ІСТОРІОГРАФІЯ ТА ДЖЕРЕЛОЗНАВСТВО

Ю. С. Ганжуров (Київ)

Кількісні методи в історичних дослідженнях: досвід і деякі питання методики

Нині, коли наша держава переживає складний період свого становлення, набуває особливої ваги неупереджене вивчення історичного досвіду України. Проте саме на цьому етапі, в умовах помітної руйнації системи ціннісних орієнтирів тоталітарної доби і формування нових концептуальних підходів до аналізу суспільних явищ і процесів, на перший план висувається проблема запобігання новим кон'юнктурним оцінкам, хибному збоченню в наукових дослідженнях на догоду тим, хто, як й в минулі часи, хоче видавати бажане за дійсне. Однією з запорук того, щоб кожного року не переписувати чергову «концепцію» історії України, є застосування надійних методів дослідження. Не випадково М. А. Берг зауважував: «Процес історіописання слід розуміти в двох смислах: як процес сприйняття, «дешифровки» і впорядкування досвіду минулого з метою тлумачення його в світі досвіду сучасності та як метод реалізації такої програми» [1].

Багато у чому комплексному розв'язанню цього завдання сприяє кліометрика, спеціальна історична дисципліна, що має на меті формалізацію якісних ознак суспільних явищ і процесів у кількісні показники з наступною інтерпретацією отриманих результатів.

Одним з найважливіших стимулів у поширенні математико-статистичних методів є прагнення дослідників дати більш широке і об'єктивне тлумачення фактів, що містяться в джерелах, зробити аналіз джерел на більш високому науковому рівні, точніше охарактеризувати тенденції та закономірності розвитку історичних явищ й процесів.

Як відомо, обсяг інформації в галузі управління подвоюється кожні 12—15 років, а в архівах — кожні 20—30 років, причому темпи зростання збільшуються [2]. Через це складність, яку відчувають історики при вивченні масових сукупностей текстів як історичних джерел, призводить у разі їхнього неповного опрацювання до чисто ілюстративного підходу, до фрагментарності у висвітленні подій.

В зв'язку з цим набувають особливої ваги кількісні методи аналізу історичних джерел (історичні хроніки, історико-публіцистичні твори, періодична преса, документи різних організацій, мемуари, щоденники, автобіографії, приватне листування тощо). Звичайно, їх доцільно використовувати поряд із традиційними, чисто описовими методами вивчення джерел. Як зазначає І. Д. Ковальченко, «сильний бік описових методів — їх конкретність і образність, а кількісних — їх глибина і точність» [3]. Інакше кажучи, якщо кількісні методи виступають дійовим інструментом виявлення певних закономірностей історичних процесів і через це мають велику силу доказовості, описові (в кращому розумінні) методи сприяють ідіографічному відображенню в історичних дослідженнях окремих конкретно-історичних явищ. До речі, за кордоном вже тривалий період досить помітну увагу приділяють застосуванню кількісних методів, розуміючи їхню ефективність. Зокрема, розділ «Контент-аналіз» включений у склад підручників з кількісних методів в історичних дослідженнях. Деякі журнали в США регулярно публікують статті, в яких «контент-аналіз застосовується для вив-

чення змісту історичних джерел»[4]. Незважаючи на вельми обмежений масив оригінальної західної літератури з питань кліометрики, певне уявлення про перспективи цієї спеціальної історичної дисципліни за кордоном дають історіографічні огляди, аналітичні матеріали на зазначену проблематику[5]. Аналіз відповідних джерел свідчить, що діяльність західних істориків дозволила значно збагатити зміст економічної, політичної та соціальної історії, сприяючи створенню умов для подальшого розвитку кліометрики від обгрунтування загальних правил використання математико-статистичних методів до конкретизації їх використання в певних галузях історичного пізнання, таких, як, приміром, «нова історія міста», історія сім'ї і демографічної поведінки, вивчення «колективних біографій» різних соціальних, професійних і етнічних груп населення тощо*.

Що стосується східних істориків-кліометристів, то вони здебільшого застосовували кількісні методи в дослідженні джерел, котрі відкривали особливості аграрної історії Росії, оскільки саме в цій галузі є найбільше дискусійних питань, котрі можна було розв'язати завдяки вдосконаленню кількісних методів опрацювання нових комплексів джерел. Зокрема І. Д. Ковальченко, вивчаючи систему експлуатації російських кріпаків, відзначив, що без математичних методів було б неможливо здійснити зрівняльний аналіз факторів розвитку землеробства і соціально-економічних відносин в російських селах першої половини XIX ст. Визначивши за допомогою коефіцієнта лінійної кореляції залежність продуктивності різних категорій кріпаків у 40—50-х рр. XIX ст. від інтенсивності їхньої експлуатації, І. Д. Ковальченко зробив висновок про те, що розвиток кріпацького господарства у першій половині XIX ст. було обумовлено передусім розміром феодальної ренти[6].

