

# ГОСУДАРСТВЕННАЯ ВЛАСТЬ И МЕСТНОЕ САМОУПРАВЛЕНИЕ

DOI: 10.17803/1994-1471.2020.118.9.019-031

С. С. Зенин\*,  
Д. Л. Кутейников\*\*,  
И. М. Япрынцеv\*\*\*

## Большие данные в законодательном процессе<sup>1</sup>

**Аннотация.** В статье представлены возможные направления использования технологии больших данных в рамках законодательной деятельности. Обозначены существенные характеристики технологии больших данных, которые выступают предпосылкой ее внедрения в сферу публичного управления. Авторами описаны существующие практики имплементации этой технологии в сфере юриспруденции. С учетом отлаженных процессов использования больших данных в рамках частного сектора авторами оценены перспективы использования этого опыта в процессе разработки нормативных правовых актов, в том числе в корреляции с конкретными стадиями законодательного процесса. Особое внимание уделяется анализу индивидуализированного регулирования и гранулярных норм, а также основаниям и особенностям использования микродиректив как результата использования больших данных при формировании норм права. В заключении авторы приводят ряд проблемных аспектов (проблема правовой определенности, обеспечение принципа равенства), которые обостряются в связи с использованием технологии больших данных в законодательной деятельности, а также обосновывают ряд дополнительных требований, способствующих минимизации приведенных угроз.

**Ключевые слова:** законодательный процесс; большие данные; база данных; гранулярные нормы; персонализированное правовое регулирование; микродирективы.

---

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-29-16214.

---

© Зенин С. С., Кутейников Д. Л., Япрынцеv И. М., 2020

\* *Зенин Сергей Сергеевич*, кандидат юридических наук, доцент, директор НИИ Московского государственного юридического университета имени О.Е. Кутафина (МГЮА)

Садовая-Кудринская ул., д. 9, г. Москва, Россия, 125993

zeninsergei@mail.ru

\*\* *Кутейников Дмитрий Леонидович*, кандидат юридических наук, старший преподаватель кафедры конституционного и муниципального права Московского государственного юридического университета имени О.Е. Кутафина (МГЮА)

Садовая-Кудринская ул., д. 9, г. Москва, Россия, 125993

kuteynikov@me.com

\*\*\* *Япрынцеv Иван Михайлович*, кандидат юридических наук, советник судьи Конституционного Суда РФ

Сенатская пл., д. 1, г. Санкт-Петербург, Россия, 190000

imihyapryncev@msal.ru

**Для цитирования:** Зенин С. С., Кутейников Д. Л., Япрынтцев И. М. Большие данные в законодательном процессе // Актуальные проблемы российского права. — 2020. — Т. 15. — № 9. — С. 19—31. — DOI: 10.17803/1994-1471.2020.118.9.019-031.

## Big Data in the Legislative Process<sup>2</sup>

**Sergey S. Zenin**, Cand. Sci. (Law), Associate Professor, Head of the Kutafin University Research Institute (MSAL)  
ul. Sadovaya-Kudrinskaya, d. 9, Moscow, Russia, 125993  
zeninsergei@mail.ru

**Dmitriy L. Kuteynikov**, Cand. Sci. (Law), Senior Lecturer, Department of Constitutional and Municipal Law, Kutafin Moscow State Law University (MSAL)  
ul. Sadovaya-Kudrinskaya, d. 9, Moscow, Russia, 125993  
kuteynikov@me.com

**Ivan M. Yapryntsev**, Cand. Sci. (Law), Advisor to the Judge of the Constitutional Court of the Russian Federation  
pl. Senatskaya, d. 1, St. Petersburg, Russia, 190000  
imihyapryntsev@msal.ru

**Abstract.** The paper elucidates possible directions of the use of the big data technology in the framework of legislative activities. The authors determine essential characteristics of the big data technology, which act as a prerequisite for its implementation in the field of public administration. The authors describe the existing practices of implementation of this technology in the field of jurisprudence. Taking into account the well-established processes of using big data in the private sector, the authors evaluate the prospects of using this experience in the process of developing regulatory legal acts, including their correlation with specific stages of the legislative process. Particular attention is paid to the analysis of individualized regulation and granular norms, as well as to the grounds and features of the use of microdirectives as a result of the use of big data in rules of law formation. In conclusion, the authors enumerate a number of problematic aspects (the problem of legal certainty, ensuring the principle of equality), which are exacerbated by the use of the big data technology in legislative activities, and substantiate application of a number of additional requirements that help to minimize these threats.

**Keywords:** legislative process; big data; database; granular norms; personalized legal regulation; microdirectives. Cite as : Zenin SS, Kuteynikov DL, Yapryntsev IM. Bolshie dannye v zakonodatel'nom protsesse [Big Data in the Legislative Process]. *Aktualnye problemy rossiyskogo prava*. 2020;15(9):19—31. DOI: 10.17803/1994-1471.2020.118.9.019-031. (In Russ., abstract in Eng.).

Применение различных цифровых технологий по своей направленности выступает одним из этапов трансформации института государства, его перехода к цифровому состоянию, что предполагает формирование необходимой и достаточной нормативно-правовой основы такой трансформации. При этом очевидно, что качественно иной (более высокий) уровень

развития и изменения государственных институтов, обусловленный внедрением современных технологий, имманентно связан с проблемой создания механизма, позволяющего адекватно нормировать эти процессы.

Соответственно, одной из первоочередных задач в процессе становления цифрового государства является развитие процесса разработ-

<sup>2</sup> The reported study was funded by RFBR according to the research project № 18-29-16214.

