**9. előadás: Dokumentum- és tartalomelemzés**

**I. Mi is az a tartalomelemzés?**

1.1. Tartalomelemzés általánosságban

A tartalomelemzés tudatos, módszeres formája mellett mindenki végez ösztönös tartalomelemzést, amikor újságot olvas, tévét néz, vagy rádiót hallgat, a másik ember közléseit hallgatja. A közlések értelmét kategóriákba rendezi: „ez most jó”, „ez rossz”, „ez kedvező rám nézve”. Az ösztönös tartalomelemzés igazi kvalitatív elemzés. Újságolvasásnál, hírek hallgatásánál sokszor a „sorok között” olvasunk, vagy a mondottak mögött is hallunk egy „néma szöveget”. Sokszor a szövegösszefüggésekből érteni véljük a közlő beállítódását, hovatartozását, meggyőződéseit, értékrendszerét, bár valósában semmi ilyen nem szerepel a szövegben. Az ösztönös tartalomelemzést nem szabad lebecsülnünk, hiszen ennek a bennünk rejlő „logikai rendszernek” a formalizálásából született meg a tartalomelemzés módszere.

Másfelől az is nyilvánvaló, hogy az „ösztönös” formája alapvetően szubjektív, esetleges és nem tesz eleget a tudományos kutatás kritériumainak.

A tudományos igényű tartalomelemzés viszonylag új a társadalomtudományok módszerei között. Alkalmazása az utóbbi két évtizedben a társadalomtudományokban, mindenekelőtt a szociológiában, a pszichológiában, a politológiában, a médiakutatásban, de a pszichiátriában, a néprajzban, a nyelvészetben, illetve az irodalom- és történelemtudományban vált ismertté hazánkban is. Szélesebb körű érdeklődésre a közgazdaságtudomány vagy speciálisan a piackutatás részéről csak az utóbbi időben tart számot.

Természetesen egy-egy regény, politikai beszéd tartalmának elemzéséről már régebb óta beszéltek anélkül, hogy ez körülírt vizsgálati módszerré vált volna. A tartalomelemzés több tudományterületből származó, interdiszciplináris módszer.

A tartalomelemzés – klasszikus formájában – a kutató emberi produktumok sajátos osztályait, a közleményeket vizsgálja. Közleményeknek nevezünk minden szó- vagy írásbeli kommunikációt. Tágabb értelemben ide sorolható minden olyan emberi megnyilatkozás, amelynek mondanivalója „dekódolható”, mint pl. egy festmény, a tánc, a mozgások és gesztusok, az álmok. A klasszikus tartalomelemzés sajátos megközelítésmódot választ az adatok elemzésére azzal, hogy a közlések (közlemények) tárgyának a tartalmat tekinti.

A tartalomelemzés segítségével a piackutató azonosítja és feltárja a fókuszcsoportok, az interjúk szövegében, a reklámokban vagy akármilyen más szövegben a szavak jelentéseit, a fogalmak értelmét, illetve a használt fogalmak összetett kapcsolatait. A szavak, szófordulatok gyakoriságából és a fogalmak egymáshoz viszonyított kapcsolataiból következtetéseket von le – vagyis hipotéziseket vet fel – a közlés, közlemény forrásáról. Az azonosított, illetve feltárt tartalmak lehetnek nyilvánvalóak - manifeszt tartalmak - vagy mélyen rejlő, lappangó, nem a szöveg elsődleges jelentésében szereplő tartalmak – látens tartalmak. Ezek olyan információk, amelyek a közleményben nincsenek nyíltan kimondva. A közleményeknek nem egyetlen jelentésük van, hanem mindig többféle nézőpontból vizsgálhatók, a jelentésük egyénenként is különbözhet, különösen akkor, ha a vizsgálat témája az emberi kapcsolatok vagy a termékekhez való viszonyulások értékelése. Ahhoz, hogy a tartalomelemzés értelemmel telítődjön, és valamilyen rendszer szerint működjön, a fogalmak kapcsolódási mintázatainak vizsgálata előtt néhány előkészítő műveletet kell elvégezni. Ezek mind a kvantitatív, mind a kvalitatív tartalomelemzés esetében a mintavétel, az elemzési egységek meghatározása és a tartalmak kódolása.

Tartalomelemzésnek nevezzük azt a kutatási eljárást, amely segítségével a szövegadatokból olyan következtetések vonhatók le, amelyek a közleményekben nyíltan nincsenek kimondva, de a szöveg szerkezetéből, az elemek együttes előfordulásából és azok törvényszerűen visszatérő sajátosságaiból kiolvashatók. Bármilyen társadalommal kapcsolatos közlés elemezhető ez által, úgy, mint könyvek, levelek törvények gyűjteménye, stb. Alapkérdései: ki mit mond, kinek, miért, hogyan, és milyen hatással?

1.2. A tartalomelemzés tulajdonságai

* •nem igényel sem sok pénzt, sem sok időt
* •nem jár kockázattal, mivel nem avatkozik be a jelenségbe, ill. szakaszokra bontható a kutatás
* •lehetővé teszi hosszú idő alatt lezajló események vizsgálatát is
* •hátrány, hogy csak rögzített kommunikációkon végezhető
* •ha egy kommunikációs folyamat a vizsgálódás tárgya, nincs gond az érvényességgel és a megbízhatósággal
* •az anyagok konkrétsága növeli a megbízhatóság esélyét
* •a kódolás bármikor megismételhető

**2. Történeti áttekintés**

**2.1. Kvantitatív újságelemzés**

A tartalomelemzést először a sajtóval foglalkozó kutatók, majd később a szociológusok alkalmazták, amerikai újságok tartalmának tanulmányozására. Ez a munka a Columbiai Egyetem Újságíró Iskolájában összpontosult. Ezek a korai tanulmányok csaknem kizárólag határozott témakategóriákat alkalmaztak, mint pl. családi ügyek, politika, munkaügy, bűnözés, sport, stb...