Методика використання кількісних методів у зазначеній галузі, започаткована цим вченим, набула помітного поширення серед істориків. Так Ю. П. Бокарєв зробив вдалу спробу простежити динаміку співвідношення розвитку промисловості і сільського господарства Росії в середині 20-х рр. XX ст. Для з'ясування багатьох аспектів взаємовідносин промисловості і сільського господарства автор вдало використав математичні методи. В результаті всебічного вивчення бюджетів сільського господарства, обстежень промислової документації, інших джерел він дійшов висновку, що диспропорція між розвитком сільського господарства і промисловістю виникла в результаті розпорошеності сільського господарства[7].

Посилений інтерес істориків-аграрників до проблеми використання математичних методів, широке обговорення цієї проблеми відклалося й в матеріалах відповідних наукових конференцій. Так, наприклад, на XIX сесії Міжреспубліканського симпозіуму з питань аграрної історії Східної Європи (Кишинів, 1976 р.) було представлено понад 10 доповідей і повідомлень, в яких математичні методи дозволили отримати важливі історичні висновки: це свідчило про те, що вони посіли в арсеналі історика помітне місце[8].

Аналіз відповідної галузевої літератури з питань застосування кількісних методів при дослідженні історичних джерел[9] свідчить, що помітне місце в арсеналі кліометристів належить контент-аналізу як вельми ефективному прийому підвищення інформативної віддачі історичного джерела. Проблемі використання зазначеного методу в історичній науці присвячений чималий масив літератури, в якому містяться як публікації із власне методики контент-аналізу[10], так і роботи, підготовлені за результатами використання цього методу в конкретних дослідженнях[11]. Звичайно, що в процесі його використання науковці не тільки опрацьовують отримані результати, а й прагнуть також до вдосконалення методики та до упорядкування понятійного змісту, конкретизації сутності цього методу. Л. Д. Дергачова, посилаючись

на методичні розробки новосибірських фахівців у галузі соціології, розглядає «контент-аналіз як метод формалізованого спостереження і статистичних процедур при аналізі масових сукупностей текстів» [12]. Це визначення в принципі не викликає заперечень, хоча з боку сутнісної конкретизації у даній редакції нічим не відрізняється від, наприклад, кореляційного методу. Це пояснюється тим, що у представленому формулюванні не зафіксовані основні функціональні ознаки контент-аналізу, механізм його реалізації на практиці.

З огляду на це заслуговує на увагу стисле визначення техніки застосування методу контент-аналіз до матеріалів преси, зроблене Л. І. Буряк, яка зазначає, що «процедура цього методу полягає у розбивці тексту на окремі змістові одиниці рахунку і в наступному визначенні площі, що вони займають». На думку дослідниці, основним ланцюгом є складання категоріальної схеми, що віддзеркалює основні аспекти проблеми, визначення одиниці обчислення для наступного опрацювання формалізованого масиву інформації на ЕОМ з метою наукової інтерпретації отриманих результатів [13].

Що стосується визначення самої сутності контент-аналізу в найбільш узагальненому вигляді, то він в принципі дійсно полягає у підрахунках певних ознак джерела, котрі мають відбивати суттєві якості його змісту [14]. У такому разі якість тексту стає вимірюваною, доступною для точних обчислювальних операцій. Контент-аналіз застосовують переважно при наявності великого за обсягом і несистематизованого матеріалу, коли його безпосереднє використання ускладнено. Такий підхід корисний в тих випадках, коли категорії, важливі для цілей дослідника, характеризуються частотністю появи у текстах, що вивчаються.

На першому етапі контент-аналізу складають список (перелік) символів (первинних понять), що цікавлять дослідника, а потім вводять категорії — більші за обсягом змістовні одиниці, що являють собою класи символів. Індикатори (символи) за своїм характером можуть бути неоднорідними: окремі слова чи словосполучення, які відносяться до теми, терміни, імена людей, назви організацій, географічні назви, згадки подій тощо. При вивченні текстів, пов'язаних, наприклад, з політичними проблемами, змістовні одиниці можуть включати внутрішні й зовнішні міжнародні події, осіб, котрі є їхніми учасниками і ініціаторами, відношення до подій в термінах: «за — проти», «вигідно — невигідно», «добре — погано» і т. п.; спосіб досягнення мети (переконання, примус, економічний тиск, моральний чи політичний вплив); характеристики соціальних прошарків, суспільних груп тощо. Суттєвими характеристиками в контент-аналізі є поряд з врахуванням частотності змістовних одиниць також й відношення до цих одиниць комунікаторів (знак висловлювання).

Характеризуючи можливості контент-аналізу, варто зазначити, що широке розповсюдження цей метод отримав у дослідженнях двох типів:

1 — співставлення текстів одного автора (чи, наприклад, однієї організації, друкованого органу), що відносяться до різного часу, з метою виявлення тенденцій до змін поглядів, позицій респондента;

2 — співставлення текстів, що належать різним авторам (організаціям, друкованим органам тощо) з метою виявлення різниць, що властиві цим текстам.

У працях, пов'язаних із застосуванням контент-аналізу, містяться різні оцінки рівня об'єктивності таких досліджень [15]. Проте, безумовно, перевагою контент-аналізу є відтворюваність результатів, отриманих на його основі. Адже, якщо історик зафіксував набір змістовних одиниць тексту, то висновки, що він їх отримав у процесі роботи, можуть бути перевірені на основі частотних даних, видобутих із тексту. У цьому плані контент-аналіз вигідно відрізняється від ілюстративно-

го підходу, коли дослідник підтверджує свою точку зору звичайними витягами з тексту.

