

# **Metodología para el análisis de la opinión pública y la comunicación política en Internet**

Miguel Castillejo Sierra

**Tutor:** Joan Corbella

**Curs:** 2009/10

**Treballs de recerca dels programes de postgrau del  
Departament de Comunicació**

**Departament de Comunicació**

**Universitat Pompeu Fabra**

**Abstract:** Para muchos tipos de investigación en ciencias sociales existe la necesidad de obtener información sobre la opinión pública. Internet puede ayudar a satisfacer esta necesidad porque es una fuente de información fácilmente accesible. Este trabajo pretende establecer una metodología de análisis completa y coherente y para ello desarrolla un análisis de las herramientas, métodos y teorías utilizadas en el análisis de la opinión pública en Internet. Ya que la idea de opinión pública en que me baso se fundamenta en la organización temática de la misma, analizo las teorías clásicas de la *agenda-setting* y de la tematización. Sin embargo, la característica complejidad de la sociedad actual y la fragmentación de la agenda pública hacen necesario estudiar el nuevo sistema comunicativo desde perspectivas menos deterministas. La teoría de redes complejas permite analizar los fenómenos de emergencia temática y autoorganización desde una perspectiva más amplia que tiene en cuenta los numerosos factores que determinan el funcionamiento del sistema comunicativo actual.

**Keywords:** Public Opinion, Political Communication, Issue Emergence, Agenda Setting, Social Network Analysis, Complex Systems, Self Organization, Semantic Web, Issue Crawling, Infography.

## Indice

1.- Presentación.....	3
2.- Estado de la cuestión.....	10
2.1.- Tematización vs. <i>Agenda-Setting</i> .....	10
2.2.- La teoría de la tematización aplicada al sistema comunicativo de Internet.....	17
Web semántica.....	17
Lenguajes controlados.....	18
Redes temáticas.....	20
2.3.- Sistemas complejos, autoorganización y emergencia.....	23
Propiedades de las redes complejas.....	25
Autoorganización y emergencia.....	28
Sistemas adaptativos y no adaptativos.....	31
2.4.- Visibilidad y otros procesos de selección de la información en Internet.....	33
Aproximaciones al concepto de visibilidad.....	34
Procesos de selección de la información.....	39
Procesos de selección vertical.....	40
Procesos de selección horizontal.....	43

3.- Objetivos y preguntas de investigación.....	47
Objetivo teórico general.....	49
Objetivo de investigación aplicada.....	51
4.- Metodología y técnicas de investigación.....	53
Propuesta metodológica para la identificación de redes temáticas...	53
Propuesta metodológica para el análisis de subtemas y/o marcos de referencia en una red temática.....	57
Propuesta metodológica para la identificación de tipos de actores en una red temática.....	61
Propuesta metodológica para la representación gráfica de redes temáticas.....	64
Propuesta metodológica para la evaluación del nivel de visibilidad de una red temática.....	72
5.- Estructura y principales conceptos de la investigación.....	73
Estructura y contenidos.....	73
Conceptos principales.....	75
6.- Bibliografía.....	76

## 1. Presentación

Para muchos tipos de investigación en ciencias sociales y estudios de mercado existe la necesidad de obtener información sobre la opinión pública o la reacción pública ante determinados temas. Internet<sup>1</sup> puede ayudar a satisfacer esta necesidad porque es una nueva fuente natural de información fácilmente accesible sobre las opiniones o actividades de un espectro muy amplio de la sociedad (Burnett y Marshall, 2002; Hine, 2000). Las nuevas tecnologías de la información y la comunicación han permitido desarrollar un nuevo tipo de comunicación mediada por ordenador<sup>2</sup> que está provocando una transformación en el sistema comunicativo y en los procesos de formación de la opinión pública. Los medios cibernéticos cuentan con unos atributos que empujan hacia el consumo activo en los procesos de comunicación, permiten una mayor facilidad de almacenamiento de datos que pueden ser seleccionados por los usuarios en función de sus necesidades, facilitan la existencia de mecanismos de intercambio, y ofrecen la posibilidad de integrar a un gran número de usuarios en la expresión de experiencias y opiniones sobre un tema.

Debido a estas características que provocan la aparición de un público activo muy diferente al tradicional, el pluralismo aumenta cuantitativa y cualitativamente (López-García, 2006: 231). La presunción de que existe una exposición generalizada a un contenido político relativamente uniforme, que ha fundamentado las tres principales teorías sobre los efectos de la comunicación política, *agenda-setting*, espiral del silencio

---

1 Con Internet no me refiero al conjunto descentralizado de redes de comunicación interconectadas sino al sistema de documentos hipertexto o hipermedios enlazados y accesibles a través de Internet que técnicamente se conoce como Red Global Mundial, *World Wide Web*, o WWW por sus siglas en inglés.

2 Traducción de *Computer Mediated Communication*

y teoría del cultivo, no puede seguir considerándose una certeza (Blumler, 2001: 206). Este trabajo, que es producto de un proceso de análisis de las herramientas, métodos y tendencias que se están utilizando en el ámbito del análisis de la opinión pública en Internet, tiene la intención de contribuir a establecer métodos para analizar los fenómenos de opinión en este nuevo canal de comunicación.

Los cambios señalados no son los únicos determinantes en la transformación del sistema comunicativo actual y los procesos de formación de la opinión pública. El sociólogo alemán Niklas Luhmann planteó ya a finales de los años 60 en su libro *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität* (1968) la aparición de una sociedad compleja provocada por la aceleración del desarrollo industrial y tecnológico de la sociedad y por el incremento de las organizaciones orientadas hacia la administración social o hacia la gestión económica, política, cultural o social. Esta rápida transformación conllevaba la creación de una sociedad compleja caracterizada por una diferenciación funcional y una especificación en subsistemas o sistemas parciales. En este modelo de sociedad compleja, la tematización aparece como principio organizador del sistema comunicativo, el sistema político y la opinión pública. Donald Shaw, teórico de la *agenda-setting* y colaborador de Maxwell McCombs, plantea una pregunta a la que se pretende dar respuesta en este trabajo. Según dice, en esta nueva era los ciudadanos están llegando a las noticias y a la información de los asuntos públicos a través de muy diferentes y variadas páginas web. ¿Es posible que los investigadores sigan llevando a cabo estudios para medir los efectos de los medios de comunicación cuando existe una cantidad de fuentes por las que el ciudadano es influido tan grande que hace la cuantificación y el análisis casi inabarcable? (Shaw, Stevenson y Hamm, 2001: 7-8). Según Carlo Marletti, estudioso de la teoría de la tematización y continuador

de la obra de Luhman en el ámbito del periodismo, en una sociedad de estas características las corrientes de opinión tienen un componente de indeterminación que a menudo las hace aparecer como los fenómenos climáticos (Berrio, 2001: 198).

Siguiendo esta línea de posible fractura del viejo paradigma se han llevado a cabo investigaciones para evaluar si la función de la *agenda-setting* pervive tras el desarrollo de los medios de comunicación en Internet. Los estudios de Althaus y Tewksbury (2002: 196) en que comparaban la versión impresa y la versión online de diferentes periódicos nacionales reveló que los lectores de noticias de la versión online sufrían menor influencia que los lectores de la versión impresa en el reconocimiento y memorización de hechos relacionados con temas de carácter internacional, nacional y político. Con diferente objetivo pero en la misma línea de intentar redefinir el papel de la teoría de la *agenda-setting* en el panorama actual de la comunicación se desarrolló una investigación que extendía el modelo de la *inter-media agenda-setting* incluyendo medios de comunicación en Internet. Los resultados de este estudio revelaron que el contenido de los sitios web electorales afectaba la formación de la agenda de los medios tradicionales de comunicación (Ku, Kaid y Pfau, 2003: 540).

La abundancia de canales posibilitada por las Nuevas Tecnología de la Información y la Comunicación, un diseño de Internet pensado para la descentralización donde la posibilidad de crear y publicar contenidos está al alcance de cualquiera y la comodidad y portabilidad que ofrecen los cada vez más pequeños aparatos electrónicos, están provocando una desmasificación de los medios que inevitablemente provoca la aparición de un número indeterminado de públicos. Por el papel y la influencia de los emisores en este nuevo escenario parece claro que las hegemónicas posiciones de influencia y control del sistema político y los medios de comunicación de masas pasan

por una etapa de transformación e incipiente reforma. La crisis actual de los medios tradicionales lo refleja quizá hasta el punto de poder afirmar que en un futuro próximo el tema de discusión puede pasar de ser ¿qué hacen las grandes compañías de medios con la gente? para convertirse en ¿qué hace la gente con los medios? (Chaffee y Metzger, 2001: 370). Pero ¿qué entienden Chaffee y Metzger por “gente” cuando hablan de lo que hará con los medios?. ¿Estamos hablando de la suma de los individuos que intervienen en los procesos de comunicación? ¿entienden “gente” como un ente con personalidad propia al que podemos caracterizar o es más bien la personificación de lo público como ente representante del grupo en el que todos estamos incluidos?. Según defiende Allport en su artículo “Towards a science of public opinion” (1937) todas estas posibilidades no son sino ficciones. El término opinión pública tanto coloquialmente como en la literatura o incluso en la bibliografía científica, está cargado de analogías, personificaciones y otros recursos discursivos que dificultan un pensamiento claro.

Entre las teorías que Allport critica como ficciones del entendimiento de la opinión pública existe una que él no se atreve a negar. La denomina teoría del grupo producto o emergente o, en una versión alternativa con pequeñas diferencias, teoría eulogística<sup>3</sup>. La teoría del grupo producto o emergente no se refiere a personificaciones de la opinión pública sino a ella como resultado. Según esto, la opinión pública es un nuevo producto que surge tanto de la opinión media o mayoritaria como de la opinión de cualquier individuo particular. Es un punto de encuentro del interés social en su organización hacia la acción.

La teoría eulogística entiende la opinión pública como el resultado emergente de un grupo de discusión. En ocasiones este producto no sólo es diferente de la opinión

---

3 Traducción de *Eulogistic Theory*



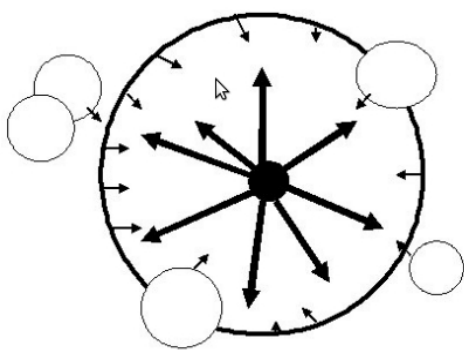
individual sino que es superior. En el proceso de interacción los errores son subsanados de tal manera que la opinión más brillante, mejorada por la discusión, prevalecerá sobre las demás. Allport no niega la existencia de este nuevo producto pero señala la imposibilidad de identificarlo, analizarlo o evaluarlo. Señala además que este proceso puede estar influido por condicionantes emocionales, estereotipos, símbolos o líderes políticos. Vemos en esta crítica de Allport ciertas similitudes con la discusión entre la teoría de democracia deliberativa de Habermas y la teoría de la espiral del silencio de Noelle-Neumann. Todas estas teorías se basan en la idea de que los procesos de formación de la opinión pública se producen en un sistema lineal, donde las relaciones de influencia se dan de forma directa o cuasi directa a través de líderes de opinión. En cualquier caso, se trata de un sistema comunicativo ciertamente alejado de la complejidad en que se basa la teoría de Luhmann.

Técnicamente, Internet parece ser un espacio comunicativo no jerárquico que permite a todo el mundo presentar y recibir información y opiniones sin depender de la selección y descripción que ejercían los medios de masas tradicionales. El espacio público que desde la aparición de la sociedad postindustrial se había concentrado en los medios de comunicación de masas se transforma con la aparición de Internet en un sistema donde cualquiera puede crear y difundir sus informaciones.

Según Sampedro (2000: 20) existen dos tipos de formas de opinión pública desarrolladas de distinta manera. La opinión pública agregada es un resultado, la suma de juicios individuales a través del voto y los sondeos. En cambio, la opinión pública discursiva es el proceso de un agente colectivo, el pueblo, que conversa en ambientes formales o informales procesando experiencias propias, conocimientos e informaciones. Así, la opinión pública agregada es un resultado donde el público que importa es la

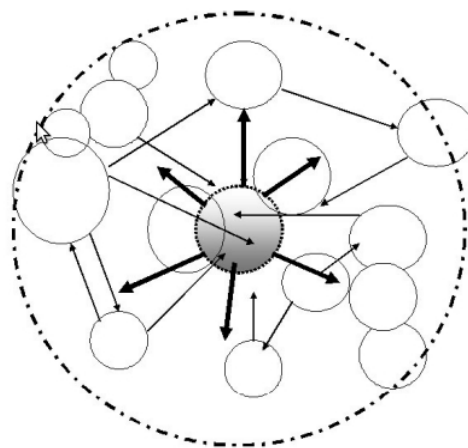
mayoría que suma voluntades ya determinadas, mientras que la opinión pública discursiva es un proceso, donde el público es un colectivo de voluntades individuales que deliberan entre sí y se condicionan mutuamente. En consecuencia, la opinión pública discursiva siempre es más amplia que la agregada, que prima las mayorías y los números sobre los argumentos.

*Figura 1: Esfera pública en la sociedad de masas.*



*Extraído de López (2006: 236).*

*Figura 2: Esfera pública en la sociedad de masas.*



*Extraído de López (2000: 240).*

A pesar de este cambio de paradigma hacia una esfera pública en red, no podemos afirmar que hayamos llegado al ideal de democracia deliberativa de Habermas, en el que todos tenemos la misma capacidad de crear y difundir opiniones. Sería ingenuo pensar que no existen procesos de selección. La cantidad de sitios web existentes incluso para temas muy concretos es tan grande que sería imposible para un usuario conseguir la información deseada sin un mapa que le guíe. La visibilidad es el principal factor discriminador en la formación de ese mapa que necesita el usuario. La teoría de las redes sociales ya planteó la importancia de la posición de centralidad en una red como

factor determinante para su capacidad de influir en el resto de los componentes de la red. Sin embargo, con la irrupción de Internet en este campo, el concepto de centralidad ha dado paso al concepto actual de visibilidad y la vieja teoría de redes sociales basada en las teorías sistémicas de la sociología ha visto como los últimos avances en la ciencia de redes complejas aportados por matemáticos, físicos, informáticos y sociólogos provocaban una revolución en el concepto de emergencia. Este concepto, aplicado a la aparición de los temas en la agenda pública, resulta central tanto en la teoría de la *agenda-setting* y la teoría de la tematización como en el trabajo que presento. Para investigar en este ámbito de la emergencia temática y los procesos de formación de la agenda pública propongo una serie de métodos de análisis que pretendo aplicar al análisis de la comunicación política de la campaña electoral de las próximas elecciones a la Generalitat de Catalunya en otoño de 2010.

## **2. Estado de la cuestión**

El estado de la cuestión que se presenta a continuación está organizado en cuatro apartados. En el primero se realiza una revisión de las teorías clásicas de los efectos de la comunicación con especial atención en la teoría de la *agenda-setting* y la teoría de la tematización. En el segundo se exponen brevemente las aplicaciones que ya se están realizando de la teoría de la tematización en el sistema comunicativo de Internet. En el tercero se estudian los modelos alternativos a las teorías clásicas de los efectos que han desarrollado sus avances más importantes en los últimos veinte años. Entre estas teorías prestamos especial atención a la teoría de sistemas complejos –y su versión más reciente denominada ciencia de las redes– y a los procesos de autoorganización y emergencia que plantean y estudian estas teorías. En el cuarto y último se estudian los factores que determinan la capacidad de influencia de una red, con aplicación a los procesos de comunicación en Internet, donde el principal factor que determina los procesos de selección de la información es la visibilidad.

### **2.1. Tematización vs. *Agenda-Setting***

Las teorías de la tematización y la *agenda-setting* son cercanas en su objeto de estudio y autores como Neumann las han considerado sinónimas (1979: 439). Sin embargo, a nuestro juicio existen diferencias importantes entre estas dos teorías de los efectos de la comunicación. Para poder entenderlas retomamos una idea que resulta fundamental en

ambas. En 1922 Lippmann, en su libro *Public Opinion* situaba la opinión pública lejos de las cuestiones éticas y políticas que caracterizan la teoría normativa, y al mismo tiempo lejos de las medidas empíricas de las opiniones individuales a que aspiraban las teorías positivistas. Su teoría centraba el tema fundamental de la opinión pública en los estereotipos con carga emocional que dominan las opiniones de las personas (Berrio, en prensa). Para Lippmann:

El entorno real resulta en conjunto excesivamente grande, complejo y fugaz para que podamos conocerlo de forma directa. No estamos capacitados para manejar tanta sutileza y variedad, ni para considerar un número tan elevado de permutaciones y combinaciones. En consecuencia, por mucho que debamos actuar en él, nos vemos en la necesidad de reconstruirlo en modelos más asequibles para poder manejarlo. Podría decirse que estos modelos son como mapas que nos guían a través del mundo (Lippmann, 1922: 33).

