**Опыт создания Базы Данных для источников личного происхождения**

Алимгазинов К.

Систематизация исторического материала для последующего его изучения, если принять во внимание, что исследуется большой массив источников, обуславливается конечными исследовательскими возможностями его исследования. Массовые исторические источники уже получили опыт систематизации при приведении их данных в унифицированные анкеты, таблицы .

Систематизация источников личного происхождения среди историков в основном носит черты классификации на уровне выделения внешних и внутренних атрибутивных признаков, определения вопросов авторства . Однако такая классификация затрудняет составление общей картины на основе содержания данных всех источников, когда исследователь намерен провести анализ всего мемуарного комплекса. Необходимость представления общей картины всего комплекса воспоминаний подводит нас к необходимости унифицирования внешней и внутренней структуры источника. Подобное возможно при соответствующей методике путем составления первоначальной базы данных (БД).

Цель такой работы – построение вариационных рядов признаков исследуемых источников, и, соответственно этому, проведение на их основе источниковедческого анализа источников личного происхождения. Опыт создания статистической базы данных применительно к историческим источникам уже не ново. Однако в нашей отечественной науке это не получило еще широкого распространения. Использование подобной методики проводится применительно к источникам, где содержатся информация, характеризующаяся ее количественным проявлением во всем массиве источников: переписи, анкеты, и т.д. . Теоретические аспекты использования математических приемов изучения источников личного происхождения , особенности составления БД рассмотрены в работах ученых дальнего и ближнего зарубежья, некоторые стороны этих вопросов затронуты в работах отечественных исследователей – сторонников квантитативного направления исторической науки.

Особенности воспоминаний, как повествовательных исторических источников, заключаются прежде всего в специфике сохранности и освещения в них информации. Как считает западный исследователь квантитативной истории, если суммировать признаки, говорящие о специфике исторических источников применительно к уровню их структурированности и совместимости с моделями, на которых основаны системы управления базами данных, то большинство часто встречающихся трудностей можно свести к различным видам нестандарности (нерегулярности) данных, вполне обычным для исторических источников .

Мемуарные источники относятся к тем источникам, где содержащимся данным приписывается свойство зависимости от личностных ориентировок автора, как по характеру изложения, так и по фактам, получившим место в источнике. Таким образом, воспоминания как ни один другой исторический источник характеризуются отрывочностью приводимых автором сведений, обусловленностью как авторскими предпочтениями перед выбором фактов, так свойствами памяти. "Большинство признаков, характеризующих исторические объекты, - писал еще Ковальченко И.Д., - являются качественными, атрибутивными. Они не имеют единиц измерения, т.е. соотвествующее свойство не может быть выражено в тех или иных метрических показателях" . Это затрудняет приведение источника к подобию унифицированной анкеты, где каждая строка и элемент будут иметь значение согласно принципам построения базы данных (БД). Поэтому, первоначальная задача систематизации заключалась к поиске таких признаков источника, которым свойственно проявление на протяжении всего массива источников, и способных создать основу для составления исследовательских гипотез.

Одним из самых распространенных среди современных отечественных и российских исследователей является методика определения количественных показателей путем подсчета выделенных свойств каждого объекта, в данном случае это будет отдельный мемуарист и его произведение, в контексте всех исследуемых объектов (общего массива всех мемуарных источников).

При построении БД был учтен опыт проведения подобных исследований в работах историков-клиометриков. Принципная основа аналогична существующим БД.

Применительно к историческим источникам западными и российскими историками разработано несколько получивших наибольшую известность прикладных реляционных программ . В основном они предназначены для исследования источников, происхождение которых не оставляет сомнений в их массовости (анкеты, статистические источники по аграрному сектору, источники по структуре народонаселения и другие).

