

УДК 303.09, 303.8

**МИКОЛА СИДОРОВ,**

*кандидат фізико-математичних наук, доцент кафедри методології та методів соціологічних досліджень Київського національного університету імені Тараса Шевченка*

## **Використання параданих у соціологічних дослідженнях**

*Анотація*

*У статті розглянуто парадані як особливий тип даних, зібраних під час опитування про сам процес збирання інформації, а саме поведінку респондентів, характеристики умов їх проживання, реакції на запитання, час відповіді тощо. Наведено приклади використання параданих у соціальних дослідженнях.*

**Ключові слова:** парадані, метадані, невідповіді

Упродовж останнього десятиліття дедалі більше опитувань проводять з використанням комп'ютерної техніки, причому не тільки для аналізу результатів опитувань, а й для збирання інформації, тобто набувають поширення комп'ютерно опосередковані методи збирання та аналізу інформації. Як результат виникає можливість фіксувати, накопичувати та аналізувати не тільки самі відповіді респондента, а й інші дані та результати інших процесів, які супроводжують власне опитування, зокрема відомості про час на роздуми над відповіддю, емоційне забарвлення, зовнішній вигляд респондента, причину відмови від анкетування тощо.

Унаслідок цього з'являються нові можливості для використання метаданих, тобто даних про дані дослідження — від формулювання запитань та альтернатив, часу, місця проведення опитування до інформації про зручний час для інтерв'ювання, відгуки сусідів про респондента тощо.

Отже, важливість використання метаданих є очевидною. На жаль, в Україні таку інформацію ще рідко використовують у соціальних дослідженнях.

Сформулюю в загальних рисах, для чого можуть бути потрібні метадані. У соціальних науках такі дані використовують з кількох причин [Gregory, 2009: р. 7]:

- для впевненості в тому, що інформація в достатньому обсязі є доступною для користувачів, для правильного розуміння та використання даних. Без суттєвої документації та опису дослідник не зможе адекватно інтерпретувати та зрозуміти сутність даних;
- для полегшення пошуку та доступу до даних за ознаками застосування: ключовими словами, темою та метою дослідження, описом вибірки і генеральної сукупності та ін. Навіть найкращі дані у світі є марними, якщо про них ніхто не знає і не використовує їх;
- для підтримки тривалого збереження даних з упевненістю, що важлива інформація залишається з ними для подальшого використання або для перекодування в інші формати;
- для забезпечення та підтримки обміну інформацією між індивідами або інститутами.

Чим детальнішою і повнішою є супровідна документація про дослідження, його методологію, методи, генеральну сукупність, вибірку, тим зручніше використовувати його результати дослідникам.

Одним із різновидів метаданих є парадані — інформація про сам процес опитування.

**Мета** даної статті — розкрити концепт параданих як особливого виду метаданих та навести приклади їх використання у соціологічних дослідженнях на основі аналізу міжнародного досвіду.

Соціальні дослідження в багатьох країнах мають добре розвинену практику використання метаданих, адже самі дані дослідження — це відповіді респондентів, зібрані, опрацьовані, проаналізовані й використані в перебігу дослідження, а метадані — це вся документація про дані. Сучасні німецькі дослідники А.Грегорі та його співавтори [Gregory, 2009: р. 2] уводять таку типологію метаданих:

**Структурні метадані** описують структуру даних: імена змінних; які значення змінних з'являються у яких стовпчиках; який рядок подає який випадок; чи є ієрархічні зв'язки тощо. Цю інформацію має містити паспорт анкети (статистичний пакет ОСА) або вкладка "Змінні" (статистичний пакет SPSS).

**Описові метадані** являють собою інформації про методологію, вибірку, якість вимірювання. Така інформація має бути в програмі дослідження.

**Адміністративні метадані** — це інформація, утворювана самим процесом управління даними під час їх збирання, опрацювання, публікації тощо. До таких даних відносять дані про права доступу до результатів, юридичні документи дослідження, інформацію про формат файла збереження даних та місце його розташування тощо.