A kezdeti tartalomelemzés olyan kvantitatív módszer volt, ami kizárólag a megjelenő, tehát manifeszt szövegtulajdonságokra volt tekintettel, a lappangó, látens jelentésre nem. Más szóval csak olyan tartalmakat elemeztek, amelyek egy dokumentumban nyíltan ki vannak fejezve. Ezt a merev álláspontot a modernebb kutatók meghaladták, hiszen a tartalmi elemzés létértelme éppen az a többletinformáció, amelyet a ki nem fejezett nyelvi szint tartalmaz.

A tartalomelemző tanulmányok másik csoportja az irodalom területén koncentrálódott és az angol költészet és próza különböző stílusjegyeit analizálta.

A századfordulón az Egyesült Államokban ugrásszerűen megnövekedett a kiadott lapok száma és ezzel egyidejűleg megjelent a kereslet a tömegpiaci igények feltérképezésére és közvélemény megkérdezésére. Erre az időszakra tehetők az újságíró iskolák létrejötte, az etikai normák születése és a sajtójelenség empirikus vizsgálatára való igény kialakulása. Ezeket az igényeket elégítette ki a kvantitatív újságelemzés.

Az újabb tömegkommunikációs eszközök elterjedésével először a rádió esetében, majd a mozi és a televízió esetében is megjelent az a megközelítés, hogy egy adott témáról megjelent szövegmennyiségét mérték. Sőt, a mai napig létezik, különböző tartalmak vizsgálatára alkalmazzák, legyen szó tankönyvről, képregényről, szónoklatról, vagy reklámról.

 2.2. Tartalomelemzés kezdetei

A tartalomelemzés színvonalának emelkedésében nagyon fontos szerepet játszott három tényező:

* •Új és az eddiginél hatalmasabb tömegkommunikációs eszközök önállósodása.
* •A nagy gazdasági válságot követő időszakban a felmerült társadalmi és politikai problémákért elsősorban a tömegkommunikációs eszközöket tették
* felelőssé.
* •A társadalomtudományokban megjelentek az empirikus kutatási módszerek.

A szociológiában például egyre inkább elterjedt a közvéleménykutatás. Amint a sajtó mellett megjelent a film és a rádió is, ezek elemzése is elkezdődött. A televízió térhódítása óta pedig igen elterjedt lett a tévéműsorok tartalomelemzése is, gyakoriak az olyan tematizált vizsgálatok, mint például „erőszak a televízióban”. Egy amerikai felmérés során megvizsgálták, hogy egy tévéműsorra átlagosan hány erőszakos cselekmény jut, ebből mennyi a tragikus, mennyi a nem végzetes kimenetelű, kik az elkövetők (fehér amerikaiak, bevándorolt fehérek, néger amerikaiak), mi a büntetésük, kik az áldozatok, és azok hogyan oszlanak meg ugyanazon három etnikai csoport között. Az ilyen kutatásokból származó eredmények lehetőséget teremtenek további szociológiai vizsgálódásokra.

A politikai szimbólumok iránti érdeklődés növekedésével a nyilvános közlemények elemzése egy új dimenzióval bővült.

Több tudományág is vizsgálatokat kezdeményezett annak kiderítésére, hogy milyen trendet mutat saját tudományának változása a szakfolyóiratok szerint.

A kvantitatív újságelemzés tehát több szempontból is különbözik a tartalomelemzés korai szakaszától:

1.Kiváló tudósok vettek részt benne, akik megteremtették az elméleti alapokat.

2.Részletes fogalmakat definiáltak és felismerték az adatokban az attitűdöket,

sztereotípiákat, stílusokat, szimbólumokat, értékeket, propagandafogásokat.

3.Hatékonyabb statisztikai elemzéseket alkalmaztak.

4.A tartalomelemzés adatai az átfogóbb kutatások részévé váltak.

2.3. Propagandaelemzés

A tartalomelemzésben fontos ugrást jelentett a második világháború, mely valós szükségletet jelentett a tartalmi vizsgálatokra. A propaganda-analízis során a propagandaelemzők a hangsúlyt az ellenséges szándékok előre megmondására helyezték, katonai és politikai hírszerzésre volt szükség. Az egyik leghíresebb előrejelzés a britek részéről az volt, mikor megjósolták a német csodafegyverek Anglia elleni bevetésének várható időpontját. Goebbels beszédeit vizsgálva a fegyverek gyártására vonatkozó célzásokat mutattak ki, és pár hetes időkeretben meg tudták mondani a támadás megindításának dátumát. A.L. George 1959-ben megjelent Propaganda-analízis című könyve mérföldkő volt a tartalomelemzés ezen ága szempontjából.

**III. Tartalom és téma**

A tartalomelemzés egy sajátos megközelítést választ az adatok elemzésére, az elemzés tárgyát, a tartalmat szemléli. A tartalom a valóság dolgainak belső, lényegi mozzanata, amely a forma révén nyer jelenségként aktuális érvényt, kifejeződést. A tartalomnak pszichológiai, logikai, általános esztétikai, kommunikációelméleti és kultúraelméleti meghatározásai is ismeretesek. Ezek szerint a tartalom az egyes művekben kifejeződő alkotói és befogadói érzelmek magatartásrendszere, a valóság jelenségeinek a művekben kifejezett lényege, az üzenet információja, a művekben szereplő kulturális értékek sokasága.

Ezek szerint tehát azt az üzenetet kell megkeresni és kifejezni, melynek információs értéke van annak a közönségnek a szempontjából, amelyet célközönségnek, az információs rendszer használóinak tekinthetünk. A tartalom a szöveg jelentése. A tudás, tartalom, téma, szöveg elválaszthatatlanok egymástól. Középpontban a szöveg áll, ebben tárgyiasul a tudás. Itt testesül meg a tartalom vagy az üzenet, a szöveg adja a tartalom vagy a téma formáját, a szöveg az információtudomány vizsgálódásának a tárgya, szöveget szolgáltatunk, amikor információs igény keletkezik, és a szöveg az , amely az eredeti tartalmat képviseli.