Ambas teorías coinciden en la necesidad que tiene nuestra mente de simplificar la realidad para entenderla y coinciden también en la función de puente entre la realidad y las imágenes de nuestra mente, que Lippmann otorga a los medios de comunicación de masas. Sin embargo, la tematización resulta muy diferente del estudio del establecimiento de la agenda temática ya que su fundamentación teórica de la sociología sistémica, su contextualización del proceso de tematización en el sistema político y en el sistema comunicativo y su vinculación con las transformaciones tecnológicas y políticas que estamos experimentando actualmente son desconocidas por la Agenda Setting (Saperas 1987: 92). El concepto de tematización apareció por primera vez en el texto de Niklas Luhmann “*Öffentliche Meinung*” (Luhmann, 1974). Para este autor, la tematización se enmarca en el análisis de la opinión pública y de la comunicación, y sobre esta base define la opinión pública como:

La opinión pública debe ser concebida como estructura temática de la comunicación pública [...]. La opinión pública no consiste en la generalización del contenido de las opiniones individuales mediante fórmulas generales, aceptables por cualquiera que tenga uso de razón, sino en la adaptación de la estructura de los temas del proceso de comunicación política a las necesidades decisionales de la sociedad y de su sistema político (Luhmann, 1978: 87).

Luhmann inscribe sus planteamientos sobre la tematización en el ámbito de la sociología y la relaciona de forma implícita con la teoría de la comunicación. El concepto será desarrollado y adaptado a la investigación comunicativa por un grupo de autores italianos entre los que se encuentran Agostini, Rosetti, Marletti y Rossi. En España, los estudios sobre la función de la tematización de Luhmann y sus posteriores reformulaciones han sido realizados por Saperas, Berrio y Badía. De este último extraemos los tres presupuestos que fundamentan la hipótesis sobre la función de la tematización, que nos servirán para diferenciar definitivamente la agenda setting de la tematización:

- a) El concepto de opinión pública entendida ésta como una estructura selectiva de temas de la comunicación política. Luhmann asigna a la opinión pública la función de mecanismo guía del sistema político que reduce la elevada contingencia de aquello que es políticamente y jurídicamente posible por la vía estratégica de adaptar la estructura temática de los procesos de comunicación pública a las variables exigencias decisorias de la sociedad y de su sistema político.
- b) De la transformación estructural del sistema político, se ha derivado la creciente centralidad de los temas y de las controversias políticas (*issues*) por lo que concierne tanto a la orientación de los procesos decisorios, como al funcionamiento de las mismas lógicas de consenso de las sociedades contemporáneas. Por consiguiente los

temas se han convertido en recursos estratégicos esenciales de la política en las sociedades complejas, lo cual ha de ser necesariamente planteado en el marco de referencia de las interacciones entre el sistema político y el sistema comunicativo.

c) La función de tematización remite finalmente a una dinámica evolutiva de diferenciación funcional de los distintos modelos profesionales de periodismo y de los distintos roles informativos ejercidos por los medios de comunicación (Badía, 1992:172).

Como vemos, la tematización, en cuanto estructura temática, se enmarca en la comunicación política y, al mismo tiempo, la opinión pública se sitúa en relación con el sistema político y a sus necesidades de decisión estratégica. Luhmann asigna a la opinión pública la función de mecanismo guía del sistema político que reduce la elevada contingencia de aquello que es políticamente y jurídicamente posible por la vía estratégica de adaptar la estructura temática de los procesos de comunicación pública a las exigencias decisorias de la sociedad y de su sistema político. De la transformación estructural del sistema político se ha derivado la creciente centralidad de los temas y de las controversias políticas (*issues*), por lo que concierne tanto a la orientación de los procesos decisorios como al funcionamiento de las mismas lógicas de consenso de las sociedades contemporáneas. Por consiguiente, los temas se han convertido en recursos estratégicos esenciales de la política en las sociedades complejas, lo cual tiene que plantearse en el marco de las interacciones entre el sistema político y el sistema comunicativo. Los temas obtienen su centralidad a partir de la creciente complejidad del sistema social y de las relaciones entre el sistema político, el sistema comunicativo y la opinión pública. Con algunas diferencias de fundamentación, Marletti desarrolla una diferenciación entre lo que define como tema, controversia e *issue* o tema político. Según esta diferenciación, el tema no es sino una forma de generalización simbólica de

situaciones singulares que posibilita la comunicación, mientras que un *issue* o tema político se forma a través de la controversia generada en torno a un tema. Por tanto, los *issues* se originan en el contexto de los procesos de interacción entre los actores sociales y políticos. Su introducción en la agenda política pasa por el reconocimiento público de su relevancia y, seguidamente, por la generación de una dinámica de tomas de posición en los ambientes políticos y sociales influyentes (Marletti, 1985).

McCombs y Shaw, en su célebre artículo “The Agenda Setting Function of Mass Media”, definían la función de la *agenda-setting* como el resultado de la relación que se establece entre el énfasis manifestado por el tratamiento de un tema por parte de los *mass media* y las prioridades temáticas manifestadas por los miembros de una audiencia tras recibir los impactos de los medios de comunicación. Según esta hipótesis, cuanto mayor es el énfasis de los media sobre un tema, mayor es el incremento de la importancia que los miembros de una audiencia ofrecen a estos temas como orientadores de la atención pública. Existe, pues, una relación entre la agenda de los media y la agenda pública, siendo la primera la que inicia el proceso (McCombs y Shaw, 1972:177).

Al comparar esta definición con la que Luhmann hace de la tematización puede apreciarse la mayor amplitud de este concepto frente al de *agenda-setting*. Una de las críticas clásicas a la teoría es la de desconocer a la audiencia, a la que se presenta con cierto carácter de uniformidad y de la que se ignora la agenda intrapersonal e interpersonal, factores de gran importancia para conocer los efectos cognitivos generados por los medios de comunicación (Saperas, 1987:86). La teoría de la tematización describe el sistemas político y el sistema comunicativo como un sistema complejo en el que, si bien los medios de comunicación tienen un papel determinante



para la formación de la opinión pública, no se excluyen diferentes tipos de relaciones e influencias entre los distintos subsistemas. En palabras de Agostini, la tematización es un proceso que se realiza en la relación establecida entre el sistema político y la opinión pública, a través de la mediación de los medios de comunicación de masas. De esta manera los medios de comunicación han sido considerados no como los protagonistas sino sólo como mediadores de esta relación” (Agostini, 1984:531).

Otra de las críticas clásicas a la teoría de la *agenda-setting* se extiende también la teoría de la tematización. Ambas propusieron en sus inicios la idea de que lo importante no es lo que se dice sino de qué se habla (Igartua y Humanes, 2004:244). Ante esto, autores como Gladys Engel Lang y Kurt Lang criticaron la ambigüedad de esa afirmación señalando que lo que la gente piensa no puede separarse fácilmente de aquello sobre lo que se piensa y que, al contrario, muchas opiniones diferentes pueden originarse mediante las diversas relevancias que la gente asocia a los elementos presentes en una situación compleja (Lang y Lang, 1981:449). Los teóricos de la *agenda-setting* y la tematización parecen haber entendido bien esta crítica ya que en ambos casos se han desarrollado teorizaciones complementarias para contrarrestarla. La teoría de la *agenda-setting* ha incorporado un segundo nivel de análisis cuyo objetivo es estudiar el encuadre noticioso o marco de referencia<sup>4</sup> desde el que se trata una noticia, pues una información no es sólo un conjunto aleatorio de hechos, sino que más bien se debería hablar de historias informativas que poseen un tema organizador que encuadra los hechos noticiosos (Igartua y Humanes, 2004:256).

Se han desarrollado, además, teorías complementarias como la *inter-media agenda-setting*, que plantea la influencia que ciertos medios de comunicación ejercen sobre

---

<sup>4</sup> De aquí en adelante utilizaremos el término “marco de referencia” para referirnos a este segundo nivel de análisis de la información.

otros medios de comunicación en la formación de sus agendas temáticas (Roberts, Wanta y Dzwo, 2002); o la teoría de la *agenda melding*, que plantea la existencia de una influencia entre la agenda del propio individuo y la agenda de los grupos en los que está inscrito con la agenda temática de los medios de comunicación de masas (Shaw et al., 1999:5-6). Con algunas diferencias, Marletti desarrolla una diferenciación entre tema, controversia y tema político o *issue*. Según esta diferenciación, el tema no es sino una forma de generalización simbólica de situaciones singulares que posibilita la comunicación, mientras que un *issue* o tema político se forma a través de la consolidación de las posturas defendidas en una controversia generada en torno a un tema. Por tanto, los *issues* se originan en el contexto de los procesos de interacción entre los actores sociales y políticos, y su introducción en la agenda política pasa por el reconocimiento público de su relevancia y, seguidamente, por la generación de una dinámica de tomas de posición en los ambientes políticos y sociales influyentes (Marletti, 1985). Este desarrollo teórico de Marletti tiene en cuenta factores como la interacción entre los actores sociales y políticos y la toma de posiciones en los ambientes políticos y sociales, y supera por tanto la crítica realizada por Gladys y Kurt Lang.

## **2.2. La teoría de la tematización aplicada al sistema comunicativo de Internet**

En lo hasta aquí señalado, se observa que la teoría de la tematización está preparada para soportar los cambios que se están produciendo en el sistema comunicativo actual. La relación que establece entre el sistema político y la opinión pública y su

fundamentación en la complejidad del sistema social actual le permite abordar de manera natural variaciones en las características del sistema comunicativo y en los procesos de formación de la opinión pública más allá de aquellos que los medios de comunicación de masas producen de forma generalizada y unidireccional. A continuación se comentará brevemente cómo la teoría de la tematización se relaciona con el proyecto de web semántica a través de los lenguajes controlados y cómo puede ser aplicarse al estudio de los fenómenos de formación de la opinión pública a través del análisis de redes temáticas.

### **Web semántica**

La web semántica supone añadir metadatos a los documentos con el objetivo de facilitar el proceso de gestión de la información en las búsquedas en la web. Para conseguir esto, los autores de las páginas web deben añadir etiquetas especiales que sigan las recomendaciones del WWW Consortium. El objetivo fundacional de la web semántica fue desarrollar una serie de tecnologías que permitieran a los ordenadores no solo entender el contenido de las páginas web sino llevar a cabo razonamientos sobre las mismas. Codina (2010: 5) señala los siguientes componentes conceptuales como principales características del proyecto de web semántica:

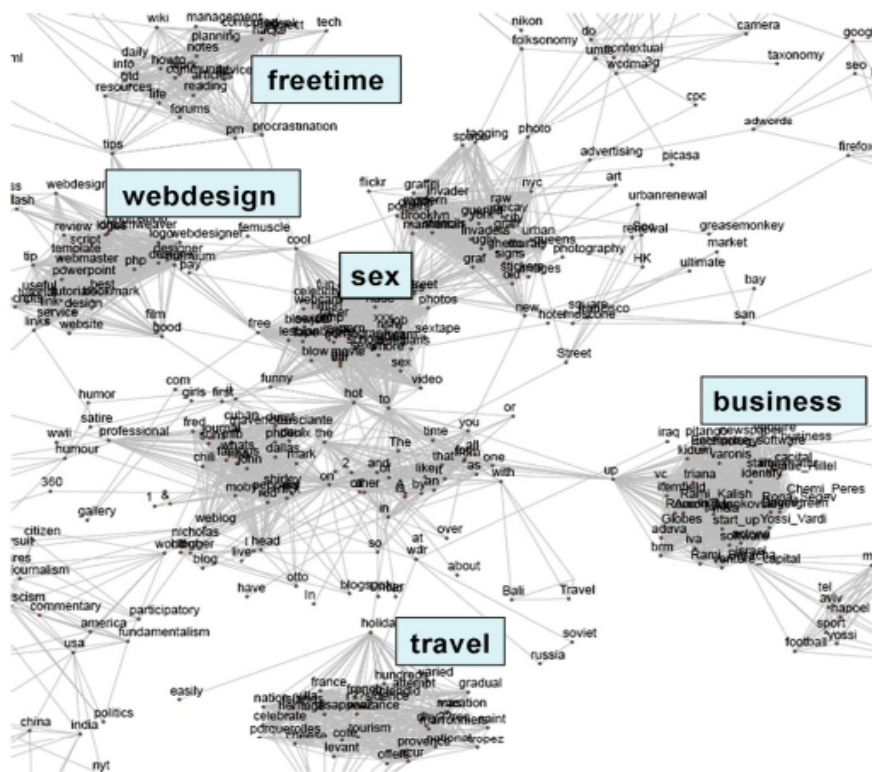
- La web entendida como una gran base de datos: la idea es conseguir que los documentos publicados en la web estén marcados de manera que sean similares a los registros de una base de datos
- Metadatos: las páginas web estarían caracterizadas con el uso intensivo de

sistemas de metadatos como parte de su marcaje

- Ontologías y lógica formal: se desarrollarían ontologías para que los ordenadores interpreten la semántica de las páginas web, y sistemas de razonamiento basados en lógica formal que podrán hacer inferencias
- Agentes de usuario: Los sistemas informáticos serían capaces de representar los intereses de sus usuarios y de interactuar con otros sistemas sin la intervención de los usuarios.

Si la comunidad de Internet adopta este formalismo como hizo con el lenguaje HTML en los años noventa, el resultado puede ser una red abstracta compuesta de redes de conceptos interdependientes (Nazar, en prensa).

*Figura 3: Etiquetas de la red social del.icio.us asociadas a su ocurrencia.*



*Extraído de Mika (2007: 9).*

## **Lenguajes controlados**

Los lenguajes controlados son mecanismos utilizados para la presentación y organización del conocimiento que tienen por objetivo controlar y normalizar la asignación de palabras clave a un documento (Vallez et al., 2010). A continuación presentaré brevemente los modelos de lenguaje controlado que considero caracterizan mejor la idea de representación del conocimiento que quiero desarrollar en este trabajo:

- **Taxonomías:** Son una forma de clasificación jerárquica del conocimiento que permite relacionar unidades de información organizándolas en clases y subclases (Chris, 2007 citado en Valle et al., 2010).
- **Tesauros:** Son listas de términos usadas para organizar el conocimiento de un ámbito y controlar la descripción temática de un documento. Se basa en términos, descriptores y relaciones semánticas y funcionales que se establecen entre los términos. Estas relaciones pueden ser de equivalencia, asociación o jerarquía (López-Huertas, 99 citado en Valle et al., 2010).
- **Ontologías:** son la especificación del conocimiento de un determinado ámbito a través de la organización de conceptos, normalmente estructurados jerárquicamente, para representar entidades, ideas o eventos además de su propiedades y relaciones en un sistema de categorías (Gruber, 2008 citado en Nazar, en prensa).

Estas son sólo algunas de las propuestas de lenguaje controlado que se desarrollan en la actualidad, pero existen otras propuestas para la representación del conocimiento entre las que destacan las redes semánticas de Quillian (1968), los gráficos conceptuales de

Sowa (1991), el análisis conceptual formal (Ganter y Wille, 1999), el marco de los espacios conceptuales (Gärdenfors y Williams, 2001), los mapas conceptuales (Novak y Cañas, 2006) o los mapas temáticos (Park y Hunting, 2003). Todas ellas comparten gran parte de las características que hemos señalado en la descripción de taxonomías, tesauros y ontologías pero quizá sean los mapas temáticos los que se estén imponiendo como forma más extendida de representación del conocimiento. En este sentido destaca la creación de TopicMaps.org<sup>5</sup>, un consorcio independiente que desarrolla la aplicación del modelo de mapas temáticos especificado en la norma ISO 13250:2000 para Internet. Este consorcio establece una serie de especificaciones para XML Topic Maps (XTM), un modelo que provee una gramática para representar la estructura de los recursos de información utilizados para definir temas (*topics*).

Figura 4: Código fuente de un mapa temático.

<pre> &lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;topicMap   xmlns="http://www.topicmaps.org/xtm/1.0/"   xmlns:xlink="http://www.w3.org/1999/xlink"&gt;   &lt;!-- 1st topic starts here --&gt;     &lt;topic id="topic_1"&gt;       &lt;baseName&gt;       &lt;baseNameString&gt;Java Island&lt;/baseNameString&gt;       &lt;/baseName&gt;     &lt;/topic&gt;   &lt;!-- 1st topic ends here, 2nd topic starts --&gt;     &lt;topic id="topic_2"&gt;       &lt;baseName&gt;       &lt;baseNameString&gt;Jakarta&lt;/baseNameString&gt;       &lt;/baseName&gt;       &lt;!-- Occurrence of 2nd topic: a HTML page --&gt;       &lt;occurrence&gt;         &lt;resourceRef           xlink:href="http://www.greatestcities.com/Asia/Indonesi             a/Jakarta/introduction/" /&gt;         &lt;/occurrence&gt;       &lt;/topic&gt; </pre>	<pre>   &lt;!-- 3rd topic starts here: the association type     "is located on" --&gt;     &lt;topic id="topic_at_1"&gt;       &lt;baseName&gt;       &lt;baseNameString&gt;is located on       &lt;/baseNameString&gt;       &lt;/baseName&gt;     &lt;/topic&gt;   &lt;!-- The association starts here, connecting the     two initial topics --&gt;     &lt;association id="association_1"&gt;       &lt;!-- Pointer to the association type --&gt;       &lt;instanceOf&gt;         &lt;topicRef xlink:href="#topic_at_1"/&gt;       &lt;/instanceOf&gt;       &lt;!-- Pointers to the two topics --&gt;       &lt;member&gt;         &lt;topicRef xlink:href="#topic_1"/&gt;       &lt;/member&gt;       &lt;member&gt;         &lt;topicRef xlink:href="#topic_2"/&gt;       &lt;/member&gt;     &lt;/association&gt;   &lt;/topicMap&gt; </pre>
---	--

Extraído de Nazar (en prensa).

