

Så görs Mediestudiers innehållsanalys

Karl-Arvid Färm*

Denna skrift baseras på en undersökning – benämnd Mediestudiers innehållsanalys – som har gjorts av 26 olika medier, såväl riksmédier som medier från fyra utvalda regioner (Stockholm, Värmland, Västernorrland och Västmanland).

Dessa 26 medier har hittills studerats under tjugoåtta slumpmässigt valda dagar, sju vardera från 2007, 2014, 2016 och 2018 i analoga respektive digitala upplagor. Alla 26 medier har om möjligt studerats i båda dessa varianter. Ett par av medierna existerade endast i antingen analog form eller digital form. De digitala upplagorna bestod i samtliga fall av hemsidor, de analoga bestod antingen av papperstidningar, TV-sändningar eller radiosändningar. Den typ av studie som görs är en kvantitativ innehållsanalys. Se kodschema i appendix för lista över ingående medier.

* Karl-Arvid Färm har haft huvudansvaret för kodningen av Mediestudiers innehållsanalys.

Kvantitativ innehållsanalys

En kvantitativ innehållsanalys används när man vill göra en systematisk analys av ett innehåll och därmed kartlägga och studera viktiga skeenden. Dessa innehållsanalyser är ofta deskriptiva: de strävar efter att spegla det faktiska innehållet så objektivt som möjligt. De är även replikerbara: baserat på samma instruktioner och tillvägagångssätt ska undersökningen kunna göras om och nå samma resultat.

En sådan undersökning genomförs genom att skapa ett kodschema med tillhörande instruktioner. Kodschemat innehåller en grupp variabler utifrån vilka det valda materialet läses och analyseras. Instruktionerna ska ge så exakt beskrivning som möjligt av vad kodarna letar efter i var och en av variablerna. Tanken är att kodinstruktionerna ska vara så utförliga att undersökningen ska kunna göras om med hjälp av dessa och nå i stort sett samma resultat. Detta för att skapa relevans och tillförlitlighet. En eller flera kodare studerar sedan systematiskt varje enhet i det valda materialet utifrån de valda variablerna.

Efter genomgång av tidigare forskning och diskussioner i forskningsprojektets arbetsgrupp konstruerades ett kodschema med tillhörande instruktioner som innehöll relevanta variabler. Dessa kan ses i detalj i appendix. För detta projekt har hittills kodats 21 033 analysenheter (vilket i det här fallet kan översättas till artiklar eller TV/radio-inslag) av tre olika kodare.

Datumurval

För varje år slumpades sju datum att studera fram. För år 2016 och 2018 visste vi datumen i förväg och samlade in materialet när datumen väl inträffade.

För att få tillgång till digitalt material från 2007 och 2014 fick urvalet förhålla sig till vad som fanns tillgängligt och ibland frångå de slumpmässigt valda dagarna, eftersom alla dagar inte fanns sparade. För att få tag på det digitala materialet användes webbtjänsten Wayback Machine (web.archive.com). Utifrån dessa fjorton datum valdes de närmaste datum som Wayback Machine hade arkiverat för varje medium. I vissa fall 2007 fanns inte sju dagar arkiverade och då användes de datum som fanns.

Följande tabell visar de tjugoåtta datum som slumpats fram:

2007	2014	2016	2018
2007-03-19	2014-02-16	2016-02-03	2018-03-01
2007-04-06	2014-04-19	2016-04-22	2018-05-14
2007-05-13	2014-05-28	2016-07-11	2018-07-13
2007-06-03	2014-06-04	2016-08-09	2018-08-27
2007-07-04	2014-08-02	2016-09-27	2018-10-21
2007-11-12	2014-10-05	2016-10-28	2018-11-18
2007-12-24	2014-11-02	2016-12-14	2018-12-06

Dessutom har vissa medier startats eller stängts ner under denna period. Sundsvalls Nyheter startades inte förrän i slutet av 2007. Nyhetssidan Nyheter24

har bara kodats 2016 och 2018 på grund av att den inte sparats från de tidigare åren av Wayback Machine och Dagbladet i Västernorrland lades ner 2015 och har därmed bara kodats under 2007 och 2014.