**IV. Dokumentumelemzés és tartalomelemzés**

A dokumentumelemzés és tartalomelemzés két különböző módszer, több szempontból is eltérnek egymástól. Közös tulajdonságaik csak annyi, hogy mindkét elemzési technika szövegek és anyagok mennyiségi és minőségi analízisét teszi lehetővé. A dokumentumelemzés legfőképp napjaink anyagaival foglalkozik, nagyobb szabadságot biztosít a kutató számára és csak az explicit módon megjelenő tartalmak vizsgálatára vállalkozik.

Ezzel ellentétben viszont a tartalomelemzés irányulhat olyan forrásokra is, amelyek a múlthoz kötődnek, a kutató számára nem biztosít akkora szabadságot, mivel meghatározott lépések egymásutánjából áll és nem korlátozódik le csupán a manifeszt tartalmak elemzésére, hanem rejtett összefüggéseket is feltár.

4.4. A tartalomelemzés alkalmazási területei:

Ezt a kutatási technikát legfőképp a politikatudományban és a szociológiában alkalmazzák. A kutatások céljai, hogy információt nyerjenek a szövegekből az üzenet megfogalmazójának szándékáról, majd további következtetéseket vonjanak le a személyekről és társadalmi jelenségekről.

Néhány alkalmazási terület:

* Nemzetközi különbségek kiszűrése a kommunikáció tartalmából.
* A kommunikációs médiumok és szintek összehasonlítása.
* A tudományosság fejlődésének kimutatása.
* Személyek vagy csoportok pszichológiai állapotának felmérése.
* Propagandacélok kimutatása.
* Kulturális jellemzők kimutatása.
* Statisztikai jellemzők feltárása.
* Kommunikációs trendek kimutatása.

**V. A tartalomelemzés lépései**

Klaus Krippendorff a tartalomelemzést hét, egymástól jól elkülöníthető részre osztotta, melyek sorrendben a következők:

•adatkészítés

•egységmeghatározás

•mintavétel

•adatrögzítés

•adatredukció

•következtetés

•elemzés

5.1. Adatkészítés

Az adatok feljegyzett, rögzített szimbólumok, melyek jelentése sajátos helyzetben derül ki. Adatnak sokféle formája lehetséges, lehet akár egy betű, vagy egy kimondott szó, viszont a tartalomelemzés szempontjából fontos, hogy rögzítsék őket. Téves az a felfogás, hogy az adat egyenlő a ténnyel, hiszen az adatnak nem biztos, hogy köze van a való világhoz (pl. propaganda). Az adatoknak információhordozóknak és reprezentatívnak kell lenniük a tényleges jelenségre vonatkozóan.

5.2. Egységmeghatározás

Az elemzési egységeknek háromféle típusa létezik: mintavételi egység, adatrögzítési egység és kontextuális egység.

A mintavételi egységek a mintavétel szempontjából fontosak, magában foglalják a sokaság elemeit és a mintát ebből veszik.

Mintavételi egységeknek tekinthetők az egymástól független megfigyelt valóság vagy a nyelvi kifejezés. Jól definiálható határai léteznek. Ezen egységeknek függetlennek kell lenniük egymástól, hogy valamely mintavételi

egységből való adat elemzésből való kizárása ne járjon más egységekre nézve

következményekkel.

Az adatrögzítési egység a mintavételi egységek részét képezik, és egy leíró munka eredménye. Az adatrögzítési egységek az elemzés alapját jelentik és a mintavételi egységeken belül az információt együttesen tartalmazzák. Az adatrögzítési egységeket a kontextuális egységek építik fel. Körvonalazzák az adat részleteit, amit meg kell vizsgálni annak érdekében, hogy az adatrögzítési egységeket jellemezni tudjuk. Ezen módszerek segítségével a tartalomelemzés során a következő féle egységeket azonosíthatjuk:

* •fizikai egységek: lehet akár könyv, plakát, költemény, stb.
* •szintaktikai egységek: a szavak a legkisebb egységei a dokumentumoknak és a legbiztosabb adatrögzítési egységei is.
* •referenciális egységek: ezeknek az egységeknek a jellemzői, hogy saját objektumukkal, történésekkel, személyekkel írhatjuk le. A referenciális egységek használhatóak a létező jelenségek ábrázolásának jellemzésére, elemzi az objektumokat és azok attribútumait.
* •propozicionális egységek
* •tematikai egységek

**5.3. Mintavétel a tartalomelemzésben**

Nagy terjedelmű közlések vizsgálatakor lehetetlen minden egyes elemet figyelembe és számba venni. A tartalomelemzésekben a mennyiségeknek legalább három különböző eredetük lehet. Először is származhatnak az ismétlődő előfordulások megszámlálásából. Ez összekapcsol egy mennyiséget az azonos adatrögzítési egységek egy osztályával. Másodszor az olyan mérték, mint egy cikk hasábszélessége, az egyik adatrögzítési egység leírójává válhat. Harmadszor az olyan mértékek, mint a példányszám egy egész mintavételi egységgel vannak kapcsolatban. Szembekerülve az irdatlan mennyiséggel, tartalomelemzés megkezdése előtt döntést kell hozni: ha a releváns anyag túl terjedelmes, akkor olyan mintát kell kiválasztania, amely elég nagy ahhoz, hogy tartalmazza a lényeges információkat, de elég kicsi ahhoz, hogy elemezni lehessen. A mintavétel lehetséges módjait több kutató közel azonos módon határozta meg.

**Véletlen mintavétel**

Ha semmiféle előzetes ismeretünk nincs a jelenségkörről, akkor az egyszerű, véletlen mintát előirányzó mintavételi terv az összes olyan releváns egység (reklám klippek, dokumentumok, az egyes megszólalóhoz tartozó megszólalások egy csoportban, szavak vagy mondatok) lajstromba vételét tartalmazza, amelyekkel kapcsolatban törvényszerűségeket kívánunk megfogalmazni. Annak eldöntéséhez pedig, hogy melyik egység kerüljön be a mintába bármiféle más módszer alkalmazható, amely egyforma esélyt nyújt minden egység számára.

