

УДК 316.774

## БАГАТОВИМІРНИЙ АНАЛІЗ СОЦІОЛОГІЧНОЇ ІНФОРМАЦІЇ: СУТНІСТЬ, КЛАСИЧНІ МЕТОДИ ТА ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ

**Кислова Ольга Миколаївна** – кандидат соціологічних наук, доцент, докторант кафедри прикладної соціології Харківського національного університету імені В. Н. Каразіна

*Статья посвящена рассмотрению сущности многомерного анализа социологической информации. Автор кратко касается истории проникновения идеи многомерности в практику исследования социальной реальности, уточняет ключевые понятия, анализирует возможности методов и технологий многомерного анализа социологических данных.*

**Ключевые слова:** многомерность, многомерный анализ, социологическая информация

*Статья посвящена розгляду сутності багатовимірного аналізу соціологічної інформації. Автор кратко торкається історії проникнення ідеї багатовимірності в практику дослідження соціальної реальності, уточнює ключові поняття, аналізує можливості методів і технологій багатовимірного аналізу соціологічних даних.*

**Ключові слова:** багатовимірність, багатовимірний аналіз, соціологічна інформація

*The article is dedicated to the essence of multidimensional analysis as applied to sociological information. The author briefly deals with the history of introduction of the multidimensionality idea to the research of social reality, specifies the key concepts, analyses the possibilities of methods and technologies of the multidimensional analysis for sociological data.*

**Key word:** multidimensional ,multidimensional analysis, sociological information

*Всі науки настільки пов'язані між собою, що легше вивчати їх всі відразу, ніж якусь одну із них окремо від інших.*

*Правильно визначайте слова, і ви позбавите світ від половини непорозумінь.*

*Рене Декарт*

Актуальність дослідження сутності та методів багатовимірного аналізу соціологічної інформації обумовлена специфікою соціальної реальності, що завжди уявляється як складний, багатогранний та багатозначний феномен, який інтегрує багатовимірність суспільства з багатовимірністю внутрішнього світу окремої людини. Соціологи, вивчаючи соціальну реальність, стикаються з необхідністю вибору адекватних підходів та методів, здатних охопити всі аспекти досліджуваних соціальних явищ, враховуючи їх цілісність та взаємозалежність. Як відомо, на теоретичному рівні вирішення цього завдання здійснювалося шляхом розробки різноманітних теоретичних підходів, які склали основу поліпарадигмальності сучасної соціології. На емпіричному рівні пошуки "кращих" методів пізнання багатовимірної соціальної реальності призвели до розмежування кількісної та якісної методологій. Результатом стало те, що на сьогодні немає єдиного бачення того, яким чином можна здійснити саме "багатовимірний" аналіз, який би дав змогу поєднати дослідження об'єктивних соціальних закономірностей з особливостями їх суб'єктивних проявів.

На жаль останнім часом ідея багатовимірності частіш за все застосовується тільки в контексті методів кількісного аналізу, хоча її евристичний потенціал значно ширший, що відзначалось ще соціальними філософами радянської школи (див. [1]), а Дж. Александер навіть сформулював основи інтегративної "багатовимірної соціології", що поєднує системний підхід з мікросоціологічним дослідженням мотивацій.

Нам здається, що настав час знову звернути увагу на можливості багатовимірного аналізу, переглянути попередні ідеї стосовно "багатовимірності" людини й соціуму та проаналізувати можливості новітніх інформаційних технологій у дослідженні багатовимірних соціальних феноменів. У рамках даної статті ми обмежимо фокус своєї уваги емпіричним рівнем, взяв за мету уточнення термінології багатовимірного аналізу та розгляд нових можливостей його реалізації, які з'явилися у зв'язку з розвитком методів інтелектуального аналізу даних.

Перш за все зазначимо, що багатовимірним аналізом називають сукупність різноманітних методів, призначених для вивчення багатовимірних явищ. У цьому визначенні поняття "багатовимірне явище" є ключовим. Проте, як в соціологічній, так й в філософській літературі слово "багатовимірний" застосовують частіш за все в якості метафори, коли автори бажають підкреслити складність, багатогранність, "глибину" людського буття, вважаючи багатовимірним все те, що має "багато" вимірів, залишаючи поза увагою відповіді на питання: "Багато – це скільки? Два, три, чотири чи ще більше? Чи можна говорити про багатовимірний аналіз, коли соціолог аналізує двовимірні або тривимірні таблиці спряженості?". Щоб відповісти на ці питання, необхідно звернутися до геометрії, де саме й виник термін "багатовимірність".

