

Randanteckningar om svensk mediepolitik 1920–1970

Ur temamodellering av 11 000 riksdagsprotokoll

PELLE SNICKARS*

Abstrakt

In an explorative manner, this article uses a data-driven digital history set-up to focus on how the Swedish Riksdag, its political parties and parliament members discussed media issues. By distant reading and topic modeling a dataset of more than 11,000 parliamentary minutes – approximately 400 million characters – the article gives a new perspective of Swedish media history in general and national media politics in particular. The article uses the computational method topic modeling to study latent themes or discourses in the dataset (by accentuating words that tend to co-occur) that together create different topics. The article includes a number of media historical observations and findings, and the frequent graphs also display interconnections between different media topics, as well as other societal topics that media were related to.

Keywords: digital history, media history, topic modeling, Swedish Riksdag.

Inledning

Vilka är de tematiska idéer och tankefigurer som styr svensk politik sedan demokratin införande under tidigt 1900-tal? En sådan fråga är så omfattande att den kan tyckas omöjlig att besvara. En statsvetare skulle möjligen peka på partiväsendets framväxt som ett sätt att få korn på övergripande politiska diskurser, medan den ekonomiske historikern nog skulle leta svar i nationalekonomins utveckling. Den genusmedvetna kvinnohistorikern skulle säkerligen framhålla centrala aktörer som exempelvis Landsföreningen för kvinnans politiska rösträtt – bildad 1902 och upplöst 1921 när den allmänna rösträtten genomfördes – medan idéhistorikern med all säkerhet skulle lyfta fram hur klassiska ideologier som liberalism, socialism och konservatism varit avgörande. Ett antal tongivande nyckeltexter skulle

*Pelle Snickars, professor i digitala kulturer, institutionen för kulturvetenskaper, Lunds universitet, pelle.snickars@kultur.lu.se

då i regel åberopas, till exempel August Palms ”Hvad vil Sosial-Demokraterne”, ett anförande han höll på Hotel Stockholm i Malmö 1881.¹ Alternativt skulle den politiskt orienterade historikern framhäva skrifter författade av någon framträdande minister eller parti-ideolog, kanhända genom ett idépolitiskt studium av högerledaren Arvid Lindmans tal och föredrag – han var statsminister 1906–1911 och 1928–1930 – eller en analys av docenten i nordiska språk, tillika finansminister Ernst Wigforss omfattande publikationsverksamhet. Det är visserligen långt ifrån alla socialdemokrater som efterlämnade femtiotalet publicerade skrifter, men för en hugad historiker skulle det likväl handla om närläsning av ett relativt begränsat antal texter. För det är ju så många historiker arbetar; även jag själv – vi kan inte läsa allt.

För den digitalt orienterade historikern öppnar sig emellertid nya möjligheter att besvara frågekomplex av makro-karaktär. Om humanister under lång tid ägnat sig åt kvalitativ närläsning och texttolkning med fokus på det innehåll som enstaka dokument rymmer, då har sådana forskningspraktiker under det senaste decenniet kompletterats med olika former av distans- eller fjärläsning utförda med datorers hjälp och språkteknologisk mjukvara. Så kallad *text mining* handlar strängt taget inte om att läsa, snarare om att mjukvara upptäcker språkliga mönster i textmassor. Lingvister och språkteknologer har länge arbetat med snarlika metoder applicerade på omfattande textkorpusar.² Enklare språkteknologiska metoder inbegriper ordfrekvenser och kollokationer, *bigrams* (ordpar som uppträder tillsammans), liksom mer sofistikerade NER-algoritmer (*Named-entity recognition*) där personnamn, geografi och tidsmarkörer automatiskt taggas upp. Den senare metoden är användbar om det inte är språkets form som är av primärt intresse utan texters historiespecifika innehåll. En annan analysmetod som blivit vanlig bland digitala historiker är så kallad *topic modeling* – temamodellering på svenska – ett samlingsnamn på en rad algoritmer som kan klassificera stora textkorpusar baserat på tematiska strukturer i textmassan. Genom temamodellering kan diskurser och begrepp urskiljas – ja, till och med tankefigurer som genomsyrat svensk politik – baserade på ords statistiska sammanhang i ett storskaligt empiriskt material.³

Arkivdriven historievetenskap har i så måtto kompletterats med data-driven *digital history*, det senare ett kunskapsfält som kommit att omfatta flera olika humanistiska discipliner. ”Digital history is an approach to examining and representing the past; it uses new communication technologies and media applications and experiments with computational methods for the analyses, production and dissemination of historical knowledge.”⁴ En återkommande frågeställning inom detta kunskapsfält är vad mjukvara kan uppfatta när algoritmisk textanalys appliceras på ett

omfattande empiriskt material. En materialkategori som lockat många digitala historiker är riksdagstryck (i olika former) – av den enkla anledningen att sådant textmaterial av demokratiska skäl inte har några upphovsrättsliga restriktioner. Inom den europeiska forskningsinfrastrukturen för språkteknologi, CLARIN, finns exempelvis 26 länders dataset av riksdagstryck öppet tillgängliga. Det har också inneburit att forskare byggt olika slags applikationer för att analysera sådana dataset. Mest bekant är förmodligen *the Hansard* – transkriberade debatter från Storbritanniens parlament från 1803 till idag, ett dataset på flera hundra miljoner ord – vilket kan analyseras via applikationer som The Hansard Viewer eller Hansard at Huddersfield. I Sverige har Centrum för digital humaniora vid Göteborgs universitet utvecklat applikationen Riksdagsmotioner.nu som möjliggör sökning bland 145 000 riksdagsmotioner från 1971 till idag, och i Tyskland ger Open Discourse access till alla debatter i den tyska Förbundsdagen, *Plenarprotokolle des deutschen Bundestages*, från 1949 till idag, ett textkorpus som innehåller mer än 900 000 anföranden fördelade på 4 000 protokoll.⁵

De flesta av dessa applikationer är tämligen rudimentära och handlar om att söka efter enstaka ord där grafer visar hur ofta de förekommer över tid. På Riksdagsmotioner.nu återfinns också en träfflista med information om vilket parti och riksdagsledamot som skrivit motionen (i vilket sökordet ingår) samt vad den handlar om. Andra applikationer som The Hansard Viewer och Open Discourse är mer avancerade och ger bland annat möjlighet att studera diskursiva politiska tematiker över tid genom temamodellering. I den senare finns 73 förvalda teman – från *Arbeit* till *Öffentliche Verwaltung* – vilka kan studeras över tid, sorteras efter parti, riksdagsledamot, kön, ålder eller (tidigare) yrkeskategori. Gränssnittet är mycket användarvänligt, det är utvecklat av företaget Limebit (för den tyska riksdagens räkning). Open Discourse ”is a database consisting of six data tables providing granular information about speeches, factions, politicians, election periods, interjections and other contribution types like applause, laughter, and cheerfulness”, kan man läsa på Harvard Dataverse där alla dessa dataset är samlade med öppen access.⁶

I den här artikeln kommer jag att fokusera på hur svenska partier och riksdagsledamöter diskuterat mediefrågor i Sveriges Riksdag. Det dataset jag arbetar med är baserat på fler än 11 000 riksdagsprotokoll – cirka 400 miljoner tecken – och omfattar en hundraårig tidsperiod (1920 till idag). Av utrymmesskal kommer jag att begränsa mig till perioden 1920 till 1970 (tvåkammarriksdagen upphörde också detta år). Jag är mediehistoriker och min tanke är att studera övergripande politiska tankefigurer kring medier genom att modellera mitt dataset. Artikeln syftar till att beskriva hur data sammanställts liksom vad metoden temamodellering innebär,