**Поведінкові метадані** (які називають *параданими*) є інформацією про реакцію та поведінку дослідників при використанні даних дослідження, а також поведінку інтерв'юєрів та респондентів під час опитування. До параданих відносять оцінки інтерв'юєром зовнішнього вигляду респондента, реакцію респондента на запитання, стан домогосподарства, враження сусідів, причини відмов.

Надалі ми детальніше спинимося лише на четвертому типі, який називають параданими, чому і присвячена стаття.

Термін “paradata (парадані)” був застосований уперше Міком Купером, американським соціологом у 1998 році у Далласі на семінарі JSMASA (Joint Statistical Meetings of the American Statistical Association) під час обговорення доповіді [Couper, 1998] “Вимірювання якості опитувальника у середовищі комп’ютерно опосередкованого збору даних” та на 53-й конференції AAPOR (American Association for Public Opinion Research) [Hansen, 1998]. Внесена Купером пропозиція про використання даних про процес опитування, автоматично згенерованих комп’ютером під час CAPI (Computer Assisted Personal Interviewing), згодом була розширена щодо інших аспектів самого вибіркового процесу та інших методів збирання даних – CATI (Computer Assisted Telephone Interviewing) та PAPI (Paper Assisted Personal Interviewing).

Використання даних про процес збирання інформації дає можливість поліпшити якість результатів та зменшити похибки дослідження, проаналізувати причини невідповідей, оцінити їх вплив на результати та визначити оптимальні характеристики для подальших досліджень, таких як час проведення дослідження, відбір респондентів тощо [O’Reilly, 2009; Kreuter, 2010(a); (b)], що було відзначено на певних міжнародних конференціях (International Workshop on Household Survey Nonresponse, European Survey Research Association (ESRA), American Association for Public Opinion Research (AAPOR) та ін.).

Стандартного однозначного визначення поняття “парадані” ще немає. Спочатку парадані визначали як дані, автоматично записані комп’ютером під час проведення інтерв’ю [Hansen, 1998], але з часом значення цього терміна було розширене і почало охоплювати дані, записані інтерв’юєром під час проведення опитування або ж отримані з інших джерел, наприклад аудіозапису телефонного інтерв’ю [Couper, 1998].

У таблиці 1 записано різновиди походження параданих залежно від типу опитування.

*Таблиця 1*

**Джерела отримання параданих**

Тип дослідження	Місце розташування параданих	Різновид запису
CAPI	Сервер, персональний комп’ютер	Автоматично комп’ютером
CATI	Сервер	Автоматично комп’ютером, частково інтерв’юєром
PAPI	Контактна форма	Інтерв’юєром

Важливо не стільки дати чітку інтерпретацію поняттю “парадані”, скільки зрозуміти сам концепт отриманих під час опитування даних про його перебіг.

Одним із типів даних, який вважають параданими, є записи дзвінків, зібрані впродовж опитування вибірки через телефонне інтерв’ю. Час дзвінка, дата, так само як і результат дзвінка (відмова, інтерв’ю, не піднімають трубку, не дійсне, не закінчене тощо), є доступними даними, автоматично записаними системою CATI під час спроби провести інтерв’ю. За цими даними можна, наприклад, скоригувати час проведення інтерв’ю для зменшення

кількості відмов або ситуацій відсутності абонента [Laflamme, 2008]. У США, наприклад, проводиться дослідження (US National Survey of Family Growth), де парадані використовують для щоденної централізованої корекції схеми опитування через телефон [Lepkowski, 2009]. Однак більшість досліджень парадані аналізують після збирання інформації.

Другий набір даних, який відносять до параданих, також збирають на початковій фазі встановлення контакту та залучення до опитування типових респондентів. Відмінність від параданих, розглянутих вище, полягає в тому, що тут сам інтерв'юєр збирає ці дані, які мають охоплювати всі випадки — як ті, коли респондент погоджується взяти участь у інтерв'ю, так і випадки відмови. Інтерв'юєри мають фіксувати спостереження щодо оточення респондентів, житлових характеристик, вказувати причини відмови брати участь у дослідженнях. Для підвищення частки відповідей були розроблені з урахуванням цих особливостей цілі схеми й системи рекомендацій інтерв'юєрам, наприклад PEDAksi (Pre-Emptive Doorstep Administration of Key Survey Items) [Lynn, 2003], IAS (Interviewer Attitudes Survey) [Campanelli, 1997] та ін.