Звичайно, висновки, одержані за допомогою контент-аналізу, можуть бути й помилковими — вони багато у чому визначаються підходом історика до формування набору змістовних одиниць тексту. Проте, як свідчить практика, цей метод, який забезпечує інтерпретацію кількісних показників у якісні ознаки, дає можливість отримати історику значною мірою неупереджену інформацію. В зв'язку з цим заслуговує на увагу вельми цікавий й конструктивний варіант адаптації методик контент-аналізу до потреб завдань вивчення історичних джерел, що викладається в праці Д. В. Деопіка [16]. Як зазначає цей автор, головна проблема полягає у тому, що рахувати і як рахувати. Адже необхідно привести фактичний матеріал, надзвичайно багатоаспектний, до суми якостей, котрі можна кількісно співставляти.

Як основну змістовну одиницю тексту стародавнього китайського рукопису «Чунцю» Д. В. Деопік визначив «найпростішу подію» (н. п.), котра характеризується через поняття «дія». З літопису «Чунцю» були виписані усі види «дій», що склали 64 н. п. Найбільш масовими виявилися 20 н. п., оскільки саме на них припадає понад 90 % згадок н. п. Результати аналізу щодо частотності «найпростіших подій», які згадувалися в тексті, були наглядно представлені у вигляді таблиць, котрі дозволяли визначити як стабільні в часі, так й еволюціонізуючі явища, що розкривали в основному різні політичні події. Контент-аналіз дозволив, фіксуючи чисельні «дрібні» факти, простежити окремі тенденції з історії китайського суспільства досліджуваного періоду.

Цікаві результати в галузі соціально-економічної історії Росії XVIII—XIX ст. за допомогою контент-аналізу отримані Б. М. Мироновим [17]. На основі грунтовного ознайомлення з матеріалами про торгівлю по відповідях на господарські анкети, а також з топографічними і математичними описами губерній XVIII—XIX ст. він склав програму систематизації й формалізації первинних даних, котра мала вигляд описуваного листа чи анкети. Результати аналізу частот зустрічаємості визначених категорій за 24 питаннями анкети, підраховані за матеріалами понад 4000 ярмарків, зведені Б. М. Мироновим у таблиці, котрі детально розкривають структуру торгівлі Росії у зазначений період.

Цікавий підхід щодо визначення символів-інформантів як сигнальних одиниць формалізації якісних ознак тексту запропоновані Л. М. Брагиною. Авторка дослідила позиції авторів філософських трактатів епохи Відродження при визначенні змісту основних етичних категорій [18]. Оригінальність авторського прийому полягала у тому, що «виявлені в тексті етичні терміни із зазначенням частоти їхнього вживання утворюють частотний словник» [19]. Досліджувалися тексти двох філософських трактатів. За допомогою групування були отримані 38 змістовні одиниці для першого трактату і 50 — для другого. Частотність використання цих термінів коливалася від 6 до 190 для першого трактату і від 6 до 429 — для другого. В першому «Про справжнє благородство» були визначені головні терміни — «благородство» та «доброчесність», частотність зустрічаємості яких виявлялася найбільшою. У другому трактаті «Про звичаї» найбільшою частотністю відзначалися терміни «доброчесність» та «щастя». Щоб з'ясувати, який зміст вкладали гуманісти XV ст. у ці головні етичні категорії, Л. М. Брагіна ввела розгляд частотності спільного використання пари термінів у рамках контексту. З цією метою вона визначила межі контексту. Виходячи з особливостей трактату, такі межі охоплювали 3 фрази — одну фразу до і одну фразу після тієї фрази, де використовувався головний термін. Таким чином, спільне використання пари термінів фіксувалося виключно у контексті трьох фраз. В процесі дослідження з'ясувалося, що, наприклад, у розмірковуваннях автора трактату про зміст «благородства» переважають поняття: доброчесність, діяльність, походження,

знання, товариство, гідність, розум. Відповідні висновки одержані й для решти головних термінів обох трактатів.

Відомо, що вельми ефективним в опрацюванні масових джерел є контент-аналіз у дослідженні матеріалів преси. Адже саме із завдань вивчення змісту масової комунікації й виник контент-аналіз. Проблемам застосування цього методу до матеріалів газетної періодики присвячені роботи Л. І. Буряк, Л. Д. Дергачової, В. П. Коцура, В. М. Мазиріна та інших істориків [20].

Джерелознавча практика свідчить, що замітки, кореспонденції, репортажі, звіти, інтерв'ю, статті (передові, редакційні, пропагандистські тощо), нариси, фейлетони різняться не тільки певними літературно-стилістичними засобами, а й переважанням різних шарів інформації. Вивчення споріднених явищ й процесів на різних рівнях публіцистичних жанрів дозволяє виявити внутрішній взаємозв'язок між первинною й вторинною інформацією, що міститься в публікаціях. Розв'язання цього завдання забезпечує зрівняльна типовість, однорідність та повторюваність газетної інформації. Визначаючи сутність такої риси, як повторюваність, Л. Д. Дергачова характеризує її «як традиційність ряда тем, наявність в періодичних органах постійних розділів і рубрик... застосування деяких жанрів у розкритті певної тематики» [21].