Efter denna process fanns upp till tjugoåtta datum valda för varje medium som sedan kodades i både dess digitala och analoga upplaga. Flera av tidningarna ges inte ut dagligen utan en eller ett par gånger i veckan, och i dessa fall valdes utgåvan som ”täckte” det valda datumet (exempel: Mitt i Huddinge ges endast ut på tisdagar. Om det valda datumet för Mitt i Huddinge var fredagen 14 februari 2014 för att den digitala utgåvan fanns tillgänglig i Wayback Machine så kodades därmed den analoga tidningen som gavs ut tisdagen den 18 februari 2014 eftersom den då ”täcker” tidsperioden onsdag 12 februari t o m tisdag 18 februari).

Samma sak gällde ifall en TV-sändning eller radiosändning inte fanns tillgänglig (exempelvis visade det sig att vissa nyhetssändningar sommaren 2014 hade ställts in på grund av fotbolls-VM). Då kodades TV-sändningen från dagen därpå.

Urvalsregler och undantag för analoga medier

Grundprincipen för kodningen av analoga medier var att välja var tredje text som fanns i tidningen eller var tredje inslag som fanns i TV- eller radiosändningarna och koda dem utifrån kodschemat. Ett antal undantagsregler skapades för att inte koda material som inte ansågs angeläget för undersökningens syfte. Nedan följer en lista med dessa undantagsregler:

- För att få en konsekvent kodning av alla analoga medier så kodades endast huvudtidningen för alla de analoga tidningarna. Alla former av bilagor valdes bort. Det finns ett antal orsaker till detta. Dels arbetsbördan, där exempelvis Dagens Nyheter med alla bilagor uppgår till närmare 100 sidor dagligen vilket skulle ha utökat projektets storlek markant. Dessutom fanns det inget säkert sätt att få tag på alla bilagor från alla datum då en stor del av materialet kodades från diverse biblioteks mikrofilmsarkiv och via PDF-scanningar.
- Alla former av annonser valdes bort.
- Diverse typer av inslag i en tidning ansågs ogiltiga för kodning. Exempel på dessa är seriestrippar, recept, ”Dagens citat”, ”Dagens fråga” (där man endast skriver ut resultat från exempelvis en webundersökning utan längre förklarande text), dödsannonser (om det inte handlar om längre texter som exempelvis beskriver personens liv), horoskop, korsord, frågesporter, vitsar, väderkartor etc.
- TV-sidorna valdes konsekvent bort, med följden att filmrecensioner och krönikor från dessa sidor inte kodades. Utöver detta var den enda formen av längre texter som valdes bort de fall där en tidning publicerade hela romankapitel som ”sommarläsning” el dyl.
- I TV och radio valdes ytterst få inslag bort, egentligen bara introt som snabbt nämner vilka inslag som kommer visas i programmet och eventuella reklam-snuttar för andra av kanalens program.

Urvalsregler och undantag för digitala medier

För att göra urvalet från de digitala medierna var grundprincipen att gå till det digitala mediets huvudsida, exempelvis www.aftonbladet.se, och klicka på var tredje länk som uppfyllde våra krav. Sedan kodades artikeln eller videoinslaget som länkades. I ett flertal fall, särskilt från 2007, hade inte Wayback Machine sparat alla länkar. I dessa fall gjordes ett försök att söka upp samma artikel på den nuvarande hemsidan, vilket oftast fungerade när medierna hade egna databaser för sitt material (som exempelvis Aftonbladet) men sällan fungerade när de inte har en liknande databas (som exempelvis TV4 som totalt har förändrat sin hemsida ett antal gånger mellan 2007 och hösten 2015 när arbetet med undersökningen startade). Om en artikel inte kunde hittas i mediernas egna sökmotorer användes Google för ett ytterligare försök. De artiklar som inte hittades med dessa sökmotorers hjälp kodades som MISSING, vilket betyder att de kommenterades med ordet MISSING och de flesta variabelvärdet för dessa artiklar lämnades tomma. Kodaren fyllde i den information kring artikeln som kunde anas från framsidan (oftast huruvida den hade en bild, huruvida ett videoklipp var del av artikeln, och i vissa fall såsom sport ansågs variabeln Ämne kunna kodas säkert även från så lite information). Av 21 033 analysenheter har 1 479 st (7 %) kodats som MISSING.