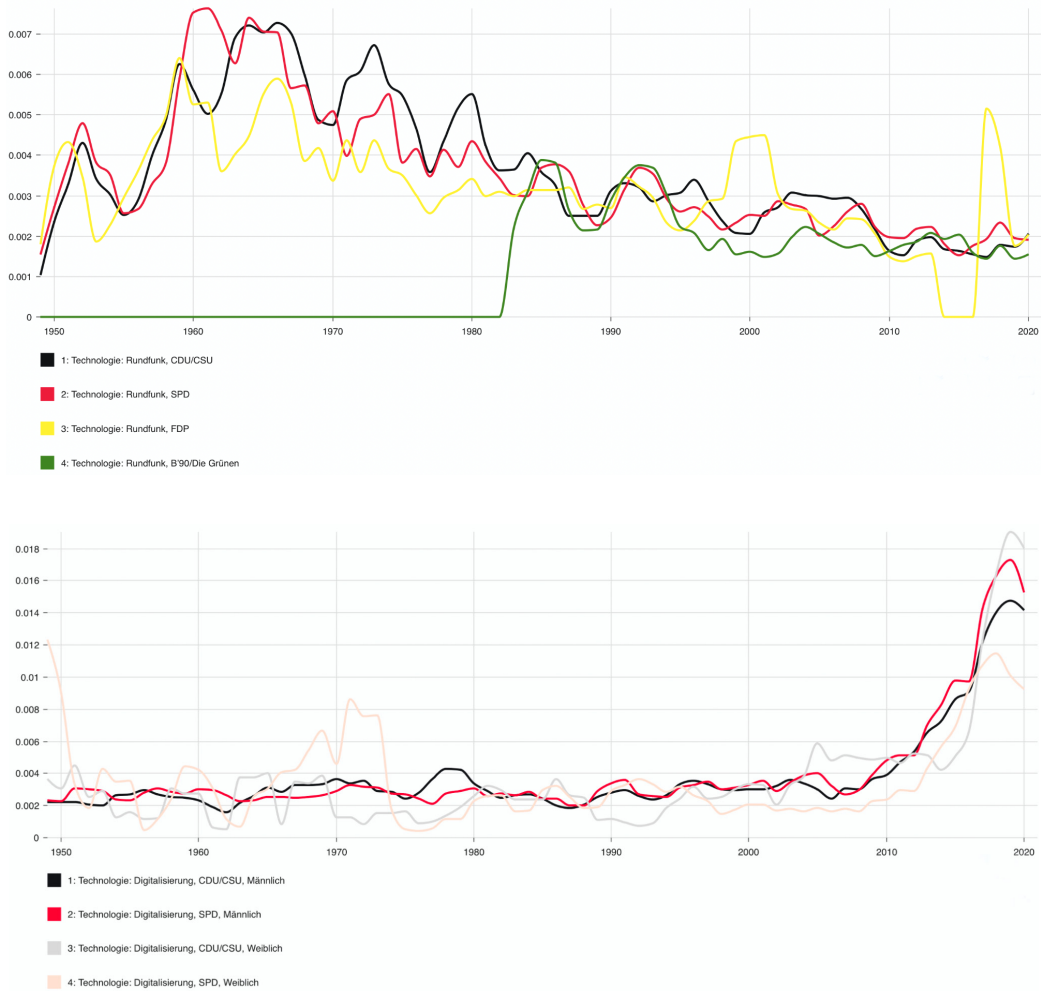


Fig. 1 & 2. Grafer från opendisourse.de baserade på ett dataset (på 200 miljoner tecken) av alla protokoll i den tyska Förbundsdagen från 1949 till idag. Den övre grafen illustrerar när fyra partier diskuterat tema *Rundfunk* – alltså etermedia. Partiet B'90/Die Grünen kom in i Förbundsdagen först 1983, därav nollvärdet före dess. Men nollvärdet för liberala FDP i mitten av 2010-talet beror sannolikt på felaktigheter i datan. Den nedre grafen illustrerar när män respektive kvinnor i konservativa CDU/CSU och socialdemokratiska SPD diskuterat temat *Digitalisierung*.

men främst gör jag ett par analytiska nedslag i en omfattande empiri – där jag också citerar flitigt från enskilda anföranden. Jag är medveten om att detta syfte är omfattande och nog egentligen hade krävt mer utrymme än en enskild artikel. Icke desto mindre är min ambition i det följande att beskriva både data och metod liksom ett par forskningsresultat. En bred ansats går med andra ord före mer djuplodande resonemang. Noterbart är att artikeln innehåller flera grafer; i det följande handlar det lika mycket om att visa som att beskriva. Arbetet är relaterat till det forskningsprojekt jag leder, Välfärdsstaten analyserad. Textanalys och modellering av svensk politik, media och kultur 1945–1989. Där ägnar vi oss åt algoritmisk textanalys av storskalig empiri från politikens sfärer (digitaliserat riksdagstryck och offentliga utredningar), dagspress och skönlitteratur.⁷ Svensk efterkrigstid är en väl utforskad period, men genom att applicera digitala metoder på stora dataset kan politikens, mediernas och kulturens sfärer granskas på nytt.

Data och metod

Svenskt riksdagstryck har under de senaste femton åren successivt digitaliserats i omgångar av Riksdagsförvaltningen. År 2010 öppnade de sina databaser för fri användning. Materialet omfattar protokoll – med debatter från både första och andra kammaren – propositioner, förslag och redogörelser, motioner, utskottsbetänkanden och riksdagsskrivelser. På riksdagen.se går det att fritextsöka i detta material från och med tvåkammarriksdagens införande 1866, och på Riksdagens öppna data och Kungliga bibliotekets sajt för öppen data kan allt riksdagstryck enkelt laddas ned – ända tillbaka till ståndsriksdagens införande 1521. Med utgångspunkt i Riksdagsförvaltningen digitaliseringsarbete har vi i vårt forskningsprojekt skapat ett kurerat dataset av alla riksdagsprotokoll från första och andra kammaren under de senaste hundra åren. Det rör sig om knappt 11 200 protokoll med cirka 750 000 anföranden. OCR-kvaliteten på de inskannade dokumenten är god, men innehållet är synnerligen heterogent. För även om de PDF:er som finns tillgängliga är sökbara så innehåller de ingen strukturerad data. Det har alltså inte tidigare varit möjligt att till exempel få fram information om vad kvinnliga folkpartister eller män i Bondeförbundet ansåg om en viss fråga under en viss period. Det är heller inte alltid så enkelt att redogöra för vem som *de facto* debatterade, för i protokollet framgår inte alltid exakt vem som förde talan. Ett exempel: protokollet från andra kammaren den 11–12 maj 1937 är 130 sidor långt. I kammaren diskuterades då bland annat undernärda barn i Norrbotten, lönereglering för befattningshavare ”vid lots- och fyrstaten”, underhåll av flottans fartyg, liksom ”åtgärder för gynnande av produktio-

nen av verkligt god folkuppfostrande svensk film” – det senare apropå den folkpartistiske riksdagsledamoten Ernst Åqvist motion i ärendet. I denna filmdebatt i kammaren deltog vid sidan om Åqvist själv även ”Herr Hagberg i Malmö”, ”Herr Bergström”, ”Herr Hoppe”, ”Herr Mosesson” – vars inlägg ”Herrar Johanson i Huskvarna och Pettersson i Lerbäcksbym” instämde i – ”Herr Gustafsson i Bogla”, ”Herr Hallén”, ”Fröken Hennings”, ”Herr Lovén” (igen), ”Herr Skoglund”, ”Herr Karlsson i Munke-dal”, ”Herr Olovson”, ”Herr Vougt”, ”Herr Hallén” (igen) samt till sist ånyo ”Herr Lovén”.⁸

Det var många herrar (och en fröken), men alla dessa riksdagsledamöter hade förstås fullständiga namn och partitillhörighet. Det säger sig självt att det är ett omfattande arbete att förse protokollen med adekvat metadata. Vi har använt oss av dokumentmallen Parla-Clarín, en europeisk XML-standard (Extensible Markup Language) som utvecklats just för annotering av riksdagsdata. Utan att gå in på tekniska detaljer innebär det att alla riksdagsprotokoll kompletterats med metadata för namn på talare, geografisk hemvist, partitillhörighet och kön – ett arbete som kan beskrivas med en mening, men som varit (och är) mycket tidskrävande. Eftersom vårt dataset är så omfattande är det i princip omöjligt att städa manuellt varför en rad skript använts för kurering. Ett speciellt bökigt problem har varit att på automatisk väg märka upp när ett anförande slutar eftersom det inte alltid framgår. Ett annat har varit att både identifiera och särskilja riksdagsledamöter från varandra, det handlar om tusentals namn (som ibland är identiska). Genom maskininlärning har olika algoritmer konstruerats som förbättrat metadata; de har tränats på mindre dataset för att göra datadrivna förutsägelser – varefter de kontrollerats manuellt. När det gäller namn på riksdagsledamöter har vi i princip byggt om all metadata från grunden genom ett samarbete med Wikidata, där alla individer (som riksdagsledamöter) har ett unikt identifieringsnummer (folkpartisten Ernst Åqvist är exempelvis Q6343614). Det har gjort det möjligt att både särskilja ledamöter från varandra (om de har samma namn), förse dem med tidsangivelser när de satt i riksdagen (ibland i omgångar), ändra deras roll i riksdagen över tid (ledamot, minister, talman etcetera), liksom (mer sällan) förändrad partitillhörighet.