**Szisztematikus mintavétel**

A szisztematikus mintavétel azt jelenti, hogy egy lista minden n-edik egységét beválasztjuk a mintába, miután véletlenszerűen meghatároztuk a sorozat kiindulópontját. A tartalomelemzésben a szisztematikus mintavételt akkor részesítik előnyben, ha az adatok rendszeresen megjelenő publikációkból, tévé-műsorokból, emberek közötti interakciók sorozatából, írásművek, rádióreklámok vagy reklámfilmek sorozatából származnak. A szisztematikus mintavétel fő problémája az, hogy az n szakasz intervalluma állandó, és ezért félrevezető mintákat hozhat létre, ha egybeesik bizonyos természetes ritmusokkal, pl. a szezonális változásokkal vagy más, ciklikus szabályszerűségekkel.

**Rétegzett minták**

A rétegezett mintavétel abból indul ki, hogy több, egymástól különböző alpopuláció (réteg) létezik egy populáción belül. Minden egyes mintavételi egység csupán egy réteghez tartozik. Véletlen mintavétel csak külön-külön, az egyes rétegeken belül folyik úgy, hogy a kapott minta a populáción belül eleve létező, ismert különbségeket tükrözze. Ez tulajdonképpen a szegmentációs kritériumoknak megfelelő mintavétel. A véletlenszerűség csak a szegmensen belüli kiválasztásban érvényesül.

**Klaszteres mintavétel**

A klaszteres mintavétel az elemek csoportjait tekinti mintavételi egységeknek. A csoportoknak természetes rendeltetése és határa van. Egy valamilyen csoport kiválasztása annak minden elemét beviszi a mintába, s mivel a csoportok ismeretlen számú elemet tartalmaznak, annak valószínűsége, hogy egy egység bekerül-e egy mintába, a csoportméretektől függ. A szegmentálás kritériumai ugyancsak követhetik a klaszterezés elveit.

A klaszteres minta kialakítása nem más, mint gyakorlati válasz arra, hogy általában képtelenség egyenként összeírni egy populáción belül az elemeket arra viszont már van lehetőség, hogy összeírjuk azokat a csoportokat, amelyekben ezek előfordulnak. A klaszteres (csoportos) mintavétel azt feltételezi, hogy egy nagyobb adathalmaz néhány kritérium szerint alcsoportokba osztható: pl. egy termék állandó fogyasztóinak csoportja, a konkurens tartalmainak hatását vizsgáljuk, mintát vehetünk a sűrűn tévézők és az alig tévézők csoportjából.

**Változó valószínűségű mintavétel**

Ebben az esetben a mintába való bekerülés valószínűségét minden egység számára előzetesen megállapított kritérium szabja meg. Leggyakrabban az ún. „arányosított alminták alapján történő mintavétellel” találkozunk, amelyeknél az almintába kerülés valószínűségét az almintához (minta százalékos arányához) igazítják. Ez azt jelenti, hogy amilyen mértékben része az alminta

annak a populációnak, amelyből a mintát képezzük, olyan mértékben kerül bele a vizsgálatba.

A változó valószínűségű mintavétel azért fontos a tartalomelemzés számára, mert épp az a feladata, hogy olyan jelenségekkel kapcsolatban fogalmazzon meg következtetéseket, amelyek sokszor torzítva jelennek meg a hivatalos közleményekben.

Ha a közvéleményt szeretnénk vizsgálni valamilyen közérdekű jelenséggel kapcsolatban és csak a tévében vagy a rádióban megjelenő közleményeket tennénk meg véletlenszerű mintavétellel a tartalomelemzés tárgyává, mintavételünk torzított lenne. A mindennapi életben megjelenő közvélemény alakításában nem ugyanaz a részesedése a médiának és a magánembernek. Ha csak a médiát figyeljük, torzítunk, ha pedig csak a magánember megnyilatkozásait elemezzük, akkor sem a teljes valóságot vizsgáljuk. Ilyen esetben más-más arányban veszünk mintát a médiaközleményekből is (média alminta) és a magánemberek vélekedéseiből is (magánemberek almintája). Ennek az arányszámnak, szorzónak (valószínűségnek) a megállapítása külön előtanulmányt igényel: különböző rétegekhez tartozó emberek véleményeinek, meggyőződéseinek feltérképezése abból a szempontból, hogy milyen forrásokból érték őket hatások a vizsgált témakörben. Általában mindenki hallgat rádiót, néz tévét, ám nem kizárólag innen szerzi információit. Ezen kívül mindenkit befolyásolnak azok az emberek (családtagok, barátok, munkatársak, ismerősök stb.), akikkel mindennapi kapcsolatban van. Így megfigyelhető, hogy átlagosan 2–4 médiaközleményre, 10–15 magánemberi befolyás esik. Ha képet akarunk kapni arról, hogy milyen összkép alakult ki a fogyasztók körében egy adott termékről, érdemes belevenni – különböző arányszámmal– az egyik almintába a reklámközlemények különböző elemeit és a másik almintába a termék fogyasztóinak közléseit. Az ilyen vizsgálat megtervezésénél a piackutatónak szem előtt kell tartania, hogy a reklámok nem mások, mint szisztematikusan eltorzított mintái a fogyasztók körében élő véleményeknek. Előfordulhat, hogy a tartalomelemzés eredményei az egyik almintában lényegesen különböznek a másik almintában kapott eredményektől. Fontos támpontot jelenthet a gyártó vagy a forgalmazó számára, hogy a fogyasztók másként vélekednek a termékről, mint ahogyan azt a termék reklámjai üzenik. A változó valószínűségek azt a gyakorlati ismeretet reprezentálják, amellyel a kutató az adatok kontextusáról rendelkezik, viszont ez az ismeret gyakran bizonytalan vagy hipotetikus, így a változó valószínűségű mintákat nehéz igazolni, és csak rendkívüli óvatossággal szabad velük dolgozni.

