

ПРОБЛЕМЫ ПРЕДСТАВЛЕНИЯ ИСТОРИЧЕСКИХ ДАННЫХ В ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ

Н.И. Быстрицкий, П.В. Кузенков

Применение новейших компьютерных технологий уже дало впечатляющие результаты во многих сферах, причем не только в естественных науках, но и в прикладных областях, связанных с хранением и систематизацией большого количества данных (управление, финансы). В гуманитарных дисциплинах эффективное использование возможностей принципиально нового качества сталкивается со слабой степенью математизации и формализация гуманитарного знания: проблемы объективности и верифицируемости информации, ее адекватной артикулируемости и трансляции. Отсутствие эффективных методов решения указанных проблем в исторической науке порождает релятивизм и скептицизм в отношении традиционных методов и результатов исторического познания. Обратной стороной того же процесса стало распространение в обществе псевдонаучных теорий, претендующих на «математическую» ревизию и отрицание складывавшейся тысячелетиями исторической традиции [1].

Между тем, современное развитие информационных технологий позволяет предложить эффективный подход к указанной проблеме, выводящий исторические исследования на новый уровень без какого бы то ни было разрыва научной традиции. Но, несмотря на многочисленные попытки энтузиастов и наличие предварительных наработок отдельных зарубежных институтов, в сфере исторических исследований до сих пор не была предпринята попытка создания интегральной информационной системы, обладающей характеристиками орудия фундаментального научного исследования. Последнее подразумевает создание и интегрирование таких необходимых для работы историка ресурсов, как библиографические, просопографические, казуальные, генеалогические базы данных, обеспечение систематизации и обработки основных групп нарративных источников (литературные сочинения, акты и документы, письма и архивные материалы), обеспечение исследований по всему спектру традиционных направлений исторических исследований. Общедоступные на сегодняшний день исторические информационные ресурсы отличаются низким уровнем системности, фрагментарностью представленных материалов и не выходят за рамки традиционных справочных ресурсов [2;3].

Современный подход к созданию интегральной информационной системы исторического знания, по нашему мнению, должен включать следующие основные работы:

3. разработку теоретических основ и принципов формализации и представления исторической информации;
4. формирование строго обоснованного объективными критериями и формализованного свода исторической информации (базы данных исторических событий);
5. создание практических и теоретических методов обработки исторической информации;
6. верификация полученных результатов.

В настоящее время основные усилия сосредоточены на решении задачи первой стадии — разработке теоретических основ и принципов формализации и представления исторической информации. Собранный по мере ее решения опыт может быть положен в основу последующих, аналитических, этапов работы с наличными историческими данными. В ходе проделанных работ выявлены определенные теоретические проблемы, которые мы постараемся обобщить далее.

Как известно, фиксация на письме исторических событий покрывает примерно 5000 лет человеческой цивилизации, с существенными неравномерностями как по региональному распределению, так и по степени детализации. Таким образом, историческая наука имеет в своем распоряжении наиболее древний и богатейший комплекс информационных ресурсов [4]. Однако сам по себе концепт исторической информации нуждается в четком определении, без которого оказывается невозможна не только ее формализация и теоретическая обработка, но и организация хранения, передачи и верификации. Данная проблема особенно остро встает на этапе преобразования сведений нарративных, документальных и иных исторических источников в системно организованную электронную базу данных. Успешное решение проблемы научной формализации исторического знания должно быть достигнуто в ходе работы по сбору и классификации исторической информации. В первую очередь следует выработать систему критериев, рангов, хронологических и географических параметров и иных базовых характеристик исторического события. Серьезной логической проблемой является выявление информационных элементов исторического процесса: событий (фактов, происшествий, деяний), их участников и обстоятельств. Необходимо строгим образом сформулировать виды взаимных связей между событиями (причинно-следственные, типологические, генетические и др.). Наконец, важными задачами являются выявление и учет как индивидуальных, так и социальных факторов — проблемы, стоящие в центре внимания, соответственно, исторической психологии и социологии. Таким образом, для этого и последующих этапов компьютерного анализа и обработки требуется сведение массива разнородной

исторической информации к универсальной форме, приемлемой для накопления и хранения в цифровом виде — насущной оказывается необходимость разработки методов формализации и представления исторических сведений.