<sup>5</sup> Proyecto Topic Maps: <http://www.topicmaps.org/>

## Redes temáticas

En el campo del análisis de redes sociales online se están utilizando conceptos asociados a la organización temática propuesta por Luhmann, como esfera web o red temática (Rogers, 2002: 198; Foot y Schneider, 2002: 226). El concepto de red temática<sup>6</sup> desarrollado por Rogers (2002) se define como una red de enlaces-hyperlink<sup>7</sup> por la que fluye información relacionada con una palabra clave. Como vemos, plantea un estudio del sistema comunicativo en Internet centrado en los circuitos de comunicación y en los flujos de información. Propone además diferenciar los diferentes circuitos en función de palabras clave y organizados, por tanto, en torno a temas. El concepto de esfera web, a diferencia del de red temática, se plantea no sólo como una colección de sitios web que tratan un tema en común sino como una serie de recursos digitales que se organizan dinámicamente en torno a un evento, concepto o tema, y a menudo están conectados a través de hyperlinks (Foot y Schneider 2002: 227).

Los conceptos de red temática y esfera web se han utilizado en numerosos trabajos con un significado común o al menos muy cercano. En algunos casos se ha considerado el concepto de esfera web como una categoría inferior a la red temática, delimitada por un mayor nivel de relación dentro de una red temática más amplia (Lee, 2006). En otros casos el concepto de esfera web se ha utilizado como categoría amplia dentro de la cual existen unidades más pequeñas organizadas en torno a temas, subtemas o marcos de referencia que tienen como resultado un mayor nivel de conectividad y relación (Foot y Schneider, 2002). En la práctica de las investigaciones el concepto de esfera web es cuasisinónimo del de red temática. En este estudio se utilizará el término red temática

---

<sup>6</sup> Traducción de *Issue Network*

<sup>7</sup> De aquí en adelante se utilizará hyperlink en referencia al término enlace –más utilizado en el castellano coloquial– ya que considero que resulta más amplio y ofrece más posibilidades al lenguaje académico.

por ser éste más cercano a la teoría de redes en la que se basa la argumentación de la propuesta metodológica. En cualquier caso, ambos conceptos comparten el planteamiento de una organización en categorías y subcategorías que se delimitan por el hecho de compartir interés por un tema u opinión común, lo cual se refleja en un mayor nivel de conectividad entre nodos. Se utilizará el término marco de referencia en los casos en que la diferenciación de una categoría responde a diferentes posturas o puntos de vista respecto a un tema común. La elección del término marco de referencia que se utiliza en la teoría del *framing-analysis* se debe a que con este concepto se supera una visión predominantemente cuantitativa del análisis de contenidos y en este caso de la representación del conocimiento. En los casos que la diferenciación entre categorías no responda a diferentes puntos de vista respecto a un tema común sino a diferencias de jerarquía dentro del mismo tema se utilizará el término subtema.

Utilizando el concepto de red temática, con o sin referencia a su creador, autores como Ackland, Thelwall o Park han desarrollado estudios con el objetivo de analizar diferentes regiones de Internet como la blogosfera de los partidos demócrata y republicano en Estados Unidos (Ackland, 2005), la visibilidad de la nanotecnología en Internet (Ackland, 2007), o los hyperlinks que reciben y realizan las páginas de los políticos en Corea del Sur (Park y Thelwall, 2008). El mismo Thelwall ha propuesto un tipo de análisis que ha denominado *–broad issue tracking–*. Según Thelwall, el rastreo amplio de un tema<sup>8</sup> es la tarea de identificar debates públicos importantes que surgen dentro de un determinado tema (Thelwall, Prabowo y Fairclough, 2004). A diferencia del análisis temático, el rastreo temático amplio es suficientemente abierto para que los debates individuales puedan ser separados a través de su identificación por una

---

8 Traducción de *broad issue tracking*.



terminología completamente diferente que permita un posterior análisis de marcos de referencia diferentes.

Esta forma de organización de la opinión pública resulta especialmente interesante para nuestro trabajo, ya que si utilizamos esta perspectiva para visualizar el sistema comunicativo en Internet veremos infinidad de esferas o redes temáticas que se dividen a su vez en subesferas o subredes, que en ocasiones se encuentran conectadas entre sí formando una red de redes temáticas que forman una macroestructura temática. Esta idea nos hace volver al concepto de opinión pública que planteaba Luhmann, aunque esta vez con una estructura mucho más fragmentada tanto en relación a la cantidad de redes temáticas existentes –opinión pública fragmentada o grupos de opinión– como a la cantidad de emisores de información que pueden influir en el proceso de formación de la opinión pública.

### **2.3. Sistemas complejos, autoorganización y emergencia**

Para entender los factores que determinan la emergencia de un tema en la esfera pública me apoyaré en el concepto de complejidad que se ha ido señalando a lo largo de este trabajo como característica de la sociedad actual. La teoría de la complejidad se utilizó inicialmente a finales de los años 60 en el ámbito de la termodinámica del no-equilibrio (Johnson, 2003). Años más tarde dio lugar a la Ciencia de los Sistemas Complejos, que en las dos últimas décadas ha irrumpido con fuerza prácticamente en todas las ramas de la ciencia. Tal y como afirman Solomon y Shir (2003) el ambicioso desafío de la

investigación en complejidad es la prospectiva, cartografía, colonización y desarrollo del territorio interdisciplinar que existe entre las ciencias tradicionales. Según Haken (2006) los sistemas complejos son ubicuos, aparecen en todas partes, y presentan características universales. El resultado de las investigaciones en sistemas complejos es la aparición de conceptos generales independientes del contexto, que tienen la característica de ser universales, es decir, comportamientos similares en sistemas muy diferentes (San Miguel, Toral y Eguíluz, 2004). Este concepto de universalidad fue introducido en los años 70 en el estudio de los fenómenos críticos y los cambios de fase, que son fenómenos particulares de los sistemas complejos. En su sentido original, los cambios de fase o transiciones de fase se refieren al cambio de un estado a otro de la materia, al punto crítico en el que se produce el paso de un nivel a otro de orden en el sistema, como por ejemplo cuando hierve el agua pasando de líquido a gas o cuando se funde el hielo pasando de sólido a líquido.

Los sistemas complejos o fenómenos colectivos presentan unas características generales, entre las cuales podemos decir que la principal es que el comportamiento global no es reducible a la suma de las partes del sistema o al comportamiento típico o promedio de un constituyente del sistema. Si tomamos como ejemplo una sociedad, diríamos que su comportamiento global no es reducible, según la visión reduccionista clásica, a la psicología individual y las acciones de sus componentes, cuyas características individuales pueden llegar a ser irrelevantes para el comportamiento colectivo. La segunda característica es la que se relaciona con las transiciones de fase anteriormente explicadas. El comportamiento complejo se refiere a una situación intermedia entre un estado bien ordenado y un estado totalmente desordenado. Por ejemplo, el crecimiento y organización espacial de una ciudad representa una situación

intermedia entre una distribución aleatoria y una que sigue las pautas con que se organiza un asentamiento militar.

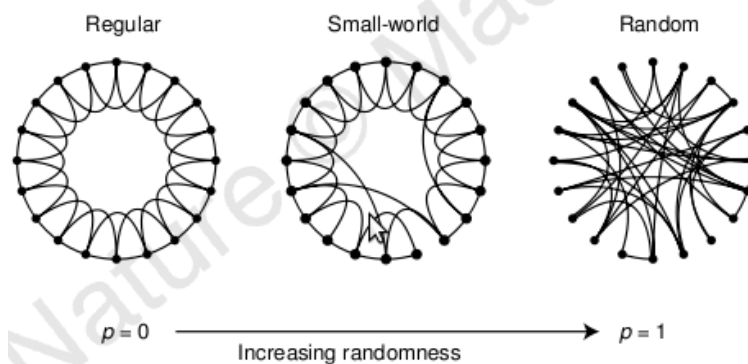
Para entender la irrupción de la teoría de sistemas complejos en un rango tan amplio de disciplinas de la ciencia es necesario señalar la que ha sido una de las mayores innovaciones científicas de las últimas décadas del siglo XX, la aparición de la Ciencia de las Redes Complejas o Ciencia de las Redes (Barabasi, 2002; Watts, 2003; Solé, 2009). Esta nueva ciencia, aunque desarrollada mayoritariamente por físicos, abarca entre sus objetivos el estudio de estructuras y fenómenos sociales tales como la cultura, la propagación de enfermedades, la comunicación electrónica, las crisis financieras, los atascos de tráfico o las organizaciones terroristas. Tradicionalmente, estas redes se estudiaban utilizando una visión estocástica, es decir, desde la concepción de que la formación y evolución de estas redes atendía a normas de aleatoriedad y probabilística. En el polo opuesto, los físicos basaban sus investigaciones de objetos en interacción en modelos de redes regulares perfectamente estructuradas y predecibles. La Ciencia de Redes Complejas ha revolucionado estos modelos al demostrar que las redes complejas pueden presentar una serie de características que se desconocían hasta el momento.

### **Propiedades de las redes complejas**

Muchos han sido los avances en la teoría de redes desde sus orígenes en 1736 con el estudio de los puentes de Koninsberg por parte de Euler, o la publicación en 1959 del revolucionario artículo de Erdos y Réyni sobre redes aleatorias. A continuación, intentaremos señalar brevemente las características más importantes que pueden presentar las redes complejas.

La primera, conocida como propiedad de mundo pequeño, indica que en una red que presenta esta característica, dos objetos pueden conectarse con pocos pasos intermedios. A medida que aumentamos el nivel de aleatoriedad en una red regular, el número de pasos para llegar de un extremo a otro de la red se reduce.

*Figura 5: Red de mundo pequeño.*



*Extraído de Watts y Strogatz (1998: 441).*

Para modelar redes con un alto nivel de conexión, Watts y Strogatz comenzaron con un círculo de nodos, donde cada nodo está conectado a su inmediato vecino y al siguiente vecino más cercano. Para hacer este mundo más pequeño se modificaron algunos de los enlaces aleatoriamente. Estos enlaces de tipo largo ofrecieron los atajos cruciales entre grandes distancias dentro de la red, acortando drásticamente la separación entre los nodos que formaban la red. Esta idea, que ya había sido planteada en la vieja afirmación de que todas las personas del planeta están conectadas por seis grados de separación, no se había teorizado de forma matemática hasta las investigaciones de Watts. Las investigaciones posteriores de Barabasi sobre el grado de separación e interconexión de

las páginas existentes en Internet han dado como resultado un grado de separación que no alcanza los 19 pasos de promedio, lo cual demuestra que Internet es una red de mundo pequeño (Barabási, 2002: 34). Esta característica ha sido descubierta en redes de citación científica, en redes de colaboración de actores y en redes sociales en general por poner algunos ejemplos. Pero lo más importante de este descubrimiento es la demostración de la afirmación que Axelrod hacía en su trabajo sobre dispersión cultural, según la cual las comunicaciones electrónicas –Internet– permiten desarrollar unas redes de interacción que son elegidas por los individuos y no impuestas por la geografía. Esto supone que con esas interacciones a larga distancia, la heterogeneidad sostenida por las interacciones locales puede desaparecer (Axelrod, 1997).

La segunda de las características que presentan las redes complejas es la conocida como curva de ley de potencia. Las redes que presentan esta propiedad tienen como característica principal que la distribución de probabilidad del número de hyperlinks de cada agente no tiene un valor típico bien definido –una escala característica– sino que se extiende hasta valores muy grandes del número de hyperlinks. En este tipo de redes la probabilidad de que un nodo de la red tenga  $K$  hyperlinks es inversamente proporcional a la potencia  $K^\gamma$ , siendo  $\gamma$  un exponente característico de la red, lo que significa que hay un número pequeño de nodos que tiene un número de hyperlinks muy alto. Este tipo de nodos que presentan un alto índice de conexión es lo que se conoce en la terminología de Internet como *hubs*. La aparición de este tipo de nodos con muchos hyperlinks supone que hay individuos muy bien conectados socialmente y cabe preguntarse por el efecto de estos nodos en la dinámica colectiva del sistema.

Figura 6: Funciones de distribución de conectividad en grandes redes.

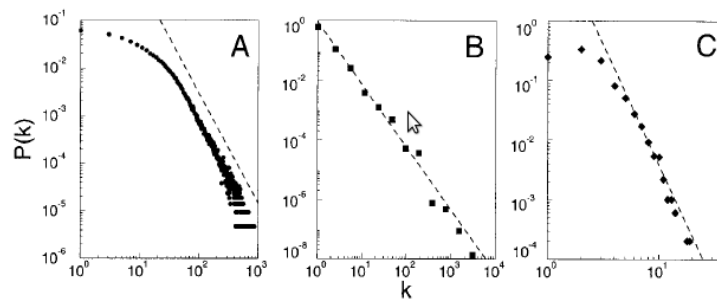


Fig. 1. The distribution function of connectivities for various large networks. (A) Actor collaboration graph with  $N = 212,250$  vertices and average connectivity  $\langle k \rangle = 28.78$ . (B) WWW,  $N = 325,729$ ,  $\langle k \rangle = 5.46$  (6). (C) Power grid data,  $N = 4941$ ,  $\langle k \rangle = 2.67$ . The dashed lines have slopes (A)  $\gamma_{\text{actor}} = 2.3$ , (B)  $\gamma_{\text{www}} = 2.1$  and (C)  $\gamma_{\text{power}} = 4$ .

Extraído de Barabási y Albert (1999: 510).

La tercera de las características que presentan las redes complejas es la conocida como **libre de escala** (Barabási y Albert, 1999). Es una característica común de muchas grandes redes como las redes genéticas o Internet. En este tipo de redes la distribución de los hyperlinks sigue una curva de ley de potencia en la que los componentes que entran a formar parte de la red tienden a conectarse con los nodos más conectados en un fenómeno que se ha denominado efecto “el rico se hace más rico”<sup>9</sup>. Este fenómeno permite que la distribución de la conectividad de la red no se vea afectada por el crecimiento o decrecimiento del número de nodos que componen la red, siempre y cuando no se rebaje un número mínimo de componentes, lo que se conoce como punto de equilibrio de la red. Tal y como afirman Barabasi y Albert, este hecho es consecuencia de dos mecanismos genéricos: (1) las redes se expanden continuamente por la suma de nuevos vértices, y (2) los nuevos vértices se conectan preferentemente a los sitios que están bien conectados. Un modelo basado en estos dos ingredientes reproduce esta distribución estacionaria libre de escala, lo cual indica que el desarrollo

<sup>9</sup> Traducción de *the rich get richer*.

de grandes redes está gobernado por un fenómeno de autoorganización que va más allá de las particularidades de los sistemas individuales.

### **Autoorganización y emergencia**

La que señalábamos como segunda característica de los sistemas complejos, es decir, las transiciones de fase que provocan el paso de un nivel de orden a otro diferente, se basan en los dos fenómenos que dan nombre a este apartado, la autoorganización y la emergencia. El proceso por el cual aparece el orden en un sistema autoorganizado se conoce como emergencia. Ya Platón, en el *Faedo*, planteaba la cuestión de la diferencia entre el todo como la suma de las partes y el todo como una entidad única que emerge de las partes y es distinto del agregado de las mismas. En términos de la dialéctica hegeliana esto significa que cualidades nuevas superan la vieja estructura del sistema total (Fuchs, 2006: 101). El sistema se transforma y alcanza un nivel superior. En las fases críticas de transición, el sistema puede expandirse e intensificarse a sí mismo por un proceso de retroalimentación positiva. Esto provoca que pequeñas causas puedan tener grandes efectos dentro del sistema. Así, los sistemas autoorganizados son su propia razón y causa, es decir, se producen a sí mismos (*causa sui*).

Paralelamente al gran desarrollo de los estudios de fenómenos cooperativos en Física en los años 70, los sociólogos abordaban los mismos problemas conceptuales, que en este ámbito recibieron el nombre de problemas “micro-macro”. En el famoso artículo de T. Schelling *Dynamic Model of Segregation* (1971), se lee: “Este trabajo se refiere a los mecanismos que trasladan el comportamiento individual desorganizado a resultados colectivos”. Al hablar sobre variables irrelevantes dice que hay una serie de

proposiciones importantes “...que son verdad para el agregado pero no en detalle, y que son verdad independientemente de cómo se comporta la gente”. Existen ejemplos de la aplicación de la teoría de la complejidad en los procesos de formación de movimientos sociales (Fuchs, 2006), de opinión colectiva en los mercados financieros (Eguíluz y Zimmermann, 2000) o de diseminación cultural (Axelrod, 1997). La motivación principal para aplicar el concepto de sistema autoorganizado al ámbito de las ciencias sociales es que el mundo moderno es inherentemente complejo y dinámico y sus fenómenos pueden ser mejor explicados con conceptos que hacen hincapié en los cambios permanentes y en las formas de organización en red (Fuchs, 2006: 102). Los planteamientos de Barabasi sobre las redes libres de escala defienden que las posibles interacciones entre los individuos que forman parte en un fenómeno colectivo vienen determinadas por la red que define las relaciones sociales entre ellos. No obstante, la estructura de las redes sociales no viene dada y determinada para siempre como una realidad externa sino que, de la misma manera que las acciones de los individuos están determinadas por la red, la red se crea a través de las acciones de los individuos y por tanto existe una coevolución de la red y las relaciones de los individuos (Lazer, 2001 citado en Eguíluz et al., 2005: 133).