Nästan all kvantitativt orienterad digital historieforskning är beroende av datakurering. För det är ofta först med städad och strukturerad data som det går att applicera olika algoritmiska analysmodeller. Samtidigt kommer den riksdagsdata vi arbetar med aldrig att bli komplett, i ett så stort dataset finns det alltid förbättringar att göra.⁹ På det digitala humaniora-centrat Humlab på Umeå universitet har utvecklare i vårt projekt först kört all riksdagstext genom Språkbankens korpuspipeline Sparv för att annotera och märka ord i texten med dess ordklass och

grundform. Resultatet av både annotering och lemmatisering (sammanföring av olika böjningsformer och varianter av ett ord) lagras som csv-filer, och de är från dessa som modelleringen utgår. Inom temamodellering finns flera modeller (och olika algoritmer) att arbeta med. När det gäller riksdagsdatan har vi använt så kallade LDA (Latent Dirichlet Allocation) Mallet-modeller om 100, 200 och 500 teman. Det är alltså valet av modell som bestämmer vilket antal teman som ett dataset ska delas upp i (inte forskaren). Teman listas genom en serie av ord som är tematiskt snarlika: radio, tv, tidning, program, reklam och presstöd är till exempel några av de mest frekvent förekommande orden i ett medietema som återfinns i 100-modellen av riksdagsdatan. Det är främst denna modell som jag använder som fallstudie i denna artikel. Man bör notera att teman alltid enbart listas i stigande nummerordning (från 0 till 99 eller 0 till 499), varför all namngivning av teman baserar sig på subjektiva tolkningar.

Grundfrågan inom all temamodellering handlar om vilka diskurser, motiv eller tematiker som kan tänkas återfinnas i det dataset som forskaren intresserar sig för. Temamodeller arbetar med statistisk analys av distribution av ord i form av sannolikhetsfördelning över alla ord i en textmassa. Modellerna kan urskilja vilka ord som ingår i ett tema (baserat på ords statistiska fördelning), hur det fördelar sig över tid, liksom hur temat – eller ord i ett tema – anknyter till andra teman.¹⁰ Grundprincipen är att ord samförekommer. Det hela är inte konstigare än att det är mer sannolikt att ord som biograf och spelfilm förekommer i en debatt om filmpolitik än termer som traktor eller jordbruk. Men all temamodellering handlar om probabilitet. I modeller som listar 500 teman är sannolikhetsvärden mycket låga (i promille, eftersom det handlar om många teman), medan de i modeller om 100 teman är något högre (ofta i procent).

Samtliga temamodeller har gjorts åtkomliga för mig i utvecklingsmiljön Jupyter Lab; via ett användarvänligt gränssnitt kan man där köra (och ändra på) Pythonkod direkt i webbläsaren. I Jupyter Lab finns kodceller där jag som forskare med knappar och manuella reglage kan styra parametersättningen av koden; de kan modifieras och exempelvis plotta valfritt antal ord i varje tema (från 25 till 200), lista teman över tid (eller en specifik tidsperiod), jämföra eller vikta teman med andra snarlika tematiker etcetera. Resultatet förevisas som tabeller, diagram eller enklare nätverk. Noterbart är att Jupyter Lab också tillåter närläsning av de anföranden som ingår i varje tema – även de statistiskt listade i frekvensordning. Alla enskilda anföranden återfinns i de 11 200 protokollen, vilka är länkade från Jupyter Lab-miljön och tillgängliga i XML-format på forskningsprojektets Githubsajt. Det gör det möjligt att zooma in och ut ur riksdagsdatan, från övergripande diskurser till de enskilda yttranden som alla teman är uppbyggda av.

Tema medier

I 100-modellen av riksdagsdatan återfinns bredare samhällseliga teman och politiska frågeställningar som diskuterats i de bägge kamrarna – skatter, infrastruktur, utbildning, försvar, sjukvård. Men bland dessa teman kan man också hitta två som handlar om medier: nummer 77 om press, radio och television, samt nummer 87 om kultur, film och litteratur. Det anförande i riksdagsprotokollen som enligt temamodellerings beräkningar allra främst innehåller tema 77 (under perioden 1920–70) är ett kort genmäle av högerns dåvarande partiledare Yngve Holmberg från mars 1966: ”Herr talman! Det finns ingen risk för pressmonopol i Sverige, men vi kan konstatera att det finns ett statligt radio- och TV-monopol.” Faktum är att samme Holmberg också återfinns högt upp i det mediekulturella temat 87, detta i ett polemiskt anförande från oktober 1964 riktat till den socialdemokratiska ecklesiastikministern Ragnar Edenmann (som året före hade varit ansvarig för filmreformen 1963 och etableringen av Svenska Filminstitutet): ”Anser Herr Statsrådet att det ingår i Svenska Filminstitutets uppgifter att anordna fester av så påkostat slag som skedde vid filmfestivalen i Venedig i år?”¹¹ Temamodellering som metod gör det alltså möjligt att i en mycket omfattande empiri tydligt precisera när mediepolitiska diskussioner förts i riksdagen. I Jupyter Lab-miljön går det att bokstavligen lista tiotusentals sådana anföranden. Dessa är naturligtvis också möjliga att lokalisera genom fritextsökning, men fördelen med temamodeller som 77 (om press och etermedia) och 87 (om mediekultur) är att de ger många fler semantiska ingångar än enbart enstaka sökord. För som framgår av nätverket nedan inbegriper varje tema hundratals ord som modellen listar.

Beträffande de medier som de två modellerna inkluderar kan man också initialt göra den mediehistoriska observationen att film samförekommer med kulturella verksamheter. I riksdagsprotokollen förefaller film att främst betraktats som en kulturpolitisk angelägenhet, medan press, radio och tv mer tycks handlat om kommunikationstekniska frågor, framför allt när det handlar om etermedia. Ord som biograf, filmpolitik och filmproduktion samsas i tema 87 med begrepp som konstnär, kulturliv och kulturskapare. Samma sak gäller även i institutionellt avseende; Svenska Filminstitutet figurerar i temat sida vid sida med Dramaten, Skansen och Operan. Iakttagelsen kan tyckas marginell, men det är precis den sortens oväntade resultat som gör temamodellering till en spännande metod. Givet att det mediekulturella temat 87 ingriper en rad kulturpolitiska områden är det emellertid inte förvånande att det förekommer i långt större utsträckning under perioden än det tema som främst handlar om etermedia. Av stapeldiagrammen nedan framgår att det senare temat är