Историческая наука строит свои исследования на основе изучения событий прошлого, которые определяли состояния и изменение участников (объектов) исторического процесса (под субъектом исследования понимается в данном случае сам ученый-исследователь). Историческое событие (eventum) выступает как проявленные и зафиксированные состояния объекта исследования и отражает результат целенаправленных действий или нецеленаправленных процессов. Событие характеризуется конкретным значением пространственно-временных координат и текущим состоянием объекта. Как и в физике, принимается, что событие не обладает длительностью, либо длительность его пренебрежимо мала. Для однозначного описания временных характеристик исторических событий признано целесообразным использовать Хронологический юлианский день (последовательный счет суток с 0:00 ч. 1 января 4713 г. до н.э. юлианского календаря), широко применяемый в астрономии и компьютерных науках [5,594]; для описания пространственных данных применяются традиционные географические координаты. События могут находиться друг с другом в различных, в том числе и каузальных (причинно-следственных) отношениях. Одной из основных задач исторического исследования является выявление каузальных связей между событиями [6,110], осуществляемое на основе установления темпоральных и иных типов отношений.

Исследуемые объекты осуществляют самостоятельные действия или подвергаются внешним действиям или процессам. Действие (actus) по определению носит целенаправленный (целеполагающий) характер и осуществляется так называемым действующим объектом (objectum activum — A) в непосредственном отношении к пассивному объекту (objectum passivum — P); опосредованные отношения к действию характеризуют относительный объект (objectum relativum — R). Действие обладает определенной пространственно-временной длительностью и может быть представлено в виде вектора, начинающегося событием начала действия (initio) и заканчивающегося событием окончания действия (finis). Последнее является определяющим событием, так как именно оно характеризует анализируемое состояние объекта по результатам действия. Процессы определяются аналогично действиям, однако принципиально не обладают целенаправленностью (например, природные или случайные явления).

Онтологически основные виды действий/процессов можно разделить на основные категории согласно их форме (вещественной или информационной) и содержанию (перемещение, обработка, хранение). На основе такой классификации устанавливаются операнды действия, которые объединяют сходные виды действий (синонимы). Операнд действия определяет группу действий, которая может лингвистически выражаться через лексические единицы (глаголы) с семантически пересекающимися или включающими одно другое полями. Границы между полями являются, как правило, размытыми. Например, многие глаголы не могут быть однозначно расклассифицированы как глаголы познания или коммуникации, и т.п. Таким образом, приходится описывать глаголы как полисемантические, если обнаруживается, что они могут быть отнесены одновременно к разным семантическим полям. В этой связи текущей проблемой является создание полного списка глагольных лексем с непересекающимися семантическими полями.

По направленности действия подразделяются на типы:

- активные — один или несколько объектов, действующих вовне ($A > P$);
- обоюдные — взаимодействующие объекты ($A/P <> P/A$);
- возвратные — один объект, действия которого направлены на самого себя ($<A/P <$);
- пассивные — претерпевание внешнего действия ($P < A$).

Также действия могут быть представлены в различных формах:

- дуальная (стандартная): действующий объект — действие — пассивный объект;
- унарная: действующий объект — действие; действие — пассивный объект;
- плюральная: действующий объект — действие — пассивный объект — относительный(ы/е) объект(ы)

Возможность действия обуславливается наличием определенного отношения (relatio) между объектами. Вид этого отношения устанавливает некоторое множество возможных видов действий. Таким образом, особого рода задачу представляет собой выявление и фиксация различных видов отношений между объектами.

Ряд важных теоретических проблем связан с вероятностной природой исторической информации. Необходимо, в частности, формализация достоверности тех или иных источников. В частности, каждое нарративное сообщение имеет определенную достоверность, связанную с субъективностью автора, степенью его информированности, зависимостью от другого источника, жанровых особенностей и проч. В отношении археологических и иных «немых» источников на первый план выходит оценка убедительности исследовательских гипотез и интерпретаций. Данную проблему теоретически можно решать при помощи экспертного ранжирования и вероятностных методов, однако ее практическое решение нуждается в серьезной апробации на примере конкретного исторического материала.

