# ПРОЕКТ ЭКСПЕРТНОЙ СИСТЕМЫ «ВИЗАНТИЙСКОЕ ПРАВО И АКТЫ»\*

# THE PROJECT OF THE EXPERT SYSTEM "BYZANTINE LAW AND ACTS"

## Вин Юрий Яковлевич

Кандидат исторических наук, старший научный сотрудник Института всеобщей истории PAH E-mail: hkn@igh.ras.ru

## Кондратьев Дмитрий Евгеньевич

Специалист ООО «Колби-И», инженер по знаниям научного проекта «Экспертная система "Византийское право и акты"» E-mail: bandey@mail.ru

#### Ямщиков Ярослав Анатольевич

Студент Московского государственного технического университета им. Н.Э. Баумана E-mail: xardas90@mail.ru

Yury Ya. Vin,

**Dmitry E. Kondratjev,** 

Yaroslav A. Yamschikov

Авторы предлагают к рассмотрению проект создания экспертной системы (ЭС) «Византийское право и акты» как наиболее эффективного средства научного исследования с помощью специальных баз знаний и аналитических программных приложений. Сфера применения ЭС — византинистика и смежные научные отрасли. ЭС включит эксклюзивные интервью их ведущих специалистов.

*Ключевые слова*: проект, экспертная система, византийское право, византийские акты.

The authors of the article consider the project of Expert System (ES) "Byzantine Law and Acts" as the most effective mean of scientific study by knowledge-bases and analytic program applications. The application' sphere of ES is specialized by Byzantine studies and related branches. ES is to include the exclusive interviews of their specialists.

*Keywords*: Project, Expert System, Byzantine studies, Byzantine acts.

В своем выступлении на XXII международном конгрессе византийских исследований авторы предложили сообществу византинистов общими усилиями построить универсальную экспертную систему (ЭС), которая охватила бы весь спектр византиноведческих изысканий<sup>1</sup>. Как бы ни были они широки в своем размахе, общеотраслевая ЭС, безусловно, может стать вполне надежным средством профессиональной ориентации для каждого византиниста. Посылкой тому служит признание того факта, что потребности совре-

менной исторической науки вряд ли удовлетворит веб-сайт с отсылками на соответствующие интерактивные мультимедиа, к которым византинисты наравне с медиевистами все чаще и чаще, чем прежде, обращаются сегодня за информацией<sup>2</sup>. Другое дело — современная ЭС, являющаяся неотъемлемым компонентом систем искусственного интеллекта<sup>3</sup>. Вместе с тем предшествующие десятилетия, когда такого типа программные средства получили широкое распространение, показывают, что наиболее эффективное применение находят прежде всего

<sup>\*</sup> Исследование выполняется при финансовой поддержке РГНФ (проект № 13-01-12002) и РФФИ (проект № 13-06-00236).

узкоориентированные ЭС. Даже с учетом «традиционности» византинистики, конечно, вряд ли оправданно ограничиваться созданием новых баз данных и банков данных — программных средств, предназначенных обеспечить рациональное хранение и аналитическую систематизацию уже полученной исследователем информации (данных). В отличие от них ЭС как система, основанная на знаниях (базах знаний, содержащих данные и правила их применения), должна стать средством создания новых методов изучения и новых точек отсчета в процессе получения новых знаний<sup>4</sup>.

В этом свете успешный опыт построения информационно-аналитического комплекса «Византийское право» заложил основание для начала нового этапа проектных работ по модернизации названной программной системы с целью создания ЭС. Поэтому следует обсудить именно данный проект исходя из следующих принципов конструирования ЭС:

- 1) пошаговое конструирование «от простого к сложному»;
- 2) модульные конструкции;
- 3) машинный анализ текстуального, аудиои видеорядов;
- 4) интеграция информации;
- 5) универсальность;
- 6) искусственный интеллект.