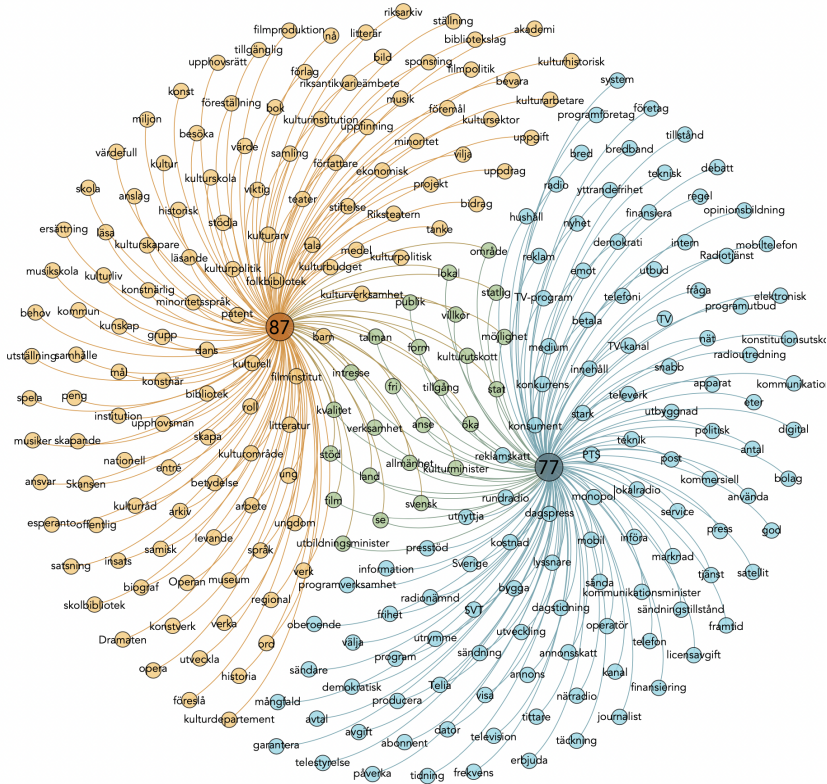


Fig. 3. Nätverk i Gephi med algoritmen Fruchterman Rheingold med cirka 260 ord-noder från 100-modellen av riksdagsdatan med de mest frekvent förekommande termerna i tema 77 (om press, radio och television) och tema 87 (om kultur, film, litteratur, bibliotek, museer och teater). De gröna noderna i mitten utgörs av de begrepp som sammanbinder dessa bägge medieteman – bland dem passande nog kulturminister.

koncentrerat kring televisionens genombrott och utbyggnad från mitten av 1950-talet, även om en uppgång i början av 1930-talet också är skönjbar, kopplad till radiomediets ökade popularitet. Det mediekulturella temat 87 är långt mer närvarande i riksdagsprotokollen under hela tidsperioden, möjligen med en viss nedgång under krigsåren.

Eftersom kultur- och mediepolitik ofta handlat om statligt stöd så återkommer termer som kulturbudget, reklam, sponsring och presstödd i bägge teman. Framför allt i tema 87 antyder ekonomiska begrepp som stöd, anslag och bidrag att den mediekulturella politiken ofta handlat om kärv ekonomi. Anföranden från den tidigare nämnda filmdebatten i andra kammaren våren 1937 är exempelvis starka i temat, bland dem folkpartisten Beth Hennings plaidoyer för ökat statligt stöd till filmen:

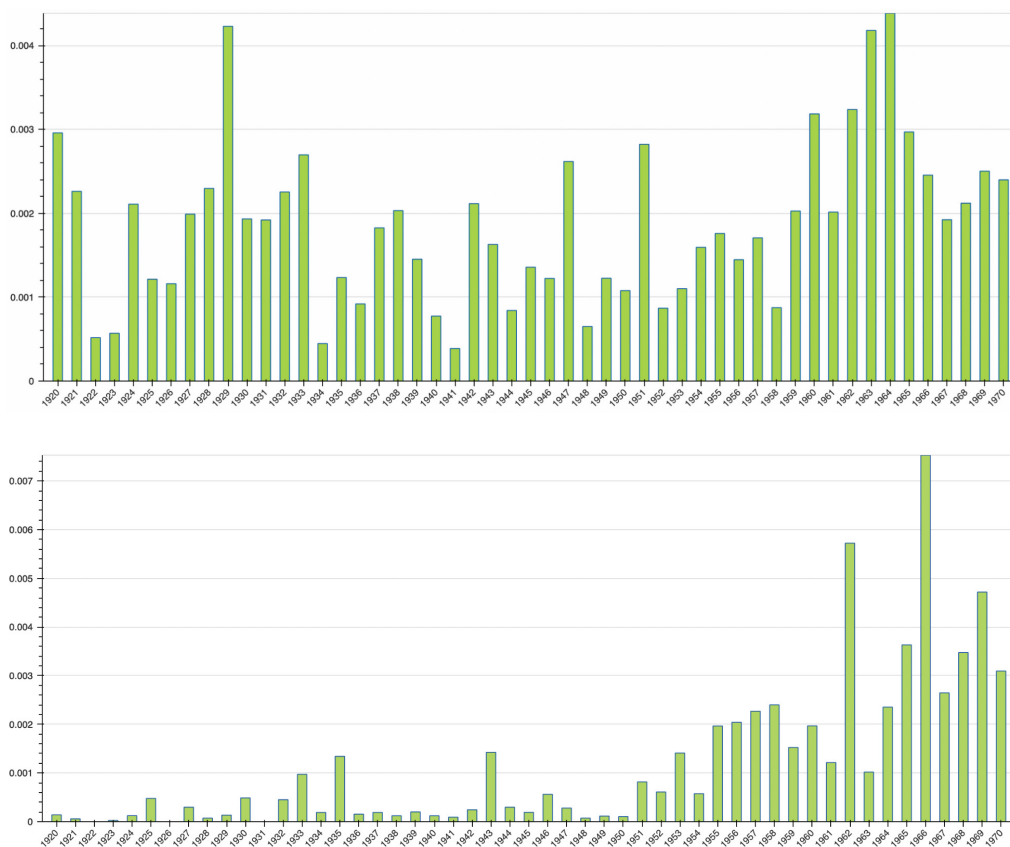


Fig. 4 & 5. Stapeldiagram från utvecklingsmiljön Jupyter Lab av 100-modellen av riksdagsdatan med press och etermedietema 77 (ovan) och det medicokulturella tema 87 (under).

Är det icke i alla fall märkligt, att medan svenska staten faktiskt delvis under århundraden har understött teater, musik, konst och litteratur, så finner man det nu vara så egendomligt, att staten skulle taga ett steg vidare på denna linje och understödja den femte konstarten, som man kan säga är typisk för vår egen tid? Det borde väl vara naturligt, att när staten offrar stora summor på dessa andra konstarter, man nu borde vara mogen att gå in för ett understöd åt den konstart, som nu är, kan man väl säga, i så särskilt hög grad det svenska folkets. Det är faktiskt så, att filmen i vissa fall tränger djupare ned och vidare omkring än någon annan konstart. Men hittills har staten icke gjort något som helst för den svenska filmen.¹²

Beth Hennings filmpolitiska inspel kan tas som intäkt för att se närmare på hur just folkpartistiska kvinnors anföranden figurerar inom det kulturpolitiska tema 87 över en längre period. En av temamodelleringens styr-

kor som metod är möjligheten att närmare analysera ett specifikt tema, i detta fall genom att undersöka hur tid, genus och/eller partitillhörighet är faktorer som påverkar – med förbehållet att riksdagens genushistoria är synnerligen ensidig. 1949 var mandaten i första kammaren fördelade på 144 män och sex kvinnor, och efter valet 1952 var 79 procent män i andra kammaren.¹³ Kvinnor var representerade i riksdagens kammare – men de var mycket få. Som genushistorikern Camilla Norrbin påpekat så tog det ett halvt sekel ”innan kvinnornas andel hade vuxit till knappt femton procent i Sveriges riksdag”.¹⁴