Elvileg rengeteg lehetőség van a különféle mintavételi eljárások kidolgozásában, de az a lényeg, hogy olyan mintákat kapjunk, amelyek reprezentatívak a vizsgált jelenségkört illetően.

Miután meghatároztuk a mintavételi eljárást, meg kell határoznunk a minta nagyságát. A mintanagyság tesztelésére szolgáló technika az ún. felezéses technika. Ennek során a mintát két egyforma részre osztjuk. És ha valamelyik rész ugyanazokat a statisztikai következtetéseket támasztja alá, ugyanazon a megbízhatósági szinten, akkor az egészet el lehet fogadni megfelelő méretű mintának.

**Véletlen és rendszeres tévedés**

A mintavétellel kapcsolatban beszélhetünk még a véletlen tévedés és a rendszeres tévedésről. A különbség köztük az, hogy a véletlen tévedések többnyire kiegyenlítik egymást, mivel nem ugyanabba az irányba hatnak, és a vizsgálat végeredményét nem hamisítják meg. Ellenkező esetben, a rendszeres tévedés esetében, a tévedések mindig egy irányba hatnak, és a vizsgálat nagysága nemhogy eltüntetné ezeket a tévedéseket, hanem csak fokozza és ezáltal a végeredménybe is belekerül. Példa erre Hatch kutatása 1932 és 1942 között, amely arra vonatkozott, hogy a New York Times társasági rovata milyen házasságkötésekről adott hírt. A teljes 10 évnyi anyag nagyon sok lett volna, ezért minden évből a június havi példányokat választotta ki mintául. Így történhetett, hogy a vizsgálat eredménye szerint 10 év alatt egy zsidó pár sem házasodott össze. Hatch különféle hipotézisek gyártásába kezdett a jelenség magyarázatára. A rejtélyről Cahnmann rántotta le a leplet 1948-ban, miszerint azsidó vallás tiltja júniusban a házasságkötést. Érdemes tehát nagyobb vizsgálat előtt szakértőket, statisztikusokat megkérdezni a mintavétellel kapcsolatban, főleg a nagyobb kutatások előtt.

**5.4. Adatrögzítés**

Csak azokat az adatokat tudjuk megfelelően elemezni, amelyek le vannak rögzítve, hiszen általában tele vannak szimbolikus formákkal. Ezért ezeket a szimbolikus közleményeket mindig olyan alakba kell átalakítani, amelyek már lehetővé teszik az adatfeldolgozást és a következtetések levonását.

*5.4.1. Kódolás és kategóriák*

A kódolás fázisában a szöveg egyes részeit (pl. a fókuszcsoport vagy az egyéni mélyinterjú szöveges átiratának „kódolási egységeit”) előre megállapított kategóriákhoz soroljuk. A kvalitatív tartalomelemzés esetében a kódolás nem előre meghatározott kategóriák szerint történik, hanem ezek a kategóriák az elemzés folyamatában alakulnak ki. Csak olyan szavak, mondatok jelentése kódolható, amelyek a szövegben ténylegesen megjelennek. A jelen nem levőt nem lehet sem megszámolni, sem kódolni.

*5.4.2. A kódolás helye a tartalomelemzésben.*

A kódolás az a folyamat, amelynek révén a nyers szövegadatokat módszeresen nagyobb egységekbe soroljuk, s így lehetővé tesszük a tartalom

valóban lényeges leírását és összefüggéseinek megragadását. Azokat a szabályokat nevezzük kódolási szabályoknak, amelyek a szövegelemek és a kódnévben jelölt kategória közötti kapcsolatot írják elő (másként kódolási előírásoknak is nevezik ezeket). A kódolási szabályok tehát minden tartalomelemzés központi részét képezik, s ezek előkészítése egy sor lényeges döntést igényel. Többek között az alábbi kérdésekre kell válaszolni:

•Milyen kategóriákon keresztül határozzuk meg a kutatási problémát, azaz milyen kategóriákat vegyünk fel?

•Milyen rendszert képezzenek a felvett kategóriák?

•A tartalom milyen egységeit osztályozzuk?

*5.4.3. A kategóriák*

A tartalomelemzés egyik alapelve szerint ez a típusú vizsgálódás abból áll, hogy „a sokból keveset csináljunk”. A közlemények igen nagy számú elemét (szavait, mondatait) viszonylag kevés számú elemre, kategóriára csökkentjük. Ezzel a kevesebb számú adattal, amely adatok a jelenségek egy-egy fogalmi kategóriáját képezik már könnyebben tudunk bánni. Egy ilyen egyszerűsítő művelettel a szöveg már átláthatóvá és rendezetté válik, s az elemezhetőség szempontjából a „kevesebből több lett”.

Lényegében a kategóriák két típusát különböztetjük meg:

•az előre meghatározott kategóriák,

•a vizsgálat során kibontakozó kategóriák.

A kategóriák egyik esetben sem állnak önmagukban vagy felsorolásszerű rendezetlenségben. Mindig valamilyen logikai rendszerbe rendezik a kategóriákat, s így kategória-rendszerekről beszélünk.

*7.4.4. A kategória-alkotás szabályai*

A kategória-alkotásnak vannak szabályai, mint például, hogy a kategóriák tükrözzék a kutatás célját, legyenek kimerítőek, egymást kölcsönösen kizárják és származzanak ugyanazon osztályozási elvből. Tükrözniük kell a kutató problémáját, azaz a kutatásnak világosan meg kell határoznia azokat a változókat, amelyekkel foglalkozik, valamint pontosan körül kell írnia, hogy melyek azok az ismérvek, amelyek eldöntik, hogy egy konkrét tartalmi adat az adott kategóriákhoz tartozik-e. Ez két követelményt von maga után: az érvényességet és a megbízhatóságot.

