An Annotation Model for Making Sense of Information Quality in Online Video

Nicholas Diakopoulos and Irfan Essa

Georgia Institute of Technology | School of Interactive Computing | GVU Center

85 5th St. NW Atlanta, GA 30332

nad@cc.gatech.edu, irfan@cc.gatech.edu

ABSTRACT

Making sense of the information quality of online media including things such as the accuracy and validity of claims and the reliability of sources is essential for people to be well-informed. We are developing Videolyzer to address the challenge of information quality sense-making by allowing motivated individuals to analyze, collect, share, and respond to criticisms of the information quality of online political videos and their transcripts. In this paper specifically we present a model of how the annotation ontology and collaborative dynamics embedded in Videolyzer can enhance information quality.

Categories and Subject Descriptors

H.5.4 [Hypertext/Hypermedia]: User issues, H.5.3 [Group and Organization Interfaces] Collaborative computing

General Terms

Design, Human Factors

Keywords

Collaborative Annotation, Tagging, Information Quality, Video Analysis, Journalism

1. INTRODUCTION

In traditional journalism practice, information quality is managed on the side of the reporter with the help of expert editors. But an alternate model based on the distributed and collaborative evaluation of information quality is being explored by such communities as Wikipedia [18]. In the collaborative paradigm, the process of quality evaluation and assurance is apportioned across many individuals with disparate expertise and interests. Individual knowledge leads people to have different predictions, expectations, and evaluations of information on the web [15]. Within this model we are developing Videolyzer, which seeks to harness the power of many people to negotiate a patina of annotations relating to the quality of the information in an underlying video. The tool is designed to aid politically astute people (such as political bloggers) analyze the quality of online political video and then aggregate and share these analyses.

Videolyzer is an application which incorporates elements of automatic content analysis (both textual, audio, and video) into a

larger semi-structured manual analysis system that lets people rate both objective and subjective elements of quality such as accuracy, credibility, expertise and trust at a high granularity (i.e. anchored to specific time or transcript markers). Collecting this layer of annotations will enable better searchability of video as well as visualizations which can cue end consumers to the credibility and trustworthiness of sections of video.

A number of challenges are associated with sense-making related to information quality including how to define information quality and how to operationalize it as a set of annotations that are available to jumpstart people's analyses. And while information quality has been studied extensively in many domains, simplifying its presentation and making it flexible in the user interface is both a hard problem and essential for adoption. In this paper we present a simplified model and an operationalization of information quality as a set of annotations. We detail how the annotations and collaborative dynamic around these annotations are expected to enhance information quality understanding in the resulting annotated videos.

2. THE NEED FOR BETTER TOOLS

We have initially chosen to focus the content domain of the application on online politically oriented video, although we could also expand this to include other types of informational video such as documentaries or news video. In fact, the annotation model that we present here can also be applied to other forms of media such as a purely textual article, which is a degenerate case of the video plus transcript system we are building. The content focus on political video allows us to hone in on some of the needs of a subset of users interested in and motivated to critique this content, namely political bloggers and viewers of online political video. To understand how this niche group might use our application we interviewed three local political bloggers and surveyed 27 people who were either viewers of online political video, political bloggers, or political forum readers. The interviews and survey included questions about online video consumption and commenting behaviors and needs.

We found there to be a recognized need for better ways to filter and categorize the numerous comments posted in response to videos online. Several survey respondents voiced the need for more structured commenting mechanisms, including the categorization of comments and the credibility of comments, the ability to provide links to supporting or refuting sources of information, the ability to see different points of view on a complex issue clearly and easily, and facilities for tagging the quality of comments (e.g. relevant, off topic etc.). While many respondents (89%) indicated that they read comments "sometimes" or "often," fully 63% found them *not* to be useful or

valuable. In short, they need better support for semi-structured hypermedia discourse around the video to aid in sense-making. Videolyzer addresses these needs by supporting highly granular annotation of a video, its transcript, and its annotations and by leveraging the information quality annotation framework we have developed to scaffold the sense-making process.

3. ANNOTATION SYSTEM

3.1 Design Goals

Taking a cue from the practicality of journalism, the design of the annotation system in Videolyzer strives to reduce the complexity of the information quality model for users [21]. The goal is to operationalize analysis of *most* of the aspects of information quality in as *simple* an interface as possible. For reference, a screenshot of the application interface is shown in Figure 1., however the interface is not the focus of this paper. The goal of the annotation system is to provide a way for people to analyze, collect, share, and respond to criticisms of information quality. This involves support for arriving at a consensus about the quality of a source as much as it does for providing a description of the strife or contentiousness of a claim. Ultimately we want a discourse ontology that is expressive and flexible enough to represent most information quality issues without being overly restrictive or abstract.

Annotation systems can be characterized by the degree to which the annotation vocabulary is restricted. At one end of the spectrum there are completely uncontrolled annotation systems such as collaborative tagging and at the other end are tightly controlled vocabularies used by information architects. The advantages of using an uncontrolled approach are that no knowledge of a special vocabulary is needed, people can describe content in whatever way is most meaningful to them in their context and for their use, and there is a low human cost for making a contribution. On the other hand, controlled vocabulary annotation facilitates structured organization in the interface and more formal processing capabilities based on the typed relationships in the ontology [1]. Using controlled vocabulary also avoids some of the difficult issues surrounding inferring meaning from free text comments [2].