Прикладом збирання параданих може послугувати дослідження National Survey of Family Growth, яке проводиться у США. Це багатокрокове дослідження домогосподарств, у яких для опитування відбирають особу віком від 15 до 44 років, з якою проводять інтерв'ю впродовж 60–80 хвилин з використанням комп'ютера. Дослідження сфокусоване на аспектах народжуваності та сексуального досвіду. Інтерв'юєрам доводилось спостерігати за домогосподарствами й оцінювати наявність сексуального партнера протилежної статі у респондента, наявність дітей у домогосподарстві. Разом із тим в анкеті теж фігурували запитання на такі теми. Таким чином, парадані використовували для зважування результатів опитування [Groves, 2007].

У Європейському соціальному дослідженні (ESS) у всіх хвилях використовується контактна форма, що передбачає заповнення інтерв'юєром [Соціальне дослідження ESS (a), 2008] і має забезпечити досить повну картину стосовно самого респондента, його місця проживання, готовності взяти участь в опитуванні та ін. Ці дані відносять до параданих і використовують як для коригування самого процесу збирання інформації, так і для корекції впливу невідповідей. У документі “Response-based Quality Assessment” [Соціальне дослідження (b), 2009] описано роль параданих, отриманих із контактних форм інтерв'юєрів. Спочатку збирають відомості про всі типові одиниці, включно з тими, хто відповів на анкету, і тими, хто відмовився від анкетування (така контактна форма [Соціальне дослідження ESS (a), 2008] також включає інформацію про візуальні спостереження інтерв'юєра, його оцінку середовища перебування). Це вможливило опис людей, які відмовляються від опитування, і тих, хто готовий його пройти. У разі відмови інтерв'юєр має оцінити готовність респондента брати участь (або ж знову відмовитися) у наступних або повторних опитуваннях. Саме інформація про “невідповіді” і дає змогу надалі коригувати модель вибірки. Після цього щодо всіх респондентів збираються оцінки їх помешкання (домогосподарств) стосовно умов проживання. Крім того, ці дані збираються впродовж усього дослідження, тому можна простежити ще й певну динаміку.

Контактні форми постійно модифікуються і поліпшуються задля більшої зручності їх використання інтерв'юєром.

Ці приклади ілюструють можливість як поточних коригувань програми дослідження для зменшення невідповідей, так і модифікації програми наступних досліджень. Умови появи невідповідей можна відстежити як за респондентами, які погодилися брати участь у дослідженні, так і за тими, хто відмовився, адже парадані фіксують і для тих, і для тих. За ними можна моделювати процес опитування таким чином, щоби зменшити появу невідповідей.

В автоматизованих опитуваннях є можливість записувати всю розмову, включно з інтонацією, мімікою, емоціями, а не тільки фіксувати контактну форму. Ці дані віднесемо до третього типу параданих. Дослідження підтверджують, що реакція респондента на запрошення взяти участь в опитуванні свідчить про його рівень потенційної згоди на інтерв'ю; опріч того фіксується, як голосові характеристики інтерв'юєра впливають на згоду на опитування [Groves, 2007]. Аудіозапис (або ж відеозапис) та отримані в такий спосіб парадані становлять інтерес не лише на етапі рекрутування респондентів для проведення опитування, а й для оцінювання похибки вимірювань для кожного із запитань, адже емоційний стан впливає на розважливість респондента. Навіть коментарі респондента до питань можуть вплинути на його вибір відповідей — наскільки він переймається проблемою та наскільки він емоційно ставиться до варіантів відповіді та самого запитання. Отже, з аудіозапису можна отримати такі парадані: зміна інтонації та міміка респондента при прослуховуванні альтернатив, зміни варіантів своїх відповідей на запитання, коментарі до запитань та альтернатив, час на роздуми, уточнення запитання; з відеозапису — ще міміка, сконцентрованість чи розсіяність, рух очей, жести тощо. На перший погляд, ці додаткові дані не є суттєвими, але насправді вони містять багато об'єктивних характеристик (тобто зумисно не контрольованих ані респондентом, ані інтерв'юєром), про що свідчать деякі дослідження (наприклад, [Congrad, 2010]). Скажімо, за рухом очей можна оцінити правдивість відповідей респондента, за мімікою — рівень збудженості, який впливає на емоційність відповідей, тощо.

