

Methods for Coding Television, and Comparing to Internet Content

Jason is interested in the way that symbols that can interact with words, interacting and conflicting with them. He thinks it's under-theorised and is wondering if he should design something for every study, or if we need rules for doing this. Nathan asks if Jason has seen Otto Santa Anna's work, and Jason responds that most people doing critical discourse analysis have a different methods.

Deen argues that we need textbook-style standards which evolve, so we can evaluate the quality of work. Maybe it's a methods wiki, or a set of pitfalls. Nathan mentions the [Data Journalism Handbook](http://datajournalismhandbook.org/1.0/en/). Deen suggests that methods guides should try to be platform independent.

Nathan asks Floriana if she has caption data. They recorded video, then used a coding frame from another situation and adapted it to their new data. They have the captions, and quantitatively measure how they are used in different kinds of frames. Nathan suggests that it should be possible to add caption data to MediaCloud.

Floriana: manual content analysis of TV is useful and can provide broader context. Deen agrees. We don't just want to only grasp the low hanging fruit, just doing automated analysis because that's what we do. The methods should be guided by the research questions which we have. Nathan suggests that this distinction may not last much longer-- things like semi-automated techniques or machine learning techniques can integrate human coding and automated approaches.

Jason: all of these new approaches make it harder to do the textbook methods approach. Why can't we just talk about a toolbag and evaluate people based on the methods they actually use?

Rob: people get used to certain things. If you develop a new method, it's hard for people to review your work.

Fred: are there cases where people have decided to venture and have faced obstacles in the review process? Fred mentioned a Master student's project which combined human and automated analysis. The article was well justified, but it has taken a long time to be reviewed.

Fred: how true is the perception that American researchers tend to go more for the quantitative and automated approach, compared to the British and the European.

Jason: You can see this difference in the training in America versus Britain. Deen: if you're publishing something with a methodological bent, look at their back issues and point out that they have done this work. A lot of Comm journals like survey analysis, with advanced statistics on surveys. But you don't get much automated textual analysis. But where you do tend to see it is HCI, social computing, and computational linguistics.

Nathan says that this is a great opportunity-- applying technologies to questions in other fields.

One participant worries that methods then lead people to ask the questions, which might not be the right approach. He mentions a content analysis study which he did to assess public service TV stations. Tech might have made things easier for him. But what if the methods led his questions rather than the other way around?

Deen points out that on one hand, HCI people are trying to push forward the math, and the Communications people often are unable to adopt those methods. As long as people feel like they have to learn a new PhD to do this stuff, there's going to be a huge block.

In a discussion of what role people are likely to play in research, Deen reminds us that we don't want to lose out

What methods are people using?

- Multi-site methods
 - Mediamaking
 - Facilitated meetings between groups that don't share ontologies
 - (look up Ramesh Srinivasan paper)
- (+1) Automated Content Analysis
 - Centering Resonance Analysis
 - Topic Modeling
 - Sentiment Analysis
 - Keyword Analysis
 - Similarity
- (+1) Drone(balloon) Researchers
- (+1) Social Network Analysis
- (+1) Digital Data Extraction
- (+1) Behaviour Mining
- (+1) Critical Discourse / Multimodal Analysis
- (+1) Tablet & Mobile & SMS
 - Surveying
 - Data Collection
- Virtual Ethnography / Netnography / Online Ethnography
- (+1) The Crowd

Automated content analysis

Deen is looking at the persistent of Internet/Facebook/Twitter revolution memes around the Arab Spring and Occupy. He's looking at traditional media (top 10 news sites in the US, and top 10 tech sites) Taking blog data out of Google Blog Search (using search terms). Posts have to have "Egypt" "Tunisia" followed by "Facebook" "Google" or "Twitter." There is a lot of noise in the webpages, and you have to re-filter it to take out false positives, after extracting text from each page. Then, he searched to see if blog posts contained certain words or not, to include

certain words in the dataset. It's possible to talk about the way he did this specific project, but it's hard to communicate a general approach to students. With techniques like this, it's important to check the dataset to make sure it meets your sampling requirements.

Deen reflects on the kinds of knowledge needed for research in Communications and compares knowledge of computing to a basic knowledge of statistics.

Centering Resonance Analysis

Using proprietary research software: <http://www.crowdadtech.com/software.htm>

Crowd words together, trying to make sense of the words based on the recurrence of those words. The discussion comes back to the need for qualitative, interpretive research in order to understand frames and context.

Deen puts on his conservative reviewer hat. He would ask: what is it about this particular dataset that justifies deviation from standard practice? If you were to say "this method helps tell us things we wouldn't otherwise know" then perhaps. But the novelty of the method may not justify publication.

Jason thinks that the real value comes from a mixture of critical discourse and quantitative methods. Critical Discourse Analysis, unlike traditional frame analysis, is an open, fluid, and developing discipline. Looking at a text, we can take a series of steps in the analysis process, looking at headline aspects of the text, then the language of the text, perhaps take it out of context to look at the type of words and their influence on its meaning. You can then count things within the text. You can then look at a larger sample of texts, taking those features and comparing the sample. Once you move to multi-modal analysis, you focus on sounds, images, borrowing things like content analysis-- building up a picture which Jason thinks is ideal for looking at power and discourse.