В нашем случае мы имеем перед собой источник, который при комплексном подходе к нему как к общей единой массе, несет в себе статистическую информацию (деление по годам создания, авторскому составу, по периодам освещения событий). Кроме этого, в источниках содержится значительный элемент текстовой информации, которому свойственна как неоднородность, так и нерегулярность проявления в текстах мемуарных источников, названная клиометриками как "свободный текст". Жесткая конструкция реляционной модели БД значительно затрудняет использовать эти сведения. Механическое перенесение элементов текста в структуру БД, во-первых, невозможно ввиду нерегламентированности первого, а во-вторых, могло бы привести пользователя БД к искаженным впечатлениям.

С другой стороны, пренебрегать этим родом сведений, тем самым занижая преднамеренно источниковедческую ценность воспоминаний, было бы непростительно. Актуальность разрешения этих проблем в контексте применения историками математических приемов изучения источников, уже затронута в исследованиях историков-клиометриков. Концепция создания реляционной модели, проведение статистического анализа для массива источников личного происхождения исходила, в нашем случае, из постулата проблемно-ориентированного подхода при изучении исторических источников, предложенным западным историком Манфред Таллером .

Затронутые выше проблемы "свободного текста" мы попытались решить при помощи проведения над воспоминаниями апробированного уже и зарекомендовавшего себя с положительной стороны метода контент-анализа. В результате были выявлены 6 признаков, дающих нам представление о социально-психологических процессах в сознании участников гражданской войны в Казахстане. В совокупности эти признаки составили группу, именуемую "образ врага". Ниже же мы остановимся на особенностях составления применительно к мемуарам БД. Особенности исторических источников частного происхождения подводят нас к проблеме выбора адекватного программного обеспечения для возможности при его использовании достижения поставленных исследовательских целей. Проблемы выработки специализированного программного обеспечения с учетом специфики исторических источников, содержащих "свободный текст", судя по дискуссиям на международных семинарах "История и Компьютер" , актуальны. Распространение же существующих прикладных программ носит коммерческий характер. Таким образом, исходя из объективно сложившихся обстоятельств, нами была проведена попытка построения реляционной модели БД на основе имеющегося пакета программ, работающих в среде WINDOUS. Большинство историков, как считает Гарскова И.М., поступают подобным образом, довольствуясь имеющимися прикладными программами как вполне отвечающими требованиям при создании, управлении, обработке БД в соответствии с текущими исследователькими задачами . Основные аспекты построения модели соответствуют конструктивным основам имеющейся системы SOCRATES, разработанной авторитетным специалистом из Голландии в области применения технологии баз данных в исторических исследованиях – Л.Брере.

Модель БД состоит из нескольких таблиц, представляющих из себя двумерный массив типа "объект – признак". Другими словами, это – созданные для фиксации различных признаков объектов унифицированные анкеты, которые заполнялись нами по мере изучения исторического материала.

Всего было изучено 42 признака, дающих представление об изучаемом мемуарном комплексе. Из них 6 выделенных признаков обнаружены путем проведения над воспоминаниями контент-аналитической процедуры работы.

Массив выделенных признаков вошел в 13 групп, на уровне которых проводились исследования коэффициентов корреляции. При определении групп признаков мы придерживались основных классификационных приемов систематизации (однородности признаков, знания закономерности связей), а также следуя конечным исследовательским целям.

Итак, выделенные признаки составили 13 рядов: - Дата рождения - Дата смерти - Пол - Социальный статус накануне событий - Общественный статус в событиях - Национальность - Образование - Идейная (политическая) ориентация мемуариста - Географический срез: территориальная представленность событий в воспоминаниях - Хронологические рамки событий - Редакции воспоминаний (во временном аспекте) - Форма источника - Содержание источника.

Таким образом, содержательная структура данных рядов позволяет нам выделить 5 групп признаков. В дальнейшем, характеристика взаимозависимости проводилась в плоскости межгрупповых ее проявлений:

1. личность автора по 4 признакам (дата рождения, дата смерти, пол, национальность, образование)

2. роль автора в событиях по 3 (соц.статус, общест.статус, идейная ориентация)

3. представленность фактов по 2 (географический срез, хронологич.рамки событий)

4. атрибутивные свойства произведений по 2 (редакция воспоминаний и форма источника)

5. Содержание источника по 6 (признаки категории "образ врага"). Вариационный же ряд был представлен числом мемуаристов.