Різноманітність матеріалів преси робить актуальним їх аналіз з точки зору наявності в них неприхованої і латентної інформації. Перший тип інформації переважає передусім в аналітичних жанрах: замітці, кореспонденції, звіту, які найбільш насичені фактами. Другий — в синтетичних, в першу чергу — статтях, інтерв'ю. Як вважає М. Я. Варшавчик, статті викликають інтерес саме наявністю в них латентної інформації — «мотивами і метою публікації, концепцією автора і газети» [22].

Зрозуміло, що реалізація цих завдань неможлива без застосування до газетних джерел методу контент-аналіз, що дозволяє розкривати якісні ознаки не лише окремих джерел, а й їхніх тематично споріднених сукупностей. Саме на цьому наголошується в роботі І. Д. Архангельської, котра, критикуючи ілюстративний підхід до вивчення періодики, підкреслює важливість формалізовано-системного аналізу публікацій [23].

При розробці методики джерелознавчого вивчення матеріалів «Торгово-промышленной газеты» за 1910—1914 рр. І. Д. Архангельська за змістовну одиницю обрала конкретну проблему, а за одиницю рахунку частотність її зустрічаємості, котра може слугувати покажчиком вагомості проблеми з точки зору друкованого органу, а, відповідно, й охарактеризувати його позицію щодо даного питання. В результаті опрацювання в тексті було виявлено понад 30 проблем, зокрема інформації про торговельні зв'язки Росії з Німеччиною, законопроекти і реформи, їх підготовка та обговорення, фінансові проблеми Росії тощо. Для кожного матеріалу «Торгово-промышленной газеты» була складена первинна картка обліку інформації, в якій фіксувалися категорії аналізу: дата, рубрика, назва інформації, автор. Водночас спеціальним індексом позначалася проблема, до якої належав матеріал, а також індекс спорідненої проблеми. Відправним моментом другого етапу формалізації газетної інформації стало визначення стислого змісту статті. В процесі «розшифровки», інтерпретації формалізованих даних використовувся комплекс прийомів, який в практиці отримав назву методу джерелознавчої критики.

Наукова доцільність використання цього методу обумовлюється тим, що він дозволяє отримати й відповідним чином проаналізувати не окремі факти, а оптимальну їхню сукупність. На відміну від вельми обмежених можливостей «внутрішньої» і «зовнішньої» критики, що зосереджуються в принципі на вивченні окремих джерел, кожний з яких може дати дослідникові лише фрагментарний науковий факт, аналітико-синтетична критика обумовлює якісно новий етап у комплексно-

му дослідженні джерел. Як вважає М. Я. Варшавчик, уся структура джерелознавчої критики складається з «аналітичної, що має об'єктом окремі джерела і мету — отримання окремих фактів, і синтетичної, що має об'єктом комплекси джерел і мету — отримання сукупності фактів» [24]. Ключовим питанням аналітичної критики є оцінка достовірності, повноти і, таким чином, об'єктивної наукової цінності джерела. В зв'язку з цим актуалізується проблема відображення історичного явища, що мало місце в дійсності, і рівень об'єктивності відображення цього явища в публікації. Розв'язання цього завдання залежить й від з'ясування рівня повноти джерела, ступеня проникнення в сутнісну глибину явища автором публікації. Таким чином, головною ознакою повноти джерела вважається при аналізі газетних публікацій відображення суттєвих рис описуваних подій. При визначенні таких рис використовуються й методи, запозичені в галузі інформатики, суть яких полягає у методиці виділення в тексті релевантної (конкретної) і нерелевантної (супутньої) інформації [25]. Пропорційне співвідношення цих видів інформації значною мірою забезпечує повноту джерела. В дійсності, як відомо, повнота й правдивість інформації нероздільно пов'язані.

Проте в практиці досліджень історичних джерел поряд із основною проблемою виявлення якісних ознак й підвищення інформаційної віддачі джерела за допомогою кількісних методів, зокрема контент-аналізу, історикам часто доводиться вирішувати питання, пов'язані із атрибуцією різних джерел. В сучасній історіографії останніх років у пошуках шляхів виявлення об'єктивних критеріїв авторського стилю творів став помітним внесок істориків, що використовували з цією метою кількісні методи. Цій проблемі присвячені спеціальні публікації [26].

Достатньою грунтовною працею, що мала на меті довести доцільність методики формалізованої фіксації авторських особливостей в різних текстах, є, наприклад, робота Л. І. Бородкіна, Л. В. Мілова, Л. Ю. Морозової [27]. За одиницю обчислення й формалізації змісту текстів, що досліджувалися, автори обрали найбільш характерні слова, конкретні словосполучення, специфічні мовні конструкції. В процесі підготовки текстів для відповідного опрацювання автори визначили основні етапи своєї програми: обрали систему граматичних класів, що достатньо детально описувала граматичні особливості мови періоду, що вивчається; прокодували послідовність слів тексту, що аналізується, у послідовність відповідних позначень граматичних класів. В основу цієї методики атрибуції давньоруських текстів була покладена гіпотеза про те, що стиль автора виявляється в його звичці до певних граматичних зв'язків, частотність яких в тексті є високою. Також враховувалася й та обставина, що внаслідок досить поширеного у той час втручання у давні тексти переписувачів, редакторів, присутність в авторських текстах традиційних конструкцій, пов'язаних із тогочасним літературним етикетом тощо, частина граматичних зв'язків із високими показниками частотності буде підвищувати питому вагу так званих «неавторських елементів» у тексті.