Визначення багатовимірного явища тісно пов'язано з поняттями багатовимірного простору та багатовимірної геометрії, уявлення про які зародилися під впливом роздумів відомих математиків (І. Кант, Ж. Д'Аламбер, А. Келі, Г. Грассман, Л. Шлефлі) про можливість приєднання часу в якості четвертого виміру до тривимірного світу, в якому ми існуємо [2].

Людина звикла орієнтуватися у тривимірному просторі, четвертий вимір є складним для людського сприйняття, що й обумовило визначення всіх просторів, що мають число вимірів (розмірність) більше трьох, у якості багатомірних. Елементарна геометрія вивчає звичайні (евклідові) простори, розмірність яких не перевищує трьох. Виникнення поняття "багатовимірний простір" пов'язано з процесом узагальнення предмета геометрії, яке стало можливим завдяки відкриттю відносин і форм, подібних із просторовими, для численних класів різноманітних об'єктів (найчастіше таких, що не мають геометричного характеру). У ході цього процесу поступово викристалізувалася ідея абстрактного простору як системи елементів будь-якої природи, між якими встановлені відносини, подібні відносинам між крапками звичайного простору, що вивчався елементарною евклідовою геометрією. Аналогічні міркування дали можливість П. Бурдье поширити поняття абстрактного геометричного простору на соціальні феномени та ввести категорію "соціальний простір" [3]. Розвиток математичних абстракцій та впровадження їх у практику соціального пізнання також склали основу дуже модної нині традиції говорити про кіберпростір, медіапростір, простір культури, простір сприйняття тощо, причому всі ці простори апріорно вважаються багатовимірними.

Визначення багатовимірного простору як простору, розмірність якого більша за три, є загальноприйнятим в математиці, що створює передумови для екстраполяції такої дефініції на розуміння багатовимірного об'єкту (явища, процесу) та багатовимірного аналізу як таких, що мають мінімум чотири виміри. Ми вважаємо важливим таке уточнення, оскільки деякі автори пропонують вважати багатовимірним аналіз трьох перемінних [4], що відкидає попередні наукові здобутки в галузі математики та не відповідає науковому тлумаченню, яке викристалізувалось у тривалому процесі становлення ідеї багатовимірності.

Щоб експліціювати сутність поняття "багатовимірний аналіз соціологічної інформації", необхідно розглянути також, що розуміється під соціологічною інформацією.

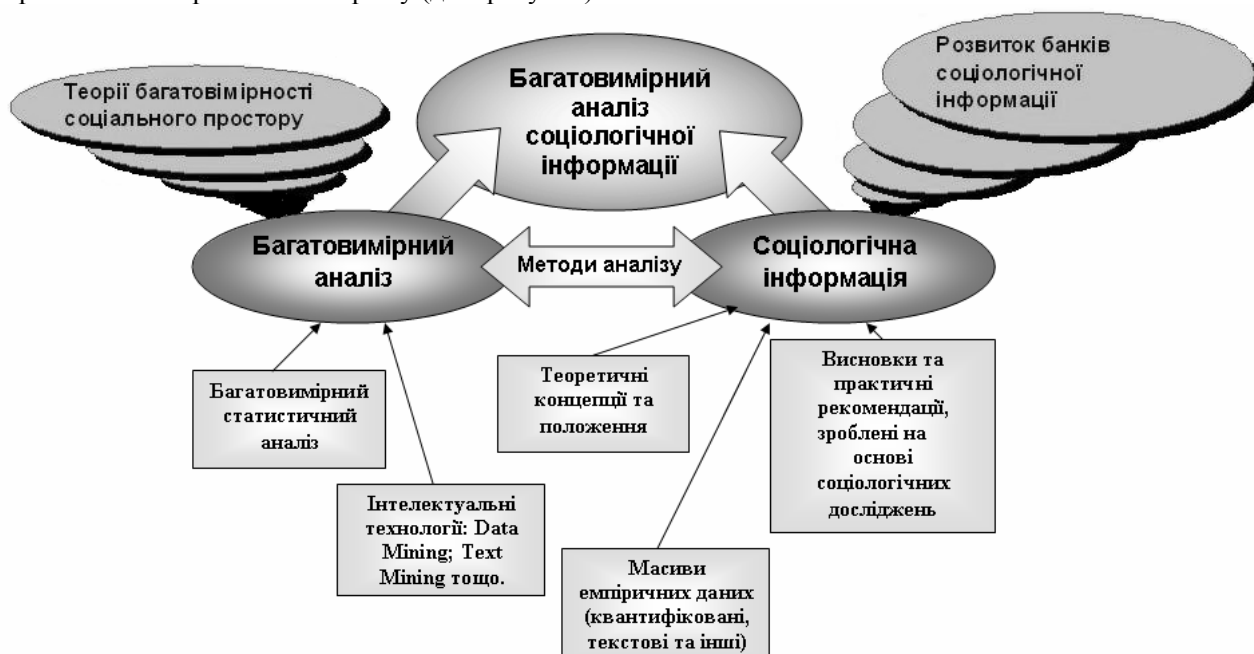
У Законі України "Про інформацію" дається таке визначення: "Соціологічна інформація – це документовані або публічно оголошені відомості про ставлення окремих громадян і соціальних груп до суспільних подій та явищ, процесів, фактів. Основними джерелами соціологічної інформації є документовані або публічно оголошені відомості, у яких відображено результати соціологічних опитувань, спостережень та інших соціологічних досліджень" [5].