Так, "форма источника" включала в себя три составляющих переменных: 1 – собственно воспоминания, 2 – воспоминания в форме историко-критических трудов, 3 – воспоминания служебной формы назначения. Таким же образом каждая группа состояла из "атомарных" составляющих.

Исследование БД проводилось на основе анализа парных связей вариационного ряда с последующим графическим представлением и интерпретацией. Работа преследовала цели определения характерных свойств признаков, тенденций внутри вариационного ряда. Говоря словами И.Д. Ковальченко, "установить абсолютную и относительную меру рассматриваемых черт и свойств объектов и явлений и выявить интенсивность их проявления..." .

Исследование количественных показателей вариационного ряда для последующего процесса бифуркации значений на основе результатов анализа бинарных и тернарных связей проводилось при визуальном отображении логических структур зависимостей, т.е. графически . Это позволило создать первоначальное впечатление от прослеженных связей и позже сравнить его с полученными результатами анализа.

Анализ вариационного ряда парных зависимых признаков раскрыл социальный состав мемуаристов, половозрастную структуру авторов, полноту географической представленности сведений, характерные тенденции авторов для освещения отдельных периодов гражданской войны.

К примеру, изучение возрастной группы мемуаристов показало черту активности молодой части населения края в событиях: к 1918 году среднестатистический возраст участника событий составил 26,7 лет. "Весь комполитсостав, включая комдива... был очень молодой (20-25 и 30 лет)." , - вспоминал позже один из мемуаристов.

Национальный состав мемуаристов, хотя и представлен разнообразно, однако исследование всего массива источников и определение процентных соотношений каждого из признаков к общему количеству выявило характерное превалирование источников авторов некоренных национальностей: 65% из всего состава изученных мемуаристов – выходцы из регионов России и 3,4% относятся к бывшим европейцам-военопленным и ссыльным. Далее при освещении общественного статуса мемуаристов была определена степень значимости для ее характеристики этого удельного веса, а также зависимость ее от занимаемого мемуаристом общественного положения в период гражданской войны. Так, подавляющая часть выделенного состава авторов значилась в рядах "лидеров вооруженной борьбы локального характера" и "идейных вдохновителях". Сопоставление полученных данных наряду с исследованием фактора географической представленности мемуаристов показало, что инициатива общественно-политических преобразований в крае носила характер привнесенного элемента – идейными вдохновителями революционых идей выступило в 28 % случаях элемент демобилизованных недавно из царской армии мемуаристов. Эти сведения, полученные путем статистических вычислений, не противоречат другим историческим документам. Выступая на первой краевой конференции мусульманских организаций РКП (24-30 мая 1919 год) в Ташкенте, один из организаторов советской власти в крае отметил факт неучастия коренных наций к происходящим событиям в крае. "В Туркестанской республике, - сказал П. Кобозев, - 95% мусульман, и только 5% русских, и эти 5% держат в своих руках всю власть".

Построение логической структуры связи между тремя приизнаками практическое воплощение получило при определении тенденции складывания мемуарного комплекса в промежуток выбранных хронологических рамок, анализ ее зависимости от формы источника. При обнаруженных последовательных связях трех признаков промежуточное звено принималось нами либо как интерпретирующий фактор, либо как объясняющий.