Окремо слід сказати про збирання параданих під час проведення опитувань у середовищі Інтернет, де є можливість здійснити не тільки анкетування, а й аудіо- та навіть відеоопитування, про ефективність та можливості одержання інформації в межах яких уже говорили чимало дослідників (див.: [Fuchs, 2007]). Передусім автоматично збирається інформація про підключення респондента до Інтернету (IP-адреса, провайдер, регіон, тип підключення) та деякі параметри комп'ютера (тип браузера, розподільна здатність екрана, операційна система, деякі додаткові характеристики комп'ютера, які вільно транслюються під час передавання запитів). Автоматично фіксується час початку та закінчення опитування, крім того, більшість систем для проведення опитувань он-лайн мають можливість фіксувати не тільки час, що його витратив респондент на кожну з відповідей, а й навіть рух мишкою по екрану під час роздумів для визначення зосередженості на відповіді та активності вікна з опитувальником.

При проведенні досліджень у web (електронна анкета, розташована на web-сторінці) парадані можна отримувати як із серверної частини опитувальника (server side), так і з клієнтської (client side). Із сервера отримують інформацію про те, звідки зайшов респондент до опитувальника (IP-адреса), коли зайшов, о котрій почав опитування, коли його закінчив, відповів на всі запитання за один раз чи переривав процес заповнення анкети, персональний пароль доступу, адресу, з якої було здійснено перехід на анкету, тип операційної системи, тип браузера. За останніми двома параметрами можна оцінити швидкодію комп'ютера та здійснити поправку на швидкість завантаження для поправки оцінки часу, витраченого на відповіді. Ці парадані зберігаються на сервері або у вигляді таблиці бази даних, або у вигляді текстового файлу з протоколом. Значно цікавіші парадані можна отримати з клієнтської частини опитувальника. Клієнтська частина опитувальника є програмою, вбудованою в саме тіло анкети. Ця програма, як правило, написана на Java, php або іншою мовою програмування у web і покликана збирати додаткову інформацію про перебіг опитування. Вона може працювати лише тоді, коли це дозволено комп'ютером клієнта і не є вірусом, шпигунською системою чи будь-яким іншим небажаним компонентом. З допомогою клієнтської частини опитувальника можна отримувати інформацію про: час, який витрачає респондент на відповіді з кожного запитання, інформацію про те, чи змінював відповіді респондент під час опитування, його навігацію за запитаннями опитувальника, якщо це дозволено самим опитувальником, оцінити ці параметри стосовно різних типів запитань та різновидів подання альтернатив (кнопки, галочки, випадаючі меню альтернатив, табличні запитання з ранжованими сумісними альтернативами тощо). Наприклад, виявилось, що випадаючі меню альтернатив (drop-down boxes) є менш зручними для респондентів, ніж кнопки radio button (набір альтернатив, з якого можна обрати лише один варіант відповіді, причому всі альтернативи записані в рядок або стовпчик) [Sowan, 2010]. Використання цих client side параданих має досить широкий спектр. Бельгійський соціолог Д.Геервег у [Heerwegh, 2004] наводить чотири приклади ефективного використання параданих, отриманих з клієнтської частини web-опитувальника: калібрування індикатора рівня заповненості анкети як одного з чинників, що заважає вичерпному заповненню анкет респондентами; оцінка ефекту використання запитань у різних форматах для підвищення зручності фіксації відповідей для респондента; тест ефективності різних формулювань запитань для виявлення найбільш однозначного й неупередженого формулювання як самого запитання, так і альтернатив; оцінювання компетентності та рівня ставлення до досліджуваної проблеми через оцінювання часу відповіді та зміни відповідей (траєкторію відповідей).

Очевидно, що цей перелік прикладів не є повним, але він дає напрям для виявлення можливостей використання параданих у соціологічних дослідженнях.

Отже, ми можемо сформулювати кілька головних причин застосування параданих:

- поточний аналіз процедури збирання даних та її корекція, що сприяє економії коштів;
- використання параданих для ремонту вибірки після збирання інформації, що вможливорює підвищення якості результатів дослідження.

