

ВЪЗМОЖНО ЛИ Е КОМПЮТРИЗИРАНЕТО НА КАЧЕСТВЕНИЯ АНАЛИЗ НА ДАННИ

ЕМИЛИЯ СТ. ТОДОРОВА, ДИМО М. МИЛЕВ, ИВАЙЛО П. ДОНЧЕВ

ABOUT THE POSSIBILITY TO COMPUTERIZE QUALITATIVE DATA ANALYSIS

EMILIA ST. TODOROVA, DIMO M. MILEV, IVAYLO P. DONCHEV

***ABSTRACT:** Qualitative methods of data collection and analysis produce large amounts of data. Naturally the question arises whether it is possible and how these activities can be automated. In this study we examined and principles of the various methods of qualitative data analysis and summarizes the opportunities for computer support of these methods. We tried to identify some activities that can be assisted by computer software.*

***KEYWORDS:** Qualitative Data Analysis (QDA), Computer Assisted Qualitative Data Analysis Software (CAQDAS), Grounded Theory (GT), qualitative research*

1. Увод

Компютърният софтуер все по-често се използва за подпомагане на анализа на качествени данни. Сред изследователите в тази област се налага мнението, че той е от голяма полза при организирането и управлението на данни. Те твърдят, че по този начин изследователят е "по-близо" до данните. Основната функционалност на тези програми се състои в кодиране и извличане. Ако веднъж текстът се кодира от изследователя, търсенето е бързо и всеобхватно. Има възможност да се реализира и сложно логическо търсене.

Една част от изследователите, занимаващи се с качествен анализ на данни, отричат възможността компютърните програми да бъдат нещо повече от "механична помощ". Те заявяват, че въвеждането на компютърни програми за анализ (както за количествени, така и за качествени данни) позволява на потребителите да правят комплексни анализи, без да разбират принципите на самия анализ [1], [2], [3].

Сред друга част обаче се налага тенденцията, че софтуерът може действително да направи повече от това просто да подпомага кодирането. В [4] внимателно и ясно се посочва, че Computer Assisted Qualitative Data Analysis Software (CAQDAS) не може да прави анализ и изследователят все още има главна роля за него. Въпреки това тези автори описват софтуера за код-базираните методи като "позволяващ на изследователя да тества взаимоотношения между проблеми, концепции, теми (например да развият по-широки или от по-висок ред категории)", или в другата крайност, "да се развият по-подробни специфични кодове, когато определени условия се съчетават в данните". Те твърдят, че с помощта на софтуера за търсене на връзки между кодове (например, съвместно наличие или близост) "ще се създаде вторично ниво на аналитично кодиране". В допълнение заключават, че "когато се използват систематично и в рамките на зоната на комфорт на отделния изследовател", CAQDAS пакетите могат да способстват за увеличаване на прозрачността и методологическата строгост на провежданите изследвания.

В хода на нашето проучване на възможностите за софтуерно подпомагане на качествените научни изследвания бяха разгледани голяма част от методите за осъществяване

на подобни изследвания и седем CAQDAS пакета: ATLAS.ti; Dedoose; Digital Replay System (DRS); HyperRESEARCH; QSR Nvivo; MAXQDA; QDA Miner.

В настоящия доклад се правят някои изводи относно възможностите за използване на софтуер при част от методите за качествен анализ на данни.

2. Обоснована теория (Grounded Theory) и компютърен софтуер

Обоснованата теория (ОТ) е една от най-широко използваните методологии в качествените научни изследвания. Това е системна методология в областта на социалните науки, включваща формирането на теория чрез анализ на данни [5], [6]. Този изследователски метод работи по противоположен начин на традиционните изследвания в социалните науки. Вместо да започне с хипотеза, първата стъпка е събиране на данни чрез различни методи. От събраните данни основните точки се маркират с поредица от кодове, които се извличат от текста. Кодовете се групират в сходни концепции с цел да бъдат данните по-лесно използвани. От тези концепции се формират категории, които са в основата на създаването на обяснителен модел. Това противоречи на традиционните научни изследвания, при които изследователят първо избира теоретичната рамка и едва след това прилага избрания модел към проучваното от него явление [7]. От първоначалната си поява в здравеопазването [8] тази теория започва да се използва в различни области като езикознание, мениджмънт, производство и образование.

Обоснованата теория не е създадена като една всеобхватна теория. Тя е по-скоро методология за подпомагане на създаването на обяснителен модел, основан на емпирични данни. Перспективите за развитието на модел по този начин прави обоснованата теория атрактивна за следване методология при качествения анализ на данни. Въпреки това се твърди, че много изследователи се ограничават и правят подробно описание единствено на областта на научни изследвания или просто количествен анализ на съдържанието, а не обяснителен модел.