ки и принятия нормативных правовых актов: законодателю следует корректно оценивать действительные сложности и вызовы, которые предопределены внедрением различных цифровых технологий, и осознавать, что правовые инструменты, которые ранее использовались для регулирования общественных отношений (писаное право, кодексы, своды законов), не могут рассматриваться в качестве достаточных в нынешних условиях<sup>3</sup>.

Большие данные (Big Data, далее — БД) как одна из разновидностей современных технологий, позволяющая не только аккумулировать, но и обрабатывать существенные информационные потоки (причем не только постфактум, но и в реальном времени) во всё большем объеме, используется в правовой сфере, хотя и отмечаются сложности и противоречия этого процесса<sup>4</sup>. Тем не менее плюсы от их внедрения — в том числе в части осуществления публичных функций и предоставления услуг гражданам — общепризнаны<sup>5</sup>.

Если проблемам внедрения и использования этой технологии в соответствии с действующим регулированием защиты персональных данных внимание уделяется в достаточной степени, то тем новым возможностям для публичного (государственного) сектора, в частности в сфере законодательного процесса, с которыми связано использование БД, в литературе отводится меньшее внимание.

## Большие данные и право: сферы применения

БД по своей сути раскрывается через систему определенных признаков — большой объем (Volume); разнообразие данных (Variety); высокая скорость их изменения (Velocity); достоверность (Veracity)<sup>6</sup>, совокупность которых и гарантирует положительные результаты от внедрения БД в правовую сферу.

Общепризнанными являются следующие направления использования БД в правовой сфере:

**1. E-discovery.** Поиск электронной информации, представленной на цифровых носителях, которые могут быть использованы в качестве доказательств в судебных спорах<sup>7</sup>. В этом аспекте технология e-discovery представлена множеством уже действующих ресурсов, среди которых: <https://www.relativity.com/>, <https://www.symantec.com/>. Анализ такого рода может быть полезен при оценке поступивших предложений по проектам нормативных правовых актов. Привлечение заинтересованных лиц к обсуждению концепций нормативных актов позволит на самой начальной стадии предотвратить возможное недовольство определенных социальных групп. Формируемые предложения (после их перевода в машиночитаемый вид, что возможно в рамках использования машинного обучения при помощи обработки естественного языка (Natural Language Processing, NLP<sup>8</sup>) могут быть учтены в законодательном процессе.

---

<sup>3</sup> Lenz R. Big Data: Ethics and Law // SSRN Electronic Journal. 2019. P. 33. URL: <https://ssrn.com/abstract=3459004>.

<sup>4</sup> См.: Mark A. Cohen. Why Is Law So Slow To Use Data? // URL: <https://www.forbes.com/sites/mark-cohen1/2019/06/24/why-is-law-so-slow-to-use-data/#14ffc709b8eb>.

<sup>5</sup> Sloat B. van der, Schendel S. van. International and comparative legal study on Big Data // wr. The Hague, 2016. P. 25.

<sup>6</sup> Laney D. 3-D Data Management: Controlling Data Volume, Velocity and Variety // Application Delivery Strategies. META Group. February 6, 2001. URL: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (дата обращения: 07.05.2020); Савельев А. И. Проблемы применения законодательства о персональных данных в эпоху «больших данных» (Big Data) // Право. Журнал Высшей школы экономики. 2015. № 1. С. 46—47; Дэви С., Арно М., Мухамед А. Основы data science и Big Data/Python и наука о данных. СПб. : Питер, 2017; Федосеев С. В. Применение современных технологий больших данных в правовой сфере // Правовая информатика. 2018. № 4. С. 51.

<sup>7</sup> E-Discovery — технология, используемая в рамках ознакомления с документами на начальных стадиях судебных споров. См.: What is eDiscovery? // URL: <https://www.aiim.org/What-is-eDiscovery#> (дата обращения: 10.05.2020).

<sup>8</sup> См.: Dale R. Law and Word Order: NLP in Legal Tech // Natural Language Engineering. 2019. 25 (1). Pp. 211—212; Using NLP and ML to Analyze Legislative Burdens Upon Businesses // URL: <https://medium.com/@>

**2. Legal Research.** Анализ поступающей информации для юриста является одним из ключевых навыков, предопределяющих результативность его работы. В настоящее время существуют и активно используются ресурсы, которые предоставляют любую статистическую информацию по заранее определенным критериям. В частности, ресурс *Ravel* (<https://home.ravellaw.com/>) совместно с Harvard Law School позволяет анализировать прецедентное право (в том числе в контексте использования судьями тех или иных аргументов, речевых оборотов и т.д.) и визуализировать результаты этого анализа, а также предоставляет услуги по анализу собственной практики юридических фирм.

Еще одним примером использования возможностей БД в сфере Legal Research может служить ресурс *Lex Machina* (<https://lexmachina.com/>). Lex Machina также предоставляет возможность сбора судебной информации на основании находящихся в открытом доступе информационных ресурсов органов публичной власти<sup>9</sup>, а затем анализирует полученную информацию при помощи специально разработанных алгоритмов.

Использование подобного рода функционала в рамках законодательной деятельности (в частности, при разработке нормативных правовых актов) связано, например, с возможностью по-

иска дублирующих норм в действующем законодательстве, что в условиях наличия нескольких уровней правового регулирования (федерального, регионального, муниципального) крайне важно и позволит избежать противоречивости всей системы правового регулирования, учитывая, помимо прочего, сложности разграничения вопросов ведения федерального и регионального уровня<sup>10</sup>.