Кожна соціологічна компанія намагається підвищити якість зібраної соціологічної інформації з мінімальними фінансовими витратами. Впливовим чинником зменшення репрезентативності вибірки є рівень невідповідей як в анкетуванні, так і в телефонному інтерв'ю. Національна канадська статистична компанія Statistics Canada [Statistics Canada, s.a.] виступає одним з ініціаторів пошуку різних додаткових можливостей для поліпшення якості результатів дослідження без зростання його вартості. Одним із таких шляхів є використання параданих при проведенні САТІ [Laflamme, 2008]. До статистичної бази записують усі події, що відбуваються під час опитування: дата першого та всіх подальших дзвінків, час початку дзвінка, час закінчення дзвінка, час на кожну відповідь, стан контакту (відмова без погодження на повторне опитування, відмова з перенесенням часу опитування та ін.), відповідь на всі запитання впродовж одного сеансу чи кількох, місце перебування респондента (у разі дзвінка на мобільний телефон), емоційний стан респондента тощо. За цими даними коригується оптимальний час контакту, манера звернення інтерв'юєра до респондента [Conrad, 2010] для встановлення контакту тощо. Використання параданих дало змогу зменшити час, витрачений інтерв'юєром на опитування, та кількість невідповідей.

Ще одним застосуванням параданих є спроба знайти зв'язок між параданими й відповідями й, відповідно, здійснити подальше коригувальне зважування даних або невідповідей [Little, 2005; Maitland, 2008]. У доповіді Крепера 18 лютого 2011 року на “Research Agenda for the Future of Social Science Data Collection Workshop” [Kreuter, 2011] наведено кореляційні залежності між параданими й деякими ознаками в різних соціальних дослідженнях. У [Kreuter, 2010(a)] та [Maitland, 2008] також виявлено залежності між рівнем відмов і характеристиками респондентів. Наприклад, у [Maitland, 2008] викладено результати за дослідженням, яке проводилось у США 2006 року; обсяг вибірки становив 44264 сім'ї, з яких було відібрано 31142 (ті, в яких було знайдено контакт бодай з одним із членів родини). Розглядали три типи змінних: парадані ( $z$ ), участь в опитуванні ( $p$ ) та змінні інтересу ( $y$ ). До параданих віднесли три типи змінних, які гіпотетично могли б вплинути на корекцію невідповідей: одна група змінних описує зусилля зі встановлення контакту із членом сім'ї (наявність завад для доступу до помешкання, відсутність членів родини вдома тощо); друга група вимірює рівень співпраці в домогосподарстві (респондент заявляє, що йому не цікаве це дослідження, він хворий, втомився тощо); до третьої групи належить відмова від інтерв'ювання впродовж самого опитування (респондент перериває інтерв'ю, посилаючись на здоров'я або на інші причини).

Спочатку аналізувалися кореляції між змінними груп  $z$ ,  $p$  та  $y$ , і  $z-p$  кореляції виявилися в межах 0,02–0,41 із середнім значенням 0,14. Найсильнішим прогнозом незгоди відповідати на запитання анкети була змінна параданих, яка показувала відсутність зацікавленості та зайнятість респондента. Натомість кореляція між рівнем контактності та згодою на анкетування коливалася між 0 і 0,17 із середнім 0,07, а найсильнішим індикатором був факт, коли в домогосподарстві не відповідали на дзвінок (нікого не було вдома або удавали відсутність мешканців). Найбільшим коефіцієнтом кореляції між змінними  $z$  і  $y$  був 0,1 із середнім значенням 0,05.

Далі було застосовано факторний аналіз для змінних, які відображали співпрацю та контактність. У таблиці 2 наведено результати зв'язку між факторами, з одного боку, та участю в опитуванні і змінними зацікавленості (інтересу) — з іншого.