Los sistemas sociales tienen un carácter dinámico que surge de la interacción entre actores individuales o grupos de actores y las estructuras sociales. Las sinergias que surgen de los procesos de comunicación entre actores humanos tiene como resultado la producción y reproducción de las estructuras sociales que, a su vez, permiten nuevas prácticas y comunicaciones por las que las estructuras sociales pueden volver a reproducirse. Este proceso es el que se conoce como autoipoesis social, según el cual tenemos dos niveles de autoorganización social:



- (1) En un nivel sincrónico un sistema complejo se produce y reproduce por autoipoiesis permanentemente.
- (2) En un nivel diacrónico el orden emerge del desorden en puntos críticos de desarrollo del sistema, o como hemos definido anteriormente, en los puntos críticos de las fases de transición.

Los sistemas complejos se mantienen y producen permanentemente. Ahora bien, debido a contradicciones en el sistema o a posibles entradas de energía en caso de no tratarse de sistemas cerrados –como de hecho sucede en el sistema social– se producen fases de transición en las que aparecen nuevas cualidades del sistema. La forma en que emergen estas nuevas cualidades del sistema no es arbitraria ni está determinada sino que surge a través de un proceso, podríamos decir que dialéctico, entre posibilidad y necesidad. Para el estudio de este tipo de sistemas es necesario un periodo de tiempo suficientemente extenso que permita no sólo fotografiar el sistema en un momento dado sino estudiar la dinámica del mismo y poder así reconocer patrones de comportamiento que nos permitan, si no hacer predicciones exactas sobre el comportamiento del sistema, al menos reconocer los puntos críticos en las fases de transición y los estados de equilibrio.

### **Sistemas adaptativos y sistemas no adaptativos**

Como explicábamos en la presentación de este trabajo, en el sistema comunicativo actual hay muchos más agentes que actúan como canales de comunicación, hay mucha más cantidad de información en circulación y ésta es más variada (López García, 2006:

231). En un sistema complejo densamente interconectado como éste se crean lo que se conoce en la teoría de redes complejas como circuitos de retroalimentación y reverberación. Las transiciones de fase entre un estado de equilibrio del sistema y otro, que hemos comentado en el apartado anterior como característica de los sistemas complejos, dependen de esos procesos de retroalimentación y reverberación del sistema.

En un sistema complejo se pueden producir dos tipos de retroalimentación, negativa y positiva. La retroalimentación negativa supone una autorregulación del sistema. A través de diversos mecanismos, un sistema complejo puede detectar las variables externas y autorregularse para llegar a un punto de equilibrio en el que las condiciones son adecuadas para el sistema<sup>10</sup>. Este tipo de sistema con procesos de retroalimentación negativa es lo que se conoce como sistemas adaptativos. Ante una variable exterior se produce un fenómeno emergente que adapta el sistema a las nuevas condiciones. Los sistemas en los que se producen procesos de retroalimentación positiva y reverberación son aquellos en los que una variable externa produce una reacción en el sistema que se retroalimenta y provoca una reverberación<sup>11</sup>. Este tipo de sistemas sin mecanismos de retroalimentación negativa o autorregulación son los que se conocen como sistemas no adaptativos. Ante la aparición o modificación de una variable exterior se produce un fenómeno emergente en el sistema, que, en este caso, no alcanza un nuevo punto de equilibrio sino que produce una reverberación que sólo se detendrá cuando desaparezca la variable que lo produjo, o bien, cuando aparezca otra variable que lo contrarreste.

El sistema comunicativo actual responde a la descripción de sistema complejo con

---

10 Un termostato es el ejemplo más simple de retroalimentación negativa. Cuando la temperatura sube, el termostato lo detecta y regula la fuente de calor para conservar la temperatura de equilibrio en el sistema.

11 El ejemplo más simple de este tipo de retroalimentación positiva es el de la guitarra eléctrica que al ser acercada al altavoz por el que sale su sonido capta dicho sonido creando un bucle que provoca una reverberación o acople acompañado de un sonido ensordecedor.

retroalimentación positiva. En relación con el funcionamiento de los circuitos comunicativos, un ejemplo que caracteriza este comportamiento, aunque fuera del universo de Internet, es el que plantea Johnson en su libro *Sistemas Emergentes* (2003). En los años 80, la cadena de noticias estadounidense CNN decidió liberar los contenidos que no iban a ser emitidos. Este hecho que puede parecer insignificante provocó que noticias que los directores de la cadena no querían emitir fuesen emitidas por cadenas locales y alcanzaran la opinión pública. La publicación de ciertas noticias por otras cadenas obligaba a CNN a emitir informaciones que en un primer momento habían desestimado<sup>12</sup>. Si este hecho revolucionó el sistema televisivo estadounidense a principios de los años 90, resulta pertinente plantearse qué puede ocurrir en el sistema comunicativo actual, donde Internet esta cobrando el máximo protagonismo.

### **2.3. Visibilidad y otros procesos de selección de la información en Internet**

Como señalábamos en apartados anteriores, los trabajos de análisis de redes sociales online utilizan conceptos como esfera web o red temática asociados a la organización temática de la información que proponía Luhmann. Sin embargo, la simple organización en temas no aporta toda la información necesaria para conocer el sistema de comunicación en la web. Aunque puede decirse que esta red de conexiones entre sitios web muestra la estructura del sistema comunicativo, es necesario conocer igualmente otros factores que determinan los procesos de selección de la información.

---

<sup>12</sup> Se trata de las informaciones relativas a la pregunta que Jim Wooten hizo a Bill Clinton sobre una presunta relación con la cantante Gennifer Flowers en 1992.

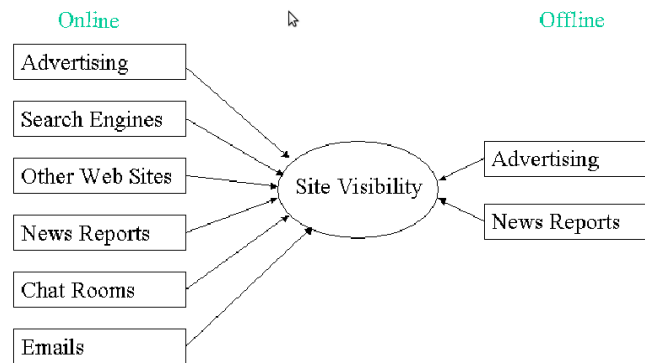
Aunque a primera vista Internet parece ser un espacio comunicativo no jerárquico que permite a cualquiera presentar y obtener información y opiniones con independencia de los procesos de selección de los medios de comunicación de masas tradicionales, existen otros procesos de selección propios de Internet. Entre ellos, la visibilidad es el principal factor que determina la capacidad de influencia de los actores que defienden una postura en la discusión sobre un tema. Pero al mismo tiempo, los factores que influyen en el nivel de visibilidad pública en Internet son complejos y necesitan un gran conocimiento del sistema y recursos para modificarlo a nuestro favor.

### **Aproximaciones al concepto de visibilidad**

La visibilidad tiene diferentes definiciones dependiendo del ámbito desde el que se estudia. Así, en el ámbito del marketing se entiende por visibilidad la presencia de una marca o producto en el medio del consumidor (Dreze & Zufryden, 2003: 25), mientras que en el ámbito del análisis de redes sociales, se entiende por visibilidad el potencial para comunicar opiniones (Freeman, 1978: 219-220) y con ello influir en otros actores.

En el estudio sobre la visibilidad que Dreze y Zufryden realizan desde la perspectiva del marketing, se distinguen dos formas de visibilidad para una empresa u organización. Estas son: la visibilidad online y la visibilidad *offline*. Cada una de estas formas de visibilidad está influida por una serie de factores, representados en la figura 7. Para realizar un cálculo de la visibilidad, los autores desarrollan una serie de representaciones matemáticas que permiten otorgar un valor numérico a cada uno de estos factores y relacionarlos después en una fórmula absoluta que representa la visibilidad total.

Figura 7: Factores de visibilidad.



Extraído de Dreze y Zufryden (2003:37).

Desde esta misma perspectiva de la gestión empresarial y el *marketing*, Schmitz-Manz y Gaul (2005) consideran que la visibilidad está compuesta de diferentes partes. Estas partes son la visibilidad obtenida a través de hyperlinks que vienen de otras páginas, la visibilidad conseguida a través de la aparición en directorios online y la visibilidad proveniente de los motores de búsqueda. Todas estas formas de visibilidad y cualquier otra que pueda surgir, se pueden clasificar, según los autores, en dos categorías de factores que la determinan.

(1) Los factores psicológicos de la visibilidad online. Esto significa que la forma en que las personas buscan, navegan e interactúan con Internet –por ejemplo el hecho que sólo se consultan los 30 primeros resultados de un buscador– tienen que ser tomadas en consideración.

(2) Los factores físicos de visibilidad online. Estos son los enlaces a una página, los anuncios de publicidad o la aparición en los listados y buscadores de Internet.

Ambos factores, físicos y psicológicos, provocan cambios en la visibilidad online. Para determinar el impacto real en el nivel de visibilidad se deben identificar los

solapamientos que se producen entre los factores que la determinan. Sólo de esta manera se puede obtener una medición precisa de la visibilidad real de un sitio web.

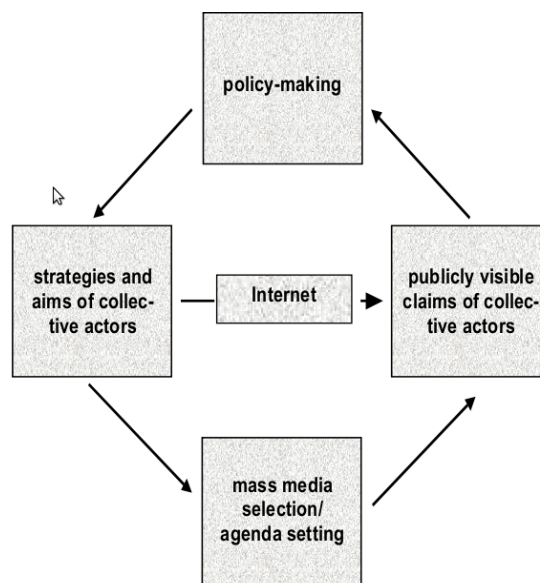
En el ámbito del análisis de redes sociales, y tal y como indica Freeman en su artículo “La centralidad en las redes sociales”, podemos suponer que los escritores que han definido el punto de centralidad en términos de grado se refieren a la visibilidad o potencial para comunicar sus opiniones. La idea de centralidad aplicada a la comunicación humana fue introducida por Bavelas en 1948 y según esta idea –tal y como funcionan los procesos de comunicación en una red social– una persona que está en una posición que le permite tener contacto directo con muchas otras debería comenzar a verse y a ser vista por otros como un canal principal de información. De alguna manera es un foco de comunicación, al menos respecto a aquellos con los que está en contacto, y está en posición de considerarse parte del canal principal del flujo de información dentro de la red (Freeman, 1978: 219-220). Estas conexiones pueden reflejar corrientes cognitivas, culturales y sociales y formaciones que incluyen la emergencia de redes y la difusión de ideas sobre ciertos temas (Bjorneborn en Lee, 2006: 51).

El modelo teórico que se muestra en la figura 8 comienza con actores colectivos que quieren influir en el proceso político. Mientras algunos de los actores con más recursos pueden ser capaces de ejercer cierta influencia sin emplear un gran nivel de visibilidad en la esfera pública –hablo por ejemplo del modo en que un lobby ejerce su influencia sobre otros actores– la mayoría de los intereses sociales no están en posición de afectar el proceso político de forma directa sino que deben hacerse públicamente visibles y conseguir el apoyo de otros actores sociales. Es a través de la movilización de este apoyo público que pueden ejercer presión sobre los representantes políticos. Todo esto

hace que los actores colectivos dependan de los medios de comunicación de masas ya que en las modernas sociedades democráticas sólo a través de ellos se puede obtener visibilidad y apoyo público.

Sin embargo, la irrupción de Internet en el sistema comunicativo ofrece a los actores colectivos un nuevo potencial para evitar a los medios de comunicación de masas tradicionales y conseguir de manera directa cierta visibilidad pública a través de su presencia en Internet. En la figura 8, el camino que conecta de izquierda a derecha, pasando por el centro, muestra esta nueva posibilidad. Con una inversión mínima de recursos cualquiera puede crear una página web y, a través de ella, hacer sus opiniones accesibles al público de todo el mundo. De esta forma, usuarios individuales o sitios web con ideas y puntos de vista similares pueden crear ciertos flujos de información y opiniones en Internet mediante la creación colectiva de enlaces a otros sitios web (Koopmans y Zimmerman, 2003: 5-6).

*Figura 8: Modelo del sistema de comunicación política actual.*



*Extraído de Koopmans y Zimmerman (2003: 5)*

Si bien el número de enlaces entrantes de un sitio constituye un indicador básico de la visibilidad y la prominencia de un sitio web, el problema con esta medida, es que no tiene en cuenta quien realiza el hyperlink. Una medida más sofisticada de la visibilidad es la que aporta el algoritmo HITS de Kleinberg (1999). Este algoritmo se basa en la premisa de la existencia de dos tipos distintos, pero relacionados, de sitios web. Los sitios de autoridades son sitios muy referenciados, es decir con muchos enlaces entrantes relacionados con un tema. Los *hubs*, en cambio, son páginas que apuntan a las páginas de autoridades y con ello les otorgan autoridad. Se da una relación de refuerzo entre autoridades y *hubs*. Un buen *hub* apunta a muchos sitios de autoridades y un buen sitio de autoridad es apuntado por muchos *hubs*. Se han desarrollado numerosos algoritmos para evaluar de la manera más precisa posible la visibilidad de una página o nodo en Internet. Dada la complejidad matemática de este tipo de algoritmos no me detendré a discutir la superioridad de una u otra variante y dejaré esta tarea a los matemáticos e informáticos interesados en este campo de investigación.

Según la teoría de redes sociales, los dos conceptos básicos que determinan la capacidad de influencia de un actor son:

- (1) La centralidad, es decir, la posición que ocupa un actor en la red.
- (2) La densidad, es decir, la cantidad de relaciones que existen entre los diferentes nodos que componen la red.

En el caso del análisis de redes temáticas, el concepto de centralidad, que resulta útil para la evaluación del nivel de visibilidad de un actor, no puede utilizarse cuando analizamos una red temática. Cuando analizamos la visibilidad de una red en su conjunto, el estudio de la posición debe ser sustituido por el estudio del tamaño de la



red, pues esta variable es la que determina que aumente la capacidad de hacerse visible a una audiencia mayor. Para determinar el tamaño de una red temática nos valemos del número de actores que comparten un espacio comunicativo (Lee 2006:72). Como ocurre en las redes que presentan autoorganización, las redes temáticas necesitan atraer un número suficiente de actores para poder construir una red de comunicación suficientemente sólida y autosostenible. Si es posible que un grupo pequeño de actores con gran interés y algunos recursos pueda iniciar las interacciones que provoquen la aparición de una red de comunicación, para que aparezca una red temática y se consolide su autoorganización y desarrollo es necesario que una gran cantidad de actores formen parte de ella. Esta es la razón que obliga a sustituir la clásica medida de la centralidad que se utilizaba en la teoría de redes sociales por la medida del tamaño utilizada en el estudio de grandes redes de comunicación.

*Tabla 1: Resumen comparado de las teorías de análisis de redes sociales, de análisis hyperlink y de análisis de redes temáticas.*

<i>Ámbito de investigación</i>	<i>Análisis de redes sociales</i>	<i>Análisis hyperlink</i>	<i>Análisis de redes temáticas</i>
Actores/nodos	Individuos/contactos	Sitios web/hyperlinks	Sitios web/hyperlinks
Objetivos de investigación	Relaciones sociales, posición social entre actores	Patrones de relación/posición estructural de sitios web	Naturaleza de las relaciones/Formación de subgrupos en una red
Comunicación/Nivel de análisis	Individual/ contactos individuales	Individual/relaciones individuales	Colectivo/ Recopilación de enlaces (subred/esfera)
Canal de comunicación	Comunicación Interpersonal	Comunicación Interpersonal	Comunicación de masas
Conceptos principales	Centralidad/ densidad	Centralidad/ densidad	Densidad/ tamaño

*Traducido de Lee (2006:71).*

## **Procesos de selección de la información**

Desde la perspectiva del estudio de la opinión pública y concretamente de la comunicación política, Koopmans y Zimmerman (2003) defienden que los principios sobre los que se asienta el establecimiento de la visibilidad pública en Internet operan mediante los dos mecanismos principales de búsqueda de información en Internet: “vertical” –portales y buscadores– y “horizontal” –redes de hyperlinks–. En el caso de los mecanismos de selección vertical, los portales y buscadores guían al usuario presentando una selección jerárquica de sitios web que trata de ser relevante para el tema que el usuario ha definido mediante los criterios de búsqueda introducidos en el buscador. De esta manera, se puede decir que los buscadores actúan como *gatekeepers* al ofrecer una selección de la realidad de acuerdo con unos criterios propios. En el caso de los mecanismos de selección horizontal, una red de selección formada por hyperlinks ayuda al usuario a encontrar sitios web siguiendo los enlaces que se ofrecen en otros sitios web. A través de hyperlinks los actores públicos actúan ellos mismo como *gatekeepers* en la esfera online al ofrecer enlaces a unos actores determinados y excluir a otros de su selección. (Zimmermann y Koopmans, 2007, 219-221).