Иногда для анализа парной зависимости нами вводился третий признак. К примеру анализ связи при автономном изучении признаков "уровень образованности" и "форма источника", естественно, невозможен ввиду отсутствия логической взаимосвязанности признаков. Для обнаружения возможной связи был введен третий признак, связь которого с двумя первыми очевидна – "мемуарист". Теперь предстояло путем проведения статистических процедур определить степень возможной зависимости, проявившиеся тенденции вариационного ряда или же констатировать отсутствие таковой. Изучение показало, что степень образованности мемуаристов оказывает значительное влияние на форму изложения материала воспоминаний, и соответственно, на степень доказательности фактологической стороны текста. Дальнейший анализ парных корреляций признаков показал, что воспоминания участников с начальным образованием в основном сложился в годы 1920-1930-е гг. и в период 1957-1962 гг. Далее был проведен историко-сравнительный анализ полученных данных, с историческими условиями появления мемуарного комплекса за данные промежутки времени. Именно в исследуемые хронологические рамки 1920-1930-х гг. и 1957-1962-х гг. сбор воспоминаний осуществлялся органами истпарта и позже научными экспедициями и носил характер устного опроса участников. Далее, характеризуемые промежутки времени обусловлены общественным и политическим интересом к появлению воспоминаний – проведением, в частности, юбилейных мероприятий. Таким образом, последние два фактора могут объяснить появление именно этой формы воспоминаний.

Изучение комплекса этих источников показало отсутствие у авторов литературных навыков к написанию, чем и объясняется в значительной степени появление воспоминаний в форме литературной записи. Историко-критические очерки и мемуары в служебных документах прежде всего предполагают иницитиву со стороны самого участника событий к возникновению воспоминаний: преследование определенных целей, будь то научных, общественно-политических или социально-практических. Характеристику данных источников следует производить с позиции мировозренческих и с точки зрения возможности автора к осмыслению происходящего, где регулятором уровня осознания и освоения жизненных реалии, способности самоидентификации в общественном пространстве выступает образовательный уровень автора.

Итак, определение свойства зависимости "формы источника" и "уровень образования" позволяет говорить о существовании опосредованной (непрямой) структуры зависимости между ними, где "мемуарист" выступает в качестве интерпретирующего звена структуры связи.

Изучение и анализ связей между признаками подводит нас к обобщению источникового материала на более высоком уровне – на уровне выделенных нами ранее 4 основных групп категорий. Количественные показатели на данном этапе переходили в процессе бифуркации в качественные признаки – интерпретация численных показателей географического плана – в степень региональной представленности событий в мемуарах, половозрастной – определение характерных личностных и общественных мировозренческих концепций мемуаристов, количественные показатели связи формы источника и индивидуальных особенностей автора – претерпевали процесс интерпретации в аспекте определения наметившейся тенденций в призме историко-сравнительного анализа.

Главный принцип, которого мы придерживались при исследований источников частного происхождения – оперирование только фактическими данными. Данными, которые не оставляют сомнений и в некотором роде выступают в качестве аксиоматической парадигмы. Общеизвестна, например, структура зависимости года рождения и национальности с конкретным (в нашем случае) мемуаристом. Не подвержены различным толкованиям также и выделеные из содержания текста мемуаров такие элементы, как хронология изложения фактов, география событий.

**Список литературы**

Моисеенко Т.Л. Методы анализа маркетизации крестьянского хозяйства России // ЭВМ и математические методы в исторических исследованиях. М., 1994. – С.74-93., Моисеенко Т.Л. Методы типизации крестьянских хозяйств по сводным данным земской статистики // Математические методы и ЭВМ в историко-типологических исследованиях. М.,1989., Литвак К.Б. Вопросы методики изучения социально-экономического строя крестьянского хозяйства по данным земских переписей // Применение количественных методов в исследованиях по аграрной истории. М.,1984., и другие. ?i?aoiaoia I. ?aca? iaioa?uiu» aaio oa?eou. // Известия АН КазССР. Серия общественная. – 1966. - №3. – С.38-44., Тартаковский А.Г. 1812 год и русская мемуаристика. Опыт источниковедческого изучения. М., Наука, 1980. – 312 с.

Ковальченко И.Д. Методы историчекого исследования. – М., Наука, 1987. – 439 с. – С. 336 и далее.