**3. Case Strategy.** Еще одним направлением в правовой сфере, использование БД в котором в настоящее время достаточно распространено, является анализ и оценка перспектив рассмотрения конкретного судебного спора. Существующие на рынке технологии позволяют проанализировать множество аналогичных или сходных дел по заранее определенным параметрам (ключевые слова; сфера спора; судья, рассматривающий дело, и т.д.) и формировать прогнозы будущего результата судебного рассмотрения. К основным субъектам, предлагающим такого рода технологии, можно отнести, в частности, французский стартап *Predictice* (<https://predictice.com/>)<sup>11</sup>; *LexPredict* (<https://www.lexpredict.com/>), одним из ответвлений которого является ресурс *LexSemble* (<https://lexsemble.com/>). Отдельно следует обратить внимание на использование возможностей БД при анализе решений Европейского Суда по правам человека<sup>12</sup>.

---

ODSC/using-nlp-and-ml-to-analyze-legislative-burdens-upon-businesses-e5cc106b85b0 (дата обращения: 23.05.2020).

<sup>9</sup> В основном это Public Access to Court Electronic Records (PACER) — электронный ресурс, позволяющий получить онлайн-доступ к делам из федеральных апелляционных, окружных судов и судов, рассматривающих дела о банкротстве (<https://www.pacer.gov/>); официальные сайты Управления по патентам и товарным знакам США (<https://www.uspto.gov/>) и Комиссии по международной торговле США (<https://www.usitc.gov/>).

<sup>10</sup> См.: Крусс В. И. Конституционный федерализм и состоятельность субфедерального законодательства // Государственная власть и местное самоуправление. 2019. № 12; Куракин А. В., Карпунин Д. В., Попова Н. Ф. Принципы разграничения предметов ведения и полномочий между органами государственной власти Российской Федерации и ее субъектами // Административное и муниципальное право. 2018. № 11.

<sup>11</sup> Пример этого ресурса крайне важен еще и потому, что считается, что подобные технологии в большей степени применимы к системам общего права (сфере case-law).

<sup>12</sup> См.: Medvedeva M., Vols M. & Wieling M. Using machine learning to predict decisions of the European Court of Human Rights // URL: <https://link.springer.com/article/10.1007/s10506-019-09255-y#citeas> (дата обращения: 17.05.2020).

Если говорить о применимости и результатах использования подобных технологий в законодательной деятельности, следует обратить внимание на ресурс Legislative Explorer (<http://www.lexisnexis.com>), который позволяет отслеживать и визуализировать процесс разработки, обсуждения и принятия законов в Конгрессе США начиная с 1973 г. Это позволяет выявить наиболее проблемные процедурные моменты в законодательном процессе (например, как отмечается авторами, к таким относятся разногласия между палатами Конгресса) и заранее их минимизировать. Кроме того, функционал этого ресурса в дальнейшем планируется расширить до отслеживания влияния принятого закона на те или иные сферы (на деятельность публичной власти, обязательства иных органов и агентств, а также поведение обычных граждан)<sup>13</sup>. В законодательной деятельности это может быть использовано в контексте анализа на начальных этапах оценки регулирующего воздействия тех или иных нормативных правил, вводимых в действующее регулирование.

Представленные направления использования БД в сфере юридических услуг по своей сути представляют возможные варианты имплементации БД и в законодательную деятельность. Тем не менее в литературе представлены и такие направления использования БД, которые характерны исключительно для правотворческого процесса.

#### **Большие данные и разработка нормативных правовых актов: дополнительные инструменты использования**

БД могут влиять на процесс законодательной деятельности по нескольким направлениям<sup>14</sup>. В частности, последствия принятых решений могут быть оценены на основании выходных данных (покупательская активность, количество совершенных правонарушений, количество зарегистрированных сделок или новых юридических лиц и т.д.), которые показывают влияние нового регулирования на микроуровне (отдельные граждане, юридические лица и т.д.).

Более того, использование БД позволяет в рамках законодательной деятельности моделировать результаты введения новых норм и оценивать их последствия на основании определенных показателей<sup>15</sup>.

Описанные примеры использования БД представляют собой общие направления их имплементации в сфере правотворчества. Говоря же о конкретных примерах, которые могут рассматриваться как действительная трансформация подготовки нормативных правовых актов, следует в первую очередь остановиться на категории гранулярных норм (*granular legal norms*) и персонализированном (индивидуализированном) регулировании (*personalized law*).

#### **Гранулярные нормы и персонализированное регулирование (*granular legal norms; personalized law*)**

Основной посыл персонализированного регулирования<sup>16</sup> основывается, в частности, на использовании потоков информации касательно поведения конкретных субъектов как в прошлом, так и в режиме реального времени, что

---

<sup>13</sup> См.: Online «Legislative Explorer» uses big data to track decades of lawmaking // URL: <https://www.washington.edu/news/2014/04/25/online-legislative-explorer-uses-big-data-to-track-decades-of-lawmaking/> (дата обращения: 21.05.2020).

<sup>14</sup> См.: *Zódi Z.* Law and Legal Science in the Age of Big Data // Human Rights and EU Conditionality in the Western Balkans. 2017. Vol. 3. No. 2. P. 77.

<sup>15</sup> См. подробнее: *Byers A.* Big Data, Big Economic Impact // I/S: A Journal of Law and Policy for the Information Society. 2015. Vol. 10. No. 3. Pp. 757—764.

<sup>16</sup> См., например: *Busch C.* Implementing Personalized Law: Personalized Disclosures in Consumer Law and Data Privacy Law // The University of Chicago Law Review. 86:309.2019. Pp. 309—331 ; *Ben-Shahar O. and Porat A.* Personalizing Negligence Law // New York University Law review. 2016. Vol. 3. No. 3.



