2 Tracing Practices within a Web Sphere

In this chapter, we explain further the notion of a Web sphere introduced in chapter 1. We discuss several dimensions of Web spheres, as well as methods for demarcating them and identifying the constituent elements within them. After introducing the set of techniques through which we analyze Web campaigning practices, we then lay out the specific approach we employ to analyze campaign Web practices as they occur within electoral Web spheres. We describe our methods of collecting and observing Web objects, and the interviews, focus groups, and surveys we conducted with campaign site producers and Web users. We conclude the chapter with an overview of the data corpus we compiled in our research on Web campaigning in the 2000, 2002, and 2004 electoral Web spheres in the United States.

Web Sphere Analysis

As explained in chapter 1, we conceptualize a Web sphere as not simply a collection of Web sites, but as a set of dynamically defined, digital resources spanning multiple Web sites deemed relevant or related to a central event, concept, or theme. In this book, we employ the electoral Web sphere as a framework within which we focus on the Web practices of electoral campaigns. Web sphere analysis is a conceptual lens for this and other studies of Web phenomena, enabling analysis of communicative actions and relations between Web producers and users developmentally over time. Web sphere analysis provides an approach for investigating relations between producers and users of Web materials as potentiated and mediated by the structural and feature elements of Web sites, hypertexts, and the links between them.

In the first stage of the multimethod approach of Web sphere analysis, researchers identify Web objects (usually either sites or pages) related to the topical focus of the sphere. These objects are archived at regular intervals during the period of the study in a manner that maintains their hyperlinked context, allowing the researchers to reproduce at some point in the future the experience users might have had (Arms et al. 2001). This enables both contemporaneous and retrospective analyses of developments in the Web sphere over time. Then metadata are created through the annotation of archived objects (again, usually sites or pages) with human or computer-generated notes and/or codes of various kinds. These metadata correspond to the units and levels of analysis anticipated by the researchers. Finally, interviews, focus groups, experiments, or surveys are conducted with site producers and Web users and triangulated with Web annotations and objects to interpret the sphere.

December's (1996) typology of units of analysis for Internet-related research is useful for understanding the nature of a Web sphere as a unit of analysis. The five types of units of analysis December identifies are:

(1) a media space, consisting of the set of all servers of a particular type that may provide information in one or more protocols, the corresponding clients that are capable of accessing these servers, and the associated content available for access on these servers;

(2) a media class—a particular set of content, servers, and clients;

(3) a media object—a specific unit in a media class with which the user can observe and interact;

(4) a media instance—a media object at a particular time; and

(5) a media experience—a particular user's perception of a set of media instances.

In correspondence with December's definitions, a Web sphere could be considered a subset of an Internet media space, constituted by a set of Web sites (a single media class), each of which are composed of elements or objects such as links, features, and texts. Boundaries based on a thematic or event orientation and a temporal framework differentiate our concept of a Web sphere from December's definition of a Web space.

The Web sphere can function as a macro, aggregate unit of analysis, by which historical and intersphere comparisons can be made. For example,

the Web sphere of the 2000 elections in the United States can be comparatively analyzed with the electoral Web sphere of 2004 and those that will develop in later years. Similarly, comparative analyses can be conducted on electoral Web spheres internationally. A significant element in our conceptualization of a Web sphere is the dynamic nature of the sites that constitute the sphere. This dynamism comes from three sources. First, researchers involved in identifying the boundaries of the sphere are likely to continually find additional sites to be included within it. Second, new sites related to the topic of the sphere may be produced as it develops. Third, the process of defining a Web sphere is recursive, in that pages that are referenced by included sites, as well as pages that reference included sites, may be considered part of the sphere under evaluation. Thus, as a Web sphere is analyzed over time, its boundaries may be dynamically shaped by researchers' identification strategies, the production of new sites, and changes in the sites themselves.