Проектируемая ЭС объединит материалы полнотекстовой БД «Византийское право» и информационно-аналитических приложений (на сегодняшний день их объем составляет около 100 тыс. учетных записей, которые репрезентируют порядка 1000 печатных листов текстов на латинском, греческом и славянском языках) со вновь формируемыми БД «Византийские акты» и базами знаний (БЗ). Их планируется выстраивать путем систем-

ного развертывания уже существующих программных приложений и построения новых программных функций и аналитических средств (рис. 1).

Отличие ЭС «Византийское право и акты» будет, безусловно, заключаться в сочетании собственно экспертных суждений и оценок, представленных в аудио- и видеоформате, равно текстуальных интервью, а также квазиэкспертных сведений печатных публикаций, с одной стороны, с другой — данных исторических источников, введенных в БД. Сочетание различных по своей ориентации способов и приемов реализации информационного подхода к объектам изучения обозначит пути к открытию новых знаний, ранее недоступных исследователю в ходе эмпирических изысканий в сфере византинистики и славистики. Соответственно тому планируется разработка программных средств обеспечения, которые будут поддерживать хранение и использование данных экспертного характера, представленных как в текстовой, так и аудио- и видеоформе. В этом плане главное возможность проведения информационно-аналитического поиска и систематизации актуальных научных сведений информационного массива, которая бы отвечала современным классификационным нормам научно-технического аппарата критического издания исторических источников. С учетом разнообразия содержания и научного описания памятников права и актов объектом информационной аналитики минимально станет примерно полутора-два десятка позиций.

Информационный массив БД охватывает сведения нескольких десятков крупнейших и важнейших правовых памятников VI–XVII вв. — от Юстинианова законодательства до Старопечатной кормчей 1653 г. Их содержание, представленное

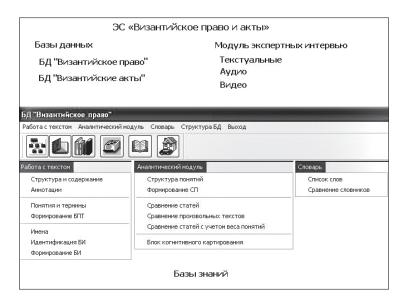


Рис. 1. Структура экспертной системы «Византийское право и акты»

в наглядном виде, открывает прямой путь к изучению эволюции отдельных положений права, а также лексики и понятий. Они воплощают в концептах и терминологии правовые идеи, отображающие широкий спектр научных интересов современного исследователя.

Средоточием экспертной информации станет БЗ «Аннотации», где сейчас аккумулируются сведения о памятниках права и связях их отдельных разделов. Ее основное предназначение — поддерживать на основе сформулированных в ходе выполнения проекта правил обработки данных информационный поиск и системный анализ научных сведений, как обнародованных ведущими специалистами в виде специально подготовленных интервью и результатов их научных исследований, так и данных самих изучаемых источников. Разработанная в ходе выполнения проекта система программного обеспечения явится посылкой для формирования информационного массива данных, в первую очередь представляющих в наглядном виде взаимосвязь памятников византийского и славянского, в частности российского, права и их отдельных разделов с целью воссоздания целостной картины правовых компиляций и рецепций, которая образована тысячами связей отдельных разделов (рис. 2). Примерно то же самое, наверное, будет сказано и об актах, когда содержащая их электронные версии БД будет введена в строй.

Таким образом, на очереди стоит задача разработки и внедрения средств поддержания свобод-

ного доступа к информации экспертных данных, включая как текстуальные, так и видео- и аудиоинтервью ведущих отраслевых специалистов (византинистов, славистов, историков права и специалистов по дипломатике). Научный анализ содержания и систематизации византийского права и актов с позиций информационного подхода требует внедрения специализированных средств навигации и информационно-аналитического поиска. Его объектом должны стать как византийское право в целом, так и отдельные его памятники, их разделы, равно правовые и социокультурные концепты и термины, которые в большом количестве, о чем надо сказать заранее, использовались в широком круге документов. В этом плане главное — возможность проведения информационноаналитического поиска и систематизации актуальных научных сведений информационного массива, которая бы отвечала современным классификационным нормам научно-технического аппарата критического издания исторических источников. С учетом разнообразия содержания и научного описания памятников права и актов объектом информационной аналитики станет минимально порядка полутора-двух десятков позиций.