Волкова Т.П. Некоторые проблемы источниковедческого изучения материалов экспедиции Ф.А. Щербины. // Вопросы историографии Казахстана. Алма-Ата, 1983. – С.224., Жакишева С.А. Применение математических методов и компьютерных технологий в исследовании истории репресий в Казахстане в 20-30-х гг. ХХ века. // Культура и история Средней Азии и Казахстана: проблемы и перспективы исследования. Алматы, 1997., Она же. Использование технологии баз и банков данных в исторических исследованиях // Компьютеризация образования: проблемы и перспективы. Материалы республиканской конференции (26-28 мая, 1998 год, Алматы). – Алматы, Казак университеты, 1998.

Методологические проблемы исторической информатики и квантитативной истории. // Новая и новейшая история. 1997. - № 3. – С.87-98., Бородкин Л.И. Многомерный статистический анализ в исторических исследованиях. М., МГУ, 1986., Он же. О некоторых аспектах автоматизации текстологических исследований. // Математические методы в историко-экономических и историко-культурных исследованиях. М.,1977., Ковальченко И.Д. Исторический источник в свете учения об информации. К постановке проблемы. // История СССР. – 1982. - № 3. – С.131-145., Он же. Теоретико-методологические проблемы исторических исследований и размышления о новых подходах. // Новая и новейшая история. – 1995. - №1. –С. 5-11., Он же. Методы исторического исследования. М.,Наука, 1987. – 439 с. – С.281-435., Гуревич А.Я. Историк конца ХХ века в поисках метода. // Одиссей. Человек в истории. Ремесло историка на исходе ХХ века. М., 1996. – С.5-7., Бокарев Ю.П. Историческая информатика как новая вспомогательная историческая дисциплина. // ЭВМ и математические методы в исторических исследованиях. М., 1993. – С.4-8., и другие.

Гарскова И.М. базы и банки данных в исторических исследованиях. М., МГУ., 1994. – 215 с., Деопик Д.В. Некоторые принципы построения формализованных языков для исследования исторических источников. // Количественные методы в гуманитарных науках. М., МГУ, 1981. – С.5-10. Жакишева С.А. Баи-полуфеодалы в Казахстане на рубеже 20-30 гг. ХХ века: историко-источниковедческий аспект проблемы. Дисс...к.и.н. – А.,1996., Волкова Т.П. Некоторые проблемы источниковедческого изучения материалов экспедиции Ф.А. Щербины. // Вопросы историографии Казахстана. Алма-Ата, 1983. – С.224.

Гарскова И.М. Базы и банки данных в исторических исследованиях. М., МГУ., 1994. 215 с. – С.111.

Ковальченко И.Д. Методы исторического исследования.... – С.348. Гарскова И.М. базы и банки..... С.132.

О двух подходах при изучении исторических источников – "проблемно-ориентированном" и "источнико-ориентированном" см. ж-л "Новая и новейшая история", за 1995-1998 гг.

Ковальченко И.Д. Методы исторического исследования... – С.315.

Термин, введенный в научный оборот бельгийской школой И.Пригожина. Используется при характеристике процесса перехода количественных показателей в качественные. С квантификацией исторической науки данный термин российскими и белорусскими исследователями активно используется в исторических научных работах (Селунская Н.Б. От квантификации к исторической информатике – от исторической к виртуальной реальности? // Новая и новейшая история. – 1997. - №5. – С.85., Бородкин Л.И. Квантитативная история в системе координат модернизма и постмодернизма. // Новая и новейшая история. – 1998. -№5. – С.16, Методология истории. Учебное пособие. Минск, 1996. – С.222. ("Бифуркация – точка ветвления процесса, порождающая новую линию эволюции").

Социологические исследования: методология, программа, методы. М., Наука, 1987. – С.190-195.

Евлампиев. Последнее предложение. // Октябрь в Казахстане. Очерки и рассказы участников гражданской войны. Составитель Г.Н. Мельников. Под ред. К.С. Сарымулдаева, Н.Н. Синицына. Алма-Ата, Издание центрального совета осоавиахима КАССР и Краевой юбилейной комиссии, 1930. - 236 с. – С.67.

Мусбюро истпарта в Туркестане. Ташкент, Госиздат, 1921. - 93 с. – С.5.