В результаті застосування вдалої методики контент-аналізу до текстів з метою їхньої атрибуції було зроблено переконливу інтерпретацію отриманих на ЕОМ даних. Це був один з чисельних прикладів застосування машинної обробки бази даних. Проте, звичайно, не варто ототожнювати процес опрацювання вхідної інформації із власне методикою кількісних методів дослідження історичних джерел. Водночас роль ЕОМ при застосуванні контент-аналізу до текстів може бути вельми суттєвою при реалізації деяких особливостей цього методу:

І — визначення частотності зустрічаємості термінів у тексті (в тих випадках, коли вони є змістовними одиницями-категоріями). До речі, перший аспект, на думку деяких дослідників, є основним для визначення сутності контент-аналізу. Зокрема, В. М. Мазурін зазначає, що «плідним напрямом дослідження масової комунікації є кількісний й

якісний аналіз символів-інформантів як первинних ознак змісту. Цей напрям дістав назву «контент-аналіз»[28]; 2 — визначення частотності зустрічаємості категорій, що охоплюють групи термінів (у тих випадках, коли контекст не впливає на ідентифікацію категорій, наприклад терміни — географічні назви, а категорія відповідає групі таких назв); 3 — формування дослідником змістовних одиниць-категорій. ЕОМ може розпечатати алфавітний список всієї лексики разом із контекстом (групою, що складається із кількох слів); 4 — опрацювання отриманих частотних даних за допомогою набору програм статистичного аналізу.

Отже, контент-аналіз передбачає при дослідженні історичних джерел використання комбінаторної методики. Розкриваючи відмінності традиційного (описового) методу аналізу джерел, Т. Л. Мойсєєнко вважає його «сутнісно-змістовним» на противагу «формально-кількісному аналізу»[29]. Насправді, як вважає І. Д. Ковальченко, «фактично в науковому дослідженні має місце чи сутнісно-описовий, чи сутнісно кількісний аналіз. Ніякого чистого якісного або самостійного описового чи кількісного аналізу не існує. Може бути лише єдність у зазначеному співвідношенні»[30].

В зв'язку з цим набуває актуальності такий важливий метод визначення якісних ознак тих чи інших об'єктів при аналізі історичних джерел, як прийом «експертних оцінок»[31]. Його сутність полягає у тому, що на основі змістовного аналізу виявляється певна кількісна градація інтенсивності відповідної якості. Після цього кожному ступеню цієї градації надається числове значення у вигляді балів, котрі повинні відбивати ступінь різниці інтенсивності властивостей на певному рівні. «Експертні оцінки» все більш широко застосовуються в різних галузях людської діяльності. При науковому дослідженні явищ суспільного життя «експертні оцінки» досить помітно використовуються у галузі конкретно-соціологічних досліджень[32]. На жаль, в історичній науці вимірювання складних якісних ознак шляхом таких оцінок ще не набуло належного поширення. А між тим при всій його обмеженості воно дозволяє значно поглибити аналіз історичних явищ й процесів. Це переконливо доводить І. Д. Ковальченко, розглядаючи, як приклад, рівень класової боротьби робітників (умовний варіант)[33]. Він зазначає, що за загальною інтенсивністю виступи робітників можна розподілити на ряд категорій. Наприклад, класифікація виступів: 1 — дуже слабкі; 2 — слабкі; 3 — слабо-помірні; 4 — помірні; 5 — помірно-сильні; 6 — сильні; 7 — дуже сильні. Такою буде якісна, а саме порядкова шкала оцінок. Для спеціалістів в галузі робітничого руху не буде важко класифікувати виступи робітників за цією шкалою.

Складніше розв'язати питання про кількісну, бальну оцінку вказаних ступенів інтенсивності виступів робітників. Основна проблема полягає тут у визначенні кількісного інтервалу між нижчим і вищим рівнями інтенсивності властивості, що вивчається. Він завжди буде мати приблизний характер. Але й при цьому введення бальних оцінок має сенс, оскільки відкривається шлях до кількісної оцінки різних якостей. Зрозуміло, що, наприклад, різниця в інтенсивності боротьби робітників у двох виступах, з яких одне було «дуже слабким», а інше «дуже сильним», стане більш наявною, якщо ця інтенсивність у першому випадку буде оцінена в 1, а у другому в 7 балів.

Це, звичайно, найпростіший підхід до «експертної оцінки» інтенсивності складних якісних властивостей історичних об'єктів і явищ. Більш складним, але точним буде підхід, що заснований на «експертній оцінці» складових компонентів комплексних якісних ознак. Для цього складні властивості, що характеризуються цими ознаками, на першому етапі мають бути розділені на більш прості складові. Після цього кожна з цих частин отримує бальну оцінку, а сума балів буде інтегральним кількісним показником інтенсивності виявлення складної якісної властивості.