Таким чином, соціологічна інформація являє собою оприлюднені результати соціологічних досліджень. А що є результатом соціологічного дослідження? Зазвичай увага акцентується на висновках, зроблених соціологами. Проте, масиви емпіричних даних, які дали можливість дійти певних висновків, теж можна вважати соціологічною інформацією. До останнього часу публічно оголошувались лише висновки (теоретичні або практичні результати), але на сьогодні соціологічна спільнота усвідомила важливість оприлюднення саме масивів емпіричних даних, на основі яких соціологи робили свої висновки, що знайшло своє відображення у створенні та розвитку архівів і банків соціологічної інформації.

Можна прогнозувати, що найближчим часом публічне представлення результатів соціологічних досліджень у вигляді масивів емпіричних даних стане обов'язковим, бо це перекриє можливість маніпулювання інструментальними та інтерпретативними ресурсами задля тенденційного представлення досліджуваних соціальних феноменів. Крім того, впровадження новітніх технологій обробки даних створює додаткові умови для продовження досліджень після публікації головних результатів з метою

більш повного та глибокого аналізу проблеми, що вивчається, та пошуку латентних закономірностей, виявлення яких потребує застосування витончених та трудомістких методів. У даному контексті дуже актуальним стає акцентування саме емпіричної складової соціологічної інформації, що й робить Г. Г. Татарова, пропонуючи таке визначення: "Соціологічною інформацією називаються будь-які емпіричні дані, які містять інформацію про соціальну реальність: соціальні явища, соціальні процеси, соціальні спільноти, соціальні інститути, соціальні системи, соціальні групи та інші соціальні феномени" [6].

Ми маємо підкреслити, що терміни "соціологічна інформація" та "емпіричні дані" дуже близькі за змістом, проте вони не є взаємозамінними, оскільки під емпіричними даними часто розуміють лише ті дані, що зібрані самим соціологом. Однак у межах соціологічного знання виділяють також інформацію більш високого рівня: теоретичні концепції, висновки й положення, які виступають у вигляді повідомлень, відомостей і застосовуються людьми в практичній діяльності. Таким чином, соціологічна інформація поєднує у собі емпіричні дані (як первинні, так й вторинні), висновки й практичні рекомендації, зроблені на їх основі, а також теоретичні концепції та положення, що перевірялися й уточнювалися на основі отриманого емпіричного матеріалу (див. рисунок).



**Рисунок . Експлікація поняття "багатовимірний аналіз соціологічної інформації" методом виділення його структурних елементів**

Первинна соціологічна інформація – дані, зібрані соціологом у результаті проведення соціологічного дослідження з метою вивчення конкретного соціального феномена, для вирішення актуальної в даний момент проблеми. Такі дані одержують у процесі анкетування, інтерв'ювання, спостереження або іншим способом залежно від обраної дослідником методології.

