовых средств<sup>247</sup>.

Процесс информатизации юридической деятельности, который является базой для внедрения Big Data и других программных информационных продуктов, интенсивно развивается. Правовая информация размещается на сайтах государственных органов власти, персональная информация в системах учетов, российские суды размещают на сайтах свои решения, существует сайт-агрегатор судебных и нормативных актов Российской Федерации СудАкт (<https://sudact.ru>), аналитическая система судебной практики «Сутяжник» (<http://sutjazhnik.garant.ru>) и др.

В целом накоплен значительный материал, который может быть обработан под решение различных задач, в том числе и под лингвистическую задачу унификации и оптимизации структуры юридических текстов разных типов, устных и письменных, их синтаксического и лексического наполнения. Поскольку юридический текст, с одной стороны, определяет, «задает» процедуры и правила, применяемые при разрешении споров и иных юридических дел, а с другой стороны, отражает фактическое состояние практики правоприменения, то лингво-юридический анализ массивов Big Data позволит стандартизировать юридический текст.

К примеру, стоит задача обеспечения терминологического единообразия в практике законодательства, что позволит повысить четкость правоприменения. Также первоочередной является работа по выявлению точности юридического текста и обеспечению этой точности.

Другая важная задача – это унификация процедур, применяемых при разрешении дел в судопроизводстве. Лингво-юридический аспект этой задачи состоит в выявлении облигаторных и вариативных блоков в протоколах судебных заседаний по идентичным делам, четких стандартизированных формулировок, представляющих фактологическую составляющую ситуации и квалификации ситуации и ее компонентов. Коммуникативные блоки высказываний стандартизированной структуры являются производными от процессуальной

---

<sup>240</sup> Хорошилов А.А., Пошатаев О.Н. Методы анализа текстов в технологиях Big Data || Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Труды XV Всероссийской научной конференции RCDL'2013. 2013. С. 30-37.

<sup>241</sup> Там же.

<sup>242</sup> Богданов Ю.М., Пошатаев О.Н., Старовойтов А.В., Хорошилов А.А. Принципы создания высокопроизводительных систем обработки и анализа текстовой информации// Информатизация и связь. 2013. № 3. С. 148-155.

<sup>243</sup> Субхангулов Р.А., Филиппов А.А. Онтологически-ориентированная система интеллектуального анализа технических документов// Четырнадцатая национальная конференция по искусственному интеллекту с международным участием. КИИ-2014. Труды конференции. 2014. С. 228-236.

<sup>244</sup> Чураков В.Д. Big Data и юриспруденция: на одном ли мы пути? //Право и информация: вопросы теории и практики: Сборник материалов международной научно-практической конференции Вып. 7. - СПб. : Президентская библиотека имени Б.Н. Ельцина, 2018. - С. 136-143.

<sup>245</sup> Черных И.И. Правовое прогнозирование в сфере гражданского судопроизводства в условиях развития информационных технологий// Актуальные проблемы российского права. -2019.- №6. - С.58-72.

<sup>246</sup> Булгакова Е.В. Методы анализа больших данных в решении юридических задач//Право и информация: вопросы теории и практики: Сборник материалов международной научно-практической конференции. - СПб, 2017. - С.90-96.

<sup>247</sup> Там же. С. 61.

стороны судопроизводства. Выявление и фиксирование таких стандартизированных блоков позволит оптимизировать судопроизводство.

Для решения задачи формирования терминологического единообразия и точности в законодательной практике используется метод оценки качества стиля нормативно-правового документа. В ключевых позициях законодательного акта, должны находиться именно термины, лексемы, соответствующие критериям и требованиям, предъявляемым к терминологическим единицам. Термин употребляется для наименования понятия, он обладает содержательной точностью и имеет дефиницию, однозначен, его значение не зависит от контекста. Термины надо отличать от смежных лингвистических явлений, не обладающих всем набором необходимых признаков.

Юридическая терминосистема включает в себя корпус «эталонных» единиц, и любой юридический текст из корпуса текстов в Big Data сопоставляется с «эталонными» носителями юридических смыслов. В терминосистемах «эталонными» являются базовые термины, обозначающие базовые понятия права, а также ядерные термины, называющие основные правовые понятия внутри группы базовых: административное правонарушение, кассационная жалоба, юридический прецедент, состав правонарушения и др.