Проте, як вже зазначалося, нехтування прийомом «експертних оцінок» при використанні контент-аналізу до історичних джерел суттєво знижує конкретність отриманих результатів, як, приміром, це мало місце у відсутності вимірюваних ознак у таких складових структурах політичної ситуації, як «соціальний об'єкт», «ліберальна політична позиція» тощо в одному із сучасних досліджень матеріалів преси [34].

Дуже важливою проблемою, котру не можна обминути при висвітленні методики контент-аналізу, є питання про попередню підготовку історичних джерел для аналітико-синтетичної критики. Йдеться передусім про механізм здійснення вибірки. Адже практично завжди в дослідницькій роботі історику доводиться визначати коло джерел, що мають аналізуватися, у двох напрямах: по-перше, як на основі небагатьох даних, які збереглися, отримати широку й достовірну історичну картину; чи, навпаки, по-друге, як з величезного масиву джерел відібрати мінімальну кількість даних, за допомогою яких можна робити висновок про історичне явище, об'єкт в цілому. Обидві проблеми задовільно вирішуються за допомогою вельми ефективного методу вибірки. Це питання є предметом спеціальних досліджень науковців [35]. Щоправда, у фаховій літературі не завжди чітко розмежовуються сутність кількісних методів і зміст методу вибірки, що створює необхідні статистико-кількісні передумови для власне застосування кількісного методу. З огляду на це не можна погодитись із думкою Т. І. Славко, котра розглядає вибірку в одному логічному ланцюзі із різними видами кількісних методів аналізу джерел [36]. Проте, як відомо, про терміни не сперечаються, про них домовляються. Незважаючи на окремі розбіжності у визначенні сутнісно-понятійного змісту вибірки, цей метод має вельми широку практику в плані її застосування щодо історичних джерел.

Для того, щоб часткове обстеження можна було б вважати застосуванням вибіркового методу дослідження в математико-статистичному розумінні цього терміну, необхідне дотримання однієї умови. А саме: об'єкти для часткового обстеження мають бути обрані випадково, тобто за жеребом чи в силу природніх умов, що склалися. При поширенні результатів вибіркового дослідження на це явище, що вивчається, дослідник робить помилку, так звану похибку репрезентативності. Але передбачити величину цієї похибки можливо, якщо вона являє собою випадкову похибку. А випадковою вона може бути тоді, коли об'єкти до вибірки потрапляють випадково. Вибірка лише тоді вважається випадковою, коли формується з генеральної сукупності таким чином, щоб кожний член такої сукупності мав рівний шанс бути відібраним.

Забезпечити випадковість вибірки, якщо історик здійснює відбір даних з великої сукупності відомостей, можна за допомогою жеребкування чи таблиці великих чисел. Також можуть бути використані й інші методи відбору: механічний, типовий (районований) і серійний. Досить докладно методика зазначених методів викладається, зокрема, Г. Д. Ізместьєвою [37]. Як приклад використання вибірки при опрацюванні великих масивів історичних джерел, автор даної публікації наводить власні розрахунки, що він їх робить при аналізі матеріалів республіканської преси 1956—1964 рр.

При підготовці масиву газетних публікацій для джерелознавчого аналізу автор прагнув розв'язати такі питання. По-перше, забезпечити опрацювання переважної більшості центральних республіканських газет, по-друге, забезпечити максимально можливе представництво місцевої преси як такої, що містить в основному первинну інформацію. Визначаючи вибіркову сукупність газетних джерел, автор виходив з того, що загальна кількість газет, з одного боку, та арифметична сукупність номерів цих видань у республіканському масштабі, з другого, обумовлює необхідність застосування двофазного відбору.

З огляду на те, що у фаховій літературі 10 % вибірка при генеральній сукупності понад 1 тис. одиниць вважається достатньо репрезентативною[38], вибіркова сукупність назв республіканських газет становить 200 одиниць (згідно з методикою направленого відбору 12 республіканських, 23 обласні, 65 міських, районних, 100 назв багатотиражних газет. З метою кінцевої перевірки репрезентативності вибірки суцільному аналізу піддана «Радянська Україна», і через це її інформаційний масив вибірці не підлягав). В результаті відповідно зменшилася кількість номерів до 80 тис. одиниць. Після цього було здійснено перехід до другого етапу двофазного відбору, використовуючи для визначеня вибіркової сукупності кількості номерів таблицю достатньо великих чисел[39]. Чисельність вибірки, що робилася у такий спосіб, залежала від величини імовірності (Р), з якої робилося припущення про ступінь надійності матеріалу, і від величини припускаємої похибки (Т прип.). З таблиці достатньо великих чисел була обрана найбільш представницька для інформаційного масиву сукупність (П). Отже, мали П=27 069 при Р=0,999, Т прип.=0,01. Це означає, що при вивченні вибірки в 27 тис. одиниць у 999 випадках з 1 тис. похибка не буде перевищувати 0,01 і лише в 1 випадку з 1 тис. можуть траплятися великі похибки. При визначенні вибіркової сукупності у кількості 27 тис. одиниць був також здійснений пропорційно-диференційований відбір однорідних інформаційних масивів за масштабом видання у відповідному відсотковому співвідношенні до генеральної сукупності: — республіканські газети 540 од. (2 %); — обласні 2,430 од. (9 %); — міські районні 18,360 од. (68 %); — багатотиражні 5,670 од. (21 %).