Таблица 1. Етапи на качествения анализ при ОТ

Етап	Цел
<i>Кодиране</i>	Идентифициране на ключова точка в събраните данни.
<i>Определяне на концепция</i>	Определяне на съвкупност от кодове със сходно съдържание, позволяващо групиране на данните.
<i>Определяне на категория</i>	Определяне на група от сходни концепции, способстващи за генериране на теория.
<i>Определяне на теория (обяснителен модел)</i>	Съставяне на съвкупност от категории, определяща предмета на изследването.

Отвореното кодиране (open coding) е концептуализиране на първо ниво на абстракция. Писмените данни се анализират ред по ред. В началото на изследването всичко се кодира, за да се открие проблема и начин за решаването му. Тази фаза е трудоемка, тъй като се концептуализират всички събития в данните. Те се сравняват при по-нататъшното кодиране на все повече данни, могат да се сливат в нови концепции, и в крайна сметка да се преименуват и модифицират. Процесите на кодиране в ОТ започват с отворено кодиране или разделяне на данните на отделни части, проверка в тези части за сходства и различия, както и групиране на концептуално подобни данни, за да се образуват категории. В [9] е описано концептуализиране, или даване на концептуално име на категории (представени например от възли в подпомагащия софтуер), като първа стъпка в теоретизирането. Когато е възможно, имената на възлите се поддържат активни с цел да подпомогнат анализатора да разглежда

по-скоро процеси, а не просто описания.

Обоснованата теория е един итеративен процес, при който изследователите се движат между събиране на данни и анализ, писане на бележки, кодиране и създаване на модели. Нелинейният дизайн на избраната CAQDAS програма, улеснява такива повтарящи се подходи. Примерите, дадени в [10] се фокусират върху NVIVO, но концепциите, представени там, могат да бъдат прилагани и по отношение на използването на други CAQDAS програми. Тези примери илюстрират как се поддържат и реализират техниките на ОТ (открито кодиране, писане на бележки, аксиално кодиране и създаване на модели) в рамките на самата програма.

Последователен мониторинг на кодирането може да се постигне чрез две функции: една, която записва дефинирания от изследователя възел и втора, която му позволява да прикачва бележки директно към възела. CAQDAS програмите обикновено имат функция за създаване на списък от възли и техните описания, който може да се отпечата на всеки етап на анализа. Тези описания, заедно с бележките към възела, дават възможност за последователно използване на възли.

В ОТ записките са от съществено значение за развитието на модела. Чрез писане на записки изследователят се придвижва от описателен начин на съпоставяне на концептуално подобни пасажи към аналитичен относно възникващите понятия. След препоръките на [9] са създадени различни типове записки с цел да се улесни анализиранията на различни нива. Името на записка може да започва с префикс, обозначаващ нейния тип, така че записките да могат автоматично да бъдат подредени в документен браузър на подпомагащия софтуер. В допълнение, за да бъдат подредени по азбучен ред, те могат да се сортират по големина, брой на свързани възли, както и дата на създаване или промяна.

В таблица 2 са показани седем типа бележки, използвани в NVIVO, техните префикси, номера и цели [10].

Таблица 2. Типове бележки в NVIVO

Префикс	Тип и номер	Цел
MCn	Code note memos (n=110)	Да дефинира възела и аналитичен подход към него; да създаде връзки към други възли и записки
MTH	Theoretical memos (n=5)	Записки за развитие на понятие от по-високо ниво на абстракция; обобщаващи бележки относно селективни извадки
MOP	Operational memos (n=4)	Бележки за процедури относно следваща итерация на събиране на данни
Без префикс, диаграми и изображения	Diagrams (n=40)	Визуално представяне на връзките между категориите
nCm	News, contextual memos (n=15)	Записки и статии, представящи контекста на изследвания проблем
nVa	NVIVO memos (n=3)	Технически бележки за използването на конкретния софтуер
mEM	Executive meetings (n=2)	Бележки от срещи и дискусии с официални лица относно резултатите от изследването

Препоръчва се анализаторът, използващ методологията на обоснованата теория, да изпълнява итеративно фазите на отворено кодиране, писане на бележки, аксиално кодиране и моделиране. CAQDAS може да улесни цялостния повтарящ се процес на събиране на данни, анализ и теоретизиране. CAQDAS програмите обикновено са предназначени да улеснят итерациите в рамките на кодирането и анализа на данните. Изследователят може да преминава бързо от отворено кодиране към откриване и кодиране на различия в категорията, и отново да се върне към отворено кодиране, като през цялото време създава концептуални и теоретични бележки. Компютърното подпомагане улеснява създаването на връзки между възли, бележки и документи при този повтарящ се процес. Дава и възможност за генериране на модел, в който първичните данни са организирани и концептуално свързани от анализатора.