позволяет приспособить применяемые к этому лицу правовые предписания в соответствии (либо с учетом максимального приближения) с его требованиями и возможностями<sup>17</sup>.

С точки зрения обеспечения практической реализации персонализированных норм права на законодательном уровне необходимо закрепление определенных критериев, которые следует учитывать при этой персонификации. На основании такой детализированной информации, которая будет привязана к конкретному субъекту (либо группе субъектов), на которых будет распространяться правовое регулирование, может формироваться целая последовательность норм. В качестве примеров подобных критериев, которые могут учитываться при формировании персонализированных (индивидуализированных) норм, могут использоваться возраст, гендерная принадлежность, кредитная история<sup>18</sup>. В частности, кредитная история может быть использована при определении тех групп субъектов, которые являются наиболее уязвимыми в экономическом плане и нуждаются в поддержке со стороны государства<sup>19</sup> (если идет речь, например, о предприятиях малого или среднего бизнеса в условиях экономического кризиса).

В качестве критериев, определяющих подходы к индивидуализированному правовому регулированию, предлагают использовать персональные психологические типы<sup>20</sup>. Авторы предлагают так называемую «большую пятерку» персональных характеристик (экстраверсия, невротичность, законопослушность (доброжела-

тельность), добросовестность, открытость), учет которых позволяет предвидеть поведение человека в определенных ситуациях. В то же время следует согласиться, что использование такого рода категорий, которые крайне изменчивы и трудно определяемы, существенно усиливает важность качественных и количественных характеристик получаемой информации<sup>21</sup>, что ограничивает возможности их использования.

Наиболее полную реализацию идеи индивидуализированного регулирования приобретают в рамках использования микродиректив. Такие микродирективы будут адаптированы ко всем возможным сценариям и способны генерировать правила поведения в реальном времени (например, направлять их на персональные девайсы)<sup>22</sup>.

В то же время подобного рода микродирективы также крайне зависимы от качества и объемов поступающей информации, которая кладется в основу функционирования предсказательного механизма. Это накладывает дополнительные обязательства на публичную власть, ответственную за сбор корректной информации. В частности, при определении потенциальной сферы правового регулирования следует учитывать лишь релевантную информацию, что предопределяет необходимость использования дополнительных технологий, позволяющих осуществлять сбор и передачу такой информации (например, если речь идет о правилах дорожного движения, необходимо получать информацию о погоде, насыщенности движения, времени суток и т.д.)<sup>23</sup>.

<sup>17</sup> *Busch C. and De Franceschi A. Granular Legal Norms: Big Data and the Personalization of Private Law // Research Handbook on Data Science and Law / V. Mak, E. Tjong Tjin Tai and A. Berlee (eds). Edward Elgar, 2018.*

<sup>18</sup> *Porat A., Strahilevitz J. Personalizing Default Rules and Disclosure with Big Data // Michigan Law Review. 2014. Vol. 112. Iss. 8. 1417—1478 ; Busch C. The Future of Pre-contractual Information Duties: From Behavioural Insights to Big Data // Research Handbook on EU Consumer and Contract Law / Ch. Twigg-Flesner (ed.). Edward Elgar Publishing, 2016.*

<sup>19</sup> *Hacker P. Personalizing EU Private Law: From Disclosures to Nudges and Mandates // 25 European Review of Private Law 651. 2017. P. 661 ; Citron D., Pasquale F. The scored society: due process for automated predictions // Washington Law Review. 2014. No. 89. Pp. 14—15.*

<sup>20</sup> *Porat A., Strahilevitz J. Op. cit. P. 1417.*

<sup>21</sup> См.: *Busch C. and De Franceschi A. Op. cit.*

<sup>22</sup> *Casey A. J., Niblett A. The Death of Rules and Standards // Indiana Law Journal. 2017. Vol. 92. Iss. 4. Pp. 1401—1447.*

<sup>23</sup> *Casey A. J., Niblett A. Op. cit.*

При этом в литературе отмечается, что установление и использование большого количества подобных микродиректив, несмотря на их индивидуальную направленность и применимость, обеспечивает большую эффективность в достижении общих целей правового регулирования: снижение уровня преступности, уменьшение количества автомобильных аварий, снижение неравенства в экономическом плане и т.д. Обеспечивается это в основном за счет возможности алгоритма анализировать и учитывать все множество предусмотренных и взаимосвязанных микродиректив, применение которых в каждом конкретном случае будет способствовать достижению общей цели<sup>24</sup>.

Рассматривая персонализированное регулирование, следует отдельно остановиться на вопросе его применимости к различным сферам (отраслям) правового регулирования. Отмечается, что наибольшее распространение такой подход может получить в сфере частного права (договорное право, деликтные отношения, семейное и наследственное право)<sup>25</sup>. Так, например, в сфере деликтного права это позволит индивидуализировать требования к должной осмотрительности<sup>26</sup>, которые будут основываться на индивидуальных характеристиках субъекта<sup>27</sup>.

Следует согласиться, что сфера применения персонализированных норм ограничена. Персонализация подобного типа неприменима и к наиболее абстрактным нормам, нормам-принципам, нормам-целям и т.д., которые в большинстве своем содержатся в конституциях. Это создает своего рода предметный предел рас-

пространения микродиректив в сфере регулирования общественных отношений.