Вторинна (вже наявна) інформація – дані, зібрані раніше для цілей, відмінних від розв'язуваної у даний момент проблеми. Вторинна соціологічна інформація є основою розвитку так званого вторинного аналізу, який стає все більш затребуваним засобом одержання соціологічного знання. Його наростаюча популярність обумовлена перш за все тим, що завдяки застосуванню вторинних даних дослідник має можливість уникнути труднощів та матеріальних витрат, пов'язаних зі збором первинної соціологічної інформації. Крім того, інформаційна цінність вихідних даних зростає при повторному застосуванні (в інших дослідницьких контекстах та у сполученні з даними інших емпіричних вимірів), а коректно проведений синтез первинних даних власного дослідження з аналізом вторинної інформації дає можливість вийти на більш високий рівень узагальнень та висновків, одержати нетривіальні теоретичні результати. Розвиток банків соціологічної інформації, які дозволяють одержати доступ як до результатів аналізу, так і до вихідних масивів даних, на основі яких цей аналіз проводився, робить вторинні дані

доступними широкому колу дослідників, завдяки чому створюються передумови для більш повного використання соціологічної інформації.

Багатовимірний аналіз включає широкий спектр різноманітних методів: математичних, статистичних, графічних, інтелектуальних. На сьогодні не існує загальноприйнятого визначення границь цієї області, оскільки розуміння багатовимірного аналізу зазнає змін, які пов'язані з усвідомленням того, що його можливості не обмежуються застосуванням статистичних методів для дослідження кількісних даних, хоча історично саме для таких цілей його застосовували в першу чергу.

Багатовимірний статистичний аналіз (у широкому значенні) - розділ математичної статистики, що поєднує методи вивчення даних, які характеризують багатовимірні об'єкти. Багатовимірний статистичний аналіз (у вузькому значенні) поєднує ті багатовимірні статистичні методи, які засновані на припущенні, що результати окремих спостережень незалежні й підлеглі багатовимірному нормальному розподілу. Звичайно саме до цієї частини математичної статистики застосовують термін "багатовимірний статистичний аналіз".

Сьогодні частіше застосовують термін "багатовимірний аналіз", оскільки більшість методів, що застосовуються для аналізу навіть квантифікованих даних, не є статистичними в точному значенні (тобто не вимагають виконання умов незалежності й нормальності).

Традиційно багатовимірний статистичний аналіз соціологи застосовують для вирішення таких завдань (див. [7]):

1. Аналіз багатовимірних розподілів, який передбачає такі процедури: статистичне оцінювання досліджуваних багатовимірних розподілів та їх основних параметрів; дослідження властивостей статистичних оцінок, що використовуються дослідником; дослідження розподілів ймовірностей для ряду статистик, за допомогою яких будуються статистичні критерії перевірки різних гіпотез про імовірнісну природу аналізованих багатовимірних даних.

2. Вивчення характеру й структури взаємозв'язків компонентів досліджуваної багатовимірної ознаки проводиться завдяки застосуванню наступних класичних статистичних методів: регресійного, дисперсійного, ковариційного, факторного, латентно-структурного, логлінійного аналізів. Методи, що належать до цієї групи, включають як алгоритми, засновані на припущенні про імовірнісну природу даних, так і методи, що не укладаються в рамки якої-небудь імовірнісної моделі (останні частіше відносять до методів аналізу даних).

3. Дослідження структури сукупності багатовимірних спостережень здійснюється методами дискримінантного аналізу, кластерного аналізу, багатовимірного шкалювання. Вузловим для цих методів є поняття відстані, або міри близькості між аналізованими елементами як крапками деякого абстрактного простору.

Безумовно, статистика та методи багатовимірного статистичного аналізу застосовуються в соціології й будуть застосовуватися далі, але вони мають певні обмеження. Саме тому сьогодні й актуалізується необхідність поповнення методного арсеналу соціолога, яке відбувається, зокрема, за рахунок розвитку методів та технологій інтелектуального аналізу даних, який (у широкому сенсі) є сучасною концепцією соціального пізнання, що заснована на багатовимірному аналізі інформації. Новітні інтелектуальні інформаційні технології дають можливість обробляти та аналізувати дані, навіть якщо вони мають гігантські обсяги, є неточними, неповними, суперечливими, різнорідними або непрямими. При цьому застосовуються не тільки класичні статистичні методи, а й такі алгоритми аналізу, що володіють "елементами інтелекту", зокрема здатністю "мислити абдуктивно", робити загальні висновки на основі окремих спостережень (прикладом слугують алгоритми побудови дерева класифікації та нейронні мережі, які вже активно застосовуються в соціологічних дослідженнях). Головним припущенням інтелектуального аналізу даних є неможливість реалізації багатовимірних методів без нетривіальної автоматизації, оскільки всі ці методи дуже складні з обчислювальної точки зору й тому потребують застосування спеціалізованого програмного забезпечення.