Равным образом данные блока понятий и терминов (БПТ) будут использованы для построения соответствующего электронного банка и тезауруса<sup>5</sup>. Решению специальных исследовательских задач полностью отвечает разрабатываемый в настоящее время концептологический подход.

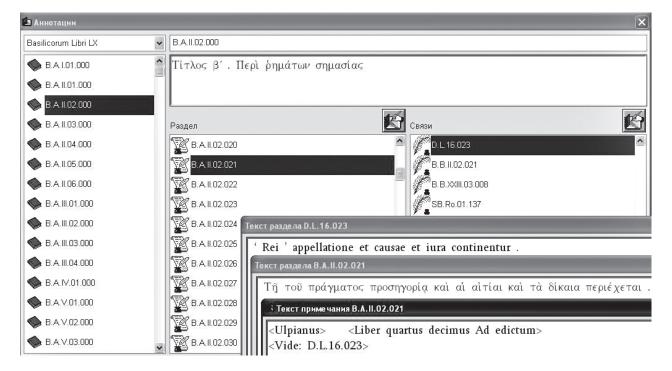


Рис. 2. Программное приложение «Аннотации»

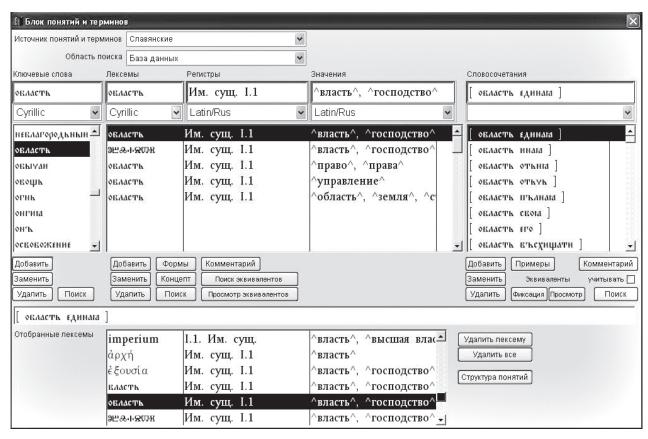


Рис. 3. Функция Блока понятий и терминов «Концепт». Создание структуры фрейма

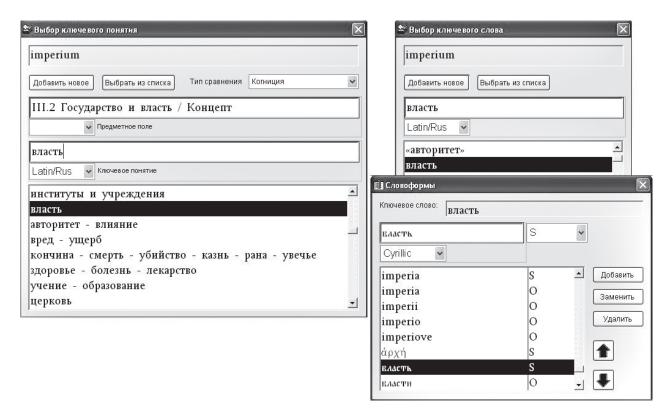


Рис. 4. Функция Блока понятий и терминов «Концепт». Создание структуры концепта

Он претворяется в функциях названного приложения «Концепт» и «Эквиваленты». Они обеспечивают формирование и анализ лексических концептов посредством семантических и межъязыковых сопоставлений латинских, греческих и славянских понятийно-терминологических эквивалентов, наделенных адекватными значениями. Так, функция «Концепт» в БПТ позволяет, исходя из значения слова, путем выбора адекватных латинских, греческих и славянских понятий и терминов, сформировать новые или дополнить существующие ключевые понятия, ключевые слова и главное фреймы программного приложения «Структура понятий» — лексико-понятийно-концептуальной иерархии, находящей применение в процедурах сравнения контекстов источников, в том числе написанных на разных языках (см. рис. 3–4). В свою очередь, функция семантических и межъязыковых сопоставлений «Эквиваленты» призвана поддерживать процедуру определения и поиска почерпнутых в текстах источников, адекватных по своему содержанию и смыслу латинских, греческих