Användandet av Wayback Machine ledde även till ett flertal fall där artikeltexten fanns tillgänglig, men bilder och videoklipp inte längre fungerade. I dessa fall kodades variablerna om Bild- och Videoförekomst som

att bild eller video förekom, men det gjordes också en kommentar om att de inte längre fanns tillgängliga och att den informationen inte hade kunnat tillgodogöras för kodningen. Wayback Machine sparar aldrig videofiler och för många medier är det praxis att radera videoklipp från hemsidan efter en viss tidsperiod för att spara utrymme.

I de fall där ett videoklipp eller ljudklipp är inbäddat tillsammans med en textnyhet kring samma ämne har de två kodats som en enhet.

Nedan följer en lista med undantagsregler för de digitala medierna. I alla andra fall har den artikel som har länkats från huvudsidan kodats:

- Alla former av navigeringslänkar som endast för användaren till en annan sida med länkar (Sport, Kultur, Arkiv, Sök o s v) har valts bort. Detta inkluderar även länkar till exempelvis informationssidor eller kontaktlänkar som sällan ändras och som inte kan anses tillhöra nyhetsrapporteringen.
- Alla länkar som markerats som någon form av annons har valts bort.
- Alla länkar till en epost-adress har valts bort.
- Om mer än en vald länk leder till en och samma artikel så har den kodats första gången som den dyker upp och sedan valts bort (samma artiklar har kunnat dyka upp både i det vanliga flödet, samt flera gånger under rubriker som ”Senaste nytt”, ”Mest delade nyheter”, ”Mest lästa nyheter” o s v). Om en och samma artikel har dykt upp på flera olika datum har den kodats flera gånger, men en kommentar har skrivits om att den har kodats tidigare

(till exempel har Aftonbladet på sin huvudsida en lista på länkar till alla krönikörers senaste text. En av dessa krönikörer skrev inte en ny krönika under en lång tid, så därför länkades det till hans senaste text på flera datum).

- Liksom i de analoga medierna har länkar till horoskop, korsord, seriestrippar, olika former av tester (exempelvis av typen ”Hur mycket vet du om...”) o s v valts bort.
- Öppna chattar som används av mediets användare för att kommentera nyheter har valts bort. Frågespalter som använder sig av chattar, där användare exempelvis har fått ställa frågor till en expert som mediet anlitat, har däremot kodats.
- Alla länkar till hela videoprogram eller hela poddavsnitt (audio) har valts bort. Många av medierna har börjat producera egna podcasts eller egna program av typen panelprogram där avsnitten kan uppgå till 30 minuter eller längre och dessa har valts bort. Detta gäller även länkar på exempelvis SVT och TV4 som länkar till avsnitt från kanalens andra, icke nyhetsrelaterade, program. Om däremot klipp från ett panelprogram på Aftonbladet eller i SVT har klippts ut och valts ut för att det exempelvis kommenterar en nyhet eller specifik händelse så har det varit med i undersökningen.
- Som ovan nämndes har ett inbäddat videoklipp eller ljudklipp vanligen kodats som del av en artikel. Undantaget till detta har varit i vissa fall, oftast i Aftonbladet, när samma videoklipp har bäddats in om och om igen i alla artiklar som berör en och samma händelse. Exempelvis inbäddades samma

videoklipp där Charlotte Kalla intervjuades efter att ha vunnit sitt OS-guld i Sotji med alla artiklar som rörde ämnet. Det gav en felaktig bild att låta samma klipp påverka kodningen gång på gång, så därför valde kodaren en av dessa artiklar (den där klippet ansågs passa bäst) och kodade videoklippen som en del av just denna artikel och ignorerade dess påverkan på resten av artiklarna som visade samma klipp.

- Det är viktigt att påpeka att länkar från en nyhetspartner som för användaren vidare till en annan hemsida, exempelvis när Svenska Dagbladet och Aftonbladet ibland länkar till varandras hemsidor, INTE har valts bort. En variabel har handlat om artikelns ursprung, där ett av alternativen har varit "Insamlad/inbäddad från partner". Flera av medierna har dessutom samarbeten med diverse fristående bloggar eller specialsidor och dessa har behandlats på liknande vis (även om dessa oftare än andra artiklar har varit MISSING då inte alla liknande sidor upprätthåller egna arkiv).