### ***Procesos de selección vertical***

Dentro de la jerarquía vertical que constituyen los buscadores en la selección de sitios web, los criterios que utilizan para llevar a cabo la selección varían de uno a otro. Dada la importancia que tienen para entender correctamente el apartado dedicado a los procesos de selección horizontal, considero imprescindible dedicar unas líneas a explicar las técnicas y variables en las que se basa su funcionamiento.

Algunos buscadores ofrecen la posibilidad de comprar un lugar con gran visibilidad dentro de su lista de resultados. Como sucede en el mercado tradicional, este mecanismo de selección tiende a reproducir las diferencias de poder existentes: el rico puede comprar un alto nivel de presencia para su sitio web, mientras que aquellos con pocos recursos se mantienen en un nivel bajo de visibilidad. No obstante, la mayoría de los motores de búsqueda comúnmente utilizados utilizan un conjunto de criterios más “democráticos” –es decir, menos dependientes de la capacidad económica de la organización que pretende mejorar su posición en los *rankings* de búsquedas–. Generalmente los buscadores no buscan en Internet directamente sino que extraen resultados de los contenidos de sus bases de datos. Existen tres formas diferentes de construcción e indexación de una base de datos: automática, manual o mixta.

Los llamados directorios dependen de manos humanas para la construcción de sus bases de datos. Los propietarios de los sitios web aportan pequeñas descripciones de sus sitios a los buscadores o editores profesionales escriben descripciones para sitios que revisan. Los buscadores que crean sus bases de datos automáticamente son conocidos como buscadores verdaderos. Rastrean la web usando un tipo de software conocido como *spiders*, *robots* o *crawlers*. Cuando este software encuentra una página la transmite a otro programa para su indexación. Este programa identifica los textos, hyperlinks y otros contenidos de la página y los almacena para que puedan ser buscados por palabras clave. Los buscadores que mantienen sus bases de datos de ambas formas, automática y manualmente, se conocen como buscadores híbridos.

Existen también buscadores que buscan en otros buscadores para compilar sus listas de resultados; se les conoce como metabuscadores. Los diferentes modos en la creación e indexación conducen obviamente a diferentes resultados; de hecho, algunos buscadores

indexan más páginas que otros o se centran en tipos diferentes de páginas. Así, no hay dos buscadores que cubran exactamente la misma colección de páginas web y cada buscador utiliza criterios diferentes para crear su ranking. El objetivo general es devolver las páginas más relevantes en los lugares más altos de la lista. Para determinar la relevancia, los buscadores basados en el rastreo (*crawler*) de páginas, según un conjunto de reglas también conocidas como algoritmo de búsqueda.

El funcionamiento del algoritmo de cada buscador se mantiene en estricto secreto, pero la mayoría siguen algunas reglas generales. Uno de los más importantes criterios es la localización y frecuencia de las palabras clave en la página. Las páginas con palabras de búsqueda en el apartado de título del código HTML suelen ser consideradas más relevantes que otras que puedan aparecer en otras partes de la página. Los buscadores también comprueban que las palabras de búsqueda aparezcan en posiciones como el encabezamiento o en los primeros párrafos del texto. Este método de localización y frecuencia es susceptible de ser manipulado por los propietarios que intentan mejorar su posición en la lista de resultados. Repitiendo una palabra cientos de veces en una página (*spamming*) intentan aumentar su frecuencia y así conseguir situar su página más alto en los listados pero los buscadores persiguen estas formas comunes de *spamming* por diferentes medios y desarrollan los llamados ranking “*off page*”, que a diferencia de los primeros, no pueden ser fácilmente manipulados por los gestores de páginas web o webmasters.

El criterio más utilizado y valorado por los diferentes buscadores es el análisis hyperlink. Analizando cómo las páginas se enlazan unas con otras, un buscador puede determinar tanto el tema que trata la página como si la página en cuestión está bien considerada y por tanto debe ocupar una mejor posición en el ranking de la lista de

resultados. Por ello, las empresas de buscadores han desarrollado sofisticadas técnicas para detectar intentos por parte de los *webmasters* de crear hyperlinks artificiales que buscan alterar sus rankings. Otro de los factores “*off page*” es la medida del tráfico que tiene la página. Esto significa que los motores de búsqueda monitorizan los resultados que los usuarios seleccionan para una búsqueda determinada y usan esta información para bajar en el ranking a páginas que no están atrayendo clicks y promocionar aquellas que interesan a los visitantes. Los criterios de medición de hyperlinks y de recuento de clicks son fenómenos emergentes que no están impuestos desde arriba ni obviamente dependen de la cantidad de recursos controlados por un actor. Desde el momento en que la popularidad de una página determina su prominencia en la lista de resultados, el efecto teórico de este criterio de búsqueda es el de reforzar la visibilidad de la páginas que ya son populares o prominentes. Es el efecto “*the rich get richer*” que se comentaba en el apartado dedicado a las redes libres de escala. Un efecto que introduce desigualdades en el espacio de Internet haciendo más visibles y más fácilmente accesibles unas páginas que otras.

Teniendo en cuenta lo dicho respecto a los criterios que utilizan los buscadores y las posibles formas de influir en los rankings que crean, considero que si se quieren desarrollar investigaciones fiables sobre los fenómenos que se producen en Internet es necesario buscar métodos que no se basen únicamente en los resultados que ofrecen estos buscadores. Por esa razón en esta investigación se presta más atención a los procesos de selección horizontal de la información que no dependen de la jerarquía establecida por las distintas empresas de buscadores.

### ***Procesos de selección horizontal***

En los procesos de selección de la información que hemos denominado horizontal, el hyperlink supone la principal tecnología que determina este tipo selección. El hyperlink es una capacidad tecnológica que permite tanto la navegación interna dentro de un sitio web, como la navegación externa estableciendo enlaces entre diferentes sitios web. Wilkinson et al. (2003:49) definen el hyperlink interno como aquel que tiene su fuente de origen y su objetivo de destino dentro del mismo sitio web. Los enlaces externos proporcionan acceso directo desde una página web de origen a una página web de destino que tiene una dirección URL diferente y que normalmente está gestionada por un autor diferente. Estos enlaces externos no están concebidos simplemente como una herramienta tecnológica sino como un canal emergente de comunicación social completamente nuevo (Park, 2003:47). A partir del trabajo de Park et al. (2005), Ackland y Gibson (2006) defienden que los enlaces hyperlink sirven para promover variadas funciones dentro del ámbito de las organizaciones. En esencia caracterizan el hyperlink como:

- (1) Provisión de información: los enlaces hyperlink pueden dirigir a los visitantes hacia fuentes de información adicionales.
- (2) Construcción de redes y refuerzo: los enlaces hyperlink permiten a las organizaciones crear alianzas virtuales y redes y mantener y reforzar las relaciones existentes en el ámbito offline a través de los lazos creados online.
- (3) Construcción de identidad: la creación de conexiones indirectas o implícitas entre grupos puede ayudar a fortalecer la identidad de los componentes de esos grupos a través del fortalecimiento de la identidad del propio grupo.
- (4) Intercambio de audiencia: la integración en un grupo puede conseguir de manera más eficiente e inmediata que la audiencia y los potenciales apoyos se compartan entre los integrantes del grupo especialmente entre los grupos de opinión.
- (5) Amplificación del mensaje o fuerza de multiplicación: –especialmente útil para

organizaciones marginales o grupos con una baja visibilidad offline— a través de los enlaces hyperlink transmite a los usuarios una imagen distorsionada de la extensión del apoyo con que cuenta el mensaje.

Los hyperlink entre sitios web implican que el autor que realiza el hyperlink atribuye cierta relevancia a la página que recibe el hyperlink. De esta forma podemos explicar las redes creadas a través de hyperlinks en términos de credibilidad, prestigio y confianza (Park, 2003:48; Kleinberg, 1999:623). Una página creíble, prestigiosa y confiable recibirá más enlaces que otra en un proceso de selección natural que tendrá como resultado inevitable que unas páginas tengan un gran número de enlaces mientras que otras no cuentan apenas con ninguno.

Cuando entendemos los enlaces hyperlinks como un tipo de relación, podemos utilizar los métodos de análisis de redes sociales para describir los patrones de conexiones o desconexiones observados en una red de sitios web (Park, 2003:48; Jackson, 1997:292). Desde los años 60, los sociólogos han estudiado las propiedades de las redes sociales y los conceptos propuestos por ellos están siendo utilizados con algunas adaptaciones al nuevo sistema comunicativo que supone Internet. No pretendemos hacer un repaso exhaustivo de las propiedades estudiadas por la teoría de redes sociales pero sí señalar las adaptaciones más importantes que se están aplicando al estudio de redes sociales online.

Como ya señalábamos antes, el concepto de centralidad es uno de los conceptos aplicados al estudio de redes sociales online. La medición de la centralidad puede indicar el grado de concentración de hyperlinks en determinados sitios web o determinados grupos de sitios web dentro de una comunidad más amplia. El análisis hyperlink puede aportar también una medida de la posición de influencia de un sitio

web, identificando sitios que aportan conexiones cruciales entre grupos de sitios web, incluso en casos en que estos sitios web conectores no tienen una gran cantidad de enlaces apuntando hacia ellos. Buscadores como Google utilizan o han utilizado la medición de la posición estructural de un sitio web a través de la estructura hyperlink que rodea al sitio web como criterio para determinar cual es el sitio web más influyente dentro de un grupo de sitios web (Henzinger, 2001: 384). De esta manera se ha utilizado la medición del número de enlaces como medida de la visibilidad o centralidad de un sitio o un grupo de sitios web.

Los conceptos de lazos fuertes y lazos débiles<sup>13</sup> que desarrollara Granovetter (1973) indican que una red social puede presentar diferentes tipos de relación, siendo los lazos fuertes los que presentan un nivel de relación más fuerte y por tanto con mayor capacidad de influencia. En relación con los lazos fuertes y débiles de Granovetter, el concepto de propagación hace referencia al alcance y velocidad de difusión de la información entre redes de comunicación. Dicho en otras palabras, la propagación es la medida de la capacidad de contagio que tiene una información (Bonchek, 1997: 33). La fuerza de relación hace referencia a la densidad de la conexión entre diferentes redes temáticas o esferas web. Esta densidad, también llamada conectividad<sup>14</sup>, influye en la propagación de un tema u opinión entre dos redes temáticas o dentro de una propia red (Jackson, 1997).

Estos son algunos de los conceptos fundamentales en el análisis de redes sociales online y procesos de formación de la opinión pública en Internet. Los estudios más avanzados incluyen además el análisis estadístico de sistemas, pero por el momento ese campo resulta completamente fuera de mi alcance y por ello he preferido no incluir nada al

---

<sup>13</sup> Traducción de *strong and weak ties*

<sup>14</sup> Traducción de *connectedness*



respecto. No obstante considero que resulta fundamental para una completa comprensión tanto del funcionamiento de la redes como de los estudios que se llevan a cabo.

### 3. Objetivos y problema de investigación

El trabajo que presento nació como un intento de establecer las posibles relaciones existentes entre los procesos de comunicación que se producen en Internet y los procesos de formación de la agenda temática que la teoría de la *agenda setting* viene estudiando desde los años 70. Con este objetivo en mente se planteó un problema de investigación inicial que relacionaba estos dos ámbitos de la comunicación, Internet y los medios tradicionales de comunicación de masas. La pregunta que nos hicimos tras un proceso de depuración en el que se precisaron los conceptos clave que debían ser estudiados, es la siguiente:

*¿Existe una correlación entre el nivel de visibilidad de un tema en Internet y su emergencia en los medios de comunicación?*

Ante esta pregunta decidí profundizar en la emergencia temática y la visibilidad online como conceptos claves a analizar antes de seguir adelante con este problema de investigación. Para el estudio del concepto de emergencia temática tuve en cuenta las teorías de los efectos de la comunicación que establecen una organización temática de la información. Principalmente la teoría de la *agenda-setting* y la teoría de la tematización.

Tras estudiar las distintas teorías de los efectos de la comunicación tradicional e intentar asimilarlas al nuevo sistema comunicativo que supone Internet, consideré que era necesario abordar el tema desde una perspectiva, si no totalmente nueva, al menos alejada de las teorías clásicas. Traté de abandonar el paradigma según el cual la formación de la opinión pública sigue una dirección descendente, es decir, se crea en

función de la transmisión de los temas que formarán la agenda pública desde la cúpula de los grandes medios de comunicación –en connivencia con el poder político– al ciudadano de a pie. Como ya he explicado, diversos autores habían planteado la aparición de un nuevo paradigma en el sistema comunicativo del que la fragmentación de la agenda pública es su característica principal y en el cual la tradicional posición hegemónica de los grandes medios de comunicación se está viendo socavada por la aparición de fenómenos de formación de opinión pública de abajo a arriba –desde la población y los medios alternativos hacia los grandes medios de comunicación de masas y poderes políticos–. Teniendo en cuenta esto, decidí recurrir a las teorías de la comunicación que habían abordado los fenómenos de formación de la opinión pública desde la perspectiva de los sistemas complejos, ya que en éstas ya se planteaban y estudiaban este tipo de fenómenos colectivos como fenómenos emergentes.

Al utilizar esta perspectiva no pretendo negar que existan formas de selección reguladoras de la agenda pública ni que el sistema tradicional de formación de la opinión pública se haya derrumbado para dar paso de forma revolucionaria al modelo que se utiliza en esta investigación. El estudio de la visibilidad como principal factor determinante de los procesos de selección en Internet, incluido en el tercer capítulo de nuestro marco teórico, responde a la aparición de estas nuevas formas de selección de la información que regulan la emergencia de los temas en la agenda pública de Internet. Tampoco quiero afirmar con ello que la visibilidad es ya la única forma de selección en Internet y que los grandes medios de masas no ejerzan una influencia fundamental en los procesos de formación de la agenda pública. Simplemente considero que para establecer los métodos y construir las herramientas necesarias para el análisis de la opinión pública en Internet es más urgente y necesario centrarse en las características

propias del sistema.

Por todo esto, las preguntas y objetivos que persigue esta investigación se centran en el sistema comunicativo que se ha generado en torno a Internet y los fenómenos de formación de la opinión pública que se producen como fenómenos emergentes dentro de un sistema complejo. Para ello, nos apoyaremos en la tematización como forma de organización y clasificación de los temas que componen la agenda pública; en la teoría de redes complejas como base teórica para el estudio del sistema comunicativo propio de Internet; y en la visibilidad como forma característica de selección de la información en este nuevo sistema comunicativo. A continuación presentamos los objetivos y preguntas de investigación de este trabajo.

### **Objetivo teórico general**

El objetivo central de esta investigación es realizar una propuesta metodológica que permita mejorar los métodos de análisis de los principales factores que determinan el funcionamiento de los procesos de formación de la opinión pública en Internet. El estudio de los objetivos específicos que me marco en este trabajo es secuencial, esto es, la consecución de los últimos depende de la consecución de los primeros.

1. Proponer métodos para identificar redes temáticas en Internet. Es necesario identificar los nodos o puntos de recepción y creación de enlaces que establecen relaciones en torno a un tema común.
2. Proponer métodos para identificar los distintos tipos de actores que participan en una red temática. Este segundo nivel de análisis de la red nos permitirá clasificar

los distintos tipos de actores que participan en ella –sean estos genereadores o recopiladores de contenidos– y descubrir patrones de conducta característicos de uno u otro tipo de actor.

3. Proponer métodos para identificar, clasificar y analizar la estructura de temas, subtemas, y/o marcos de referencia que aparecen en una red temática. Una vez identificada una red temática y los tipos de actores que la componen es necesario analizar dicha red en un nivel mayor de profundidad que nos permita identificar los subtemas y/o marcos de referencia que aparecen en la red. Según la teoría de la tematización aplicada al análisis de redes online, los distintos subtemas y marcos de referencia serán identificables por la utilización de palabras clave y expresiones semánticas diferentes, que, además, coincidirán con una subestructura identificable por la aparición de grupos de nodos con un alto grado de interconexión.
4. Proponer métodos para visualizar y monitorizar redes temáticas. Una vez identificada la red, los diferentes tipos de actores que la componen, los subtemas y/o marcos de referencia y los grupos de nodos altamente interconectados (*clusters*) que aparecen en la red, es necesario poder visualizar dicha red de manera continuada a fin de poder analizar el dinamismo característico de las redes online. El periodo de tiempo durante el cual se deberá monitorizar la red variará dependiendo del tema en torno al que se haya creado la red.
5. Proponer criterios para determinar el nivel de visibilidad de una red temática en Internet. El análisis del nivel de visibilidad de una red en su conjunto resulta de especial interés si se quiere conocer la capacidad de influencia de dicha red.

## **Objetivo de investigación aplicada**

Una vez establecidos los objetivos específicos de este trabajo, se plantean en forma de problema de investigación concreto, con una serie de preguntas que se corresponden con estos objetivos, para poner a prueba si su aplicación sería no sólo posible sino de utilidad.

Siendo muchos los posibles campos de aplicación, queremos proponer aquí su uso para el estudio de los fenómenos de opinión en Internet durante los procesos electorales, y más concretamente, para analizar la evolución de los temas que conforman la agenda temática en Internet en las elecciones catalanas de 2010.

En este caso, las preguntas pertinentes, en aplicación de todo lo dicho, son:

1. ¿Qué tipo de actores componen una red temática?
2. ¿Qué subtemas o marcos de referencia defienden cada uno de los actores que componen una red?
3. ¿Qué relaciones existen entre los diferentes actores de una red?
4. ¿Qué relaciones existen entre los diferentes subtemas o marcos de referencia que aparecen en una red?
5. ¿Qué nivel de visibilidad tiene cada una de las redes temáticas?

