Litigators, Wrangle That Data With Content Analysis

By **Lisa K. Tichy and Anna Shakotko**
November 2, 2017, 12:24 PM EDT

In recent years, the quantity of rich data generated by the internet and social media has exploded, leaving courts with an unprecedented volume of potential evidence to consider. A research field that is well-established in academia can help make sense of this wealth of information: the field of "content analysis."

Content analysis is a set of scientific methods that can be introduced as evidence by a qualified expert witness. These methods can be used to systematically and reliably address even apparently nebulous topics such as the tone of consumer comments or the messages communicated by an online video.

Lisa Tichy

What Is Content Analysis and How Is It Performed?

Content analysis is defined as the "systematic, objective, quantitative analysis of message characteristics."[1] A "message" can be as varied as a Twitter post, a news article or a television advertisement, while "message characteristics" can similarly encompass a wide range of features of varying complexity.

In analyzing a set of newspaper articles, for example, one can consider more basic characteristics, such as the length of the article or the number of times a certain word is mentioned, or more complex characteristics, such as the tone of the article, its main theme or the nature of any images that accompany it.[2]

Anna Shakotko

In practice, content analysis can be performed by human coders, who evaluate each piece of content independently and according to well-specified rules, or by computer-aided text analysis (CATA). While CATA is efficient and gives the researcher complete control of the coding algorithm, even sophisticated computer software is still not as skilled as humans at executing certain tasks.

Text interpretation occasionally requires human judgment. A Twitter post as simple as the exclamation "Yum!" could, for example, express enthusiasm or be a sarcastic expression of disgust, depending on the context. Computers are also limited in their ability to process and interpret images or video. Some limitations of computers can be overcome by machine learning strategies that "train" computers using samples of the relevant content. But for now, human coding still tends to be the preferred method for complex content analyses that center on themes or tone, or that involve images or video footage.

Content Analysis in Expert Testimony

Content analysis is suitable for expert testimony because it is a well-established research discipline that follows the scientific method, is replicable and (at least in some applications) has an error rate that can be measured. It has a decades-long history as a research field and has been widely used in disciplines including marketing, economics, psychology, finance, business and law.

As of this writing, a Google Scholar search for the term "content analysis" yields about 1.9 million hits. Content analysis has been used to analyze the impact of nutrition or food safety information on consumer behavior and product demand,[3] the impact of consumer reviews on product prices[4] and the impact of the sentiment of financial news on asset prices,[5] to mention just a few examples.

As required by the scientific method, content analysis relies on a set of rules, which the researcher specifies before any content is coded. It is not an ad hoc process, in which the rules of the game are defined as the coding proceeds. Instead, the researcher relies on an "a priori" design, which guides the selection of the content and its analysis.

In the case of human coding, this a priori design includes a detailed codebook of instructions for the coders as well as a training regimen to ensure that the coders are able to follow these instructions. In the case of CATA, much of the design centers on the choice or development of the coding algorithm.

Content analysis is objective in the sense that the researcher typically does not perform any of the actual coding of content.[6] In the case of human coding, it is considered best practice for the coders to be "blind" to the purpose of the project. Further, because a content analysis has an a priori design, and because codebooks and training materials, or computer algorithms, are disclosed, it is possible for an opposing party to replicate a given content analysis, or to perform sensitivity analyses by changing certain instructions or assumptions.

Importantly, it is possible to assess the reliability or error rate of a coding project. In the case of human coding, each piece of content should be coded independently by two different coders. The researcher can then compare the coders' respective decisions to ensure that the coding instructions were clear and did not allow for substantial discretion, which would cause the coders to deviate from each other.

Often, the researcher will report certain inter-coder reliability metrics, which are standard in the field. In a CATA project, a researcher or opposing party can review and evaluate the computer algorithm itself for accuracy. It can also be helpful to have human coders code a sample of the same content to determine whether the computer algorithm worked reliably, particularly on more complicated projects where theme or tone evaluations are at issue.

Examples of Content Analysis in Litigation

A search of case law on Google Scholar reveals that expert witnesses have provided testimony that includes content analyses in a few different contexts. Change of venue argument is one notable area where content analysis has provided insight into whether and how media coverage of a case in a local area differs from coverage in an alternative trial location.[7]

With respect to liability issues, content analysis has been used the context of false advertising and product cases.[8] Content analysis has also been accepted as evidence in trademark infringement cases,

where consumer confusion can be ascertained from customer call recordings, consumer commentary on websites or postings on social media.[9] In light of the proliferation of such digital content and social media in recent years, the use of content analysis in expert testimony is on the rise.

The value of content analysis is perhaps most readily apparent in the context of false advertising and product misrepresentation cases, where it has been used with some regularity. In these cases, content analysis can be used to systematically review relevant television, print or online advertisements and track any themes or claims. For example, in an automobile product misrepresentation case involving claims related to safety, independent coders can analyze relevant commercials to determine whether safety is indeed a feature highlighted by the manufacturer.

In a more novel application, content analysis can provide direct market-based evidence on consumer opinion. It is, for example, possible to analyze media articles or online consumer reviews about a product and assess whether consumers find at-issue product features to be relevant or important.

While consumer reviews or media articles do not provide a representative sample of consumer opinion, they do offer direct evidence from a real-world setting. This market-based evidence can then be used to confirm the results of a consumer survey whose data were collected in a more controlled setting.

