

ВОЛКОВ ЮРИЙ ВИКТОРОВИЧ

кандидат юридических наук, доцент кафедры информационного права, Уральский государственный юридический университет, г. Екатеринбург;
e-mail: yuriiivolkov@yandex.ru

ВЕКТОРНЫЙ АНАЛИЗ ЮРИДИЧЕСКИХ ТЕКСТОВ

Целью, описанного в статье исследования, является поиск и разработка способов замены трудоемкого и длительного анализа текста его аналогом на основе информационных технологий. Основным методом такого анализа может быть векторный анализ. Показано, что при соответствующем развитии и доработке векторный анализ может быть применен для анализа всего текста. Дальнейшее развитие метода приведет к замене длительной процедуры документального (устного, письменного) согласования текста документа в ходе принятия решения на быстрый технологический.

Ключевые слова: вектор; закон; информация; компьютер; семантика.

Volkov Yu. V.

PhD in Law, Assistant Professor of the Information Law Department. Ural State Law University, Yekaterinburg; e-mail: yuriiivolkov@yandex.ru

VECTOR ANALYSIS OF LEGAL TEXTS

The goal described in the research paper is to find and develop ways to replace time-consuming and lengthy text analysis with its counterpart based on information technology. The main method of such analysis can be vector analysis. It is shown that with appropriate development and refinement, vector analysis can be used to analyze the entire text. Further development of the method will lead to the replacement of a lengthy procedure of documentary (oral, written) agreement of the text of the document in the course of making a decision on a fast technological one.

Key words: computer; law; information; semantics; vector.

Влияние информационных технологий во все сферы жизни уже никем не оспаривается. Средства массовой информации периодически информируют о скорой массовой замене роботами отдельных категорий специалистов (бухгалтеров, преподавателей, юристов). Проникновение информационных технологий в разные сферы (литература, управление право), становится частой, если не сказать постоянной, темой семинаров и конференций. По мнению многих экспертов, особенно правоведов, Машина еще долго не сможет заменить человека. Однако это мало кого убеждает, свидетельством тому ряд предстоящих международных конференций (например: 30 ноября

2018 г. в г. Минске; 4 декабря 2018 г. в г. Лондон) предмет которых информационные технологии в управлении и электронном правосудии.

Современная ситуация сложилась не в один день. Еще в 60-е гг. XX в. чешский исследователь В. Кнапп отмечал необходимость «логико-математического описания волевых процессов для дальнейшего продвижения кибернетики» [Кнапп 1965, с. 68]. Аналогичные исследования проводились и в СССР, однако непреодолимыми препятствиями на пути внедрения компьютеров в правовую сферу стали: вопросы ««техники» составления программ»; «необходимость максимальной конкретизации права» [Вопросы кибернетики... 1967, с. 85–89]. Очередная «волна» интереса к использованию компьютерной техники в творческом процессе хронологически соответствует перестройке и развалу СССР. В это время специалист по математической лингвистике А. М. Кондратов констатировал создание музыкальных произведений с помощью ЭВМ «Урал» [Кондратов 1987, с. 121], а также сложных стихов скандинавского типа при помощи программы «Скальд» [Кондратов 1987, с. 133]. Позднее Ю. М. Батурич практически «по полочкам» разложил проблематику соотношения права и компьютерной техники [Батурич 1991]. Что изменилось за прошедшие десятилетия и почему именно сейчас внедрение информационных технологий в право и лингвистику актуально, как никогда ранее?

В наши дни достижения искусственного интеллекта в шахматах, в игре Го, в карточном покере, в иных сферах, – свидетельствуют об изменении приоритетов использования вычислительных мощностей. Если ранее они использовались преимущественно для хранения информации, то в настоящее время их постепенно ориентируют на операции принятия решения. Для сферы государственного управления и правоправедения основные изменения незаметны, так как они происходят в рамках компьютерной лингвистики. Последние достижения лингвистики (языковые корпуса и векторный анализ) позволяют давать текстам формально определенную, математическую оценку. Именно они повлияют, а во многих зарубежных странах (Великобритания, Европейский Союз) уже влияют на правовую сферу. Правоправедение также, как лингвистика основным предметом труда имеет текст. Литературные и правовые тексты имеют много общего. Язык, как средство коммуникации, имеет строение и определенные черты, в том числе: нормативность (правила: орфографические, синтаксические

и т. д.); логичность; формальность. Право также содержит несколько сходных признаков, в том числе: нормативность; формальную определенность; объективность. Это делает изоморфным (подобным), если не полностью, то в основной части, процесс работы с текстом (художественным, официальным, новостным, правовым и т. д.). Критерии оценки в каждом случае разные. Они могут быть заданы для конкретного типа текста: делового; художественного; правового.

На фоне ускоряющихся социальных процессов способы работы с текстами оставались, а во многих местах остаются и по ныне, прежними (чтение, анализ, выводы). Эти традиционные «медленные» способы работы с юридическими текстами уже не удовлетворяют многих. Любой комментарий, интервью, сообщение, в том числе и о ненадлежащем применении закона, могут стать достоянием социальных сетей в считанные секунды. С другой стороны, анализ, проверка фактов, толкование положений закона требуют значительного времени. Ситуация осложняется постоянным ростом объемов законодательства и снижением его качества. Решением для некоторых вопросов могут быть разработанные в лингвистике математические методы анализа текстов [Кодухов 1974, с. 87, 96]. Важную роль в продвижении компьютерной лингвистики сыграли также работы Ноама Хомского и Джорджа Миллера. Но особенный интерес представляет метод векторного анализа, ставший возможным в связи с развитием современных информационных технологий. Он, описан исследователями [Kutuzov, Kuzmenko 2016] и реализован на общедоступном сетевом ресурсе (rusvectors.org/ru). Следует отметить, что данный метод на фоне иных аналогичных методов корпусной лингвистики является, если не самым, то одним из самых передовых. Он позволяет представить результат литературного анализа термина в числовом формате. Таким образом, любой индивид, который просто знает числа, сможет наглядно представить общность и/или различие литературных структур и текстов.