#### **4. Metodología y técnicas de investigación**

Para dar respuesta a cada una de las preguntas propuestas propongo una serie de métodos que se aplican a cada uno de los objetivos específicos que forman parte del objetivo general de la investigación.

##### **Propuesta metodológica para la identificación de redes temáticas**

Para dar respuesta al primer objetivo que se plantea este trabajo recurriré al concepto de red temática –*issue network*– que proponía Rogers (2002). El objetivo inicial es llevar a cabo un mapa que represente las redes que, entorno a un tema y a través de enlaces hyperlink, se crean en Internet.

En el trabajo de Koopmans y Zimmermann (2003) sobre la comunicación política en la Unión Europea, se seleccionan siete temas –política monetaria, agricultura, inmigración, envío de tropas, jubilación y sistema de pensiones, educación e integración de la UE– que representan los principales temas de la agenda política europea. Para realizar la selección de temas que componen la agenda temática durante el periodo de elecciones en Catalunya me remito a los temas que se analizaron en el último estudio sobre las elecciones catalanas llevado a cabo por el departamento de comunicación de la Universitat Pompeu Fabra (Pericot y Capdevila, 2009).

Para llevar a cabo la identificación de las redes temáticas relacionadas con los temas escogidos, primero es necesario seleccionar una serie de sitios web prominentes

relacionados con el tema estudiado. A la hora de compilar un corpus de información, distintos investigadores han utilizado diferentes tipos de documentación: sitios web seleccionados por buscadores (Koopmans y Zimmermann, 2003), blogs (Ackland, 2005) o sistemas RSS de sindicación (Thelwall, Prabowo y Fairclough, 2006).

En este trabajo se propone no centrarse en un tipo concreto de documento y seleccionar una serie de palabras claves para introducirlas en el buscador más utilizado, Google. A través de varias frases de búsqueda diferentes para cada tema se extraerán los resultados relevantes que aparezcan en los diez primeros puestos de la lista. Para determinar los resultados relevantes propongo tener en cuenta tres criterios pudiendo eliminar alguno de ellos según el estudio que se lleve a cabo. Estos tres criterios son, en primer lugar, el contenido que aparece en la página web, que deberá ser relevante para la búsqueda introducida realizada. El segundo criterio es el lenguaje. Deberán eliminarse la páginas que no coincidan con el idioma utilizado en la búsqueda realizada en el buscador. A pesar de que normalmente los resultados obtenidos coinciden con el idioma de la búsqueda pueden producirse resultados inesperados entre otros motivos por la similitud de las palabras claves utilizadas entre varios idiomas. El tercero y último de los criterios es la localización de las páginas web. Si es necesario estudiar la red temática que se limita a un territorio geográfico deberán identificarse los sitios web cuyo autor pertenece o hace referencia al territorio en cuestión. No obstante, por la característica deslocalización de Internet es recomendable aplicar este criterio sólo cuando sea estrictamente imprescindible.

El conjunto de sitios web que obtendremos y que denominaremos semillas, será posteriormente introducido en alguno de los programas informáticos denominados *crawler*, que nos permitirán visualizar los enlaces entrantes y salientes de estos sitios y



así dibujar un gráfico de la red temática que estamos estudiando. Existe una gran variedad de *crawlers* que realizan esta tarea. Quizá el más conocido es *Issue Crawler*<sup>15</sup>, pero existen otros como el que se desarrolla en el proyecto WIRE<sup>16</sup> y que ofrece el *Center of Web Research* (CWR) encabezado por Ricardo Baeza Yates, (UPF). Una vez identificada la red temática a estudiar y dependiendo de la investigación que se lleve a cabo podrán desarrollarse diferentes acciones para identificar subredes o grupos de nodos altamente conectados. Para llevar a cabo esta tarea existen programas de rastreo que permiten configurar los criterios de selección de los sitios web a buscar. Un filtro que extraiga únicamente las páginas que presentan un número determinado de enlaces entrantes permitirá reducir el número de nodos de nuestra red y visualizar con mayor facilidad los grupos diferenciados de nodos altamente conectados para así poder identificar diferentes subredes dentro de la red temática general. SocSciBot<sup>17</sup> es un programa que genera información básica sobre los enlaces entrantes, salientes y los enlaces internos de una web. SocSciBot fue creado por Thelwall y sus colaboradores y está diseñado para rastrear hasta 9000 páginas y analizar la estructura de sus enlaces. Actualmente los desarrolladores de este tipo de programas se plantean como uno de sus objetivos prioritarios desarrollar sus programas de tal manera que sean libres de escala. Como se explicaba en el apartado dedicado a las redes libres de escala, esta característica, aplicada a los rastreadores automáticos y otras herramientas propias del ámbito de la minería de datos, conlleva que las herramientas de análisis puedan desarrollar las acciones programadas en un corpus de datos de gran volumen y así se capaces de analizar grandes secciones de Internet. En el caso de SocSciBot, el software proporciona datos numéricos a partir del rastreo de todos los enlaces internos que hay en

---

15 Issue Crawler: <http://issuecrawler.net/> [accedido el 15 de mayo de 2010]

16 The WIRE proyecto: <http://www.cwr.cl/projects/WIRE/> [accedido el 15 de mayo de 2010]

17 SocSciBot3: <http://cybermetrics.wlv.ac.uk/socscibot> [accedido el 17 de mayo de 2010]

un sitio web y de los enlaces entrantes y salientes de ese sitio. Una vez obtenido un archivo con los datos básicos sobre los patrones de los enlaces, *Colink Analysis*, –otro programa desarrollado por Thelwall– genera un archivo que permite a Pajek<sup>18</sup> –un programa de visualización de redes– construir un gráfico a partir de los datos obtenidos con SocSciBot.

El mapa inicial que incluye todos los nodos existentes tiende a no permitir apreciar de manera visual los patrones existentes. Entre los nodos, muchos de ellos no tienen ningún hyperlink o tienen muy pocos. Para optimizar la visualización de la red y poder observar los patrones existentes en ella, es necesario simplificar la red. K-core, un programa creado por Nooy, Mrvar y Batagelj (2005), permite visualizar las diferentes subredes que aparecen en la red temática general filtrando los nodos con un nivel de conectividad determinado que se fija dando un valor a la variable K en la que se basa el programa. Al refinar nuestra red con K-core, Pajek, el programa de visualización de redes, elimina los nodos que no alcancen un umbral de conectividad K, haciendo que surjan de red general los componentes más densamente conectados.

El campo de la minería de datos está en pleno desarrollo y todas las universidades e institutos que trabajan en este ámbito desarrollan sus propios programas y mejoran constantemente las funcionalidades de este tipo de software. Se han analizado algunos de los programas más populares como ejemplo de las posibilidades que ofrecen estos programas informáticos pero con ello no quiero decir que estas sean las mejores opciones. La elección de los programas a utilizar en futuras investigaciones dependerá tanto de los recursos con los que se cuenta como de las características propias de la investigación a llevar a cabo.

---

18 Pajek. *Program for Large Network Analysis*: disponible en <http://pajek.imfm.si/doku.php>

## **Propuesta metodológica para el análisis de subtemas y/o marcos de referencia en una red temática**

Para proponer los métodos que permitan alcanzar este tercer objetivo se utilizará el concepto de análisis temático –*issue analysis*– de Thelwall, y su posterior rastreo temático amplio –*broad issue tracking*– (Thelwall, Prabowo, Fairclough, 2004). También se utilizarán los diferentes métodos de representación del conocimiento, es decir, las diferentes propuestas de redes semánticas, mapas temáticos u ontologías tratadas en el apartado 2.2.

Existen diferentes metodologías para analizar los subtemas o marcos de referencia que aparecen en una red temática y desarrollar representaciones gráficas que permiten visualizar la organización de la información. En todos los casos es imprescindible llevar a cabo un análisis del contenido de las páginas que van a formar el corpus; la diferencia reside en cuando y como va a realizarse este análisis. Puede llevarse a cabo a través de un equipo humano que analice manualmente los documentos extraídos con la ayuda de un *crawler*. Otra opción es compilar el corpus a través de una *crawler* que sea capaz de filtrar los sitios web en función de las palabras clave que las caracterizan o incluso en función de los textos completos que aparecen en esos sitios web. En este sentido existen varios proyectos para obtener ontologías por medio de herramientas automáticas<sup>19</sup> pero se encuentran actualmente en fase de desarrollo. Habrá que permanecer atento a los resultados de ese nuevo tipo de rastreadores temáticos pues reducirían notablemente el trabajo de este tipo de investigaciones enfocadas al estudio de la opinión pública y la comunicación política en Internet.

Una vez identificada la red temática en cuestión y las subredes que aparecen en ella, se

---

19 CYC Ontology Project: <http://www.cyc.com> [accedido el 20 de abril de 2010].

deberán extraer los textos que aparecen en los sitios web seleccionados. Para llevar a cabo esta tarea propongo utilizar Webivore. Un programa que permite extraer automáticamente diferentes tipos de texto de un corpus de sitios web. Webivore es un rastreador orientado a la investigación diseñado para recopilar código fuente, títulos, direcciones URL, textos completos, cabeceras, *cookies*, palabras clave, información para los buscadores e incluso errores. Entre los datos que extrae Webivore los títulos y las palabras clave resultan de gran interés para analizar el contenido de cada página. Estos textos deberán ser adecuadamente codificados para luego introducirlos en un programa de análisis de contenidos automáticos.

VBPro<sup>20</sup> es un programa de análisis de contenidos automático capaz de analizar una cantidad enorme de textos a través de palabras clave y de marcos de referencia, es decir, a través de un conjunto de palabras clave relacionadas de una determinada manera. Este programa examina la frecuencia en que aparecen las palabras clave o los marcos de referencia en los textos. Aunque el análisis de contenidos asistido por ordenador no permite un análisis basado en marcos de referencia muy complejos por el momento, su gran velocidad de análisis permite aumentar el número de textos analizados incrementando con ello la validez estadística de los resultados.

Una vez realizados los análisis necesarios y extraído la información relevante del corpus de sitios web, el siguiente paso es construir el mapa temático a través del cual se organizará la información. Para hacer esto existen diferentes modelos de representación del conocimiento que se han señalado en el apartado relativo a la tematización. Estudiaremos con más atención el modelo de los mapas temáticos<sup>21</sup> por ser el que está más extendido dentro del ámbito de la organización de la información digital. No en

---

20 VBPro: <http://www.yoshikoder.org/vbpro.html> [accedido el 10 de mayo de 2010].

21 Traducción libre del término en inglés Topic Map

vano existe un un consorcio independiente que desarrolla la aplicación del modelo de mapas temáticos especificado en la norma ISO 13250:2000 para la World Wide Web<sup>22</sup>. Este consorcio establece una serie de especificaciones para XML Topic Maps (XTM), un modelo que provee una gramática para representar la estructura de los recursos de información utilizados para definir temas (*topics*).

La posibilidad de utilizar los *topic maps* en el análisis de redes sociales online se debe a que en realidad estos ya se encuentran implementados en Internet a través de las etiquetas que describen los diferentes documentos. Un mapa temático permite diferentes tipos de filtrado. Por un lado, el “*scope*” filtra diferentes *topics* o temas para una determinada comunidad o conjunto de documentos. Por otro, las facetas se encargan de filtrar recursos de información, por ejemplo, ofrecer a determinado usuario sólo los recursos informativos que estén en español. Los “*scopes*” y los “*themes*” representan el ámbito en el que las afirmaciones realizadas sobre algún “*name*”, “*occurrence*” o “*association*” de un “*topic*” tienen coherencia y validez. Un “*theme*” es la forma en que se denomina un determinado “*scope*”. Las “*facets*” proporcionan un mecanismo para asignar pares de propiedad-valor a los recursos de información, por ejemplo, considerar sólo documentos en un idioma determinado con una aplicación determinada (experimental, práctica, etc.).

La aplicación y características de los *topic maps* son complejas y permiten un nivel de descripción de la información ciertamente complejo pero, en resumen, sus características principales son las siguientes:

1) Su denominación (“*name*”). Aunque un tema puede presentar varias denominaciones, el nombre debe estar representado obligatoriamente por una forma

---

<sup>22</sup> Topic Maps Consortium <http://www.topicmaps.org/>

base “*base name*”. El “*base name*” representa la forma normal de referirse al tema. Es complementario del identificador interno (id). Los “*base name*” deben de ser únicos en un dominio determinado bajo un ámbito (*scope*) determinado. Cuando se fusionan dos mapas temáticos, se tratan los distintos “*base names*” de los “*topics*” que hacen mención a un mismo “*subject*” (concepto). Y para ello, se normalizan los “*base names*” mediante los “*public subject*”, un recurso que identifica al “*subject*” de forma no ambigua. Además, el tema puede tener otras denominaciones (*alternative names*), como el “*display name*”, que es la forma en la que se mostrará al usuario, y el “*sort name*”, que es como se ordenará alfabéticamente cuando se solicite una lista.

2) Sus apariciones (“*occurrence*”). Son enlaces a recursos de información, como por ejemplo, un sitio web. Las “*occurrences*” son elementos opcionales que son relevantes para un “*topic*” determinado. Un topic puede tener ninguna o muchas “*occurrences*”, y éstas pueden ser de muchos tipos: una cita de texto, una definición, un artículo, etc. La mayoría de estas “*occurrences*” suelen ser externas al mapa temático, como ocurría entre los índices documentales y los tesauros. A cada “*occurrence*” le corresponde un único enlace. Existen dos tipos de “*occurrences*”: un dato expreso (*resource data*) y un hyperlink a un recurso externo (*resource ref*).

3) Sus roles de asociación (“*association*”). La “*association*” es un enlace que establece una relación entre dos o más topics.

Una vez realizado el análisis de contenido de los textos que forman el corpus se extraerá el nombre de los temas y subtemas (“*basename*”) que los describen de manera más concreta. A continuación se contabilizará el número de ocurrencias de cada uno de los temas y, si es necesario para la investigación, se establecerán los diferentes tipos de

asociaciones existentes entre los diferentes temas. A través de esto se obtendrá un gráfico que represente temáticamente el corpus de sitios web.

### **Propuesta metodológica para la identificación de tipos de actores en una red temática**

Una vez identificada la red temática en cuestión, el siguiente paso será identificar los diferentes tipos de actores que participan en la red. En caso de estudiar fenómenos de opinión pública, entendiéndola siempre en relación con el sistema político, los actores relevantes serán aquellos que defienden una posición determinada respecto al tema general de la red temática que se estudie. Este tipo de actores se caracterizan por llevar a cabo una acción estratégica en la esfera pública (*claims-making*) (Koopmans y Statham, 1999). Este tipo de acción consiste en la expresión de una opinión política a través de una acción física o verbal independientemente de la forma de esta acción – comunicados, violencia, represión, decisión, manifestación, norma jurídica, etc.– o del tipo de actor que la lleva a cabo –gobiernos, movimientos sociales, organizaciones no gubernamentales, individuos, actores anónimos, etc.–. Los defensores de una postura política pueden aparecer por sus propios medios en Internet, es decir, a través de sus propios sitios web, o pueden hacerse visibles a través de su aparición en los sitios web de otros actores (Koopmans y Zimmermann, 2003). En función de la forma utilizada para hacerse visible en Internet, los actores políticos se clasifican en:

- 1) Actores con presencia autónoma en la WWW: el actor político es a la vez autor del sitio web y autor del texto que contiene la postura política defendida.

- 2) Actores con presencia dependiente en Internet: el actor político depende de otros actores para hacer visible públicamente su postura política.

En cualquiera de los dos casos anteriores existen además cinco tipos de actores en función de su papel dentro de una página web:

- 1) *Gate*: el actor que dirige el sitio web que proporciona acceso a otro sitio web o texto a través de un hyperlink externo.
- 2) Medio: el actor que dirige la página web que contiene la información relevante.
- 3) Fuente: el actor que publicó originalmente el texto que contiene la información relevante.
- 4) Autor: el actor que escribe el texto.
- 5) Actor político: el actor cuya opinión política se cita en el texto.

Según el tipo de red temática que se estudie, se utilizará esta clasificación o se deberá atender a otros criterios que aporten información más relevante para los objetivos del estudio. Así, si se estudia la red temática relacionada con los alimentos transgénicos, sería interesante identificar los diferentes tipos de actores atendiendo al tipo de organización en que se inscriben, como organizaciones no gubernamentales, empresas y entidades públicas y gubernamentales. En caso de estudiar la red temática de las elecciones catalanas sería interesante identificar los diferentes tipos de actores atendiendo a la formación política en que se inscriben o al medio que publica la información. Para poder llevar a cabo esta identificación y clasificación, se deberán utilizar diferentes métodos. En el caso de los alimentos transgénicos probablemente será



suficiente analizar el tipo de dominio que aparece en las direcciones URL de cada página. El dominio “.com” permitirá identificar las empresas, el dominio “.org”, las organizaciones no gubernamentales y el dominio “.gov” las entidades gubernamentales. En el caso de estudiar la red temática de las elecciones catalanas deberá analizarse el contenido de la página pero generalmente el grupo parlamentario será identificable por el nombre que aparece en el dominio. Así, la página del Partit Socialista Català será fácilmente identificable por su dominio: [www.socialistes.cat](http://www.socialistes.cat). En caso de que el dominio no nos permita identificar el subgrupo al que pertenece deberá analizarse el contenido del sitio web siguiendo las técnicas que se han propuesto en el apartado 4.2.