The more micro and/or molar units—such as texts, features, links, sites, or even the multisite Web presence of an actor—can be employed in analyses simultaneously within a Web sphere. Defining any of these units operationally can be challenging, particularly when the temporal and malleable aspects of Web objects are considered. For example, any Web text or feature can appear stable but actually be modified by its producer or rendered differently by technologies such as Web browsers employed by users at a particular moment. Therefore, the point in time and the way in which a Web object is observed must be part of the unit's definition for research purposes. Units such as an actor's Web presence must also reflect the potential for change over time by being situated in a particular temporal period. For instance, the Web presence of a political party might be appropriately specified by the particular week or month within an election cycle. The hyperlinked and multilevel nature of the Web makes the identification and demarcation of units of analysis a critical but difficult task, even within a Web sphere. Seemingly straightforward questions, such as what constitutes a Web site and from what or whose perspective (robot, browser, or human) that question will be framed, require careful consideration. In addition, the coproduced nature of the Web—evidenced in the joint production by multiple actors of many features and much content—makes the attribution of agency to producers of specific bits problematic.

Bounding Web Spheres

In Web sphere analysis, the process of demarcating boundaries, that is, identifying elements such as sites or pages to be included, is part discovery and part creation. Three dimensions of Web spheres have bearing on how researchers go about this process: thematic anticipatability, actor predictability, and the stability of constituent Web objects. These three dimensions can each be viewed as a continuum. The position of a Web sphere of interest on these three continua may help researchers develop strategies for identification of elements to be included in the analysis.

The degree to which it is possible to anticipate the emergence of a Web sphere is the measure of *thematic anticipatability*. Some Web spheres can be easily anticipated in view of prior and current dynamics on the Web, and others cannot. Part of the anticipatability of a Web sphere is dependent on the extent to which it is defined by a specific event. Events such as elections and the Olympics are regular occurrences and thus anticipated, while others, such as accidents, tragedies, and scientific discoveries, are less likely to be anticipated. Triggering events (Gamson and Modigliani 1989) such as major new discoveries, which in general are unanticipated, may also provide the stimulus for a thematic Web sphere. For example, a Web sphere focused on cloning may have emerged in response to the announcement of Dolly, the first cloned mammal. Web spheres that emerge in the aftermath of an event may be triggered by but not focused on that specific event. In the absence of generalized and systematic Web archiving (Kahle 1997; Schneider et al. 2003), anticipatability is often a crucial factor in whether a Web sphere is researched. Web spheres emerging quickly after an unanticipated event may be challenging to study because a rapid investment of resources (time, money, topical expertise) is required and the boundaries are difficult to demarcate.

Actor predictability is concerned with the ability of researchers to predict the types of actors who will produce materials encompassed within a Web sphere in advance of its emergence. Some Web spheres will be produced by a highly predictable set of actor types. Web spheres organized around electoral campaigns, for example, are highly actor-predictable because they may (depending on the localized political context) include sites produced by parties, candidates, press organizations, advocacy groups, citizens, and government agencies. Web spheres emerging around unanticipated natural

disasters and accidents are likely to include sites produced by a predictable set of actors: government agencies, relief and charity organizations, press organizations, and citizens, for example. Other Web spheres will be produced by a less predictable set of actor types. Following the terrorist attacks of September 11, we observed significant and unpredicted activity on Web sites produced by corporations and businesses, along with more predictable activity on sites produced by religious organizations, educational institutions, and government agencies (Foot and Schneider 2004). Actor predictability can greatly affect the thoroughness with which relevant sites can be identified. A predictable set of actors enables researchers to identify a universe of sites to examine for evidence of Web activity within a demarcated Web sphere. A less-predictable set of actors makes this task more difficult and requires additional searching and identification activities.

A third dimension of Web spheres is the *level of stability* in the development of sites, links, and other objects within the Web sphere. The level of stability has an impact on how frequently the boundaries of the Web sphere need to be reconsidered and how often the sites within the Web sphere need to be examined. We have identified three determinants of the level of stability within a Web sphere. First, consideration should be given to the frequency of entry and exit of new producers. The extent of new entrants into the Web sphere (represented by specific producers, rather than types of producers) and the extent to which producers stop updating, maintaining, or serving their sites, are measures of the stability of the sphere: Highly stable spheres have less entry and exit than unstable spheres. Second, stability is measured by the degree to which sites being analyzed add or remove links to other sites within the Web sphere. Frequent additions or deletions of links reduce stability. Third, the frequency and breadth of changes to content and features within the Web sites under examination contributes to stability. Highly stable Web spheres consist primarily of sites with infrequent and narrowly focused changes to content and features.