In the context of securities fraud cases, content analysis has attracted some attention, as reflected in the Daubert decision in Bricklayers et al. v. Credit Suisse. In that case, the court excluded an expert and suggested that the expert's analysis could have been improved with content analysis.[10]

In a more basic content analysis in a securities case, counts of relevant keywords in news headlines or news articles can serve as proxies for the relative importance of new disclosures to the market and inform materiality analyses — assuming that more important issues are discussed more frequently in the media, which may not always be the case. For example, an expert can compare counts of different sets of keywords to parse the relative materiality of different pieces of news released on the same day.[11] Content analysis can also help determine when a particular piece of information first became available in the public domain, or where the news originated.

In a more complex scenario, content analysis in a securities case could code news, analyst reports or even relevant social media content in detail to understand what information about a company was in the public domain at a certain date. If a disclosure about a clinical trial of a pharmaceutical was at issue, a content analysis could, for example, identify any media discussions of the trial's success probability and understand the market's predictions.

Content analysis featured prominently in BPI v. ABC News, one of the largest defamation cases in U.S. history, which went to trial in South Dakota in the summer of 2017.[12] BPI's content analysis expert, Professor Kimberly Neuendorf of Cleveland State University, testified at trial about a number of different content analyses that contributed to BPI's causation evidence. Her analyses showed that ABC News's coverage of BPI's beef product, lean finely textured beef or LFTB, was unprecedented both compared to other prior media coverage of the product and ABC's own coverage of food-related events.

Professor Neuendorf also performed an attribution analysis of Twitter posts on the term ABC News used in its reports to refer to BPI's product: "pink slime." Human coders reviewed various characteristics of the Twitter posts, such as the text, any hyperlinks and the presence or absence of certain themes, to determine whether a given Twitter post was responding to ABC News coverage. This attribution analysis showed that ABC News drove the conversation about the term "pink slime" in the wake of its

broadcasts, and did not merely report on an already trending topic.

Conclusion

While content analysis is a nascent tool in litigation, its use by expert witnesses has the potential to transform the type and scale of evidence that can be considered by courts.

Content analysis techniques can allow the court to rely on real-world data to determine the materiality of key case issues to consumers or investors. As data become richer and more abundant, and computer analysis becomes more sophisticated, content analysis promises to become a key technique to help courts navigate the copious available evidence.

Lisa K. Tichy is a principal and Anna Shakotko is a senior manager with Cornerstone Research in New York.

Disclosure: Cornerstone Research was retained by counsel for BPI in the BPI v. ABC News matter.

[1] K. Neuendorf, The Content Analysis Guidebook, 2nd Edition (Thousand Oaks, CA: Sage Publications, 2017), p. 1.

[2] Note that more straightforward content analyses based on word counts are sometimes described as "quantitative," while more complex analyses of tone or themes are described as "qualitative." This distinction is incorrect. All content analyses, simple or complex, described in this article are quantitative in that they summarize content in the form of counts, amounts, ratings or rankings. Researchers in content analysis use the term "qualitative content analysis," primarily for more subjective analyses carried out by the researcher himself or herself — for example, in studying the rhetorical or narrative structure of a text.

[3] D. Fan and C. Shaffer, "Effects of the Mass Media Newson Trends in the Consumption of Caffeine-Free Colas," Advances in Consumer Research 17 (1990): 406–414; N. Piggott and T. Marsh, "Does Food Safety Information Impact U.S. Meat Demand?," American Journal of Agricultural Economics 86, no. 1 (2004): 154.

[4] N. Archak, A. Ghose and P. Ipeirotis, "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," Management Science 57, no. 8 (2011): 1485–1509.

[5] D. Garcia, "Sentiment During Recessions," Journal of Finance 68, no. 3 (2013): 1267–1300.

[6] The researcher does, of course, shape the content analysis by creating the research design. One can imagine instances in which the results of the project could change based on the precise content selected for analysis, and the instructions conveyed in a codebook. However, as in any scientific research project, these choices are disclosed and can be examined and critiqued by an opposing party.

[7] See for example, Denial of motions to dismiss in U.S. v. Lindh, No. Crim. 02-37-A, U.S. District Court,

Eastern District of Virginia, Alexandria Division, 2002.

[8] Order to partially admit the opinion of Dr. Yoram Wind in LG Electronics USA, Inc. v. Whirlpool Corp., No. 08 C 242, U.S. District Court, Northern District of Illinois, Eastern Division, 2010.

[9] Order granting E. Mishan & Sons summary judgment in Hearthware Inc. v. E. Mishan & Sons Inc., Civil Action No. 11 C 5233, U.S. District Court, Northern District of Illinois, Eastern Division, 2012.

[10] Decision on the admissibility of the opinion of Dr. Scott D. Hakala in Bricklayers and Trowel Trades v. Credit Suisse LLC, No. 12-1750, U.S. Court of Appeals, First Circuit, 2014.

[11] G. Breton and R. J. Taffler (2001), "Accounting Information and Analyst Stock Recommendation Decisions: A Content Analysis Approach," Accounting and Business Research, Vol. 31, No. 2, pp. 91–101, at p. 95; F. C. Dunbar and A. Sen, (2009), "Counterfactual Keys to Causation and Damages in Shareholder Class-Action Lawsuits," Wisc. Law Review, Vol. 2009, No. 2, pp. 199–242 at 242.

[12] Beef Products Inc. v. American Broadcasting Companies Inc., No. 12-cv-292, South Dakota Circuit Court, First Circuit, Union County.