Головна ідея інтелектуального аналізу даних має глибокі філософські коріння й може бути виражена крилатою фразою "ніщо не нове під луною", інтерпретація якої в сучасному комп'ютеризованому контексті полягає в наступному: дані, крім явної інформації, містять "знання" (тобто актуальну інформацію, але в неявному вигляді); нові знання про досліджуваний об'єкт, що породив аналізовані дані, можна витягати безпосередньо із цих даних, застосовуючи сучасні інформаційні технології. Інтелектуалізація інформаційних технологій сприяла зміні пріоритетів інформаційної діяльності за рахунок переносу акценту з нарощування обсягів інформації на способи одержання

актуальних знань із безмежних інформаційних масивів, а розвиток і впровадження в практику соціологічного аналізу методів інтелектуального аналізу даних актуалізували необхідність пильної уваги до нових можливостей обробки соціологічної інформації, серед яких особливо корисними, за нашою думкою, є такі: 1) поява сучасного програмного забезпечення дозволяє "одним щигликом миші" виявляти в даних приховані закономірності, досліджувати їх різні аспекти; 2) стала можливою автоматизована обробка нових типів даних (web-контенту, фотографій, аудіо- й відеоматеріалів, мультимедіа та т.п.), що сприяє використанню таких даних як носіїв соціологічної інформації. Таким чином, методи та технології інтелектуального аналізу даних поступово перетворюються у незамінний інструмент соціолога-аналітика.

Наприкінці ми зацентруємо увагу на тому, що, хоча поняття "багатовимірність" виникло в геометрії, але ні математикою, ні фізикою воно повністю не вичерпується. Багатовимірність соціального простору ще чекає своїх дослідників: як теоретиків, так і емпіриків. Сподіваємося, що розрізнення понять, близьких за змістом, але не тотожних, стане повсякденною практикою уникання непорозумінь, а молоді соціологи будуть коректно застосовувати звучні терміни саме в тому сенсі, який вкладали в них автори. Ми хочемо вірити, що сучасні дослідники плідно застосують здобутки у сфері теоретичних уявлень про багатовимірні простори (абстрактні, фізичні, соціальні) та зможуть дати чіткі визначення таких модних нині словосполучень, як "простір політики", "простір постмодерну", "медіапростір" тощо, зробивши їх не просто прикрасами мови, а термінами, придатними для практичного застосування в процесі видобутку соціологічної інформації та прирощення соціологічного знання.

### Література:

1. Келле В. Ж. Идея многомерности в познании социальной реальности // Методология гуманитарного знания в перспективе XXI века. К 80-летию профессора Моисея Самойловича Кагана. Материалы международной научной конференции. 18 мая 2001 г. Санкт-Петербург. Серия "Symposium". – Выпуск №12. – СПб.: Издательство Санкт-Петербургского философского общества, 2001. – С. 53-62.
2. Александров А. Д. Многомерная геометрия // Математическая энциклопедия. – М.: Советская энциклопедия, 1982. – Т. 3. – С. 730- 731.
3. Бурдьё П. Физическое и социальное пространства // Социология политики. Пер. с фр. / Сост., общ. ред. и предисл. Н. А. Шматко. – М.: Socio-Logos, 1993. – С. 33 - 52.
4. Терещенко О. В. Анализ многомерный статистический // Социология: Энциклопедия / Сост. А.А. Грицанов, В.Л. Абушенко, Г.М. Евелькин, Г.Н. Соколова, О.В. Терещенко., 2003 г. – Доступно на: <http://slovari.yandex.ru/dict/sociology>.
5. Закон України "Про інформацію". – Доступно на: <http://zakon.rada.gov.ua/cgi-bin/laws/main.cgi?nreg=2657-12>.
6. Татарова Г. Г. Методология анализа данных в социологии (введение): учебник для вузов. – М.: NOTA BENE, 1999. – 224 с.
7. Толстова Ю.Н. Анализ многомерный статистический. – Доступно на: <http://www.ecsocman.edu.ru/db/msg/6606>.

© О. М. Кислова, 2009