In order to get the best of both worlds Videolyzer supports both unconstrained text comments as well as typed annotations regarding information quality. Typed annotations can always be accompanied by free text reasons or rationale to explain the formal rating. Allowing for free text rationale in the interface is important since there are several subjective ratings for which the rationale aid in understanding. Also, research in psychology indicates that overconfidence bias in ratings can be mitigated when people are asked to write down their rationale [5].

Issues of literacy explored in related research suggest that by focusing the analytic activity around some key constructs from information quality which are prominent in the interface we can *guide* users towards a better analysis of information quality [16]. In the process, the value of the annotations, both to aid in sense-making and as computable resources, will be enhanced from the structure of the ontology.

3.2 Annotation Model

We have synthesized information quality descriptions and ontologies from journalism and communications [9, 12-14], information science [7, 15, 17, 20], social science [22], health

Figure 1. A screenshot of Videolyzer showing video, transcript, timeline, and annotations.

information systems [3, 11], and semiotics [10, 21]. We subsequently pruned our description of information quality to make it less complex and more tractable as an interface aimed at political bloggers, while still covering as much ground as possible. Our ontology is focused predominantly on issues espoused in journalism and information science and is meant only as a starting point for the analysis and sharing of quality judgments, which can always be augmented by free text descriptions. In this section we will (1) define some of the key facets of the ontology and (2) discuss how these facets are operationalized through annotation in Videolyzer. Again our goal is not to comprehensively define information quality (a classic problem) but to come to a practical understanding of it as it relates to sense-making and analysis of online video.

The term *information quality* suggests a degree of excellence in communicating knowledge or intelligence. This degree of excellence can be analyzed along different dimensions such as *accuracy*, *reliability*, *validity, comprehensiveness,* and *currency* [20]. In news production, where the conveyance of quality information is paramount, principles such as *accuracy*, *thoroughness*, *transparency*, and an awareness of *bias* guide journalists in communicating quality information [8, 12]. Journalism is after a practical or functional form of truth which strives to provide the reader with enough information to make his or her own decision [12]. Videolyzer has a similar goal, but approaches it using a pragmatic methodology involving *collaborative* analysis of information quality.

3.2.1 Information Quality Facets

The key facets of our simplified model are shown in Figure 2. Validity refers to the well-groundedness, justifiability, or logical correctness of a claim. To a large extent validity has to do with the degree to which the interpretations and inferences about information are sound. Journalists and social scientists practically achieve validity using techniques such as triangulation and
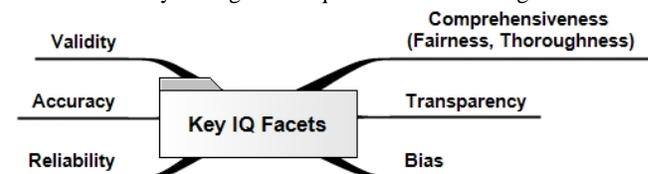

Figure 2. Key information quality facets.

corroboration of information sources or by developing a chain of evidence to support inferences [12, 22]. Accuracy is defined as "freedom from mistake or error (correctness)" or as "conformity to truth or to a standard or model (exactness)". Accuracy dictates that facts or the version of events presented have been corroborated by multiple reliable sources.

Reliability can be defined as dependability or consistency. For practical purposes we focus on the dependability definition since consistency requires analysis over time. Credibility is closely linked to reliability through the notion of trustworthiness (a synonym for dependability), and is a perceived, subjective quality of an information source or object. Trustworthiness is a measure of the truthfulness, fairness, and lack of bias of a source or communication. The other component of credibility is expertise, which is defined as the perceived knowledge, skill, and experience of the source [6, 20].

Comprehensiveness refers to the "completeness of coverage" of a particular subject or issue [20]. In journalism this is analogous to the principle of thoroughness, which involves learning and gathering as much as possible about a topic before sitting down to produce a story [8]. Thoroughness speaks to the degree of comprehensiveness of research on the side of the information producer, which allows for informed information selections. Comprehensiveness also relates to the journalistic principles of fairness and balance which call for the inclusion of different and diverse sources, perspectives, and opinions [12].

Transparency and bias are concepts that arise as key components of the journalism of verification [12]. Transparency involves disclosing and describing sources such as how they know what they know, what biases they might have, and whether or not there are conflicting accounts. Bias too has a big connection to the sources that are used, but can also include aspects of the communicator's bias or relationship to institutions and sponsors.