Приведем пример анализа юридических терминов на его основе. Два термина *распространение* и *передача* являются ключевыми в характеристике движения информации и определяющими в информационной сфере. Законодательной дефиниции *передачи* нет, а *распространение* толкуется как разновидность *передачи*. Конкуренция терминов особенно очевидна в рамках уголовного законодательства, когда *распространение* выступает элементом диспозиции многих

составов преступлений, а *передача*, в них же, таких составов не образует. Определенный интерес представляют дела в иных отраслях права, например: дело по жалобе А.И.Сушкова, рассмотренное Конституционным Судом РФ. В сердцевине конфликта разное толкование сторонами терминов *передача*, *разглашение* и *распространение* в случае отправки сообщения электронной почты с «рабочего» на «домашний» адрес. По данному делу Конституционный Суд РФ принял «соломоново решение». Он не дал ответ на вопрос что имело место распространение (разглашение) или передача, но при этом констатировал факт нарушения (практически приравнял термины друг к другу) и отправил дело на пересмотр [Постановление КС РФ № 25-П]. Наш акцент на термины *распространение* и *передача* обусловлен тем, что они имеют место в качестве прав каждого лица (ст. 29 Конституции РФ). Термин *разглашение* в данной статье отсутствует. Проведенный векторный анализ терминов с математической точностью показал безусловное различие терминов *распространение* и *передача*, их семантическую «близость» лишь на 18-м из 23-х возможных уровнях [Волков 2018, с. 32]. Существенное отклонение от полученных традиционным (догматическим) анализом результатов позволяет утверждать необходимость повторного анализа как закона, так и принятого судом решения.

Вариантом применением данного метода является анализ юридических тестов (законов, постановлений, решений и т. д.) в целом. Модель данного процесса можно представить следующим образом. Текст закона подвергается анализу на предмет установления всех возможных «векторов», ключевых терминов. Затем текст переводится в векторную модель, обозначим ее – корпус закона. Дальнейшее применение закона рассматривается через формирование векторных моделей подзаконных и правоприменительных актов. Любой акт, приведенный в векторный вид, может быть проверен путем сравнения с векторами корпуса закона. Наличие отклонения покажет степень соответствия/несоответствия эталону. Первым текстом, базовой моделью («точкой отсчета») может послужить Конституция. На основе ее текста можно сформировать конституционный корпус. Все последующие, принятые и принимаемые законы могут служить, с одной стороны, дополнением законодательного (конституционного) корпуса, с другой – предметом векторного анализа.

Несомненную пользу данный метод может представлять при отправлении правосудия и иной правоприменительной практики.

Точный и быстрый анализ ситуаций изложенных в векторном (исначально в текстовом) формате позволит сэкономить время для судьи и обеспечит более высокую гласность процесса. Естественно, что первое, на чем будут тренироваться, отрабатывать технологию, экспериментаторы – это существующий банк судебных решений. Прогнозируем множество вопросов. Еще один фактор – возможность в автоматическом режиме работать с актуальной версией юридического документа (закона, постановления, решения и т. д.). Следующим фактором является скорость принятия и объемы текстов законов разработанных весьма ограниченным составом экспертов. Есть основания полагать, что Британский закон о защите персональных данных (354 страницы), принятый в течение одного года, разработан с применение аналогичной или подобной технологии.

Полагаем, что часть решения данной проблемы будет генерироваться частным сектором, заинтересованным в контроле законотворческих, управленческих и судебных процессов. Важное значение для анализа имеет также наличие свободных вычислительных мощностей.

СПИСОК ЛИТЕРАТУРЫ

- Батурин Ю. М.* Проблемы компьютерного права. М. : Юридическая литература, 1991. 272 с.
- Волков Ю. В.* Анализ юридического термина «передача информации» // Премственность и новации в юридической науке: сборник ст. Международной науч.-практ. конф., г. Самара, 04 марта 2018 г. Стерлитамак : АМИ, 2018. С. 29–32.
- Вопросы кибернетики и право. М. : Наука. 1967. 312 с.
- Кнап В.* О возможности использования кибернетических методов в праве / пер. М. П. Лебедева. М. : Прогресс. 1965. 288 с.
- Кодухов В. И.* Общее языкознание. М. : Высшая школа, 1974. 303 с.
- Кондратов А. М.* Электронный разум. Очерк исследований по проблеме искусственного интеллекта. М. : Знание. 1987. 187 с.
- Постановление Конституционного Суда РФ от 26.10.2017 № 25-П.
- Хомский Н., Миллер Дж.* Ведение в формальный анализ естественных языков / пер. с англ. 3-е изд. М. : ЛИБРОКОМ, 2010. 64 с.
- Kutuzov A., Kuzmenko E.* WebVectors: A Toolkit for Building Web Interfaces for Vector Semantic Models. In: Ignatov D. et al. (eds) Analysis of Images, Social. Networks and Texts. AIST 2016. Communications in Computer and Information. Science, vol 661. Springer, Cham, 155–161 (2016).