3. Най-добре софтуерно поддържани качествени методи

Факт е, че CAQDAS пакетите са особено полезни за тези методически подходи, които софтуерните специалисти са имали предвид при проектирането на тяхната програма. Затова най-добре поддържани се оказват:

- Обоснованата теория [9], [11];
- Качествения анализ на съдържанието [12], [13];
- Етнографските изследвания.

Най-уникалната характеристика на CAQDAS програмите със сигурност е способността им за ефективно кодиране и поясняване на голямо разнообразие от данни, по-специално текстови. За разлика от стандартния количествен анализ на съдържанието, CAQDAS-подпомогнатото кодиране не се базира единствено на механично извършване на низови търсения, но задължително включва и тълкуване на аналитика.

В [14] се обсъжда увеличаване на използването на компютърно подпомаган качествен анализ на данните за етнографска работа. Авторите твърдят, че широко разпространеното влияние на CAQDAS пакетите постига единен режим за събиране, съхранение, анализ и представяне на данните.

В действителност, в основата на кодирането е стратегия, която по своята същност се поддава на количествен анализ. Без значение колко са сложни кодовете, те разслояват данните в стандартизирани категории, които след това могат да се анализират със статистически инструменти. За разлика от Обоснованата теория, голяма част от етнографските изследвания не са толкова строги в изискванията си към кодирането на данни. Вместо това те често изискват ефективно търсене чрез различни типове данни и са особено улеснени от възможностите за аотиране на данни. Различните програми се справят по различен начин с кодиращите и търсещите функции.

4. Методи, които частично могат да бъдат софтуерно поддържани

CAQDAS програмите често се описват като инструмент, който може да се използва за всички качествени изследвания, но най-често са проектирани за категоризиране на данни чрез система от кодове, съвместими с концепцията на обоснованата теория. Налични са няколко изследвания, използващи CAQDAS като инструмент за други качествени методи, а още по-малко на брой разработки показват използването им в дискурсен анализ. CAQDAS програми са използвани в някои количествени подходи за изследване на език (например анализ на съдържанието или корпусната лингвистика). При тях интерес представлява преброяването или измерването на езиковите характеристики, поддаващи се лесно на търсене и броене. Кодиращите възможности на компютърните програми са особено подходящи за изпълнението на тези задачи. Няма примери за използването на CAQDAS за дискурсен анализ (като Critical discourse analysis), който изследва как работи един език чрез

разглеждането на използването му в контекст. В обяснение на това [15] твърди, че докато CAQDAS може да осигури средства за мащабно количествено търсене, сериозен качествен анализ дори в по-малък набор от данни "се нуждае от по-дълбоко вникване".

Компютърният софтуер има ограничени възможности за приложение при някои от методите на качествен анализ. Например не всички форми на следните методи могат да бъдат съществено подпомогнати от него:

- Дискурсен анализ (Discourse Analysis);
- Визуален анализ (Visual Analysis);
- Анализ на съдържанието (Content Analysis);
- Сравнителен анализ (Comparative Analysis).

Визуалният анализ не се поддържа добре от някои от съществуващите CAQDAS пакети, защото нито един от тях не дава възможност за ефективно кодиране на определени рамкови последователности в клипове или части от изображения. Все пак някои програми предлагат известни мултимедийни възможности в рамките на системата, други позволяват свързването с външни мултимедийни данни. Като добър пример в това отношение може да се посочи безплатния софтуер Transana [16]. Той позволява удобно транскрибиране и аотиране на видео. За него се твърди, че е по-подходящ за повечето от целите на визуалния анализ, в сравнение с някои от останалите CAQDAS пакети. Сравнителните изследвания, които се фокусират върху институционални и социални структурни променливи, изискват предимно сортиране на данни. Това често може да се направи по-бързо в редовна база данни или дори посредством електронни таблици. По-нататъшният анализ на тези данни може да се извършва с помощта на fs/QCA (fuzzy set/ Qualitative Comparative Analysis), специално програмиран за приложение на размитите множества в социалните науки [17].

5. Заключение

Качествените методи произвеждат големи количества данни, които не се поддават лесно на манипулация, анализ или ръчно редуциране. Софтуерът може да улесни процеса на качествен анализ. Такива програми могат да изпълняват някои механични задачи на съхранение и кодиране на данни; извличане и агрегиране на предварително кодирани данни, да осъществяват връзки между категории на кодиране. Те обаче не могат да "анализират" данните в смисъла, в който го прави статистическия софтуер. Всички аспекти на концептуалната и аналитичната работа по осмисляне на данните все още се реализират от изследвателя.