Тем не менее персонализация, индивидуализация правового регулирования в условиях существенно увеличивающегося объема информации, находящегося в открытом доступе, является одним из направлений развития законодательной деятельности, внедрение которого предопределено использованием БД. В то же время сама возможность имплементации БД в сферу правотворчества основана на трансформации существующей в реальности информации в машиночитаемый вид.

Перевод данных в машиночитаемый вид —  
формирование базы данных НПА

Использование БД при разработке нормативных правовых актов сопряжено с необходимостью формирования баз данных, содержащих в машиночитаемом виде существующие правила<sup>28</sup>. Непосредственно трансформация текста в машиночитаемый вид предопределена использованием машинного обучения, в частности технологии обработки естественного языка (Natural Language Processing, NLP).

В литературе описывается общий алгоритм перевода текстов нормативных правовых актов в машиночитаемый вид и формирование базы данных для анализа и использования в рамках внедрения БД<sup>29</sup>. Последовательность форматирования действующего законодательства следующая:

**Очистка данных.** Из текста исключаются слова, которые не отражают существо регулирования (например, числа, сокращения, даты и т.д.).

---

<sup>24</sup> Casey A. J., Niblett A. Op. cit. P. 53.

<sup>25</sup> Busch C. and De Franceschi A. Op. cit.

<sup>26</sup> Ben-Shahar O., Porat A. Op. cit.

<sup>27</sup> В этом случае допустимо учитывать, например, и биологические показатели человека, если речь идет, в частности, о состоянии водителя в момент совершения правонарушения либо при формировании персональных рекомендаций (микро-директив) касательно модели поведения за рулем. См.: Porat A., Strahilevitz J. Op. cit. 1433—1453.

<sup>28</sup> Francesconi E. Reasoning with Deontic Notions in a Decidable Framework // Knowledge of the Law in the Big Data Age. 2019. Vol. 317. Pp. 63—81.

<sup>29</sup> Sharyn O'Halloran S., Sameer Maskey S., Geraldine McAllister G., Park D. K., Chen K. Data Science and Political Economy: Application to Financial Regulatory Structure // The Russell Sage Foundation Journal of the Social Sciences. 2016. Vol. 2. No. 7. Pp. 103—104.

**Токенизация данных.** Этот этап непосредственно связан с использованием NLP — процесс предполагает разделение всего текста на последовательность токенов (слов). Результатом токенизации является объяснение необходимых сокращений, разделение текста на отдельные слова, выделение начальных грамматических форм.

**Нормализация данных.** Нормализация предполагает форматирование данных после токенизации для придания им единообразного вида. Так, например, употребляемые сокращения, используемые в различных вариациях, будут приведены к единому виду (например, ГК Российской Федерации, Гражданский кодекс Российской Федерации, ГК РФ в итоге будет восприниматься как Гражданский кодекс Российской Федерации). То же самое форматирование применимо к датам, валюте и т.д. Нормализация позволит стандартизировать весь спектр данных.

**Формирование словаря данных.** Предпочтительно формирование словаря, который должен в наибольшем объеме содержать информацию касательно той сферы, для которой формируются данные. Обычно словарь формируется на основе результатов токенизации данных — все неповторяющиеся слова, сочетания слов используются в качестве составных частей словаря данных.

**Векторное представление слов** — один из методов обработки языка. Он основывается на том, что все слова представляются в виде векторов в многомерном пространстве. Чем ближе слова находятся друг к другу, тем более они сходны по семантике (т.е. векторное представление слов позволяет определить «семантическое расстояние» между словами).

**TF-IDF (term frequency — inverse document frequency) трансформация данных.** Данный

инструмент позволяет оценивать важность конкретных данных в контексте всего массива представленной информации. IDF-трансформация основывается на меньшей значимости тех данных, которые повторяются с большей частотой.

В контексте формирования базы данных действующего регулирования в машиночитаемом виде следует отметить создание Национальной системы управления данными (далее — НСУД)<sup>30</sup>. Являясь частью федерального проекта «Цифровое государственное управление» национальной программы «Цифровая экономика Российской Федерации», НСУД предполагает разработку единых требований к управлению данными, обеспечение их юридической значимости, создание процессов управления и обеспечения информационного взаимодействия всех участников этой системы<sup>31</sup>.

Представляется, что одним из направлений развития НСУД должно стать формирование единой базы данных, содержащей действующие нормативные правовые акты. В этом контексте будет интересна практика функционирования Идентификатора европейского законодательства (The European Legislation Identifier, ELI), который охватывает следующие информационные базы правовых актов — Официальный журнал Европейского Союза; Законодательство Европейского Союза и связанные с ним документы; Национальное законодательство<sup>32</sup>. Помимо обеспечения общего доступа к нормативным правовым актам ЕС и стран — членов ЕС, эта платформа обеспечивает формирование унифицированных идентификаторов в отношении всех актов, а также представляет описание связанных с ними метаданных<sup>33</sup>.

<sup>30</sup> См.: распоряжение Правительства РФ от 03.06.2019 № 1189-р «Об утверждении Концепции создания и функционирования национальной системы управления данными и плана мероприятий (“дорожной карты”) по созданию национальной системы управления данными на 2019—2021 годы» // СЗ РФ. 2019. № 23. Ст. 3041.

<sup>31</sup> URL: <https://ac.gov.ru/projects/project/nacionalnaa-sistema-upravlenia-dannymi-nsud-41> (дата обращения: 24.05.2020).

<sup>32</sup> См.: Ермакова Е. П., Ситкарева Е. В. Стратегия электронного правосудия в Европейском Союзе: правосудие в сети Интернет // Юстиция. 2014. № 1.