Alaposan át kell gondolni a kategóriák számát is, mivel előfordulhat az az eset is, ha túl sok kategóriát használunk, mindegyiket csak egyszer használjuk. Mindig fel kell tenni a kérdést, hogy „Mit nyerünk azzal, ha több kategóriát veszünk fel?”. Ha több kategóriát használunk, több lehetőség nyílik az összehasonlításra, ugyanakkor a sok, finomabb kategória miatt problémássá válhat az osztályozás és előfordulhat, hogy maradnak a szövegben nyers adatok is. A kiválasztott szöveg minden fontos elemét el kell tudni helyezni valamilyen kategóriában. Ezt jelenti az a szabály, miszerint a kategóriáknak kimerítőeknek kell lenni. Egynél több kategóriához viszont egyetlen adat sem sorolható, ez biztosítja a kölcsönös kizáróság szabályát. A kategóriák függetlensége megköveteli, hogy az egyes adatoknak kategóriákba

való besorolása független kell, hogy maradjon a többi adat kategorizálásától. Az egységes osztályozási rendszer feltétele pedig azt határozza meg, hogy a kategóriák ugyanazon a rendszeren alapuljanak, ugyanabból származzanak, azaz nem keveredhetnek különböző fogalmi síkok.

Vannak az évek során jól bevált tartalomelemzési kategóriák, amelyeket használhatunk. Az amerikai Pulitzer díjas író, Frank Luther Mott az újságok tartalmát a következő kategóriarendszerbe sorolta:

1. Külföldi hírek
2. Hazai hírek
3. Közéleti problémák
4. Vezércikkek
5. Üzlet, pénzügy
6. Sport
7. Társasági hírek
8. Női rovat (pl. divat)
9. Színház, mozi
10. Rádióműsor
11. Karikatúra
12. Egyéb illusztráció (pl. fénykép)

Leginkább az értékkategóriák használatosak. White az értékek elemzésére dolgozta ki a következő kategóriarendszert. Az ilyen fajtacsoportosítás azt jelzi, hogy egy csoport értékei között nagyobb a hasonlóság, mint különböző csoportok értékei között. Szerinte például a teljesítmény és függetlenség között ugyanakkora a különbség, mint a teljesítmény és elismerés között, de a teljesítmény és pihenés közti különbség már mindkettőnél nagyobb.

Fiziológiai értékek:

1.Élelem

2.Fajfenntartás

3.Pihenés

4.Egészség

5.Biztonság

6.Kényelem

Egoisztikus értékek:

1. Függetlenség

2. Teljesítmény

3. Elismerés

4. Önbecsülés

5. Uralkodás

6. Agresszió

Gyakorlati értékek:

1. Gyakorlatiasság

2. Birtoklás

3. Munka

Kognitív értékek:

1.Tudás

Szociális értékek:

1. Szex

2. Család

3. Barátság

Játékos értékek:

1. Új élmény
2. Izgalom
3. Szépség
4. Humor
5. Alkotó önkifejezés

Különféle értékek:

1. Boldogság
2. Érték általában

**5.4.5. Kategóriarendszerek**

A kategóriarendszerek legáltalánosabb formái a kategória-szótárak és a kategóriakatalógusok. Az ilyen szótárak és katalógusok azt határozzák meg, hogy az egyes szavak milyen magasabb fogalmi csoportba sorolhatók.

**5.4.6. A kódolási vagy elemzési egységek**

Minden empirikus vizsgálat első feladata annak meghatározása, hogy mit is kell megfigyelni, rögzíteni, majd adatnak tekinteni. „Adatok”-ról beszélünk, mert egy empirikus vizsgálathoz nagyszámú információhordozó egységre „adat”-ra van szükség. Az egységképzés azt jelenti, hogy meghatározzuk ezeket az egységeket, elkülönítjük őket, majd azonosítjuk őket az elemzés számára. A kódolási (elemzési) egységek megállapítása két szempontból is lényeges előkészítő feladat:

•Az egységek meghatározásával válik egyértelművé, hogy a kitűzött kutatási cél miként és milyen alapegységeken keresztül vizsgálható.

•A kódolási egység a mintavételi eljárást lényegesen befolyásolja.

Nem elég meghatározni, hogy milyen kategóriák szerint osztályozzuk, azaz milyen kategóriákhoz soroljuk a szavakat, mondatokat, kijelentéseket.

Az elemzőnek azt is el kell döntenie, hogy milyen természetű , illetve milyen terjedelmű egységeket kódol majd a kategóriákba. E tekintetben is különböző lehetőségek állnak fenn, mert kódolási egység lehet: a fogalom, a téma, a szereplő típusok és a műfaj.

A leggyakrabban használt kódolási egység a fogalom (s ennek fizikai hordozója, a szó) volt, amely egyben a gyakorlatilag lehetséges legkisebb kódolási egység is. A múltban, főleg a nagy kiterjedésű vizsgálatok esetében, kerülték, mert igen költséges volt. A számítógépek és a tartalomelemző szoftverek megjelenése óta ez a szempont lényegében elesett.

Gyakran használt kódolási egység a téma is. A tartalomelemzésben azonban a

„téma” fogalmát sokkal szűkebb értelemben használják, mint a köznapi életben. A téma egy egyszerű értékelő megállapítás valamivel kapcsolatban. Szinte nélkülözhetetlen eleme a piackutatásban alkalmazott tartalomelemzésének. Ezeket a témákat kell azután kategóriákhoz sorolni, azaz kódolni. A téma, mint kódolási (elemzési) egység problematikus lehet, amiatt, hogy sokszor elég nehéz a témát – vagy egyszerre jelenlévő több témát – a bonyolultabb mondatokból kihámozni. Olyan tartalomelemzéseknél, ahol a közlemény történetet mutat be, pl. reklám, szerepjáték a fókuszcsoportban, fogyasztói szokások vizsgálata stb., jól bevált kódolási egység a szereplő típusok, amelyekkel a tartalom egésze jól jellemezhető.

Ha arra vagyunk kíváncsiak, hogy milyen tartalmi jellegzetességei vannak annak, hogy valamilyen közlemény egyik vagy másik műfajban jelenik meg, akkor a műfaj is lehet kódolási egység.