Таблиця 2

## Кореляційна таблиця

Набори змінних	Зв'язок між факторами та участю в опитуванні	Зв'язок між факторами та змінними зацікавленості	
		діапазон	середнє
<i>Змінні співпраці</i>			
Фактор 1: Зміст опитувальника/приватність	0,28	0,01–0,06	0,03
Фактор 2: Відсутність часу	0,24	0,01–0,08	0,04
Фактор 3: Ворожість щодо прохання пройти анкетування	0,47	0,00–0,05	0,02
Фактор 4: Проблема доступу до помешкання (консьєрж, сторож)	0,07	0,00–0,04	0,01
<i>Змінні контактності</i>			
Фактор 1: Проблеми налагодження контакту	<b>0,18</b>	0,00–0,10	0,05
Фактор 2: Місце проживання/проблема відбору	0,02	0,00–0,04	0,02
Фактор 3: Проблеми перешкод (ворота, паркан)	0,11	0,00–0,06	0,03

Через логістичну регресію зі змінною “статус участі в опитуванні” як залежною змінною та всіма індивідуальними змінними параданих як незалежними змінними було отримано оцінку схильності до участі в опитуванні для кожної особи.

Отже, можемо сказати, що використання параданих може суттєво поліпшити якість результатів соціальних (маркетингових) досліджень без збільшення вартості самого дослідження та вможливити точніше моделювання вибіркової сукупності. Особливо це стосується систем CAPI та CATI, хоча, як, наприклад, у разі ESS, використання PAPI з контактними формами також уможлиблює збирання параданих. Використання параданих під час проведення web-опитувань показує високу ефективність для аналізу як ди-зайну анкети, так і підвищення відсотка відповідей.

Таким чином парадані можуть бути використані:

- під час пілотажного дослідження для моделювання вибіркової сукупності та корекції програми дослідження;
- під час проведення дослідження для програмування подальших досліджень;
- під час проведення дослідження (завдяки аналізу невідповідей) для корекції результатів дослідження;
- під час аналізу даних для переважування відповідей.



Виникає логічне запитання: в якому форматі зберігати парадані (та інші метадані) так, щоб він був, з одного боку, універсальним, а з іншого — зрозумілим дослідникам? Існує багато стандартів для зберігання метаданих для різних галузей науки і техніки [Означення “метадані”, s.a.]. У соціальних науках використовують головним чином кілька стандартних моделей метаданих: Data Documentation Initiative (DDI) [Data Documentation Initiative, s.a.], ISO-TS 17369 Statistical Data and Metadata Exchange3 (SDMX) [Statistical Data and Metadata, s.a.], Dublin Core Metadata Initiative (DCMI) [The Dublin Core Metadata, s.a.], ISO/IEC 11179 [Metadata Standards, s.a.] та Neuchâtel models for variables and classifications [Neuchâtel Model, s.a.]. Не порівнюватиму зараз ці моделі, лише зазначу, що всі вони покликані якомога краще відповідати зберігання метаданих.

На жаль, залишаються нерозв'язаними етичні питання використання параданих, адже у них міститься досить суб'єктивна і приватна інформація, яка не має бути доступною до всіх. Отже, потрібно фільтрувати обсяг і тип параданих для загального доступу так, щоб не зашкодити респондентам та не порушити їхніх прав, тобто видаливши персональні ідентифікаційні дані.

Зрозуміло, що парадані не є універсальним механізмом для зниження вартості дослідження та підвищення його якості. Це тільки ще одна з можливостей знизити вплив невідповідей на результати дослідження, отримати додаткову інформацію для зважування самих відповідей, оцінити модель вибірки та шлях її реалізації. Далеко не завжди у дослідженнях такі дані вирізняють в окрему категорію й називають “параданими”, але тенденція до використання додаткової інформації задля поліпшення результатів зберігається.

### *Джерела*

Означення “метадані” у Вікіпедії [Електронний ресурс]. — Режим доступу: <http://ru.wikipedia.org/wiki/Метаданные>.

Соціальне дослідження ESS, контактна форма 4 раунду 2008 рік [Електронний ресурс]. — Режим доступу : [http://ess.nsd.uib.no/streamer/?module=main&year=2009&country=UA&download=%5CFieldwork+documents%5CUA%5C2009%5C04%23ESS4+-+Contact+forms%5CLanguages%5CRussian%5CESS4ContactformUA\\_rus.pdf](http://ess.nsd.uib.no/streamer/?module=main&year=2009&country=UA&download=%5CFieldwork+documents%5CUA%5C2009%5C04%23ESS4+-+Contact+forms%5CLanguages%5CRussian%5CESS4ContactformUA_rus.pdf).