<sup>33</sup> См.: Francart T., Dann J., Pappalardo R., Malagon C., Pellegrino M. The European Legislation Identifier // Knowledge of the Law in the Big Data Age. 2019. Vol. 317. Pp. 137—148.



## Заключение

Подводя итог проведенному анализу возможных направлений использования БД в рамках законодательной деятельности, следует отметить наличие как положительных последствий этого процесса (что характерно для всех цифровых технологий), так и определенных рисков, которые необходимо учитывать.

В качестве позитивных элементов в этом случае следует отметить повышение предсказательных возможностей и учет большого количества данных при правовом регулировании, что позволяет, в частности, обеспечить высокий уровень индивидуализации правового регулирования, и как итог его исполнимости — индивидуализированная норма понятнее, нежели абстрактное предписание, что также ведет к снижению трудностей в ее толковании.

Тем не менее в литературе выделяют и ряд негативных моментов, с которыми сопряжено использование БД. Например, возникает риск несоответствия требованиям законодательства о персональных данных, что предопределено самой характеристикой БД (речь идет, например, о возможности повторного использования данных в различных целях). Действительно, соотношение БД и требований законодательства о персональных данных — одна из проблем,

решение которой связано с возможностью полноценного использования данной технологии в сфере публичного управления. Не исключены трудности в этом вопросе и в рамках использования БД в законодательной деятельности<sup>34</sup>.

Кроме того, высказываются сомнения касательно соблюдения принципов правовой определенности и равенства в условиях внедрения персонализированного регулирования и микро-директив<sup>35</sup>. В действительности автоматизация принятия различных решений сопряжена с рисками дискриминации, а потому может привести к нарушению принципа равенства<sup>36</sup>.

В то же время следует согласиться, что принцип равенства имманентно связан и с требованием учета различий, возникающих в той или иной ситуации и объективно влияющих на ситуацию. Именно этому и способствует использование БД, в том числе в рамках персонализации нормативного регулирования<sup>37</sup>. Минимизация обозначенных рисков связана, помимо прочего, с формированием корректной и достаточной по объему базы данных, которая будет использоваться в рамках правотворчества<sup>38</sup>, что должно быть обеспечено государством, как видится, в рамках формирования НСУД, о чем говорилось выше.

Кроме того, в качестве одной из институциональных гарантий минимизации обозначенных

---

<sup>34</sup> В этом контексте следует отметить опыт нормирования данной сферы в Европейском Союзе, где в Общем регламенте о защите персональных данных (General Data Protection Regulation) предусмотрено осуществление полномочий на основании закона или публичных функций в общих интересах в качестве основания для использования персональных данных. См.: Big data, artificial intelligence, machine learning and data protection. The United Kingdom Information Commissioner's Office. March. 2017. Pp. 35—36.

<sup>35</sup> Busch C. and De Franceschi A. Op. cit.

<sup>36</sup> См.: Williams B. A., Brooks C. F. and Shmargad Y. How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications // Journal of Information Policy. 2018. Vol. 8 ; Lerman J. Big Data and Its Exclusions // Stanford Law Review. 2013. Vol. 66. URL: <https://www.stanfordlawreview.org/online/privacy-and-big-data-big-data-and-its-exclusions/> (дата обращения: 13.03.2020).

<sup>37</sup> Hacker P. Op. cit. P. 659.

Подтверждается это и нормативно-доктринальным пониманием обозначенного принципа в судебной практике — недопустимо лишь введение такого различия в правах лиц, принадлежащих к одной и той же категории, которые не имеют объективного и разумного оправдания. См., например: постановления Конституционного Суда РФ от 31.10.2019 № 32-П, от 26.03.2020 № 13-П.

<sup>38</sup> Scheirer W. J., Jain L. P., Boult T. E. Probability Models for Open Set Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI). 2014. № 11 (36) ; Alpaydin E. Introduction to Machine Learning. 3rd ed. / Ethem Alpaydin. Cambridge : The MIT Press, 2014.

рисков следует рассматривать судебный контроль в отношении всех норм, разработанных и принятых с использованием цифровых технологий<sup>39</sup>. Принцип судебной подконтрольности — как административной, так и конституционной — следует рассматривать как базовое начало, с учетом которого должна осуществляться трансформация законодательного процесса в рамках имплементации современных технологий.

Таким образом, обозначенные риски имплементации технологии БД в сферу законодательной деятельности не могут рассматриваться как

препятствующие этому процессу. Существующие гарантии обеспечения объективности законодательства применимы и к нормам, разработанным в условиях автоматизации правотворчества. Обозначенные же направления трансформации разработки нормативных правовых актов способны существенно повысить эффективность регулирования. В совокупности эти факторы позволяют рассматривать имплементацию технологии БД в законодательный процесс как неотъемлемый элемент перехода к цифровому законодательному процессу.