и славянских синтагм (рис. 4). Благодаря этому объектом сравнительного анализа становится грамматическая и концептуально-терминологическая атрибуция лексики.

В ходе построения названных функций предстоит интегрировать в единое информационное пространство данные таких программных приложений, как «Структура понятий», словари (лексиконы каждого памятника права и актов), «Имена» (банк имен собственных) и других информационных ресурсов<sup>6</sup>. К примеру, основу банка имен собственных составит информационный массив программного приложения «Имена», которое предназначено опосредовать репрезентацию сведений об именах собственных (рис. 6-7). В ходе модернизации программного приложения «Имена» в банк имен собственных предусматривается, с одной стороны, сопоставить его с БЗ «Аннотации», с другой дополнить структуру банка имен собственных разделом, который представлял бы в наглядном виде не просто референтные сведения о социальном и профессиональном статусе носителей личного

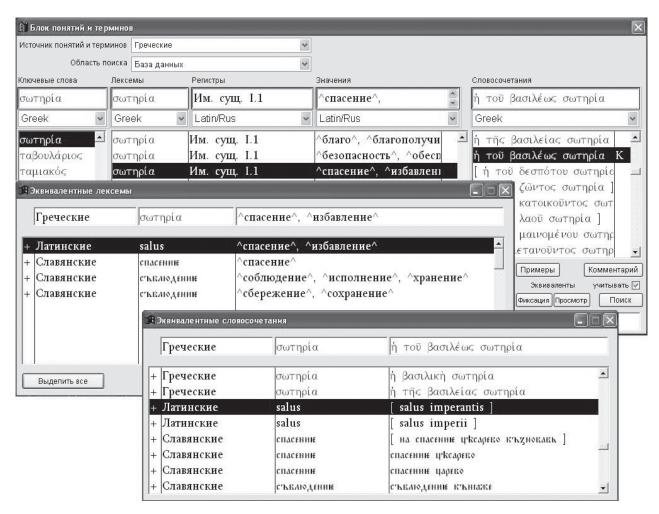


Рис. 5. Функция Блока понятий и терминов «Эквиваленты»