Reliabilitetstest

För att garantera samstämmighet mellan de tre kodarna genomfördes ett så kallat reliabilitetstest på en del av materialet. Två olika mått användes för att mäta reliabiliteten, ett snittat Holsti-mått och Krippendorffs Alpha.

Holsti undersöker i hur många procent av fallen två kodare håller med varandra i sina kodningar. Kodar två kodare exempelvis samma 100 artiklar så kan man sedan mäta i hur många procent av fallen (per variabel)

de håller med varandra. För acceptabel reliabilitet ska kodarna i de flesta fall hålla med varandra i 80 procent av fallen (vilket skrivs ut som 0.8). Det finns undantag där ett lägre mått är acceptabelt som exempelvis när en variabel har många olika svarsalternativ (majoriteten av variablerna i den här undersökningen har tre svarsalternativ, men i exempelvis variabeln om Nyhetsämne fanns femton olika alternativ). Själva Holsti-måttet kan alltså bara mätas mellan två olika kodningar, så i ett fall som detta med tre kodare tar man fram ett Holsti-mått för Kodare 1 och 2, ett mått för Kodare 1 och 3 och ett för Kodare 2 och 3. Snittet av dessa tre mått blir det slutliga, snittade måttet. I det här fallet kom samtliga variabler att ha ett mått på över 0.8 enligt det snittade Holsti-måttet.

Krippendorffs Alpha är ett mer komplicerat mått som tar in i beräkningen hur vanligt förekommande ett variabelvärde är för att försäkra att kodarna verkligen plockar upp detta variabelvärde när det väl förekommer. För att visa hur det fungerar kan man tänka sig följande exempel:

- I exemplet finns två variabler som har kodats i 100 fall av två kodare. Båda variablerna har svarsalternativen "Ja" och "Nej".
- I den ena variabeln är det "korrekta" svaret "Ja" i 50 av fallen och "Nej" i 50 av fallen. Då påverkas inte reliabiliteten nämnvärt enligt Krippendorffs Alpha om en kodare vid ett tillfälle kodar annorlunda än den andra kodaren. Ett misstag kommer inte påverka den huvudsakliga statistiken.

- I den andra variabeln är det ”korrekta” svaret ”Ja” i 98 av fallen och ”Nej” i 2 av fallen, alltså är det ena alternativet betydligt mer ovanligt förekommande. Då blir det enligt Krippendorffs Alpha-måttet extra viktigt att båda kodarna kan identifiera svarsalternativet ”Nej” när det väl dyker upp, för annars blir variabeln meningslös. Om en kodare missar en av de två ”Nej”-fallen påverkar det reliabiliteten nämnvärt enligt Krippendorffs Alpha.
- I båda dessa exempel har de kodarna alltså svarat lika i 99 fall av 100 och båda exemplen skulle få ett Holsti-mått på 0.99.
- Men på grund av att Krippendorffs Alpha beräknar variabelernas förekomst skulle de två exemplen få väldigt olika mått enligt den metoden (0.98 i det första exemplet, 0.66 i det andra exemplet).

För Krippendorffs Alpha brukar ett värde på 0.6 anses nödvändigt för att garantera reliabilitet för en variabel. I vårt interkodartest föll 9 av 39 variabler under ett Krippendorffs Alpha-värde på 0.6. I och med de höga snittade Holsti-måtten för dessa variabler har de använts i rapporten. 112 artiklar ingick i detta första test.

Kodningen av materialet från 2007, 2014 och 2016 pågick utan större avbrott under ett och ett halvt års tid och dessa tidigare nämnda reliabilitetstest ansågs vara giltiga för all denna kodning. När projektet sedan skulle uppdateras med materialet för 2018 så hade det skett en längre paus i arbetet och endast en kodare jobbade nu med projektet. Då utfördes ett ytterligare reliabilitetstest på 100 artiklar för att försäkra att

kodningen höll fortsatt god kvalitet. Resultatet av detta test mättes med Holsti och alla variabler uppnådde ett värde på över 0.8.