Другой проблемой является адекватная фиксация временных характеристик. В качестве типовой темпоральной характеристики действия выступает временной интервал. Для отношений между временными

интервалами используется аппарат интервальной временной логики, разработанный в работах Джеймса Ф. Аллена, Йохана ван Бенгема, Энтони Гальтона и др. [7-11]. Во многих случаях даты исторических событий нам известны лишь с той или иной степенью приближения, а иногда имеется несколько вариантов датировки одного события. Одним из подходов видится применение алгебры интервалов с нечеткими границами [12-14], который использует вероятностные методы. Это дает возможность создавать варианты одного и того же события с различными датами, назначая каждому варианту особую вероятностную характеристику.

В принципе, все исторические данные обладают вероятностными характеристиками. Разработка методов их определения (на основе экспертных оценок, типовых моделей и проч.), а также манипулирования и верификации представляется важной практической задачей.

НИР по разработке основ и принципов формализации и представления исторической информации, как мы убедились, необходимо проводить на трансдисциплинарной основе [15] с применением достижений различных отраслей наук: источниковедения, исторической информатики, логики, лингвистики и др.

В целях апробации намеченных теоретических подходов осуществляется создание на их основе опытной базы формализованной событийной информации для узкого хроно-топологического сегмента всемирной истории (Византийская империя, 900–950 гг.). Избранный период, с одной стороны, обеспечен достаточным, но вполне обозримым материалом нарративных и иных источников, а с другой — характеризуется наличием ряда неопределенных (гипотетичных) исторических данных (вариативность дат и проблема отождествления/разделения лиц, с вытекающими отсюда вариантами темпоральных и каузативных отношений). Это дает возможность отработать теоретические и практические решения типовых проблем на конкретном историческом материале. Кроме того, формализация событийного ряда сопровождается разработкой интегрирования разных типов научной информации (баз данных). Надеемся, что в ближайшее время можно будет говорить о достижении конкретных исследовательских результатов, которые смогут быть представлены для критического осмысления научному сообществу.

ЛИТЕРАТУРА:

1. Хвостова К.В. Математические методы в исторических исследованиях и современная эпистемология истории // Новая и новейшая история, №3, 2007. С. 66-78.
2. Куликов В.А. Пошук історичної інформації в інтернеті: Посібник для студентів історичних факультетів. – Харків: Вид-во ХНУ імені В. Н. Каразіна, 2007. – 91 с.
3. Быстрицкий Н.И. Развитие византиноведческих информационных ресурсов // Российское византиноведение: Традиции и перспективы. Тезисы докладов XIX Всероссийской научной сессии византинистов. М.: Издательство Московского университета, 2011. С. 49-54.
4. Кузенков П.В. Споры о возрасте мира в Византии VII–XI вв. (О трех мировых эрах: александрийской, "протовизантийской" и византийской) // Византийский временник. 2007. Т.66 (91). С. 93-124.
5. Herschel J.F.W. Outlines of Astronomy – Philadelphia: Leo&Blanchard, 1849. [переиздание Cambridge University Press, 2010]
6. Блок М. Апология истории. М.: Наука, 1986. С.
7. Allen J.F. Maintaining Knowledge about Temporal Intervals // Communications of the ACM, November 1983, Volume 26, Number 11. -P. 832-843.
8. Allen J.F. Time and Time Again: The Many Ways to Represent Time // The International Journal of Intelligent Systems, 6(4), July 1991. P. 341-355.
9. Benthem, J. van. The Logic of Time. Dordrecht, Boston and London: Kluwer Academic Publishers, first edition 1983 (second edition, 1991).
10. Benthem, J. van. Points on Time // Electronic News Journal on Reasoning about Actions and Change, vol. 2 1998.
11. Galton, A. Temporal Logic // The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), 1999.
12. Dubois, D., Prade, H.: Processing fuzzy temporal knowledge. IEEE Transactions of Systems, Man and Cybernetics 19, 1989. p.729–744.
13. Badaloni S., Giacomini M. A fuzzy extension of Allen's interval algebra. In E. Lamma and P. Mello, editors, AIA99: Advances in Artificial Intelligence, p. 155-165. Springer, 2000.
14. Ohlbach H. Fuzzy Time Intervals - The FuTI-Library // Forschungsbericht. Institute for Informatics, University of Munich, 2005
15. Morin E. La tête bien faite. Repenser la réforme ↔ Réformer la pensée. Paris: Editions du Seuil, 1999. P. 136.