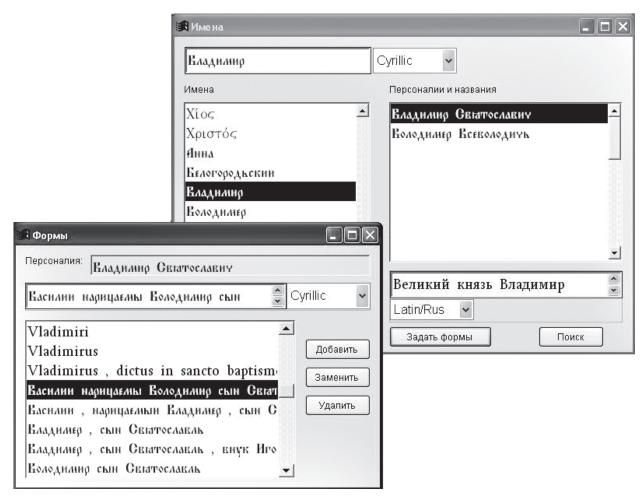


Рис. 6. Программное приложение «Имена». Структура данных

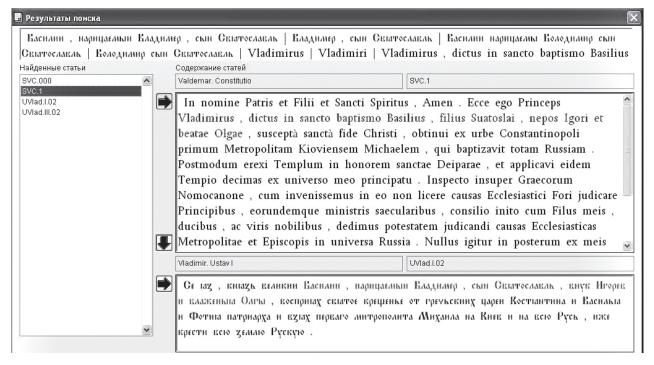


Рис. 7. Программное приложение «Имена». Результаты поиска

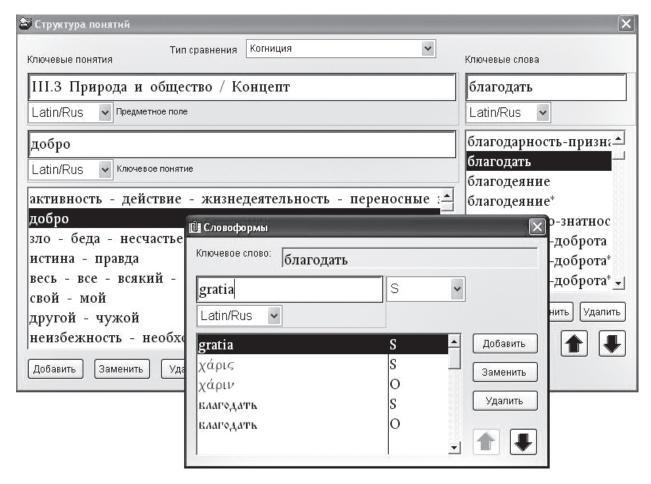


Рис. 8. Программное приложение «Структура понятий»

имени, как это делается в настоящей версии названного блока, а аутентичные данные по мере их выявления в праве и актах, соответственно требованиям современного научного уровня социального знания $^7$ .

В общем ряду указанных программных приложений нельзя не выделить особо уже упомянутый блок «Структура понятий» (рис. 8). Он поддерживает работу модуля определения информационной близости (МОИБ) сравниваемых контекстов источников, опосредуя их изучение в рамках концепции семантического поля исторического источника<sup>8</sup>.

Рассказывая о МОИБ, следует напомнить, что он выполняет статистический анализ понятий и терминов с учетом их функциональной роли по частотам употребления и дистрибуции лексики анализируемых фрагментов текста. МОИБ, следуя его названию, осуществляет определение информационной близости сравниваемых контекстов<sup>9</sup>. Практически начата разработка и построение модуля определения информационных расхождений (МОИР) и блока статистических данных (БСД). МОИР, действуя в условиях применения параметров сравнения контекстов аналогичных МОИБ, предназначен определять степень их информаци-

онного расхождения методом расчета средневзвешенного расхождения и процента невхождений лексических единиц сравниваемых контекстов. БСД же, по замыслу его разработчиков, будет решать задачи хранения и анализа статистических данных, а также метаданных (величин математически выраженных функционалов) путем расчета среднестатистических данных информационных расхождений совокупности текстов. Теоретической основой статистических счислений станут принципы предложенной ранее вниманию научного сообщества понятийно-лексической матрицы<sup>10</sup>. Анализ статистических данных, представленных в форме понятийно-лексической матрицы, нацелен на получение новых знаний.