A kódolási egység kiválasztása éppen úgy befolyásolhatja a tartalomelemzés végeredményét, mint a kategóriák megválasztása. Három amerikai tartalomelemző 1942-ben négy különböző kódolási egységgel dolgozta fel ugyanazokat a dokumentumokat. A négy módszer egyezést mutatott a tartalmi tendenciák irányai tekintetében, csupán azok terjedelmében mutatkoztak különbségek.

**5.5. Adatredukció**

A számítógépes tartalomelemzésnél használatos módszer. Egy változóról gyűjtött teljes adattömeget eredeti formában általában lehetetlen értelmezni. Az adatredukció az a folyamat, ami az eredeti adatokat jobban kezelhető alakban összefoglalja, miközben a lehető legtöbb eredeti részletet igyekszik megtartani.

**5.6. Következtetés**

Minden kommunikációs folyamat során a „ki, mit, hogyan, kinek, milyen hatással és miért közöl?” kérdésekre keressük a választ. A ki kérdése a kommunikáció forrására vonatkozik, a mit a tartaloméra, a hogyan a kódolási sajátságokra, a kinek az üzenet célba vett közönségére, a milyen hatással az üzenet effektusára, eredményére, a miért pedig az üzenetnek előzményeire és céljaira. A kutató három lehetséges célból vizsgálhatja a szöveget:

•Következtetéseket vonjon le a szöveg sajátságairól

•Az üzenet okairól vagy előzményeiről

•A kommunikáció hatásairól.

A tartalomelemzés fontos része a kutatási terv. A kutatónak a vizsgálat előtt meg kell tudni mondania, hogy milyen kérdésre várja a választ. Minden vizsgálat kiindulópontja a kutató hipotézise. Adott kell, hogy legyen egy probléma, melynek megoldására hatékonyabb és olcsóbb megoldás a tartalomelemzésnél nem létezik. A kutatási tervben a kutató hipotézise és az alkalmazandó módszerek meghatározása történik, és rögzíti, hogy a kutatásból milyen típusú következtetéseket kell levonni.

**5.6.1. A szöveg sajátságainak leírása**

Ezek a vizsgálatok a legtöbb esetben a „mit” kérdéssel állnak kapcsolatban és arra vonatkoznak, hogy valamely forrásnak valamilyen jelenséggel szemben milyen méretű a figyelme, milyen tendenciák érvényesülnek a kommunikációban, milyen különbségek vannak az egyes tömegkommunikációs eszközök között. Az összehasonlításoknak különböző módja léteznek. Ugyanannak a forrásnak a különböző időből származó dokumentumait hasonlíthatjuk össze, és így megállapíthatjuk a forrás beállítottságának, attitűdjeinek változásait, tendenciáit. Ilyen vizsgálatokat végeztek azzal kapcsolatban például, hogy miként változtak a vidéki újságok, vagy hogyan módosult a tudományos érdeklődés az utóbbi évtizedek folyamán a pszichológiában, szociológiában. Ez az elemzési módszer a leggyakoribb a tartalomelemzések közül.

A következő módszer, mikor ugyanazon forrás különböző időpontokban és különböző szituációkban létrehozott üzeneteit hasonlítják össze. Ekkor kimutatható, hogy a megváltozott körülmények milyen változást idéztek elő a szöveg jellegzetességeiben. Erre példa, mikor azt vizsgálták, hogy azok a lapok, amiknek konkurenciával, versenyhelyzettel kell számolniuk, jobb hírtájékoztatást adnak-e, mint a maguk területén monopolhelyzetben lévők.

A harmadik típus, mikor két vagy több típus összefüggéseit vizsgáljuk ugyanabban a dokumentumban, üzenetben. Ilyen kutatás volt például a háborús és népjóléti kiadás összefüggéséről az amerikai politikai irodalomban.

**5.6.2. Következtetések a kommunikáció okairól és előzményeiről**

Az üzenetet ebben az esetben azért vizsgáljuk, hogy szöveg előzményéről és okairól vonhassunk le következtetéseket. Ez a típus az üzenet küldőjére, tehát a szerzőre és az üzenet küldésének miértjére vonatkozik, éppen ezért ebben az esetben felvetődik néhány probléma. Nevezetesen, egy személy kijelentései, megállapításai és a szándékai, érzései, céljai között nem egyértelmű az összefüggés. Ugyanazt a dolgot sokféleképpen ki lehet fejezni, ezért ezek a következtetések, amelyeket kizárólag tartalmi adatokból vonunk le, sohasem meggyőzőek, külső megerősítést igényelnek. Azért, hogy ezek a következtetések megbízhatók legyenek, szükség van a tartalmi adatok összehasonlítására valamilyen független bizonyítékkal. Például irodalmi szövegek tartalomelemzését, az elemzés eredményeit az írók életrajzi adataival hasonlították össze.

**5.6.3. Következtetések az üzenetek hatásairól**

Ezek a vizsgálatok az üzenetek hatásaival kapcsolatos következtetésekre irányulnak. Itt sem lehet feltételezni, hogy egy üzenet ugyanazt a hatást fejti ki minden befogadó személyben. Különböző személyek különböző módokon fogadják be, értékelik az üzeneteket. Ez több tényezőtől függ, mint például a forrás megbízhatóságáról kialakult meggyőződésüktől vagy attól, hogy milyen helyzetben találkoztak az üzenettel. Így ebben az esetben sem lehet csupán az üzenet tartalmára koncentrálni, hanem tekintettel kell lenni a befogadóra is. A hatás megállapításának egyik lehetséges módja, ha összehasonlítjuk a befogadó személy vagy csoport azon üzeneteit, amelyeket a forrás hatása előtt hozott létre azokkal, amelyeket a forrás hatása után. Persze itt is felmerül egy kérdés, hogy biztos-e, hogy a befogadó személyre vagy csoportra nem volt hatással egy másik forrás is.