Folkpartiet är av tradition det borgerliga parti som varit mest förknipat med ett kulturpolitiskt engagemang. Bland anföranden av kvinnliga riksdagsledamöter är det också folkpartister och socialdemokrater som förekommer allra starkast i det mediekulturella tema 87 under perioden 1920-1970.¹⁵ Att det generella politiska engagemanget för mediefrågor ökade efter andra världskriget är också uppenbart, i grafen nedan framgår att före krigsåren så var det nästan enbart folkpartistiska kvinnor (som Beth Hennings) som diskuterade den här typen av frågor. Bondeförbundets kvinnor hade exempelvis inte något översvallande intresse i frågan, men partiet hade ytterst få kvinnor på mandat så deras anföranden går knappt att registrera. Efter kriget var det snarare Höger- och Folkpartiets kvinnor som debatterade med Socialdemokraterna. Ett av dessa anföranden utgjordes av folkpartisten Ingrid Andréns propå från våren 1955 för stöd till svensk barnfilm, detta apropå den höga nöjesskatt som plågade filmbranschen. Finansminister Wigforss – en man vars hjärta ”inte precis blödde för filmen”¹⁶ – hade 1948 infört en skattechock och höjt nöjesskatten till 40 procent, med påföljden att de flesta svenska filmateljéer stängde. Det var filmstoppets tid. Wigforss hustru Eva var mer filmintresserad – hon var rentav ordförande i kvinnoorganisationernas barnfilmskommitté mellan 1948-1952 – och med utgångspunkt i den typen av arbete menade folkpartisten Andréns att det var hög tid att satsa på god barnfilm, som var ”absolut beroende av ett produktionsstöd, ty i Sverige, ja t.o.m. i Skandinavien, blir dess marknad alltför begränsad”.¹⁷

Plottar man tema 87 för fem olika riksdagspartier under perioden 1920-1970 är det folkpartistiska intresset för frågan intakt, främst efter krigsslutet. Men under mellankrigstiden var det socialdemokraterna som (långt mer än andra partier) drev den här typen av medierelaterade frågor och diskussioner i riksdagens kamrar. Ett anförande där temat förekommer starkt är hämtat från en debatt i andra kammaren våren 1925 – man diskuterade anslag till en folkhögskola med musikinriktning – där socialdemokraten Anders Paulsen ivrade för svensk folkmusik, men inte genom att inrätta ordinär undervisning av densamma. ”Jag älskar denna folkmusik minst lika högt som någon av herrar reservanter, men jag anser, att

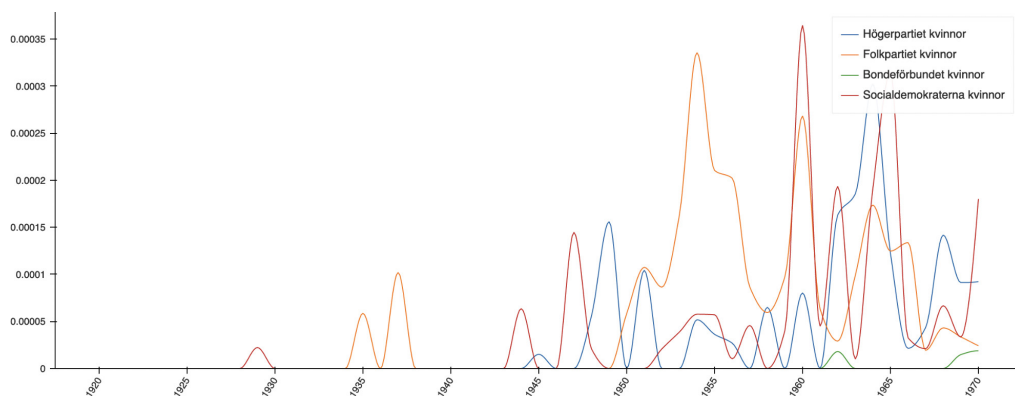


Fig. 6. Förekomst av det mediekulturella temat 87 i anföranden av kvinnliga riksdagsledamöter för fyra politiska partier under perioden 1920–70. Det kan förefalla som att Folkpartiet i mitten på 1950-talet inte diskuterade med någon alls – men förmodligen var det partiets kvinnor som då tog sig an manliga ledamöter.

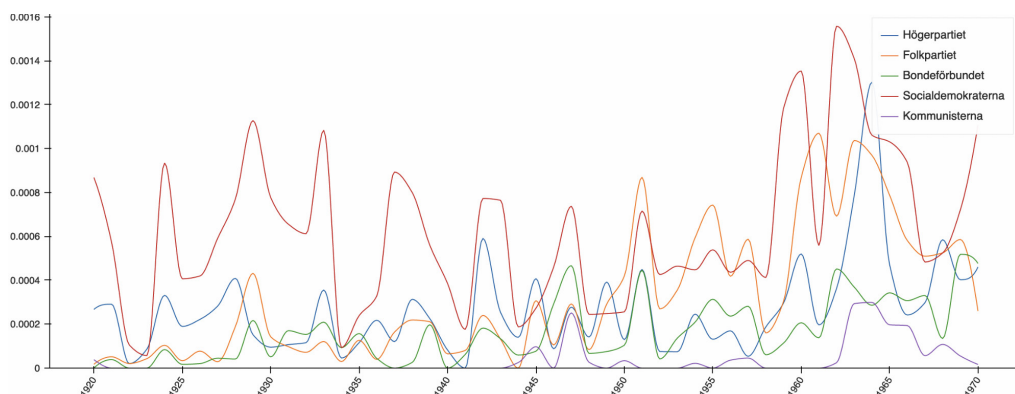


Fig. 7. Förekomst av det mediekulturella temat 87 i anföranden av samtliga riksdagsledamöter för fem olika politiska partier under perioden 1920–70. Av grafen är det uppenbart att mediekultur låg socialdemokrater varmt om hjärtat – liksom folkpartister under efterkrigstiden.

denna musikskola i stället skulle fördärva folkmusiken ... Den skulle inordnas i schablon liksom all annan musik”. Partibrodern Värner Rydén instämde, inspelad musik ”på grammfonen [var förkastlig] medan samtidigt fiolen hänger på väggen, neddamad eller måhända sprucken”.¹⁸

Modelleringen av det mediekulturella temat 87 ger klart besked om att Högerpartiet var långt mindre intresserat av den typen av mediepolitiska frågor än Socialdemokraterna och Folkpartiet. Den allmänna uppfatt-

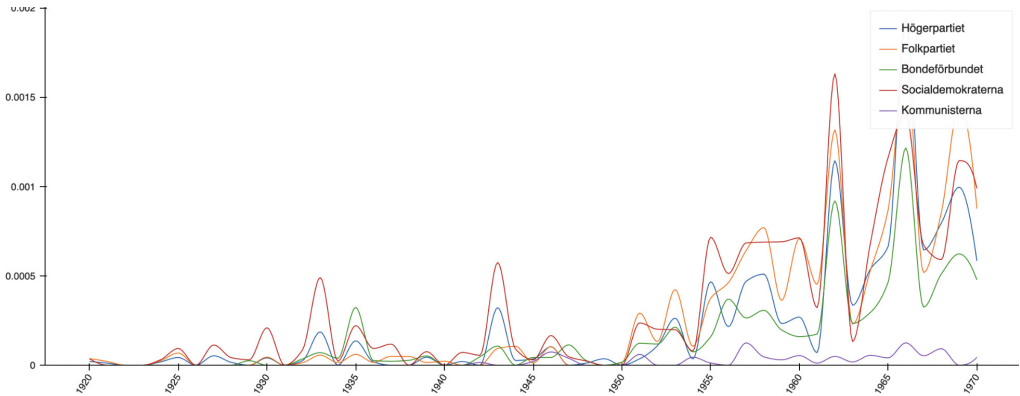


Fig. 8. Förekomst av det press och etermediala temat 77 i anföranden av samtliga riksdagsledamöter för fem olika politiska partier under perioden 1920–70. Det var med televisionens genombrott som mediala frågor av denna art på allvar började att diskuteras i riksdagens kamrar.