Равным образом нужно охарактеризовать вновь построенные информационно-аналитических средства — функцию «Когнитивное картирование» (ФКК) и блок когнитивного картирования (БКК)<sup>11</sup>. ФКК помогает фиксировать логико-грамматические и семантические характеристики лексики сравниваемых контекстов (рис. 9). Когнитивное картирование служит средством акцентуации лексико-грамматической и семантической идентичности или различий в текстологических

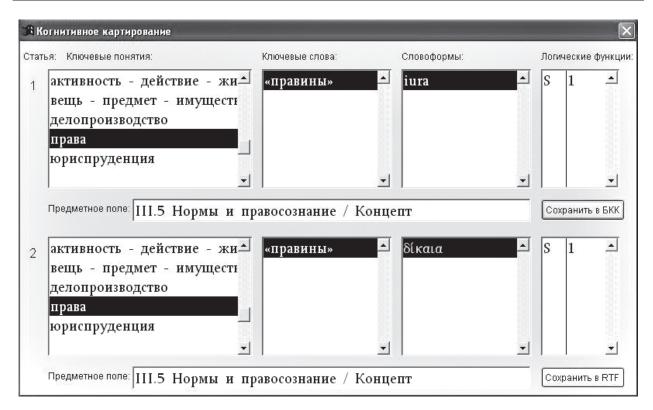


Рис. 9. Функция «Когнитивное картирование»

сравнениях путем создания когнитивных карт (КК) каждого понятия. Они наглядно представляют результаты изучения дистрибуции и семантики лексики согласно ее понятийному (в лингвистическом смысле слова) и терминологическому

содержанию, ее концептуализации и категоризации. Накопление этой информации происходит в виде интегрированных КК согласно закрепленным в их структуре правилам, что превращает БКК по сути в БЗ. Выступая в этом качестве, БКК

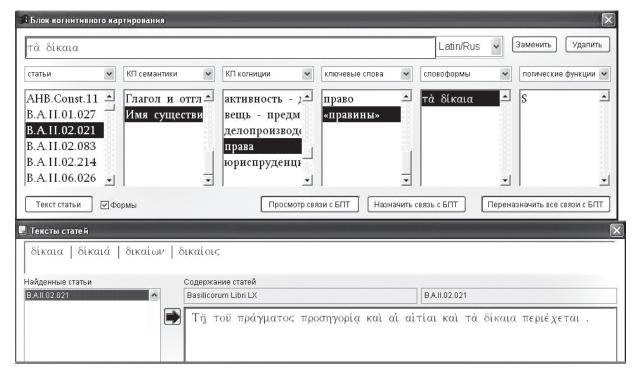


Рис. 10. Блок когнитивного картирования

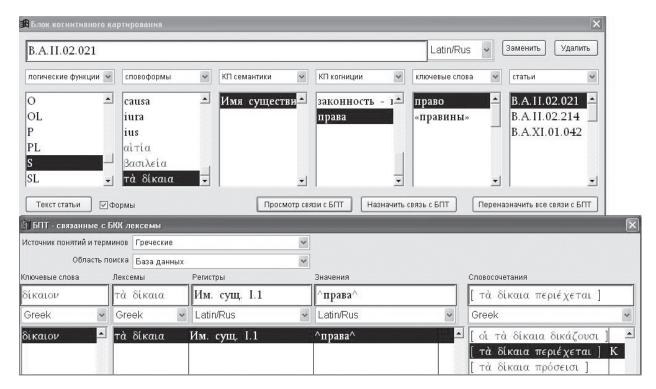


Рис. 11. Установление связи БКК и БПТ

нацелен на системное определение семантических категорий и значений вновь включаемых в него концептов или их реинтерпретацию в качестве ожидаемого нового результата познавательного процесса (рис. 10). Сопряжение данных БКК с информацией других программных приложений, прежде всего БПТ, обеспечивает последовательное наращивание «правил» анализа понятий и терминов (рис. 11).

С этой точки зрения особое значение приобретает связность понятийно-концептуальных черт лексики с присущей ей семантикой и смыслом, которым она наделена в контексте источника<sup>12</sup>. Не меньшей значимостью обладает изучение контекстуальной связанности понятийного аппарата источника, которую обусловливает частота вхождений слова в контекст. Это обстоятельство делает настоятельным применение статистического анализа и метода весовых оценок изучаемых словоупотреблений. Благодаря тому особое значение приобретает задача построения алгоритма понятийно-категориального анализа лексики исторического источника с приложением математического моделирования, которое учитывало бы, исходя из специфики дошедших до нашего времени текстов исторических источников, с одной стороны, научный потенциал современного научного исследования, с другой — особенности когнитивного развития человека прошлого.