### БИБЛИОГРАФИЯ

1. *Alpaydin E.* Introduction to Machine Learning. 3rd ed. / Ethem Alpaydin. — Cambridge : The MIT Press, 2014.
2. *Ben-Shahar O. and Porat A.* Personalizing Negligence Law // New York University Law review. — 2016. — Vol. 3. — No. 3.
3. Big data, artificial intelligence, machine learning and data protection. The United Kingdom Information Commissioner's Office. — March. 2017.
4. *Busch C. and De Franceschi A.* Granular Legal Norms: Big Data and the Personalization of Private Law // Research Handbook on Data Science and Law / V. Mak, E. Tjong Tjin Tai and A. Berlee (eds). — Edward Elgar, 2018.
5. *Busch C.* Implementing Personalized Law: Personalized Disclosures in Consumer Law and Data Privacy Law // The University of Chicago Law Abstract. — 86:309.2019. — Pp. 309—331.
6. *Busch C.* The Future of Pre-contractual Information Duties: From Behavioural Insights to Big Data // Research Handbook on EU Consumer and Contract Law / Ch. Twigg-Flesner (ed.). — Edward Elgar Publishing, 2016.
7. *Byers A.* Big Data, Big Economic Impact // I/S: A Journal of Law and Policy for the Information Society. — 2015. — Vol. 10. — No. 3. — Pp. 757—764.
8. *Citron D., Pasquale F.* The scored society: due process for automated predictions // Washington Law Abstract. — 2014. — No. 89. — Pp. 14—15.
9. *Dale R.* Law and Word Order: NLP in Legal Tech // Natural Language Engineering. — 2019. — 25 (1). — Pp. 211—212.
10. *Francart T., Dann J., Pappalardo R., Malagon C., Pellegrino M.* The European Legislation Identifier // Knowledge of the Law in the Big Data Age. — 2019. — Vol. 317. — Pp. 137—148.
11. *Francesconi E.* Reasoning with Deontic Notions in a Decidable Framework // Knowledge of the Law in the Big Data Age. — 2019. — Vol. 317. — Pp. 63—81.
12. *Gardner S.* Artificial Intelligence Poses Data Privacy Challenges // Bloomberg Law Privacy and Data Security. — 2016.
13. *Hacker P.* Personalizing EU Private Law: From Disclosures to Nudges and Mandates // 25 European Review of Private Law 651. 2017.
14. *Laney D.* 3-D Data Management: Controlling Data Volume, Velocity and Variety // Application Delivery Strategies. — META Group. — February 6, 2001. — URL: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (дата обращения: 07.05.2020).

<sup>39</sup> *Casey A. J., Niblett A.* Op. cit. P. 51.

15. *Lenz R.* Big Data: Ethics and Law // SSRN Electronic Journal. — 2019.
16. *Lerman J.* Big Data and Its Exclusions // Stanford Law Abstract. — 2013. — Vol. 66.
17. *Mark A. Cohen.* Why Is Law So Slow To Use Data? // URL: <https://www.forbes.com/sites/markcohen/2019/06/24/why-is-law-so-slow-to-use-data/#14ffc709b8eb>.
18. *Medvedeva M., Vols M. & Wieling M.* Using machine learning to predict decisions of the European Court of Human Rights // URL: <https://link.springer.com/article/10.1007/s10506-019-09255-y#citeas>.
19. Online «Legislative Explorer» uses big data to track decades of lawmaking // URL: <https://www.washington.edu/news/2014/04/25/online-legislative-explorer-uses-big-data-to-track-decades-of-lawmaking/>.
20. *Porat A., Strahilevitz J.* Personalizing Default Rules and Disclosure with Big Data // Michigan Law Abstract. — 2014. — Vol. 112. — Iss. 8. — 1417—1478.
21. *Scheirer W. J., Jain L. P., Boulton T. E.* Probability Models for Open Set Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI). — 2014. — № 11 (36).
22. *Sharyn O'Halloran S., Sameer Maskey S., Geraldine McAllister G., Park D. K., Chen K.* Data Science and Political Economy: Application to Financial Regulatory Structure // The Russell Sage Foundation Journal of the Social Sciences. — 2016. — Vol. 2. — No. 7.
23. Using NLP and ML to Analyze Legislative Burdens Upon Businesses // URL: <https://medium.com/@ODSC/using-nlp-and-ml-to-analyze-legislative-burdens-upon-businesses-e5cc106b85b0>.
24. *van der Sloot B., van Schendel S.* International and comparative legal study on Big Data // wrr. The Hague 2016.
25. *Williams B. A., Brooks C. F. and Shmargad Y.* How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications // Journal of Information Policy. — 2018. — Vol. 8.
26. *Zódi Z.* Law and Legal Science in the Age of Big Data // Human Rights and EU Conditionality in the Western Balkans. — 2017. — Vol. 3. — No. 2.
27. *Дэви С., Арно М., Мухамед А.* Основы data science и Big Data/Python и наука о данных. — СПб. : Питер, 2017.
28. *Ермакова Е. П. Ситкарева Е. В.* Стратегия электронного правосудия в Европейском Союзе: правосудие в сети Интернет // Юстиция. — 2014. — № 1.
29. *Крусс В. И.* Конституционный федерализм и состоятельность субфедерального законотворчества // Государственная власть и местное самоуправление. — 2019. — № 12.
30. *Куракин А. В., Карпунин Д. В., Попова Н. Ф.* Принципы разграничения предметов ведения и полномочий между органами государственной власти Российской Федерации и ее субъектами // Административное и муниципальное право. — 2018. — № 11.
31. *Савельев А. И.* Проблемы применения законодательства о персональных данных в эпоху «больших данных» (Big Data) // Право. Журнал Высшей школы экономики. — 2015. — № 1.
32. *Федосеев С. В.* Применение современных технологий больших данных в правовой сфере // Правовая информатика. — 2018. — № 4.

*Материал поступил в редакцию 16 июня 2020 г.*

#### REFERENCES (TRANSLITERATION)

1. Alpaydin E. Introduction to Machine Learning. 3rd ed. / Ethem Alpaydin. — Cambridge : The MIT Press, 2014.
2. Ben-Shahar O. and Porat A. Personalizing Negligence Law // New York University Law review. — 2016. — Vol. 3. — No. 3.
3. Big data, artificial intelligence, machine learning and data protection. The United Kingdom Information Commissioner's Office. — March. 2017.