ningen att kulturfrågor inte legat högern speciellt varmt om hjärtat har alltså stöd i riksdagsdatan. Om man emellertid byter tema och tittar närmare på tema 77 – med fokus på framför allt etermedia – ja, då är bilden en annan. Där är det uppenbart att högern hade ett större och återkommande intresse för public service i nästan lika hög utsträckning som Socialdemokraterna, vilka som bekant satt i regeringsställning under stora delar av den undersökta perioden (och ofta hade fler partimandat än de borgerliga partierna tillsammans). Så beklagade sig exempelvis högermannen Edvin Leffler över den föredragsserie om svenska bostadsförhållanden som Radiotjänst sände vintern 1939, varvid han ställde frågan till chefen för kommunikationsdepartementet, (Gerhard Strindlund, Bondeförbundet) huruvida det var Radiotjänst eller föredragshållare som bar ansvar för utsändningarna. ”Propaganda är nog bra i alla avseenden, men den skall föras på ett lämpligt sätt, alldeles särskilt när det gäller en såväl ekonomiskt som socialt så viktig fråga som bostadsbyggandet på landet.”¹⁹

I början av 1960-talet var det inte radio utan televisionen som engagerade de högersinnade riksdagsledamöterna. Spiken i grafen 1963 handlade om de återkommande debatter som televisionen genererade. ”Det kan kanske synas egendomligt att vi i dag tar upp frågan om televisionens ställning i Sverige, när vi hade en lång och intensiv debatt om detta ämne i riksdagen så sent som praktiskt taget på dagen för tre månader sedan”, påpekade exempelvis Högerpartiets Bo Turesson våren 1963, ett anförande som förekommer starkt tema 77:

Det föreligger dock en stor principiell skillnad mellan den fråga vi diskuterade då och den debatt som jag nu tar upp. Då var debatten koncentrerad till frågan om reklam i televisionen. [Idag] anknyter den till en motion med vidare syftning och är en följd av ett alltmer mognande demokratiskt krav på frihet i etern, d.v.s. en fri etableringsrätt för med Sveriges Radio konkurrerande televisionsföretag.²⁰

Samma typ av argumentation återkom vid flera av partiledare Holmbergs inspel (samtliga starka inom tema 77). Han ställde sig undrande inför det etermediala monopoliet, inte minst i relation till socialdemokratiska idéer på ett kommande presstödsområde som frikostigt skulle fördelas till en mängd tidningar. Holmberg trodde sig veta att det inte fanns:

... någon principiell skillnad mellan olika massmedias opinionsbildande uppgifter, d.v.s. mellan press, radio och TV samt film. Skillnaden hänförs sig egentligen bara till genomslagskraften ... Jag tror också att vanliga människor tycker att det är egendomligt att regeringen å ena sidan talar om att det behövs en stor, oberoende skara tidningar och tidskrifter för att opinionsbildningen skall bli så rik och varierad som möjlig, medan den å andra sidan säger att det skall föreligga monopol för de viktigaste opinionsbildande massmedia, radio och TV.²¹

Logiken kräver, slog Holmberg fast – inte utan klara mediehistoriska poänger – att radio och television också skulle kunna sändas i konkurrerande och oberoende former.

Närläser man de anföranden som är starka i tema 77 så handlar de övergripande flertalet om etermedia, framför allt television. Men samtidigt är ett av de ord som förekommer starkt i tema 77 *presstödsområde*. Mot slutet av 1950-talet hade flera lokala dagstidningar i Sverige lagts ned, vilket föranledde politiska diskussioner kring statliga stödformer. Det var framför allt nedläggningen av göteborgska *Ny Tid* hösten 1963 som utgjorde en sorts vattendelare i rapporteringen av tidningsdöden – som den upprepade företeelsen då började att kallas. Men tankar på offentligt ekonomiskt stöd till tidningar framstod länge som en ogörlig reform, för hur skulle pressen kunna utföra sin samhällskritiska uppgift om den var beroende av statliga subventioner? Dåtidens dagspress var politiserad på ett annat sätt än idag, med både traditionell arbetarpress (som den socialdemokratiska tidningskoncernen A-pressen) och dagstidningar med en tydlig politisk profil, exempelvis med koppling till Bondeförbundet (sedermera samlade inom koncernen Centertidningar). För att öka kunskapen om den situation som uppstått tillsatte regeringen 1963 därför en större pressutredning, vilken resulterade i betänkandet *Dagstidningarnas ekonomiska villkor* (SOU 1965:22).²²

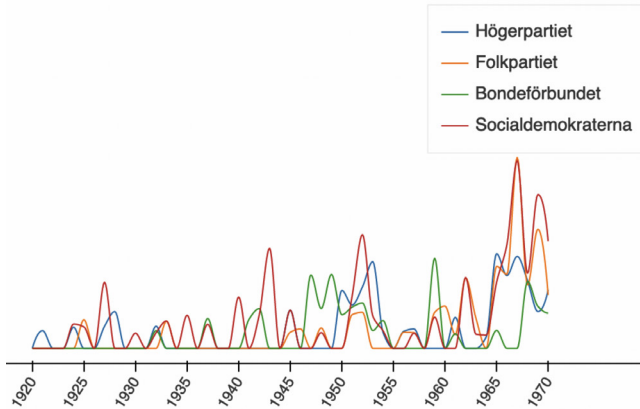


Fig. 9. Förekomst av tema 264 (om dagspress) i 500-modellen av riksdagsdatan i anföranden av riksdagsledamöter för fyra politiska partier under perioden 1920–70. I temat är termen presstöd det tredje vanligast förekommande och grafen visar med all tydlighet att den typen av stödfrågor blev alltmer centrala i svensk mediepolitik under det sena 1960-talet.

Eftersom de etermediala anförandena är så många i tema 77 är det svårt att få korn på den riksdagspolitiska diskussionen kring presstöd. För att få en mer nyanserad och detaljerad bild behöver man därför växla till en temamodell med fler teman. I 500-modellen av riksdagsdatan finns det mycket riktigt ett specifikt presstema (nummer 264), där de fem mest frekventa orden är tidning, press, presstöd, annons och stöd – ett tema som just handlar om hur staten skulle stödja tidningsbranschens ekonomiska förutsättningar. Som framgår av grafen ovan var det en diskussion som tog fart på allvar under andra halvan av 1960-talet. I första kammaren våren 1969 menade socialdemokraten Yngve Möller att en mångsidig dagspress var ”av väsentligt värde för folkstyret. Denna fråga är aktuell inte bara i Sverige utan i många länder i Västeuropa [och] åtgärder från statens sida för att stödja dagspressen diskuteras eller har beslutats i en rad länder . . . Det är verkligen en oroande utveckling vi bevittnar i Sverige . . . Vi får räkna med att tidningsdöden kommer att fortsätta”. Från borgerligt håll var man mer skeptisk, folkpartisten Ingrid Segerstedt Wiberg beklagade sig exempelvis i samma debatt över ”den tvivelaktiga förhoppningen att kunna skapa goda ekonomiska förutsättningar för pressen och jämnare konkurrensförhållanden inför ett system, som under vissa förhållanden kan leda till att pressen blir beroende av statsmakten”.²³

Om det i 100-modellen av riksdagsdatan fanns två teman som behandlade medier så återfinns det i 500-modellen många, många fler. Styrkan med temamodellering som metod är just att det går att analysera hur ett tema fördelar och förgrenar sig när riksdagsdatan modelleras i flera hundra

teman. 500-modellen har inte bara ett tydligt tema om presstödet, där finns också ett om etermedia (nummer 322) och ett om film (nummer 74). I nätverket nedan framgår att dessa tre teman innehåller ord som är långt mer mediespecifika än i de medieteman som finns i 100-modellen. De delar en hel del begrepp, och ett fåtal ord i mitten av nätverket – kulturminister, radio, allmänhet – är gemensamma för alla tre teman. I 500-modellen av riksdagsdatan är det också möjligt att mer i detalj göra mediehistoriska nedslag. Via begreppet massmedier finns exempelvis kopplingar till tema 465 (om information), vilket i sin tur innehåller en sekretessdiskurs som förgrenar sig i tema 417 (om data, datasystem och ADB-teknik), vilken i sin tur har kopplingar till ett mer samtida digitaliseringsstema (nummer 13) om bredband, digital teknik och internet. Men sådana teman ökar förstås i styrka först efter 1970.