Идеологию современной ЭС составляет свободный доступ к ознакомлению со знаниями ведущих

отраслевых специалистов, а не только посредством обращения, как уже было сказано, ко вторичной информации научной литературы или даже в виде текстуально представленной информации экспертных интервью, но и в режимах аудио-видеовоспроизведения их экспертных суждений. Решение указанных задач в форме модуля экспертных интервью ЭС, к построению которого оправданно приступить вслед за другими программными приложениями, потребует от участников проекта разработки оригинального программного обеспечения, учитывающего отраслевую специфику анализа византийских исторических источников и их славянских рецепций.

Здесь уместно заметить, что сама по себе византинистика представлена десятками выдающихся ученых, известных молодежи в лучшем случае по изданиям источников и публикациям монографий и статей преимущественно традиционного типа. Поэтому сбор и накопление текстуальных и в особенности аудио- и видеоинтервью экспертов превращает ЭС не просто в средоточие непосредственной информации, столь необходимой для осуществления аналитических процедур, а в хранилище самого ценного научного наследия — живого слова и облика ведущих ученых, большинство из которых принадлежит к старшему поколению. И если в естественных и технических отраслях знаний информация ЭС со временем в той или иной степени устаревает и теряет свою актуальность, то для гуманитарной науки, неотъемлемой часть которой является современное византиноведение, подобного рода экспертные интервью, несомненно, сохранят свою исконную значимость в качестве не имеющего преходящей цены научного наследия.

Успешная реализация проекта, конечно же, опирается на вовлечение многих специалистов различных профилей. В их сотрудничестве наиболее

трудная задача — научиться заимствовать научные достижения технических и гуманитарных отраслей знаний. Именно создание инновационной информационной технологии научного исследования раскрывает перспективы развития современной гуманитарной науки, включая византинистику ближайших десятилетий!