### **Propuesta metodológica para la representación gráfica de redes temáticas**

Una vez obtenidos todos los datos que describen las redes temáticas que se estudian, tanto respecto a los actores que las componen, como a los temas, subtemas y/o diferentes marcos de referencia que aparecen en esas redes, será necesario –o al menos aconsejable– poder visualizar todos los datos extraídos a través de algún tipo de representación gráfica que facilite su comprensión y nos permita monitorizar y hacer visible la evolución dinámica de la red. Dependiendo de la investigación que se lleve a cabo se elegirá el modelo de representación que se adapte mejor a las necesidades del estudio y a las características de la red. A continuación se repasan los diferentes modelos de representación de redes que ya se utilizan en la investigación en ciencias sociales y que propongo utilizar para el estudio de los procesos de formación de la opinión pública y la comunicación política en Internet.

Como se explicaba en apartados anteriores, la visualización de una red temática puede

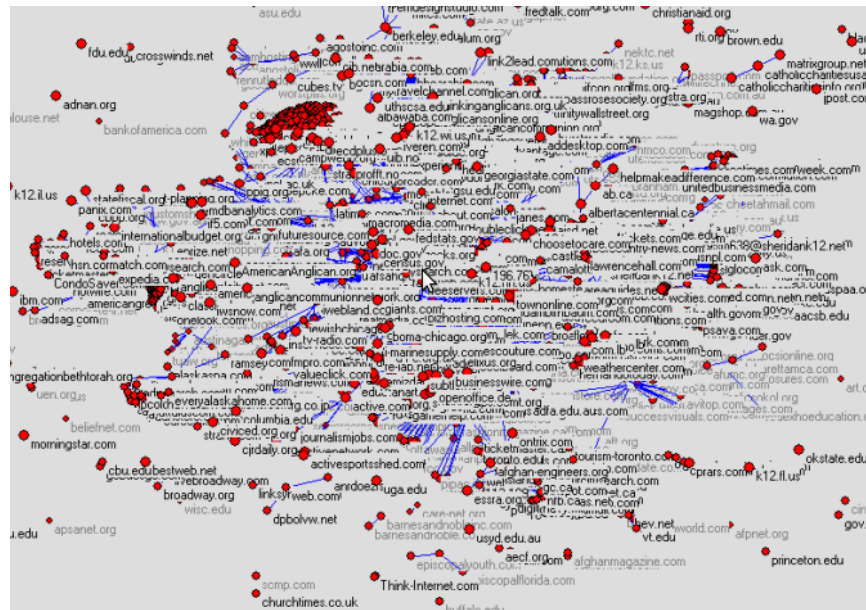
presentar inconvenientes si no se llevan a cabo procesos de refinado de los datos a través de los cuales queremos representar una red. En primer lugar, debe tenerse en cuenta el grado de profundidad con que el programa de rastreo lleva a cabo la extracción de las páginas que presentan enlaces entre sí. En un rastreo de primer nivel, se recuperan únicamente las páginas que están enlazadas directamente entre sí mientras que en un rastreo de segundo nivel se recuperan las páginas que enlazan con las páginas seleccionadas como semillas a través de una segunda página. El nivel de profundidad con que el *crawler* rastrea la *web* se puede configurar y puede llegar al nivel de profundidad que sea más conveniente para la investigación.

*Figura 9: Estructura de conectividad organizada por niveles de rastreo.*

*Extraído de Ackland y Gibson (2005:26)*

Si se quiere representar una red temática deberán seleccionarse primero los sitios web más relevantes para el tema en cuestión, que actuarán como semillas desde las cuales un rastreador extraerá las páginas que presentan algún hyperlink con estas páginas de origen.

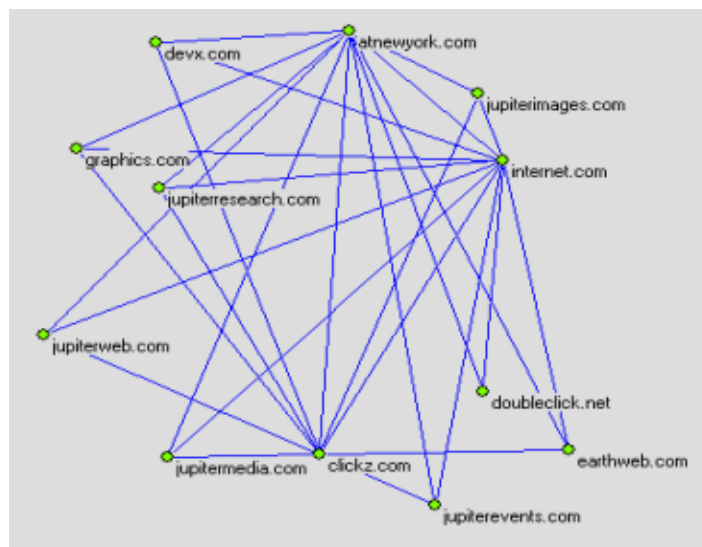
Figura 10: Representación gráfica de una red temática sin refinar.



Extraído de Lee (2006: 85)

Parece claro que un gráfico de este tipo no facilita mucho la comprensión de los datos extraídos a través del rastreador utilizado. Por este motivo es necesario llevar a cabo procesos de refinado de los datos que permitan visualizar únicamente las partes de la red temática que presentan un grado de cohesión superior.

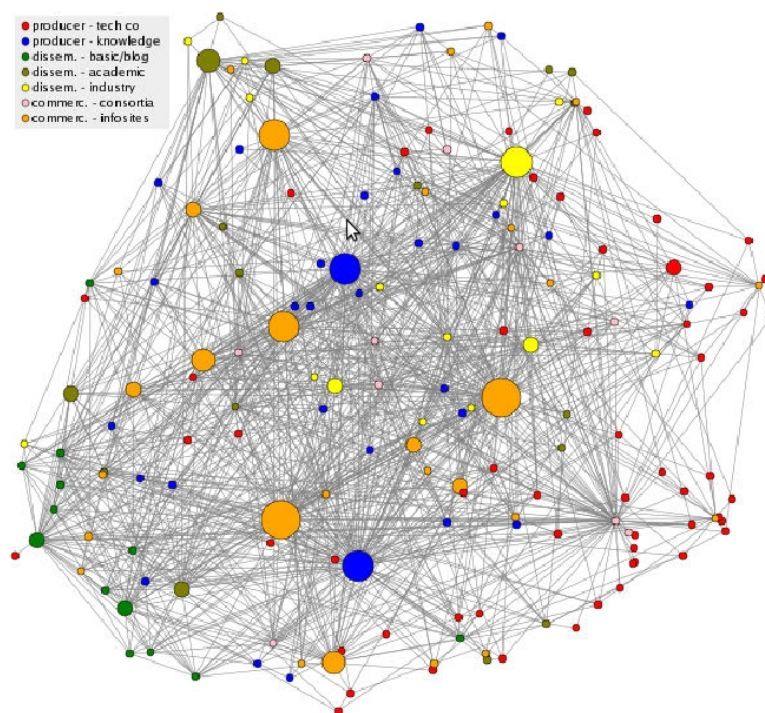
Figura 11: Representación de una Subred temática extraída de la figura 10 tras un proceso de refinación.



Extraído de Lee (2006: 85).

En determinados casos –como por ejemplo el estudio de la comunicación política europea (Koopmans y Zimmermann 2003) o el análisis de una tema concreto como la comunicación relacionada con la nanotecnología (Ackland et. al. 2007)– puede resultar interesante visualizar el tipo de actores que componen una red de manera que podamos ver el número de uno u otro tipo y el nivel de visibilidad que tiene cada actor dentro de la red. En estos casos propongo utilizar una representación del siguiente tipo en la que se asigna un color a cada tipo de actor y el tamaño del nodo que lo representa depende del número de enlaces entrantes con que cuenta:

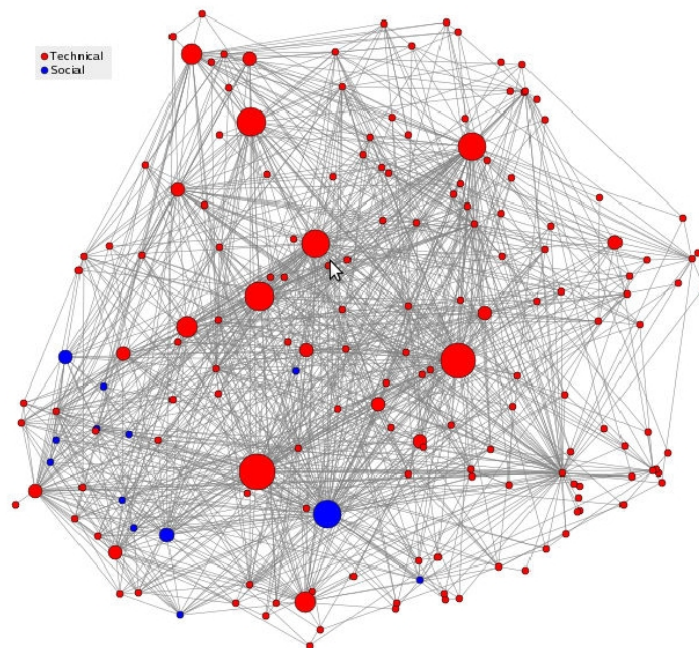
*Figura 12: Red temática en función del tipo de actores.*



*Extraído de Ackland et al. (2007: 23).*

En caso de estar interesados en la representación de los actores que componen una red en función de la postura o marco de referencia que defienden respecto a un tema común propongo una representación del tipo que se muestra en la figura 13. En este caso, cada color representa una postura diferente respecto al mismo tema diferenciando los actores que abordan el tema de la nanotecnología desde un marco de referencia tecnológico y los que abordan el tema desde un marco de referencia ético o filosófico. Igual que en el ejemplo anterior, el tamaño del nodo que representa a cada actor depende del número de enlaces entrantes con que cuenta:

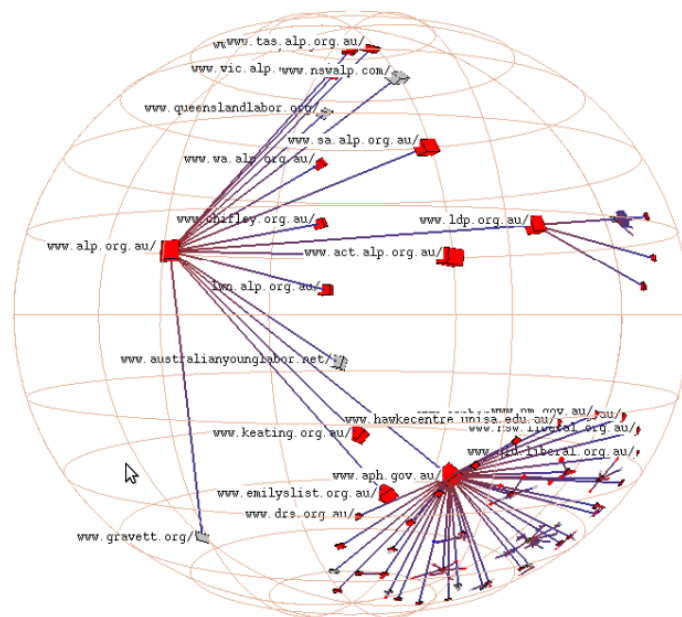
*Figura 13: Red temática en función del punto de vista de los actores respecto al tema de la nanotecnología.*



*Extraído de Ackland et al (2007: 24).*

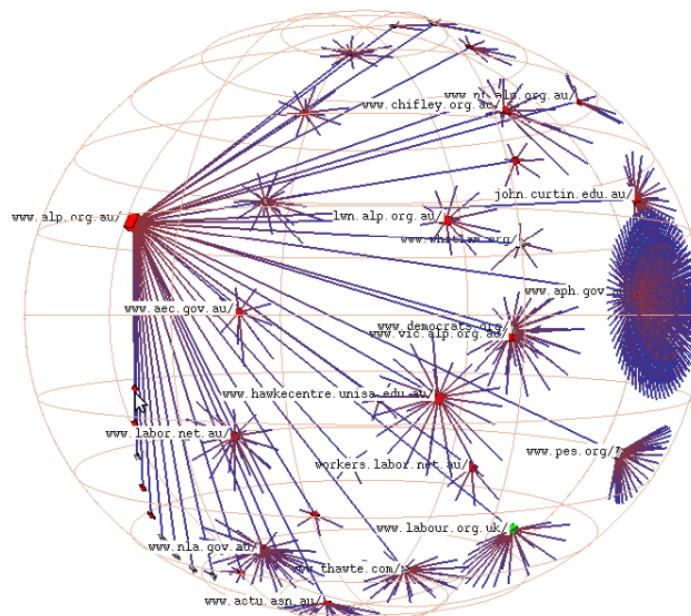
Una vez hemos sido capaces de analizar una red desde una perspectiva macro, identificando tipos de actores, y los subtemas y/o marcos de referencia, puede ser interesante analizar con mayor detalle unos de los actores que participa en la red.

*Figura 14: H3Viewer cybermap de enlaces entrantes al sitio ALP.*



*Extraído de Ackland y Gibson (2005: 11).*

*Figura 15: H3Viewer cybermap de enlaces salientes del sitio ALP.*

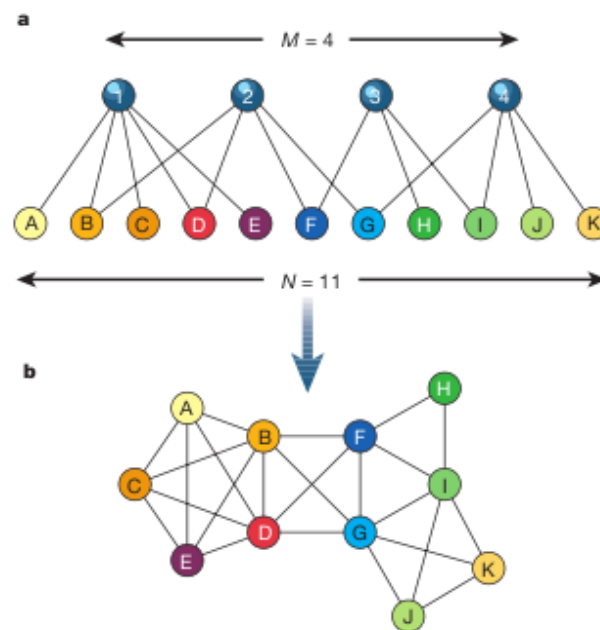


*Extraído de Ackland y Gibson (2005: 10).*

Por último, quiero proponer un tipo de representación de redes que a pesar de no utilizarse mucho en las investigaciones en ciencias sociales considero muy prometedor para analizar los procesos de formación de la opinión pública y la comunicación política en Internet desde una perspectiva temática. Las redes bipartitas permiten representar dos tipos de información diferente a través de dos grupos de nodos diferenciados en un mismo gráfico. Este tipo de representación permite visualizar en un mismo gráfico la estructura temática de una red y la estructura de actores que la componen. En la figura 16 que se presenta a continuación se analiza una red de 11 directivos teniendo en cuenta en cuál de las 4 secciones que componen la empresa trabaja cada uno. Los enlaces indican qué persona trabaja en qué sección siendo posible visualizar –y esta es otra de las ventajas de las redes bipartitas– los posibles solapamientos de actores que trabajan en más de una sección. Este tipo de representación supone sin embargo la pérdida de información relativa a las relaciones que existen entre los componentes de cada uno de los grupos de información. No podemos ver las relaciones existentes entre los distintos directivos ni las relaciones entre las diferentes secciones de la empresa. Para superar esta limitación es necesario completar el gráfico con una representación simple de cada una de las redes que componen la red como lo haríamos al representar un cuerpo sólido a través de la representación de su planta y su alzado.



Figura 16: Representación de red bipartita.



Extraído de Strogatz (2001: 274)

Propongo utilizar este tipo de gráfico para representar de manera conjunta los diferentes temas en los que un actor o actores presentan su opinión en Internet. Esta representación bipartita permite representar las relaciones que existen entre los principales temas y los actores que intervienen en cada tema; pudiendo además utilizar una representación simple de cada uno de los dos grupos que componen la red, temas y actores, para analizar con mayor detalle las relaciones que existen dentro de cada grupo; temas, subtemas y/o marcos de referencia, en la parte de la red que corresponde a los temas; y tipos de actores en la parte que corresponde a los actores. No propongo ningún software concreto para representar este tipo de redes porque todavía no se han desarrollado lo suficiente. No obstante, existen proyectos en este sentido –como *Visual Complexity* creado por Lima (2005)– que pretenden ofrecer recursos a todos aquellos investigadores interesados en la visualización de redes complejas.



## **Propuesta metodológica para la evaluación del nivel de visibilidad de una red temática.**

La evaluación del nivel de visibilidad online de una red temática resulta imprescindible para llevar a cabo estudios en los que se evalúa la correlación entre el nivel de visibilidad de un tema y su emergencia en diferentes medios de comunicación. Como exponía en el apartado 2.4, se entiende por visibilidad el potencial para comunicar opiniones y con ello influir en otros actores. En el espacio comunicativo de Internet esta capacidad de comunicación depende de dos procesos diferentes de selección de la información: los procesos de selección verticales en los que los buscadores realizan una acción de *gatekeeper* y los procesos horizontales de selección de la información dominados por los enlaces hyperlink. En el caso de los procesos verticales de selección de la información existen herramientas eficientes para evaluar el nivel de visibilidad de un tema como las que aportan los propios buscadores –Google, dispone de potentes herramientas como GoogleTrends o GoogleInsight que ofrecen datos estadísticos sobre el número de visitas que se han producido a lo largo de un periodo de tiempo determinado en relación a una frase de búsqueda concreta–. Esta información –el número de visitas que se producen en relación con un tema– se hace mucho más difícil de obtener cuando lo que quiere conocerse es el número de visitas que se producen a través de enlaces horizontales hyperlink.