**5.7. Elemzés:**

A tartalomelemzés gyakorlata igen széles körűen alkalmazza az egyes elemzési technikákat. A leggyakoribbak:

*Gyakoriság-elemzés*

A legáltalánosabb módja az adatok elemzésének. Két fajtája létezik:

•abszolút gyakoriság: A mintában a keresett szavak száma.

•relatív gyakoriság: a szavaknak a szövegben való előfordulási aránya. Ennek

nagyságát egy adott adat esetén úgy kaphatjuk meg, hogy az adat előfordulási számát osztjuk az adathalmazban levő elemek számával. Ezt a relatív gyakoriságot meg lehet adni százalékos formában is.

A gyakoriságot ábrázolja a következő diagramm is, ahol a drogproblémákkal

összefüggő cikkek számát vizsgálták négy különböző napilapban. A vizsgált időszak hossza 20 negyedév, azaz öt év volt.

**A változók közötti relációk elemzése**

Az adatok bemutatásának másik gyakori formája, az asszociációk, korrelációk és kereszttáblázatok alkalmazása. Ilyenkor az adatok egybeesésének gyakoriságát vizsgálják.

Könnyebben lehet következtetéseket levonni, ha adatokkal alátámasztható, hogy két esemény egyszerre gyakrabban fordul elő, tehát valamilyen kapcsolat van köztük. A kereszttáblázatok leggyakoribb formái két- illetve háromdimenziósak, mivel ezeket könnyebb áttekinteni, de léteznek többdimenziós kereszttáblázatok is. Ezzel a módszerrel lehetőség nyílik kérdőívek válaszainak belső összefüggés-rendszerének vizsgálatára is. Erre példa a következő táblázat, amely a hálapénz fizetésének gyakoriságát és az orvoshoz fordulás időpontja közti kapcsolatot szemlélteti.

**Imázs- vagy portré-elemzés**

Szintén egy tipikus tartalomelemzési módszer, amely olyan kérdésekre próbál

választ adni, mint például milyen képet alkot a szöveg egy személyről, közösségről, nemzetről vagy éppen egy eszméről. Ezt a technikát kétféle módon is megkülönböztethetjük. Egyrészről a személyek jellemző tulajdonságait (attribútum) vizsgálva az imázs az, amit már ismerünk, vagy amit mások mondanak valamiről, illetve valakiről. Másrészről az imázs fogalma mindazon dolgok összességét is jelentheti, amelyekkel összefüggésbe hozzák. (asszociáció).

**Kontingenciaelemzés**

A kontingencia szituációelméletet, az esetlegesség elvét jelenti. Ez az elemzési technika is az együttes előfordulásokra épít, de ezekből asszociációs hálókat szerkeszt.

**Klaszterálás**

A klaszter az elemek egyetlen összefüggő rendszerré alakítását jelenti. Ezt a

technikát akkor alkalmazzák, ha az együttes előfordulások száma nagyon nagy. A módszer lényege, hogy azokat a változókat, amelyeknek tulajdonságai közösek, egyetlen közös csoportba rendszerezze és a különböző halmazok közti különbségek tükrözzék a változók tulajdonságainak különbségét. A klaszterálás eredményét dendogrammal ábrázolják.

**Kontextuális osztályozás**

**5.8. Megbízhatóság és érvényesség**

A kódolás folyamatában fontos a megbízhatóság és érvényesség figyelembe vétele. Ahhoz, hogy vizsgálati eredményeink érvényesek legyenek, az alapjukat képező adatoknak, az elemzésbe bevont személyeknek és az alkalmazott műveleteknek megbízhatóknak kell lenni. A megbízhatóság szükséges, de nem elégséges feltétele az érvényességnek. A megbízhatóság biztosítása miatt a kategóriák felállítása után célszerű több kódolóval elvégeztetni a kategorizálást. A megbízhatóságnak 3 mutatója jut szerephez a

tartalomelemzésben:

•Stabilitás (konzisztencia): azt mutatja, hogy az idő múlásával mennyire marad változatlan a kódolási folyamat. Ez a megbízhatóság leggyöngébb

formája.

•reprodukálhatóság: azt mutatja, hogy egy folyamatot mennyire tudunk megismételni változó körülmények között, különböző kódolókkal. Ezt szokták még „kódolók közötti megbízhatóságnak” vagy a megfigyelők közt

létrejött „konszenzusnak” is hívni.

•pontosság: azt jelzi, hogy a kódolási folyamat mennyire felel meg egy ismert

standardnak (helyesnek, elfogadottnak, normának) vagy eredményezi azt, amit vártak tőle. Ez a ma rendelkezésünkre álló legerősebb megbízhatósági teszt.

Az érvényesség azt a fokot jelöli, amilyen mértékben egy eszközvalóban azt méri, aminek mérésére hivatott. Ahhoz, hogy érvényes következtetésekre juthassunk, helyes mintaválasztásra és megbízható kódolási eljárásra van szükség, azonban ezek nem elégséges feltételek. Az érvényesség megállapítására különböző módszerek alkalmazhatók.

•Tartalmi érvényesség: A legelterjedtebb és legmegbízhatóbb módszer. Lényege, hogy a kutató a vizsgálat eredményét egybeveti saját, más módon szerzett ismereteivel.

•Prediktív (előrejelző) érvényesség:Ha a tartalomelemzés alapján bizonyos feltevéseket mondunk ki jövőbeni események bekövetkezéséről, és ezek az események később valóban bekövetkeznek. Ez a típusú validitás nem korlátozódik csupán a jövőbeni eseményekre, múltbeli eseményekkel kapcsolatban is alkalmazható, ha azok bekövetkezésével kapcsolatos bizonyítékaink vannak.

•Konstrukciós (vagy fogalmi) érvényesség: Lényege, hogy a mérőeszköz valamely elmélet alapján elvárt módon viselkedik-e. Tehát nemcsak a mérési eljárás érvényességét kell igazolni, hanem a mérés alapjául szolgáló elmélet érvényességét is.

A megfigyelési eszközök esetén megbízhatóság és az érvényesség gyakran csak egymás rovására valósítható meg.