#### ПРИМЕЧАНИЯ

- <sup>1</sup> Vin Ju., Ja., Kondratjev D. E. Information Approach to Studying Byzantine Law and its Receptions // Byzantinoslavica. Praha, 2012. № 2. Р. 1–21. Также см.: Вин Ю.Я., Кондратьев Д. Е. Информационный подход к изучению византийского права: лексика и тексты // Византийские очерки. М., 2011. С. 45–66; Vin Ju. Ja. Information Approach to Studying Byzantine Law: the Lexis and Texts // Proceeding of the 22<sup>nd</sup> Intern. Congress of Byzantine Studies, Sofia, 22–27 August 2011. Vol. I: Plenary Papers. P. 595–615.
- <sup>2</sup> Быстрицкий Н.И. Развитие византиноведческих информационных ресурсов // Российское византиноведение: традиции и перспективы. М., 2011. С. 49–54. Ср.: Wechselbaumer R. Mittelalter virtuell: Mediävistik im Internet. Stuttgart, 2005.
- <sup>3</sup> См.: Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование. М., 2006 (Giarratano J., Riley G. Expert Systems: Principles and Programming. Boston, 2005); Джексон П. Введение в экспертные системы. М., 2001 (Jackson P. Introduction to Expert Systems. New York, 1998); Таунсенд К., Фохт Д. Проектирование и программная реализация экспертных систем на персональных ЭВМ. М., 1990 (Townsend C., Feucht D. Designing and programming Personal Systems. Tab Book Inc., 1986).
- <sup>4</sup> Андрейченков А. В., Егорова И. Е., Декатов Д. Е. Экспертные системы. Базы знаний. Инженерия знаний. Волгоград, 2007. С. 82–83. Также см.: Гаврилова Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. СПб., 2001. С. 19–20, 192–195; Гаскаров Д. В., Сикулер Д. В., Фомин В. В., Фомина И. К. Интеллектуальные информационные системы. Интеллектуальная информационная технология. Экспертные системы. СПб., 2004. С. 55 и далее; Рудакова Г. М. Искусственный интеллект. Экспертные системы. Красноярск, 2002. С. 10, 32, 50.
- <sup>5</sup> Подробнее см.: Вин Ю. Я., Кондратьев Д. Е. Инновационные принципы анализа лексики и текстов византийского права: когнитивный и информационный подход // Древнее право. М., 2010. № 1 (25). С. 131–132. Также см.: Вин Ю. Я., Гриднева А. Ю. Тезаурус понятийно-категориального аппарата византийских источников: правовой и религиозный аспекты. Концептуальный проект // Проблемы исторического познания. М., 2002. С. 141–196.
- <sup>6</sup> Вин Ю. Я., Кондратьев Д. Е. Специальные информационно-аналитические приложения СУБД «Византийское право»: принципы репрезентации и анализа понятий и терминов // Информационный бюллетень Ассоциации «История и компьютер» (далее ИБАИК). М.; Барнаул, 2008. № 35. С. 127–128.
- 7 Подробнее см.: Копосов Н. Е. Как думают историки. М., 2001. С. 102 и далее.
- <sup>8</sup> Вин Ю.Я., Гриднева А.Ю., Кондратьев Д.Е., Тихонова О.В. Концепция семантического поля исторического источника: основные положения // ИБАИК. М., 2003. № 31. С. 166–177; Они же. Концепция семантического поля исторического источника (проект информационного исследования) // Диалог со временем: альманах интеллектуальной истории. М., 2004. Вып. 12. С. 84–99; Вин Ю.Я. Концепция информационной близости исторического источника: анализ численных данных // Проблемы математической истории: Основания, информационные ресурсы, анализ данных. М., 2009. С. 94–115; Кондратьев Д.Е., Тихонова О.В. Алгоритм определения информационной близости исторических источников на основе статистического анализа понятий и терминов // Роль информации в формировании и развитии социума в историческом прошлом. М., 2004. С. 78–82.
- <sup>9</sup> Вин Ю. Я. Информационная технология исследований аутентичных понятий и терминов византийского права: программный модуль определения информационной близости // ИБАИК. М.; Тамбов, 2006. № 34. С. 38–39; Idem. The Information Technology of Studies on the Authentic Notions and Terms in Byzantine Law: The Programme Module of Determination of the Information Affin-

- ity // Proceeding of the 21th Intern. Congress of Byzantine Studies. London, 2006. Vol. 3. P. 111–112; Вин Ю. Я., Кондратьев Д. Е. Инновационные принципы анализа лексики и текстов византийского права... С. 132 и далее.
- Вин Ю. Я. Новые информационные ресурсы и принципы исследования аутентичных понятий и терминов византийских источников // Диалог со временем: альманах интеллектуальной истории. М., 2006. Вып. 17. С. 337–356; Он же. Информационная экспертиза аутентичных понятий и терминов исторических источников: матричный принцип // История и математика: анализ и моделирование социально-исторических процессов. М., 2007. С. 92–102.
- <sup>11</sup> Подробнее см.: Вин Ю. Я., Кондратьев Д. Е. Инновационные принципы анализа лексики и текстов византийского права... С. 135 и далее; Они же. Инновационные принципы анализа лексики византийского права и актов: когнитивное картирование // ИБАИК. М., 2012. № 39. С. 25–56, 41 и далее; Они же. Когнитивный подход смена парадигм исторического исследования: по материалам информационно-аналитического комплекса «Византийское право» // История: Электронный научно образовательный журнал. М., 2012. Вып. 6 (14) [Электронная ресурс:] URL: http://mes.igh.ru/magazine/announce/kognitivni-podhod.html.
- <sup>12</sup> Подробнее см.: Вин Ю. Я. Диалог социального и культурного в информационном исследовании византийских источников: понятийно-терминологический аспект как проблем исторического познания // Люди и тексты: исторический источник в социальном измерении. М., 2011. С. 54–109.