4. Busch C. and De Franceschi A. Granular Legal Norms: Big Data and the Personalization of Private Law // Research Handbook on Data Science and Law / V. Mak, E. Tjong Tjin Tai and A. Berlee (eds). — Edward Elgar, 2018.
5. Busch C. Implementing Personalized Law: Personalized Disclosures in Consumer Law and Data Privacy Law // The University of Chicago Law Abstract. — 86:309.2019. — Pp. 309—331.
6. Busch C. The Future of Pre-contractual Information Duties: From Behavioural Insights to Big Data // Research Handbook on EU Consumer and Contract Law / Ch. Twigg-Flesner (ed.). — Edward Elgar Publishing, 2016.
7. Byers A. Big Data, Big Economic Impact // I/S: A Journal of Law and Policy for the Information Society. — 2015. — Vol. 10. — No. 3. — Pp. 757—764.
8. Citron D., Pasquale F. The scored society: due process for automated predictions // Washington Law Abstract. — 2014. — No. 89. — Pp. 14—15.
9. Dale R. Law and Word Order: NLP in Legal Tech // Natural Language Engineering. — 2019. — 25 (1). — Pp. 211—212.
10. Francart T., Dann J., Pappalardo R., Malagon C., Pellegrino M. The European Legislation Identifier // Knowledge of the Law in the Big Data Age. — 2019. — Vol. 317. — Pp. 137—148.
11. Francesconi E. Reasoning with Deontic Notions in a Decidable Framework // Knowledge of the Law in the Big Data Age. — 2019. — Vol. 317. — Pp. 63—81.
12. Gardner S. Artificial Intelligence Poses Data Privacy Challenges // Bloomberg Law Privacy and Data Security. — 2016.
13. Hacker P. Personalizing EU Private Law: From Disclosures to Nudges and Mandates // 25 European Review of Private Law 651. 2017.
14. Laney D. 3-D Data Management: Controlling Data Volume, Velocity and Variety // Application Delivery Strategies. — META Group. — February 6, 2001. — URL: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (data obrashcheniya: 07.05.2020).
15. Lenz R. Big Data: Ethics and Law // SSRN Electronic Journal. — 2019.
16. Lerman J. Big Data and Its Exclusions // Stanford Law Abstract. — 2013. — Vol. 66.
17. Mark A. Cohen. Why Is Law So Slow To Use Data? // URL: <https://www.forbes.com/sites/markcohen1/2019/06/24/why-is-law-so-slow-to-use-data/#14ffc709b8eb>.
18. Medvedeva M., Vols M. & Wieling M. Using machine learning to predict decisions of the European Court of Human Rights // URL: <https://link.springer.com/article/10.1007/s10506-019-09255-y#citeas>.
19. Online «Legislative Explorer» uses big data to track decades of lawmaking // URL: <https://www.washington.edu/news/2014/04/25/online-legislative-explorer-uses-big-data-to-track-decades-of-lawmaking/>.
20. Porat A., Strahilevitz J. Personalizing Default Rules and Disclosure with Big Data // Michigan Law Abstract. — 2014. — Vol. 112. — Iss. 8. — 1417—1478.
21. Scheirer W. J., Jain L. P., Boulton T. E. Probability Models for Open Set Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI). — 2014. — № 11 (36).
22. Sharyn O'Halloran S., Sameer Maskey S., Geraldine McAllister G., Park D. K., Chen K. Data Science and Political Economy: Application to Financial Regulatory Structure // The Russell Sage Foundation Journal of the Social Sciences. — 2016. — Vol. 2. — No. 7.
23. Using NLP and ML to Analyze Legislative Burdens Upon Businesses // URL: <https://medium.com/@ODSC/using-nlp-and-ml-to-analyze-legislative-burdens-upon-businesses-e5cc106b85b0>.
24. van der Sloot B., van Schendel S. International and comparative legal study on Big Data // wrr. The Hague 2016.
25. Williams B. A., Brooks C. F. and Shmargad Y. How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications // Journal of Information Policy. — 2018. — Vol. 8.
26. Zódi Z. Law and Legal Science in the Age of Big Data // Human Rights and EU Conditionality in the Western Balkans. — 2017. — Vol. 3. — No. 2.

27. Devi S., Arno M., Muhamed A. Osnovy data science i Big Data/Python i nauka o dannyh. — SPb. : Piter, 2017.
28. Ermakova E. P. Sitkareva E. V. Strategiya elektronnoho pravosudiya v Evropejskom Soyuze: pravosudie v seti Internet // Yusticiya. — 2014. — № 1.
29. Kruss V. I. Konstitucionnyj federalizm i sostoyatel'nost' subfederal'nogo zakonotvorchestva // Gosudarstvennaya vlast' i mestnoe samoupravlenie. — 2019. — № 12.
30. Kurakin A. V., Karpuhin D. V., Popova N. F. Principy razgranicheniya predmetov védeniya i polnomochij mezhdou organami gosudarstvennoj vlasti Rossijskoj Federacii i ee sub"ektami // Administrativnoe i municipal'noe pravo. — 2018. — № 11.
31. Savel'ev A. I. Problemy primeneniya zakonodatel'stva o personal'nyh dannyh v epohu «bol'shih dannyh» (Big Data) // Pravo. Zhurnal Vysšej shkoly ekonomiki. — 2015. — № 1.
32. Fedoseev S. V. Primenenie sovremennyh tekhnologij bol'shih dannyh v pravovoj sfere // Pravovaya informatika. — 2018. — № 4.