Samtidigt kan man också få korn på teman i vilka somliga medier förgrenar sig in i andra sammanhang med en helt annan politisk betydelse. I 500-modellen har exempelvis filmtema 74 en hel del kopplingar till tema 321 med en betydligt mer sinister innebörd: barn, prostitution, våld, film, människohandel och barnpornografi är de sex mest frekvent förekommande orden i detta tema. Om film i vissa politiska anföranden debatterats med fokus på statligt stöd till kvalitativt, förbättrad filmproduktion, finns det med andra ord en diskussionskontext där filmmediets negativa inverkan utgjorde ett återkommande tema i riksdagens kammare. Filmcensur och (senare) videovåld är två kodord i sammanhanget. Det var förstas diskussioner som präglades av politisk oro och allvar – men även en mediemoralism som idag framstår som tämligen nattständer. I efterhand är det nu bara alltför enkelt att le åt de lätt förlegade replikväxlingarna i riksdagen, exempelvis från den filmdebatt i andra kammaren hösten 1967 som handlade om ecklesiastikminister Olof Palmes medverkan i Vilgot Sjömans, *Jag är nyfiken – gul*, en film där Palme intervjuades om det svenska klassamhället. Fler anföranden från denna debatt förekommer starkt i tema 321. Filmen är naturligtvis mest beryktad för sin frispråkiga sexliberalism och nakenscener. Med viss munterhet kan man härvidlag läsa folkpartisten Yngve Hamrins upprörda genmäle om att ”en man i herr Palmes ställning över huvud taget förekommer i denna film [med sitt] etiskt undermåliga innehåll, på grund av sina inslag av ren exhibitionism [samt] öppen och cyniskt driven sexualkult”. Jag finner det utomordentligt betänkligt, avslutade Hamrin, ”att Sveriges kultur- och kyrkominister över huvud taget agerar i en film av denna beskaffenhet” – varpå Palme replikerade: ”Vad är moral? Vad är anstötligt? Vad är hymleri? ... Jag har helt enkelt deltagit i en intervju. Det står jag som sagt för ... Jag vill för min del avsluta denna debatt, därför att jag till herr Hamrins förtyvlan måste i väg till en viktig förrättning i ett kyrkoärende.”²⁴

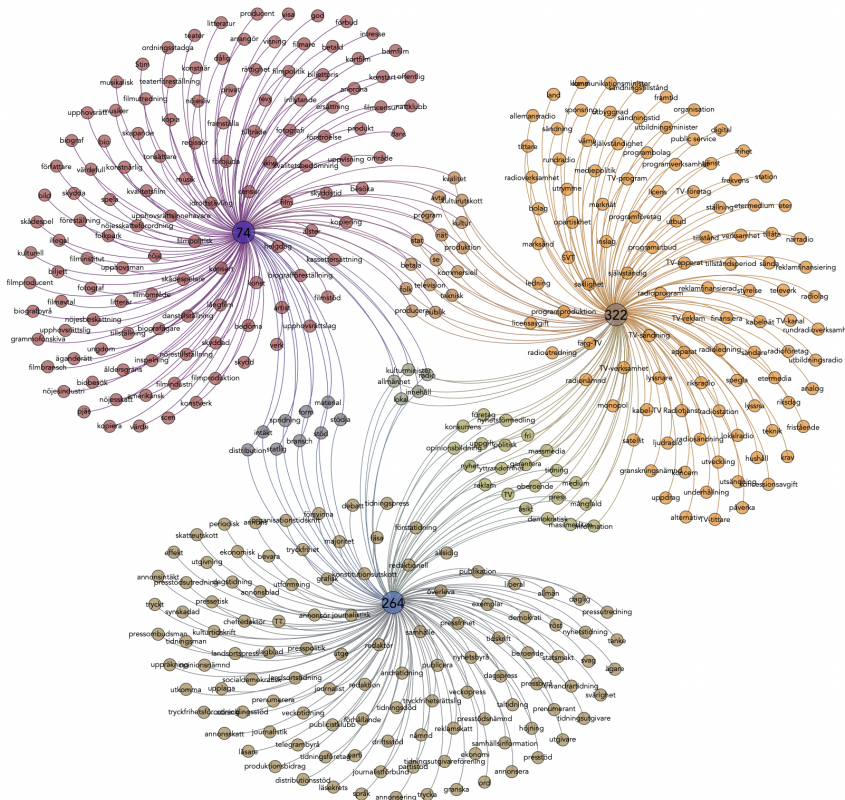


Fig. 10. Nätverk i Gephi med algoritmen Force Atlas 2 Expansion med cirka 400 ord-noder från 500-modellen av riksdagsdatan med de mest frekvent förekommande termerna i tema 74 (om film), tema 264 (om dagspress) och tema 322 (om etermedia).

Avslutning

Riksdagens mediehistoria återstår ännu att skriva. Våra politikers föreställningar, idéer och uppfattningar om medier har varit många. De har förstås också växlat över tid. Vid sidan av de politiska mediediskussioner som förts i första och andra kammaren sedan 1866 inbegriper riksdagens mediehistoria givetvis också de många medietekniker som använts på plats – från protokollförtecknarnas stålpennor och skrivmaskiner till senare perioders audiovisuella inspelningstekniker. Som det här kapitlet antytt kan en digital metod som temamodellering vara till hjälp för att skapa både överblick i en mycket omfattande empiri och spåra upp de många diskussioner som förts om medier. Det rör sig om tiotusentals anföranden bara under perioden 1920 till 1970. Min ambition har varit att visa hur

temamodellering kan ge vissa nya mediehistoriska insikter – exempelvis kring hur somliga politiska partier intresserat sig mer för vissa medier och mediefrågor än andra, liksom hur de förändrat sig över tid – men också att ge mer generella inblickar i hur en metodologiskt färgad analys av ett storskaligt dataset kan te sig. Givet alla diskussioner och debatter i Sveriges Riksdag så säger det sig självt att otaliga andra frågor än de som handlar om medier kan studeras ungefär på samma sätt, såväl historiskt som idéhistoriskt. Och för de frågor som politiker av tradition ägnar mest tid åt – arbete, skatter, sjukvård, utbildning, försvar, miljö – så finns det en ännu rikare empiri av riksdagsanföranden för temamodelleringen att utgå från. I artikeln har jag tämligen flitigt citerat från dessa anföranden; personligen ser jag inte att det finns någon motsättning mellan kvantitativ distans- eller fjärrläsning och kvalitativ närläsning. Tvärtom berikar de ofta varandra. Men jag ska heller inte hymla med att när man som mediehistoriker väl börjat att arbeta med en omfattande empiri så blir man också varse att det empiriska urvalet som i regel ligger till grund för kvalitativa närläsningar ofta är begränsat. Mitt syfte är inte att vara polemisk, eller att upprätta någon slags motsättning mellan en äldre historievetenskaplig metodik och en av mer modernt digitalt snitt. Även om jag intresserar mig för digitala sätt att undersöka historisk empiri så bedriver jag också forskning i mer traditionell mediehistorisk bemärkelse. Likväl: 11 000 riksdagsprotokoll som innehåller 400 miljoner tecken utgör empiri av ett annat slag, låt gå att det är mjukvara som tar sig an textmaterialet. Den empiriska skalan är annorlunda – fjärrläsningens kunskapsformer likaså: ”Distant reading ... *is a condition of knowledge*”.²⁵

Noter

1. August Palm: ”Hvad vil Sosial-Demokraterna” 6/11 1881, <https://www.arbark.se/dokument/temp/palm-hvad-vil-sosial-demokraterna.pdf> (senast kontrollerad 30/4 2022).