El método que propongo en este trabajo para la evaluación del nivel de visibilidad de una red temática surge de la aplicación de la teoría de análisis de redes sociales que se ha explicado en el apartado 2.4. Así, se evaluará el tamaño de una red temática teniendo en cuenta el número de actores que componen dicha red. Una vez evaluado el tamaño de la red temática se procederá a analizar la fuerza de las relaciones que se producen en

dicha red. Para evaluar esta fuerza, se medirá la densidad de los enlaces, es decir, el número de enlaces que hay entre los diferentes actores que componen la red contabilizando por separado la cantidad de enlaces entrantes, salientes y recíprocos que aparecen.

## **5.- Estructura y principales conceptos de la investigación**

A continuación expongo la estructura y los contenidos de la investigación que se pretende desarrollar y una descripción de los principales conceptos sobre los que se sustenta dicha investigación acompañada de los principales conceptos de investigación.

### **Estructura y contenidos**

- 1) Descripción del tema y justificación de su interés. Se presentará el tema de la investigación y una justificación de su interés dentro del campo de la comunicación social. Para ello utilizaré, con mínimas modificaciones, la presentación que se ha hecho en el apartado 1.
- 2) Estado de la cuestión relativo al tema de la investigación. En este segundo apartado se estudiarán los modelos teóricos relacionados con los procesos de formación de la opinión pública y la comunicación política en Internet. Dentro de este marco teórico se seguirá el mismo guión utilizado en el apartado 2 y se prestará especial atención a los siguientes campos de investigación:
  - 2.1) Teorías de investigación en comunicación social basadas en la organización temática de la opinión pública.
  - 2.2) Teoría de redes complejas aplicada al ámbito de los estudios sociales y especialmente de los procesos de formación de la opinión pública y la comunicación política en el nuevo sistema comunicativo.

2.3) Estudios de visibilidad y otros procesos de selección de la información que determinan la capacidad de influencia en el ámbito de las redes sociales online.

- 3) Problema de investigación concreto. El problema en que se centrará la investigación futura es la identificación de patrones en los procesos de formación de la opinión pública y la comunicación política en Internet. Para profundizar en este problema se plantearán una serie de preguntas de investigación que, con la posibilidad de algunos cambios, coincidirán con las preguntas que se han planteado en el apartado 3.
- 4) Metodología. La metodología utilizada para resolver el problema de investigación es la que se ha desarrollado en el apartado 4, prestando especial atención al análisis de redes temáticas en Internet.
- 5) Desarrollo del trabajo experimental. Para proceder a la aplicación de la metodología expuesta en este trabajo propongo el estudio de diferentes procesos de formación de la opinión pública en Internet. Como ya he planteado, en primer lugar se evaluará la validez de la metodología en el estudio del proceso electoral de las próximas elecciones a la Generalitat de Catalunya de 2010. Una vez confirmada esta metodología se aplicará al estudio de otros procesos de formación de la opinión pública como, por ejemplo, el estudio de redes temáticas surgidas ante hechos puntuales de la actualidad informativa o el estudio de redes temáticas relacionadas con temas de interés público como el aborto o el uso de los transgénicos por poner algunos posibles ejemplos.
- 6) Resultados y conclusiones. Se irán presentando los resultados de las diferentes investigaciones a medida que estas se vayan desarrollando para concluir con

unas conclusiones comunes a todos los procesos de formación de la opinión pública estudiados que sirvan de conclusión final de la tesis doctoral.

#### 7) Bibliografía.

### **Conceptos principales**

Los principales conceptos sobre los que se construirá la investigación futura son los que ya se han planteado en el estado de la cuestión que se presenta en esta investigación.

Estos son:

- 1) El concepto de tematización estudiado en el apartado 2.1. con especial atención a su posible aplicación en un sistema comunicativo dominado por las nuevas tecnologías
- 2) El concepto de emergencia temática estudiado en el apartado 2.3. aplicando para ello un marco teórico relacionado con la teoría de sistemas complejos.
- 3) El concepto de visibilidad estudiado en el apartado 2.4. prestando atención especial a la teoría de redes sociales y a los métodos estadísticos aplicados a las ciencias sociales.

## 6.- Bibliografia

ACKLAND, R. (2005). "Mapping the U.S. Political Blogosphere: Are Conservative Bloggers More Prominent?". *BlogTalk Downunder*. Conference, Sydney. Accesible en <http://hdl.handle.net/1885/45827> [03/06/2010].

ACKLAND, R. y GIBSON, R. K. (2006) "Hyperlinks and horizontal political communication on the WWW: the untold story of parties online". Disponible en: [http://voson.anu.edu.au/papers/hyperlinks\\_polcomm.pdf](http://voson.anu.edu.au/papers/hyperlinks_polcomm.pdf) [02/06/2010].

ACKLAND, R. et al. (2007). "Mapping 'small things' on the Web: Assessing the online presence of the nanotechnology industry". *Annual Conference of the International Communication Association*, San Francisco, Mayo 24-28.

AGOSTINI, A. (1984). "La tematizzazione. Selezione e memoria dell'informazione giornalistica". *Problemi dell'Informazione*. 9(4), 51-63.

ALLPORT, F.H. (1937). "Towards a science of public opinion". *The Public Opinion Quarterly*. 1(1), 7-23.

ALTHAUS, S. L. y TEWKSBURY, D. (2002). "Agenda setting and the "new" news: Patterns of issue importance among readers of the paper and online version of the New York Times". *Communication Research*, 29(2), 180-207.

BADIA, L. (1992). *De la persuasió a la tematizació: Introducció a la comunicació política moderna*. Barcelona: Portic.

BARABÁSI, A.L. Y Albert, R. (1999). "Emergence of scaling in random networks". *Science*, 286, 509-512.

BARABÁSI, A. L. (2002). *Linked: The new science of networks*. Cambridge, Massachusetts: Perseus Publishing.

BAVELAS, A. (1948). "A mathematical model for group structure". *Applied Anthropology* 7:16-30.

BERRIO, J. (2002). "L'Opinió Pública". *Portal de la Comunicació Incom-UAB*. Disponible en: [http://www.portalcomunicacion.com/ESP/pdf/aab\\_lec/5.pdf](http://www.portalcomunicacion.com/ESP/pdf/aab_lec/5.pdf) [02/06/2010].

BERRIO, J. (2001). "La democrazia d'opinione e l'autoreferenzialità dei media: Il punto di vista di Carlo Marletti". *Analisi : quaderns de comunicació i cultura*. Bellaterra: Servei de Publicacions de la Universitat Autònoma de Barcelona, 26, 189-206.

- BLUMLER, J. G. (2001). "The third age of political communication". *Journal of public affairs*, 1, (3), 201-209.
- BONCHEK, M.S. (1997). "From broadcast to netcast: The Internet and the flow of political information". Doctoral Dissertation. *Harvard University*.
- BURNETT, R. y MARSHALL, P. (2002). *Web theory: An introduction*. London: Routledge.
- CHAFFEE, S. H. y METZGER, M. J. (2001). "The end of mass communication?". *Mass communication and society*, 4(4), 365–379
- CODINA, L. (2010). "El futur de la informació acadèmica: ¿web semántica, web social o totes dues?" *La Terminologia i la Documentació: Relacions i Sinergies*: Actes de la VII Jornada de la Scaterm. Barcelona: Societat Catalana de Terminologia.
- DREZE, X. y ZUFREYDEN, F. (2004). "The Measurement of online Visibility and its Impact on Internet Traffic". *Journal of Interactive Marketing*, 18(1), 20-37.
- EGUÍLUZ, V.M.; TORAL, R. y SAN MIGUEL, M. (2005). "Redes complejas en la dinámica social". *Inguruak, Revista Vasca de Sociología y Ciencia Política*, 42, 127-146.
- EGUÍLUZ, V.M. Y ZIMMERMANN, M.G. (2000). "Transmission of Information and Herd Behavior: An Application to Financial Markets". *Physical. Review. Letters*, 85(26), 5659–5662.
- ERDOS, P. Y REYNI, A. (1960). "On the evolution of random networks". *Mathematical Institute of Hungarian Academic Science*, 5, 17-61.
- FREEMAN, L.C. (1978). "A Set of Measures of Centrality Based on Betweenness". *Sociometry*, 40(1), 35-41.
- FOOT, K. A. y SCHNEIDER, S.M. (2002). "Online action in campaign 2000: An explanation analysis of the U.S. political web sphere". *Journal of Broadcasting and Electronic Media*, 46(2), 222-244.
- FUCHS, C. (2006). "The self organization of social movements". *Systemic Practice and Action Research*, 19(1), 101-137.
- GANTER, B. y WILLE, R. (1999). *Formal Concept Analysis. Mathematical Foundations*. Berlin, Springer .
- GÄRDENDORFS, P. y WILLIAMS, M. (2001). "Reasoning about Categories in Conceptual Spaces". *Proceedings of the International Joint Conference on Artificial Intelligence*.
- GIRARDIN, F. (2009). "Aspects of implicit and explicit human interactions with ubiquitous geographic information". PhD thesis, *Universitat Pompeu Fabra*.

- GRANOVETTER, M. (1973). "The strength of weak ties". *American Journal of Sociology*, 78(6), 1360-1380.
- GRUBER, T. (2008). *Encyclopedia of Database Systems*, Ling Liu y M. Tamer Özsu (Eds.), "Ontology", Berlín: Springer-Verlag.
- HAKEN, H. (2006). *Information and self-organization: A macroscopic approach to complex systems*. Berlín: Springer-Verlag.
- HENZINGER, M. (2001). "Hyperlink Analysis for the Web" *IEEE Internet Computing* 5(1): 45-50.
- HINE, C. (2000). *Virtual Ethnography*. London: Sage.
- IGARTUA, J. J. y HUMANES, M. L. (2004). *Teoría e investigación en comunicación social*. Madrid: Síntesis.
- JACKSON, M.H. (1997). "Assessing the structure of communication on the world wide web". *Journal of Computer-Mediated Communication*, 3(1). Disponible en <http://jcmc.indiana.edu/vol3/issue1/jackson.html> [02-06-2010].
- JOHNSON, S. ([2001] 2003). *Sistemas emergentes: o qué tienen en común hormigas, neuronas, ciudades y software*. Madrid: Turner.
- KLEINBERG, J.C. (1999). "Authoritative Sources in a hyperlinked Environment". *Journal of the ACM*. 46 (5), 604 – 642.
- KOOPMANS, R. Y STATHAM, P. (1999). "Political claims analysis: Integrating protest event and political discourse approaches". *Mobilization: An International Quarterly*, 4(2), 203-221.
- KOOPMANS, R. y ZIMMERMANN, A. (2003). "Internet: A New Potential for European Political Communication?". *Discussion Paper SP IV 2003-402*. Wissenschaftszentrum Berlin für Sozialforschung (WZB).
- KOOPMANS, R. y ZIMMERMANN, A. (2007). "Visibility and communication networks on the Internet: the role of search engines and hyperlinks". En: *European Public Sphere: How much of it do we have and how much do we need?*. Mannheim: University of Mannheim. 213-264.
- KU, G.; KAID, L.; PFAU, M. (2003) "The Impact of Web Site Campaigning on Traditional News Media and Public Information Processing", *Journalism and Mass Communication Quarterly* 80(3): 528–47.
- LANG, K. y LANG, G. (1981). "Watergate. An exploration of the Agenda Building Process". *Mass Communication Review Yearbook*, 2, 441-456.
- LEE, J. (2006). "Computer-mediated communication as political communication: investigating the agenda-setting function". *Florida State University*, [Tesis doctoral]



inédita].

LIMA, M. (2005). "Visualcomplexity" [<http://www.visualcomplexity.com/vc/>]  
Accedido en junio de 2010]

LIPPMANN, W. ([1922] 2003). *La opinión pública*. San Lorenzo del Escorial: Langre.

LOPEZ GARCÍA, G. (2006). "Comunicación en red y mutaciones de la esfera pública". *Zer*, 20, 231-249 .

LUHMANN, N. (1968). *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*, Stuttgart: Enke (English translation: *Trust and Power*, Chichester: Wiley, 1979.)

LUHMANN, N. (1974). *Öffentliche Meinung*. In: [Wolfgang R. Langenbucher](#) (Hrsg.): *Zur Theorie der politischen Kommunikation*. Piper, München 27–54.

LUSHER, D.; y ACKLAND, R. (2009). "A relational hyperlink analysis of an online social movement: Asylum seeker advocacy groups in Australia". *Journal of Social Structure*.

MARLETI, C. (1985). *Prima e dopo. Tematizzazione e comunicazione politica*. Torino: Eri.

MCCOMBS, M.E. y SHAW, D.L. (1972). "The Agenda-Setting Function of Mass Media". *Public Opinion Quarterly*, 36, 176-187.

MIKA, P. (2007). "Ontologies are us: A unified model of social networks and semantics". *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 5, 5–15.

NAZAR, R. (En prensa). "A Quantitative Approach to Concept Analysis". Trabajo de tesis doctoral en curso. *Universitat Pompeu Fabra*.

NOELLE-NEUMANN, E. (1979). "Teoria e ricerca sui mass media in Europa e in America. L'influenza dei mass media" *Problemi dell'Informazione*, 4(3).

NOVAK, J. y CAÑAS, A. J. (2006). "The Theory Underlying Concept Maps and How To Construct Them". *Technical Report IHMC*, Florida Institute for Human and Machine Cognition.

PARK, H.W. (2003). "Hyperlink Network Analysis: A New Method for the Study of Social Structure on the Web". *Connections* 25(1): 49-61.

PARK, H.W. y THELWALL, M. (2008). "Link analysis: hyperlink patterns and social structure on politicians Web sites in South Korea". *Quality and Quantity*. Springer Netherlands, 42(5), 687-697.

- PARK, J. y HUNTINGS, S. (2003). *XML Topic Maps: creating and using topic maps for the Web*, Boston: Addison-Wesley cop.
- PERICOT, J. y CAPDEVILA, A. (Ed) (2009). *L'espectre del tripartit*. Girona: Documenta Universitaria.
- QUILLIAN, M. (1968). "Semantic Memory". *Semantic Information Processing*, 227-270, MIT Press.
- ROBERTS, M.; WANTA, W.; DZWO, T. H. (2002). "Agenda setting and issue salience online". *Communication Research*, 29, 452-465.
- ROGERS, R. (2002). "Operating issue networks on the web". *Science as Culture*, 11(2), 191- 213.
- SAMPEDRO, V. (2000). *Opinión pública y democracia deliberativa*. Medios, sondeos y urnas. Madrid: Istmo.
- SAPERAS, E. (1987). *Los efectos cognitivos de la comunicación*. Barcelona: Ariel.
- SCHELLING, T. (1971). "Dynamic models of segregation". *Journal of Mathematical Sociology*, 1, 143-186.
- SHAW, D. L. et al. (1999). "Individuals, groups, and agenda melding: a theory of social dissonance". *International Journal of Public Opinion Research*, 11(1), 2-24.
- SHAW, D. L.; STEVENSON, R. L. y HAMM, B. J. (2001). "Agenda Setting Theory and Public Opinion Studies in a Post-Mass Media Age". *WAPOR Annual Conference*. Roma, septiembre de 2001.
- SCHMIDT-MANZ, N. y GAUL, W. (2005). "Web mining and online visibility". *Classification, the ubiquitous challenge: 28th Annual Conference of the Gesellschaft. Klassifikation*. University of Dortmund.
- SOLÉ, R. (2009). *Redes complejas: del genoma a Internet*. Barcelona: Tusquets.
- SOLOMON, S y SHIR, E. (2003). "Complexity; a science at 30". *Europhysics News*, marzo-abril, 54-57.
- SONG, Y. (2007). "Internet news media and issue development: a case study on the roles of independent online news services as agenda-builders for anti-US protests in South Korea. *New Media Society*, 9(1), 71-92.
- SOWA, J. (1991). *Principles of Semantic Networks*. San Mateo: Morgan Kaufmann Publishers.
- STROGATZ, S.H. (2001). "Exploring complex networks". *Nature*, 410(6825), 268-276.

THELWALL, M.; PRABOWO, R. y FAIRCLOUGH, R. (2006). "Are Raw RSS Feeds Suitable for Broad Issue Scanning?: A Science Concern Case Study". *Journal of the American Society for Information Science and Technology*, 57(12), 1644-1654.

VALLEZ, M. et al. (2010). "Procedures for extracting keywords from web pages, based on search engine optimization". *Hipertext.net*, 8, Disponible en: [http://www.upf.edu/hipertextnet/en/numero-8/keywords\\_extraction.html](http://www.upf.edu/hipertextnet/en/numero-8/keywords_extraction.html)

WATTS, D. y STROGATZ, S.H. (1998). "Collective dynamics of 'small-world' networks". *Nature*. 393(6684), 440-442.

WATTS, D. (2003). *Six degrees: The science of a connected age*. New York: Norton.

WILKINSON, D.; THELWALL, M. y LI, X. (2003). "Exploiting hyperlinks to Study Academic Web Use." *Social Science Computer Review* , 21(3): 340-351.