Соціальне дослідження ESS, Response Based Quality Assessment [Електронний ресурс]. — Режим доступу : [http://ess.nsd.uib.no/streamer/?module=main&year=2009&country=null&download=%5CSurvey+documentation%5C2009%5C09%23ESS4+-+Response+Based+Quality+Assessment%5CLanguages%5CEnglish%5CResponse-based+Quality+Assessment+ESS4CF\\_e02.pdf](http://ess.nsd.uib.no/streamer/?module=main&year=2009&country=null&download=%5CSurvey+documentation%5C2009%5C09%23ESS4+-+Response+Based+Quality+Assessment%5CLanguages%5CEnglish%5CResponse-based+Quality+Assessment+ESS4CF_e02.pdf).

*Campanelli P.* Can you hear me knocking: An investigation into the impact of interviewers on survey response rates / P. Campanelli, P. Sturgis, S. Purdon // Technical report (Final Report for UK ESRC grant R000235776). — London : The Survey Methods Centre at SCPR (Social and Community Planning Research), 1997.

*Conrad F.* To agree or not to agree: Effects of spoken language on survey participation decisions / [F. Conrad, J.G. Broome, J. Benki, R. Groves, F. Kreuter, D. Vannette] // 63th Annual Conference of the American Association of Public Opinion and Research (AAPOR). — Chicago (ILL), 2010. — May 16.

*Couper M.P.* A Measuring Survey Quality in a CASIC Environment [Electronic resource] / Couper M.P. : Invited paper presented at the Joint Statistical Meetings of the American Statistical Association // Proceedings of the Survey Research Methods Section, ASA (1998), Achieving Quality in Surveys, Dallas, August. — P. 41–49. — Mode of access : <http://www.amstat.org/sections/srms/Proceedings/y1998f.html>.

Data Documentation Initiative (DDI) [Electronic resource]. — Mode of access : <http://www.ddialliance.org/>.

The Dublin Core Metadata Initiative [Electronic resource]. — Mode of access : <http://dublincore.org/>.

*Fuchs M.* Video Web Survey — Results of an experimental comparison with a text-based Web survey [Electronic resource] / M. Fuchs, F. Funke // Proceedings of the Association for Survey Computing's Fifth International Conference on the Impact of Technology on the Survey Process.— Southampton (Eng.). — 2007. — September. — P. 63–80. — Mode of access : <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.99.960&rep=rep1&type=pdf>.

*Gregory A.* Metadata [Electronic resource] / Gregory Arofan, Heus Pascal, Ryssevick Jostein // RatSWD, working paper № 57, march 2009 P22. — Mode of access : [http://www.ratswd.de/download/RatSWD\\_WP\\_2009/RatSWD\\_WP\\_57.pdf](http://www.ratswd.de/download/RatSWD_WP_2009/RatSWD_WP_57.pdf).

*Groves R.* Use of interviewer judgments about attributes of selected respondents in post-survey adjustment for unit nonresponse: An illustration with the National Survey of Family Growth [Electronic resource] / R. Groves, J. Wagner, E. Peytcheva // Proceedings of the Section on Survey Research Methods of the American Statistical Association. — 2007. — P. 3428–3431. — Mode of access : <http://www.amstat.org/sections/srms/Proceedings/y2007/Files/JSM2007-000782.pdf>.

*Hansen S.E.* Usability Evaluation of the NHIS CAPI Instrument [Electronic resource] / S.E. Hansen, Couper M.P., Fuchs M. // Papers Presented at the 53rd Annual Conference of the American Association for Public Opinion Research. — 1998. — May 14–17, St. Louis, Missouri. — P. 928–933. — Mode of access : [http://www.amstat.org/sections/srms/Proceedings/papers/1998\\_160.pdf](http://www.amstat.org/sections/srms/Proceedings/papers/1998_160.pdf).

*Heerwegh D.* Uses of client side paradata in Web surveys [Electronic resource] / D. Heerwegh // International Symposium in Honour of Paul Lazarsfeld. — Brussels : Belgium, 2004. — Mode of access : <http://www.websm.org/uploadi/editor/Heerwegh20Uses20of20Client20Side20Paradata20in20Web20Surveys.pdf>.