The position of a Web sphere along these dimensions influences the researchers' decision either to fix the boundaries of the sphere at the beginning of a study or to engage in a dynamic bounding process. Bounding refers to the process of identifying constituent sites or pages within the Web sphere and specifying a temporal frame for analysis. Although electoral Web spheres can be anticipated and the general categories of actors

involved in coproducing these spheres are anticipatable, electoral Web spheres can be quite unstable. Thus researchers still face significant quandaries regarding the bounding of these spheres in time and virtual space, as we explain later. Identifying constituent elements of a Web sphere to be examined may include both establishing the universe of sites or pages about which generalizations can be offered and specifying a method of sampling sites or pages to be analyzed. Constituent elements can be identified prior to analysis by following long-established practices in survey research (Hyman 1955) and content analysis (Berelson 1952). Alternatively, constituent elements can be identified as part of the analytic process, building on well-established techniques used in participant-observation (Whyte 1943) and on snowball sampling (Berg 1988; Atkinson and Flint 2001). Identifying constituent elements prior to analysis—in effect, fixing the boundaries of the Web sphere under study—offers several advantages to the researcher. A clearly defined universe of sites within a Web sphere makes representative sampling of sites possible for both archiving and structured observations. Fixed boundaries may also increase the possibility of replicating findings by subsequent analysts. Finally, fixing the boundaries of the Web sphere may enhance options for collaboration in archiving or analyzing it, particularly with entities such as libraries (Schneider et al. 2003).

On the other hand, dynamic bounding allows the researcher to be responsive to unanticipated developments and emergent trends in the Web sphere. Even within anticipated, predictable, and generally stable Web spheres, unanticipated events precipitate the production or alteration of intertextual and interlinked Web objects, sometimes in a matter of hours or over the course of a few days. We employ the concept of a *Web storm* as a unit of analysis that reflects unanticipated interactor and intersite activity over a relatively brief period of time. For instance, a political scandal is likely to result in a Web storm wherein news organizations, advocacy groups, and individuals post texts, graphics, and links regarding the scandal intensively for several days or weeks. Some Web storms develop into Web spheres that are durable on the Web over a longer period, often through the migration of individual texts and pages pertaining to an event to sites newly produced and dedicated to the event. For example, a Web storm emerged quickly in the wake of the 1988 release of the Starr report, detailing U.S. President Bill Clinton's affair with Monica Lewinsky and rais-

ing questions of perjury. The Web storm of commentaries developed into a Web sphere focused on the Clinton impeachment, as sites were produced by different types of actors advocating or opposing the removal of the president from office. Interestingly, some sites, such as moveon.org, emerged from this Web sphere and became part of other political Web spheres, including the U.S. electoral Web spheres in 2000, 2002, and 2004.

A researcher engaged in dynamic bounding is less likely to miss the opportunity to analyze these Web storms, which may actually be significant bursts of online action. Dynamic bounding as a scholarly practice is also more consistent with how the Web functions from a user perspective. It is critical, though, that dynamic bounding be implemented systematically to ensure representativeness and replicability. Researchers need to decide, preferably before beginning a study, how frequently the Web sphere boundaries will be redefined, specifying the circumstances, frequency, criteria, and techniques to be used to identify the pages, sites, or links comprising the sphere. A further factor to consider is the correspondence between research goals and the bounding strategy. For instance, if the goal is to analyze the development of a Web sphere, a fairly dynamic bounding strategy is needed. Finally, researchers should assume that increased dynamism in Web sphere demarcation will increase resource needs (time, effort, and storage), especially if systematic archiving is involved.