За алфавітною побудовою відповідно до питомої ваги кожної назви газети в загальному інформаційному масиві жеребкуванням нараховане кількісне представництво кожної із 200 газет у вибірковій сукупності. Отже, визначаючи репрезентативність вибірки, автор керувався, по-перше, тим, що всі одиниці генеральної сукупності мали рівний шанс потрапити у вибірку, по-друге, виходив із однорідності інформаційного масиву для джерелознавчого аналізу, по-третє, користувався методом механічного відбору, що гарантував необхідну випадковість вибіркової сукупності.

Отримані дані при аналізі вибіркової сукупності матеріалів республіканської преси, що відбивали зміни у політичному, соціальному, економічному, культурному житті в Україні у зазначений період, при співставленні з результатами суцільного аналізу газетних матеріалів «Радянської України» доводять ефективність обраного методу вибірки джерел для контент-аналізу.

Обмежений обсяг статті не дозволяє, звичайно, більш докладно, залучаючи для зразка різноманітні приклади використання контент-аналізу в дослідженні історичних джерел, розкрити методику цього найбільш поширеного кількісного методу. Більш докладно цю проблему можна опрацювати, використовуючи, зокрема, відповідну фахову літературу, що згадується у даній публікації. Подальше освоєння кількісних методів можливе при самостійній роботі, причому не стільки з підручниками математики та статистики, скільки з конкретним історичним матеріалом, для аналізу якого використовуються математичні методи. Це, в принципі, під силу будь-якому історику. До того ж у застосуванні кількісних методів в історії багато цікавого, бо кліометрика — це одна з порівняно нових галузей історичної науки.

[1] Берг М. А. Эпохи и идеи. Становление историзма.— М., 1987.— С. 12.

[2] Славко Т. И. Математико-статистические методы в исторических исследованиях.— М., 1981.— С. 3.

[3] Количественные методы в исторических исследованиях.— М., 1984.— С. 22.

[4] Бородкин Л. И. Многомерный статистический анализ в исторических исследованиях.— М., 1986.— С. 148.

[5] Бородкин Л. И., Соколов А. И. История и изучение социальных процессов (Об использовании массовых источников и количественных методов их анализа в новейшей зарубежной историографии) // История СССР, 1983, № 1; Количественные методы в зарубежной исторической науке (Историография 70—80-х гг.) под ред. И. Д. Ковальченко.— М., 1988; Сивачев Н. В. На 89-м ежегодном съезде Американской исторической ассоциации // Вопросы истории, 1975, № 6; Селунская Н. Б. «Количественная история» в США: итоги, проблемы, дискуссии // Математические методы в историко-экономических и историко-культурных исследованиях.— М., 1977.— та ін.

[6] Ковальченко И. Д. Русское крепостное крестьянство в первой половине XIX в.— М., 1967.

[7] Бокарев Ю. П. Крестьянское хозяйство и развитие промышленности советской России в середине 20-х годов XX в. // Математические методы в историко-экономических и историко-культурных исследованиях.— М., 1977.

[8] Славко Т. И. Указ. соч.— С. 9.

[9] Математические методы в социально-экономических и археографических исследованиях.— М., 1981; Методы количественного анализа нарративных источников.— М., 1983; Гусейнова А. С. Опыт имитационного моделирования исторических процессов.— М., 1984; Ковальченко И. Д., Бородкин Л. И. Современные методы изучения источников.— М., 1987; Проблемы применения количественных методов анализа и классификация источников.— Днепропетровск, 1988; Математические методы и ЭВМ в историко-типологических исследованиях.— М., 1989; Пащевский А. А. Математические методы в исторических исследованиях.— Краснодар, 1989; Миронов Б. Н. История в цифрах: математика в исторических исследованиях.— Л., 1991.

* В Західній Європі, США активно функціонують великі національні центри, що координують діяльність істориків-кліометристів: Асоціація соціальної наукової історії, Комітет по кількісним даним Американської історичної асоціації і Міжуніверситетський консорціум політичних і соціальних досліджень при Мічіганському університеті в США, Група по вивченню дослідження історії населення і соціальної структури при Кембріджському університеті в Великобританії, Комісія по застосуванню кількісних методів в історичних і соціальних дослідженнях в Німеччині, Центр вивчення документів за допомогою ЕОМ при Католицькому університеті в Лувені в Бельгії та ін.

[10] Методологические и методические проблемы контент-анализа.— М., Л., 1973; Проблемы контент-анализа в социологии.— Новосибирск, 1970; Ядов В. А. Социологические исследования. Методика, программа, методы.— М., 1972,; Алексеев А. Н. Контент-анализ, его задачи, объекты и средства // Социология культуры.— № 1.— М., Л., 1974 та ін.

[11] Литюшин Ю. И., Оленина Ж. А., Устинов В. А. Опыт использования контент-анализа в практике партийной работы // Социологические исследования, 1978.— № 1; Каюров В. Н., Сагаенко Г. И. Контент-анализ аттестационных характеристик как метод изучения нормативных требований к инженерам в организации // Социологические исследования, 1977.— № 1; Максимов В. В. Применение контент-анализа для исследования научных отношений // Методологические и методические проблемы контент-анализа.— М., Л., 1973, вып. 3; Хайтун С. Д. Наукометрия: состояние и перспективы.— М., 1983 та ін.