Софтуерът може да осигури:

- структуриране на работата – позволява достъп до всички части на изследователския проект веднага;
- "близост до данните" и интерактивност – незабавен достъп до изходните файлове с данни;
- достъп и разглеждане на данните – чрез инструменти за търсене на текст (дума или фраза);
- кодиращи и извличащи функционалности – създаване кодове и извличане на кодирани части от текста;
- управление на проекти и организация на данни – чрез инструменти за управление на проекти и организиране на данни;
- търсене в базата данни – търсене на отношенията между кодове;
- инструменти за запис – бележки, коментари и анотации;
- изход – доклади на хартиен носител или файл.

Софтуерът не може:

- да реализира аналитично мислене, макар че той може да има функции, подпомагащи анализа;
- да извърши кодирането вместо анализатора. Като цяло, трябва да се реши какво може да се кодира и по какъв начин. Може да се реализира автоматично кодиране в текста, но все пак е важно да се провери какво е било автоматично кодирано. Възможно е, след кодиране на един пасаж, програмата да прави предложения за това как по друг начин може да бъде кодиран. Изследователят решава дали това, което се предлага от програмата, има смисъл;
- да редуцира субективността, да подобри надеждността и/или качеството на анализа, макар че може да има функции, подпомагащи подобряването на качеството на анализа;
- да насочва как да се анализира данните.

Компютърният софтуер за качествен анализ на данни подпомага процеса на анализ. Той осигурява инструментите, но изследователят е отговорен за анализа.

ЛИТЕРАТУРА

1. Lee, R. M., and N. G. Fielding. 1991. Computing for qualitative research: Options, problems and potential. In *Using computers in qualitative research*, ed. N. G. Fielding and R. M. Lee, pp.1–13. London: Sage.
2. Weitzman, E. A. 2000. Software and qualitative research. In *Handbook of qualitative research*, 2nd ed., ed. N. K. Denzin and Y. S. Lincoln, pp. 803–20. London: Sage.
3. Bringer, J. D., L. H. Johnston, and C. H. Brackenridge. 2004. Maximizing transparency in a doctoral thesis: The complexities of writing about the use of QSR*NVIVO within a grounded theory study. *Qualitative Research* 4 (2): pp. 247–65.
4. Lewins, A. and Silver, C., 2009, Choosing a CAQDAS Package - A working paper (6th edition) <http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas/files/2009ChoosingaCAQDASPackage.pdf>
5. Patricia Yancey Martin & Barry A. Turner, "Grounded Theory and Organizational Research," *The Journal of Applied Behavioral Science*, vol. 22, no. 2 (1986), p.141.
6. Faggiolani, C. (2011). "Perceived Identity: applying Grounded Theory in Libraries". *JLIS.it* (University of Florence) 2 (1). doi:10.4403/jlis.it-4592.
7. G. Allan, "A critique of using grounded theory as a research method," *Electronic Journal of Business Research Methods*, vol. 2, no. 1 (2003) pp. 1-10.
8. Glaser BG, Strauss A. *Discovery of Grounded Theory. Strategies for Qualitative Research*. Sociology Press, 1967.
9. Strauss, A. L., and J. Corbin. 1998. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. 2nd ed. London: Sage.
10. Joy D. Bringer, Lynne Halley Johnston and Celia H. Brackenridge. *Using Computer-Assisted Qualitative Data Analysis Software to Develop a Grounded Theory Project*. *Field Methods* 2006; Vol. 18, No. 3, August 2006 245–266: Sage.
11. Glaser, Barney (2001): *The Grounded Theory Perspective: Conceptualization Contrasted with Description*, Mill Valley, CA: Sociology Press.
12. Mayring, Philipp (2000): "Qualitative Content Analysis," *Forum: Qualitative Research* 1 (2), <http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm>, last accessed: April 19, 2004.
13. Gluser, Jochen and Grit Laudel. 1999. [Theoriegeleitete Textanalyse? Das Potential einer variablenorientierten qualitativen Inhaltsanalyse](#). *WZB Working Papers* P-99-401.
14. Coffey, A; Holbrook, B and Atkinson, P, 1996, Qualitative data analysis: Technologies and representation. *Sociological Research On-Line*, 1(1).
15. MacMillan, K., 2005, More Than Just Coding? Evaluating CAQDAS in a Discourse Analysis of News Texts. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* Volume 6, No. 3, Art. 25
16. <http://www.transana.org/>
17. http://www.restore.ac.uk/lboro/research/software/caqdas_primer.php#sthash.689ocSmc.dpufRagin, Charles C. (2000): "Fuzzy-Set Social Science," Chicago, IL: University of Chicago Press.