The process of discovering and establishing the Web sphere under study also includes defining the steps or procedures necessary for identifying the specific elements to be examined. Depending on the characteristics of the Web sphere, this process can involve a number of strategies. If the producer types are highly predictable and the Web sphere itself highly anticipated, current and maintained directories of sites may be available. For example, for a researcher interested in the Web sphere related to a single season of a professional sports league, a current directory of sites representing each of the participating teams could likely serve as a starting point for identification of relevant sites. Encompassing the pre-existing directory into the research design (by specifying sites to be examined as those identified by the directory on a specific date) serves as a robust specification of identification strategy and provides a universe from which samples of sites can be reliably drawn for systematic examination.

On the other hand, some Web spheres of interest, especially those that are not anticipated or predictable, may require a more search-oriented

strategy to identify constituent elements. Using topical key words systematically in search engines may be fruitful in identifying constituent elements of a Web sphere. However, this strategy, which relies on the presence of seemingly relevant content within potential sites, may have significant drawbacks. The absence of relevant content on some Web sites may reflect action on the part of a site producer that is just as strategic as the presence of relevant Web content would be. For example, we would consider a Web site produced by a political party that provides issue information several weeks before an election—but does not mention any candidate or the election itself—a constitutive element of the electoral Web sphere in view of the party's role in the electoral process. In such a case, the absence of election-related content may reflect a strategic aim of general agenda-setting rather than specific candidate promotion. Furthermore, if explicit election-related material were to appear on a particular site at a later date, establishing its absence at the beginning of the study period would be critical for developmental analysis.

One alternative to content relevance as the primary criteria for inclusion in the Web sphere, as suggested by the previous example, is the inclusion of Web sites produced by actors identified as relevant to the sphere. Relevant actor types can be determined in several ways, including through whatever extant literature informs a particular study, as well as through methods of social network analysis. Once relevant actor types are established, the Web sites of particular actors within each type can be selected through various indices and sampling techniques. Web sites produced by relevant actors may be significant in structuring online action—or the lack thereof—within the Web sphere even if those actors have not (yet) produced Web materials relevant to the theme of the sphere at the beginning of the study.

Another strategy for identifying constituent elements in a Web sphere is to analyze patterns of linking to and from a set of URLs representing core sites in the Web sphere. For instance, the Web sphere of a sports team could be defined by tracing the links to and from the URLs of the home pages of each player and of the team itself. The linked-from and linked-to pages can be analyzed in the context of their base sites to identify their producers and thereby specify the producer type. Further analysis of these producers' Web presence may be helpful to ascertain their position or stance within the Web sphere under study.

In summary, the process of demarcating Web spheres includes establishing a thematic or topical orientation, a temporal range, and a method of identifying constituent elements. The dimensions of a Web sphere, expressed as thematic anticipatability, actor predictability, and the stability of constituent Web materials, all shape the challenge faced by researchers interested in conducting Web sphere analyses. Following an overview of some production techniques employed in the creation of online structures within Web spheres, we assess the particular characteristics of U.S. electoral Web spheres.

Web Production Techniques

We analyze Web practices through the techniques that mediate the production of Web objects or artifacts. One baseline Web technique is simply the production of any kind of content in hypertext markup language. The three techniques on which we focus in each of the following chapters are coproduction, convergence, and linking. These three techniques are particularly significant to electoral campaigning, although they are also associated with many practices and activities. Coproduction is the joint production of Web objects, whether of content, features, links, or sites. Convergence is bringing together the online and the offline realms of production, whether through media, organizations, or activity. Linking as a technique of Web production is the establishment of a mechanism that connects one Web object to another, thus enabling a user to transverse from a page produced by the campaign to a page produced by another actor. Variants of each technique are employed broadly across all four Web campaigning practices—informing, involving, connecting, and mobilizing—although not all variants are evident within each practice. General characteristics of these three techniques are described next. Specific variants of these and other techniques employed uniquely within each practice are analyzed in the following chapters.

The technique of coproducing encompasses creating and maintaining Web objects jointly between actors who are organizationally independent from each other. Collaboration between individuals who are co-located within an organization is not our focus in this technique. Evidence of a coproductive process may be textual, visual, or manifested through links. In 2000, for example, the webwhiteandblue.org site was created to provide