[12] Дергачева Л. Д. Периодическая печать как массовый источник // Массовые источники по истории советского рабочего класса / Редкол.: И. Д. Ковальченко (отв. ред). и др.— М., 1982.— С. 159.

[13] Буряк Л. И. Периодическая печать как источник изучения деятельности Коммунистической партии по патриотическому воспитанию советских людей в годы Великой Отечественной войны. Автореф. дис. канд. ист. наук.— К., 1985.— С. 13—14.

[14] Рабочая книга социолога.— М., 1976.— С. 326.

[15] Соковнин В. М. Об объективности исследования в контент-анализе // Методологические и методические проблемы контент-анализа.— М., 1973.

[16] Деопик Д. В. Опыт количественного анализа древней восточной летописи «Чунцю» // Математические методы в историко-экономических и историко-культурных исследованиях.— М., 1977.

[17] Миронов Б. Н. Формализация и генерализация содержания массовых исторических источников (на материалах анкеты о ярмарках 1779 г.) // Вспомогательные исторические дисциплины. Вып. XIII.— Л., 1981; Миронов Б. Н. Внутренний рынок в России во второй половине XVIII — первой половине XIX в.— Л., 1981.

[18] Брагина Л. М. Методика количественного анализа философских трактатов эпохи Возрождения // Математические методы в историко-экономических и историко-культурных исследованиях.— М., 1977.

[19] Там же.— С. 285.

[20] Буряк Л. И. Печать как источник изучения патриотического воспитания советских людей в годы Великой Отечественной войны (методика количественного анализа) // Научные труды по истории КПСС : Респ. межвед. сб., 1985, вып. 135; Дергачева Л. Д. Указ. соч.; Дробжев В. З., Пивоваров С. И. Массовые источники по истории советского рабочего класса и интеллигенции и количественные методы их анализа // Количественные методы в советской и американской историографии.—М., 1983; Коцур В. П. Партійно-радянська преса на початку 20-х рр. // Укр. іст. журн., 1990, № 10; Мазырин В. М. Применение методов контент-анализа

к материалам прессы // Количественные методы в гуманитарных науках.— М., 1981; Мазырин В. М. Статистический анализ материалов прессы как источника по новейшей истории Вьетнама // История восточных культур.— М., 1976.

[21] Дергачева Л. Д. Указ. соч.— С. 154.

[22] Варшавчик М. А. Источниковедение истории КПСС.— М., 1989.— С. 187.

[23] Архангельская И. Д. К вопросу изучения периодической печати методами контент-анализа // Методы количественного анализа текстов нарративных источников.— М., 1983.

[24] Варшавчик М. А. О структуре источниковедческой критики // Источниковедение отечественной истории : Сб. статей.— М., 1980.— С. 25.

[25] Звежинский С. М. Научная организация информационной деятельности.— Львов, 1974.

[26] Раскина А. А. К вопросу об атрибуции текста при помощи статистики // Межвузовская конференция по вопросам частотных словарей и автоматизации лингвистических работ.— Л., 1966; Херц М. М. Математические методы в исторической лингвистике.— М., 1974; Пиотровский В. Г. Текст, машина, человек.— Л., 1976 та ін.

[27] Бородкин Л. И., Милов Л. В., Морозова Л. Е. К вопросу о формальном анализе авторских особенностей стиля в произведениях Древней Руси // Математические методы в историко-экономических и историко-культурных исследованиях.— М., 1977.

[28] Мазырин В. М. Применение методов контент-анализ к материалам прессы.— С. 49.

[29] Количественные методы в зарубежной исторической науке.— С. 4.

[30] Количественные методы в исторических исследованиях / Под ред. И. Д. Ковальченко.— М., 1984.— С. 18.

[31] Варшавчик М. А. Указ. соч.— С. 120.

[32] Социальные исследования: построение и уравнение показателей.— М., 1978.

[33] Количественные методы в исторических исследованиях.— С. 41—42.

[34] Опыт контент-анализа альтернативной прессы Белоруссии // Социологические исследования, 1991, № 8.

[35] Гранков В. П. Выборочное наблюдение.— М., 1963; Дружинин Н. К. Выборочный метод и его применение в социально-экономических исследованиях.— М., 1970; Миронов Б. Н. Применение выборочного метода при анализе движения хлебных цен XVIII в. // Ежегодник по аграрной истории Восточной Европы.— Кишинев, 1966; Пасхавер И. С. Закон больших чисел и закономерности массового процесса.— М., 1966; Соколов А. К. Методика выборочной обработки первичных материалов профессиональной переписи 1918 г. // История СССР, 1971, № 4 та ін.

[36] Славко Т. И. Указ. соч.— С. 59—83.

[37] Изместьева Г. Ф. Выборочный метод // Количественные методы в исторических исследованиях.— М., 1984.

[38] Количественные методы в исторических исследованиях.— С. 105.

[39] Славко Т. И. Указ. соч.— С. 63.