2. Språkbanken på Göteborgs universitet erbjuder via webben analysmöjligheter av stora textkorpora genom applikationen Sparv eller konkordansverktyget Korp. Fokus är främst språketenskapligt; nyckelord, ordbild, ordattribut, ordformer etcetera. Nackdelen med Korp är att flera av de historiska textkorpora som används har bristfällig OCR. Samtidigt är det enkelt att få fram den kontext (i originaltexter) som ett ord förekommer inom. Språkbankens verksamhet är kopplad till den europeiska forskningsinfrastrukturen för språkteknologi, CLARIN (Common Language Resources and Technology Infrastructure). Genom den nationella noden Swe-Clarins kan forskare inom humaniora och samhällsvetenskap få tillgång till språkteknologiska verktyg och resurser i syfte att främja forskning med digitala verktyg (bortom den strikt språketenskapliga). Se exempelvis Swe-Clarins digitala handbok, <https://sweclarin.se/swe/handbok> (senast kontrollerad 30/4 2022).

3. För en introduktion till temamodellering, se Pelle Snickars: ”Från chiffer till

klartext? Temamodellering av statliga offentliga utredningar 1945–1989” i *Scandia* 87:3 (2022). Flera svenska historiker har skrivit om temamodellering, se exempelvis André Baltz & Fredrik Norén: ”Tematisk innehållsanalys med temamodellering” i Johan Jarlbrink & Fredrik Norén (red.): *Digitala metoder i humaniora och samhällsvetenskap* (Lund, 2021), liksom Johan Jarlbrink: ”Från masugn till mikrovågsugn. Historiografisk topic-modellering av årsboken *Daedalus* (1931–2017)” i Jenny Attemark-Gillgren & Pelle Snickars (red.): *Digitala modeller. Teknikhistoria och digitaliseringens specificitet* (Lund, 2019).

4. Hannu Salmi: *What is digital history?* (Cambridge, 2021), 7.

5. Se, <https://opendiscourse.de/>. För en vidare diskussion, se Thora Hagen: ”Inhaltliche Analyse von Plenarprotokollen des deutschen Bundestages durch Topic Modeling” 1/8 2018, <https://dhd-blog.org/?p=10307>. The Hansard Viewer – utvecklad av bland andra Jo Guldi – återfinns på <https://shinyviz.smu.edu/shiny/public/hansard-shiny/>. Hansard at Huddersfield på <https://hansard.hud.ac.uk/site/index.php>. Textkorpuset för hela Hansard finns öppet tillgängligt på, <https://www.english-corpora.org/hansard/> (alla senast kontrollerade 30/4 2022). För en diskussion om temamodellering av Hansard, se Jo Guldi: ”Parliament’s debates about infrastructure. An exercise in using dynamic topic models to synthesize historical change” i *Technology and Culture* 60:1 (2019). Om temamodellering av mer samtida parlamentsdata, se Derek Greene & James P. Cross: ”Exploring the political agenda of the European parliament using a dynamic topic modeling approach” i *Political Analyses* 25:77–94 (2017).

6. Florian Richter et. al.: ”Open discourse” 25/5 2021, Harvard Dataverse, <https://doi.org/10.7910/DVN/FIKIBO> (senast kontrollerad 30/4 2022).

7. För mer information, se forskningsprojektets webbplats, <https://www.westac.se>, liksom sajten på Github, <https://github.com/topics/westac> (senast kontrollerade 30/4 2022).

8. Riksdagens protokoll andra kammaren 1937:31, https://weburn.kb.se/riks/tv%C3%A5kammarriksdagen/pdf/web/1937/web_prot_1937__ak__31/prot_1937__ak__31.pdf (senast kontrollerad 30/4 2022).

9. I projektet Välfärdsstaten analyserad är flera maskininläringsexperter och statistiker från Uppsala universitet involverade. För dem är datakurering en kontinuerlig process, ungefär på samma sätt som utveckling av mjukvara. På Github – en plattform för lagring av data och kod – håller en versionshanterare koll på uppdateringarna av vårt dataset av riksdagsmaterialet. Den version som jag arbetat med i denna artikel har exempelvis nummer 0.4.1. Det innebär att mina resultat på ett sätt är tentativa – och kan komma att förändras när dataunderlaget blir bättre. Men dessa förändringar kommer med all sannolikhet endast att förändra mina resultat på marginalen.

10. Tidsdimensionen i temamodellering är naturligtvis av intresse för både medie- och idéhistorikern. I denna artikel fokuserar jag på åren 1920–1970, men de temamodeller jag undersöker baserar sig på data över hela tidsperioden från 1920 till idag. Det innebär att alldeles oavsett om jag filtrerar bort vissa årtal i efterhand så kan ord och begrepp dyka upp från (i mitt fall) perioden 1971 till idag – framför allt om dessa är starkt förekommande i det undersökta temat.

11. Yngve Holmberg är citerad från Riksdagens protokoll första kammaren 1964:32 och 1966:14. Alla riksdagsprotokoll i Jupyter Lab-miljön som citeras i det följande återfinns på Github, <https://github.com/welfare-state-analytics/riksdagen-corpus> (senast kontrollerad 30/4 2022).

12. Beth Hennings är citerad från Riksdagens protokoll andra kammaren 1937:31. Hennings var en pionjär inte bara som kvinnlig riksdagsledamot som uppskattade film, hon var även den första kvinnan att disputerat i historia vid Stockholms högskola med avhandlingen, *Gustav III som kronprins* (Uppsala, 1935).

13. Statistiska centralbyrån, ”Historisk statistik över valåren 1910–2018”, <https://www.scb.se/hitta-statistik/statistik-efter-amne/demokrati/allmanna-val/allmanna-val-valresultat/pong/tabell-och-diagram/historisk-valstatistik/historisk-statistik-over-valaren-1910-2018>, samt Wikipedia, ”Mandatfördelning i valet till andra kammaren, valen 1911–1968”, https://sv.wikipedia.org/wiki/Sveriges_riksdag (senast kontrollerad 30/4 2022).

14. ”Från och med 1940-talet började de andra partiernas kvinnorepresentation mer likna socialdemokraternas. Socialdemokraterna hade den högsta andelen kvinnor följt av högern. Folkpartiet låg på en mellannivå, medan bondeförbundet utgjorde ett undantag eftersom partiet hade en låg andel kvinnor tvåkammartiden igenom”. Camilla Norrbin: *Från isolering till integrering. En kollektivbiografisk studie över de kvinnliga riksdagsledamöterna under tvåkammarriksdagens tid 1922–1970* (Umeå, 2004), 14, 79.

15. Namnen på svenska politiska partier har som bekant växlat över tid, men under perioden 1920–1970 fanns det både socialdemokrater och folkpartister. Bondeförbundet bytte 1957 namn till Centerpartiet, och Högerpartiet (egentligen Högerens riksdagsgrupp) gjorde detsamma 1969 och blev Moderata samlingspartiet. För enkelhetens skull använder jag i denna artikel Bondeförbundet och Högerpartiet som namn – för att inte tala om de partisprängningar och partisplittringar som ständigt präglat Sveriges kommunistiska parti, som i det följande helt sonika kallas för Kommunisterna.

16. Leif Furhammar: *Filmen i Sverige* (Höganäs, 1991), 205.

17. Ingrid Andrén är citerad från Riksdagens protokoll andra kammaren 1955:15.

18. Anders Paulsen och Värner Rydén är bägge citerade från Riksdagens protokoll andra kammaren 1925:36.

19. Edvin Leffler är citerad från Riksdagens protokoll andra kammaren 1939:12.

20. Bo Turesson är citerad från Riksdagens protokoll andra kammaren 1963:10.

21. Yngve Holmberg är citerad från Riksdagens protokoll första kammaren 1966:14.

22. För en vidare mediehistorisk diskussion, se Mats Hyvönen, Pelle Snickars & Per Vesterlund: ”Mediestudiets formering: en introduktion” i Mats Hyvönen, Pelle Snickars & Per Vesterlund (red.): *Massmedieproblem. Mediestudiets formering* (Lund, 2015).

23. Yngve Möller och Ingrid Segerstedt Wiberg är citerade från Riksdagens protokoll första kammaren 1969:25.

24. Yngve Hamrin och Olof Palme är bägge citerade från Riksdagens protokoll andra kammaren 1967:62.

25. Franco Moretti: *Distant reading* (London, 2013), 48.