*Kreuter F.* (a) Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data [Electronic resource] / F. Kreuter, G. Müller, M. Trappmann // Public Opinion Quarterly. — 2010. — Vol. 74, № 5. — P. 880–906. — Mode of access : <http://poq.oxfordjournals.org/content/74/5/880.full.pdf>.

*Kreuter F.* (b) Paradata [Electronic resource] / F. Kreuter, C. Casas-Cordero // RatSWD, working paper № 136, april 2010 P16. — Mode of access : [http://www.ratswd.de/download/RatSWD\\_WP\\_2010/RatSWD\\_WP\\_136.pdf](http://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_136.pdf).

*Kreuter F.* Using Paradata for Post Survey Adjustments: State of the Science [Electronic resource] / F. Kreuter // Panel on a Research Agenda for the Future of Social Science Data Collection Workshop. — 2011. — February 17–18. — Washington (DC). — Mode of access : <http://www7.nationalacademies.org/cnstat/Kreuter%20Pres.pdf>.

*Laflamme F.* Data Collection Research using Paradata at Statistics Canada [Electronic resource] / F. Laflamme // Symposium 2008 : Data Collection : Challenges, Achievements and New Directions. — Component of Statistics Canada Catalogue. — № 11-522-X. — Mode of access : <http://www.statcan.gc.ca/pub/11-522-x/2008000/article/10997-eng.pdf>.

*Lepkowski J.* Use of paradata to manage a field data collection / [J. Lepkowski, R. Groves, W. Axinn, N. Kirgis, W. Mosher] // Proceedings of the Section 6 on Survey Research Methods of the American Statistical Association. — 2009. — August 2. — Joint Statistical Meetings.

*Little R.* Does Weighting for Nonresponse Increase the Variance of Survey Means? [Electronic resource] / R. Little, S.L. Vartivarian // Survey Methodology. — 2005. — 31. —

P. 161–168. — Mode of access :

<http://www.mathematica-mpr.com/publications/pdfs/weightingnonresponse.pdf>.

*Lynn P.* PEDAksi : Methodology for collecting data about survey non-respondents [Electronic resource] / P. Lynn // *Quality and Quantity*. — 2003. — № 37 (3). — P. 239–261. DOI: 10.1023/A:1024449208633. — Mode of access :

<http://www.iser.essex.ac.uk/publications/working-papers/iser/2002-05.pdf>.

*Maitland A.* An Exploration into the Use of Paradata for Nonresponse Adjustment in a Health Survey [Electronic resource] / A. Maitland, C. Casas. — Cordero, Frauke Kreuter // *Proceedings of the Survey Research Methods Section, American Statistical Association (2008). Session 381: Paradata, Data Quality, and the National Health Interview Survey (NHIS)*. — Mode of access :

<http://www.amstat.org/sections/SRMS/proceedings/y2008/Files/301425.pdf>.

Metadata Standarts ISO/IEC JTC1 SC32 WG2 [Electronic resource]. — Mode of access : <http://metadata-standards.org/>.

Neuchâtel Model — Classifications and Variables [Electronic resource]. — Mode of access : <http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=14319930>.

*O'Reilly J.* Paradata and Blaise: A review of recent applications and research [Electronic resource] / O'Reilly J. : Paper presented at the 12th International Blaise Users Conference (IBUC). — 2009. — Latvia. — Mode of access :

<http://www.blaiseusers.org/2009/papers/7d.pdf>.

*Sowan A.K.* Paradata: A New Data Source From Web-Administered Measures [Electronic resource] / Sowan A.K., Jenkins L.S. // *CIN: Computers, Informatics, Nursing*. — 2010. — Vol. 28. — № 6 (November/December). — P. 333–342. — Mode of access :

<http://www.nursingcenter.com/pdf.asp?AID=1079600>.

Statistics Canada [Electronic resource]. — Mode of access :

<http://www.statcan.gc.ca/start-debut-eng.html>.

Statistical Data and Metadata exchange [Electronic resource]. — Mode of access : <http://sdmx.org/>.