Также специального анализа требуют составные термины, включающие в себя несколько компонентов и обозначающие сложные понятия, поскольку при сочетаемости компонентов часто происходят их семантические трансформации: общедозволительный тип правового регулирования, объективно противоправное деяние, сознательно-волевой характер права и др.

Лингво-юридический анализ текста в прикладной лингвистике осуществляется с использованием многочисленных программных средств. К примеру, применяются универсальные программы GATE и Ant. так и специализированные, создающиеся в каждом конкретном случае под определенный заказ для решения ведомственных, отраслевых задач разрабатываются специализированные программные продукты.

Статистика, полученная в результате разметки корпуса юридических текстов в соответствии с маркерами терминосистемы, позволит сделать вывод о встречаемости критериев точности и единообразности юридических текстов и значимости данных критериев в оценке текста. Эта информация даёт возможность определять уровень качества правового текста, формировать рекомендации по его улучшению, моделировать оптимальные блоки для использования в юридических текстах.

Чтобы реализовать указанные функции в автоматическом режиме в рамках Big Data с использованием возможностей специализированного программного продукта, требуется специальная лингвистическая работа по структурированию юридических терминосистем. Каждая из них является многоуровневой и иерархической, с многочисленными связями по смежности.

Маркирование юридической терминосистемы по максимально возможному набору признаков позволяет произвести точную разметку корпуса юридических текстов и оценить их лингво-юридические свойства. Статистика, полученная в результате разметки корпуса текстов позволит определить уровень качества правового текста и разработать рекомендации по его улучшению.

Аналогичное маркирование правового текста следует производить для оценки его синтаксической простоты и стандартизованности структурных единиц, сложных синтаксических целых, коммуникативных блоков высказывания. В основе этой работы лежит выявление в массивах текстов «эталонных» стандартизированных синтаксических целых, которые можно соотнести 1) с конкретными видами текста, к примеру, исковое заявление, судебное решение, постановление и др., 2) со структурными частями текста, например: изложение фактов, квалификация ситуации и ее компонентов и др. В автоматическом режиме корпус текстов в Big Data маркируется на основе сопоставления с «эталоном».

Результаты позволяют выявлять уровень качества правового текста и предложить пути нормативной трансформации и оптимизации правового текста.

Все описанные процессы должны привести к формированию автоматизированной платформы создания юридических текстов разных видов. Сервис должен быть ориентирован на выполнение пользователем нескольких функциональных ролей. Одну из ролей позволит реализовать режим «Проверка», дающий возможность пользователю загрузить документ и по итогам маркирования текста произвести его редактирование путем выбора соответствий из базы «эталонных» конструкций и «эталонных» терминов. Другим режимом работы сервиса должен быть режим «Создание», позволяющий из готовых структурных единиц, из лексико-синтаксических шаблонов смоделировать правовой текст определенного вида.

Формирование правовых Big Data позволяет разрабатывать и применять системы для обработки корпусов текстов, которые ориентированы на профессионалов в области права, решать прикладные проблемы унификации юридического текста, оптимизации его создания и, как следствие, повышение качества правоприменения.

### **Список использованной литературы**

1. Хорошилов А.А., Пошатаев О.Н. Методы анализа текстов в технологиях Big Data || Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Труды XV Всероссийской научной конференции RCDL'2013. 2013. С. 30-37.
2. Богданов Ю.М., Пошатаев О.Н., Старовойтов А.В., Хорошилов А.А. Принципы создания высокопроизводительных систем обработки и анализа текстовой информации// Информатизация и связь. 2013. № 3. С. 148-155.
3. Субхангулов Р.А., Филиппов А.А. Онтологически-ориентированная система интеллектуального анализа технических документов// Четырнадцатая национальная конференция по искусственному интеллекту с международным участием. КИИ-2014. Труды конференции. 2014. С. 228-236.
4. Черных И.И. Правовое прогнозирование в сфере гражданского судопроизводства в условиях развития информационных технологий// Актуальные проблемы российского права. -2019.- №6. - С.58-72.
5. Булгакова Е.В. Методы анализа больших данных в решении юридических задач//Право и информация: вопросы теории и практики: Сборник материалов международной научно-практической конференции. - СПб, 2017. - С.90-96.
6. Чураков В.Д. Big Data и юриспруденция: на одном ли мы пути? //Право и информация: вопросы теории и практики: Сборник материалов международной научно-практической конференции Вып. 7. - СПб. : Президентская библиотека имени Б.Н. Ельцина, 2018. - С. 136-143.