3.2.2 Annotations

We have designed our annotations to address the key facets of information quality outlined in the preceding section in as simple a way as possible. The two high level annotations possible are "mark as a *claim*" and "add a *comment.*" These annotations are associated with interval anchors on the transcript or video. Claims represent statements, inferences, or interpretations made from sources. They can optionally have a *source* associated with them. Sources are origins of information, whether from documents or people, and are described according to relevant elements of the Dublin Core metadata standard [7], such as title, author, and publisher. Additionally we allow users to rate their interpretation of the *trust* and *expertise* of the source on a 3 point scale. The relationship between annotations is shown in Figure 3. The initial choice between claim or comment minimizes the complexity of the ontology at the interface level.

The rest of the elements of information quality are operationalized as a set of tags divided into two groups: *quality tags,* which are presented as semantic differentials, and *bias tags*. Each tag can have a *reason* for associating that tag. Reasons in turn can have optional supporting sources attached to them. Quality tags (e.g. invalid | valid) are presented as a 3 point scale, the mid point representing uncertainty. User defined tags are also included so as not to hem users into the ontology, however this comes with their lack of semantic interpretability in relation to the other annotations. Tags can be added to comments, claims, and reasons.
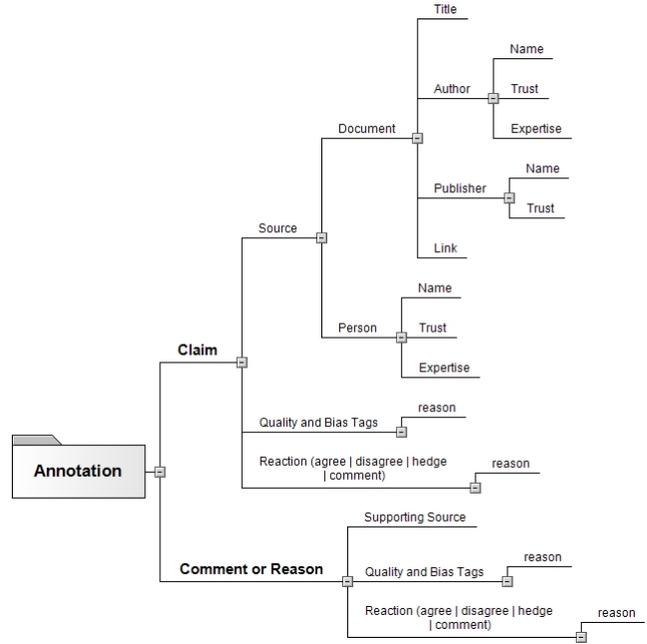

Figure 3. Relationships between types of annotations.

The various types of tags are shown in Table 1. Comments, claims, and reasons can be annotated by other people in terms of their reactivity (agree / disagree / hedge) to the comment or reason as well as by adding quality or bias tags. Thus people interact and negotiate with others through rationalized reactions and by adding tags to other comments and reasons. The net positive of this interaction is expected to not only be a rich description of the information quality, but also a better understanding of the quality by the users as they interact.

3.2.3 Effects of Annotations on Information Quality

Assessment and annotation of the validity of information hinges largely on the potential for annotators to provide *falsifying evidence* to refute the justifiability or logical correctness of an interpretation, inference, or claim. In Videolyzer, annotators are able to attack the validity of information by criticizing claims, assumptions or premises of claims, and the relevancy and adequacy of facts or opinions being presented as evidence for claims. All of these things are supported through the ability to specify tags, reasons, and supporting sources.

The annotation strategies that are expected to be most effective in addressing the accuracy facet of information quality involve

Quality Tags	Bias Tags
Invalid	Valid	Political
Inaccurate	Accurate	Omitted Information
Bad Premise	OK Premise	Under Emphasis
Opinion	Fact	Over Emphasis
Irrelevant	Relevant	Stereotyping
Unfair	Fair	Sponsorship
Missing Context	Thorough	< user defined >
Unclear	Clear	
< user defined>	

allowing people to fact check and corroborate or dispute sources

Table 1. Tags for describing information quality and bias.

by comparing to other sources. Inconsistency among reputable sources indicates an accuracy issue with the information being presented. This highlights the central importance of the source and its traits in analyzing information quality.

Reliability and credibility are operationalized in Videolyzer as a composition of *trust* (perceived trustworthiness) and *expertise* (both perceived and actual expertise). The subjectivity of these measures suggests that representations showing the distribution of ratings are important for understanding both consensus and strife. Variability or skew in these attributes will indicate which areas of a video are uncertain or are controversial. Rationale and reasons for these subjective ratings are expected to aid in sense-making.

While the validity, accuracy, reliability, and bias of claims and sources are all *directly* supported through annotation mechanisms, issues of comprehensiveness / thoroughness, and transparency are best addressed through the collaborative interplay and collection of annotations by many people. Given enough annotators with diverse backgrounds, the aggregation of their reasons and ratings will add greatly to the comprehensiveness, fairness, and balance of the information. Moreover, even though the original sources behind a video may not be known, we provide for transparency in a different way. Through the collaborative evaluation of claims and their sources, the transparency of sources, their expertise, and their trustworthiness can be elucidated. Triangulation, corroboration, and chains of evidence can arise through the comparison of diverse sources and reasons provided by different people. As individuals agree or disagree with others' ratings and reasons a rich and lasting description of discourse will have cues to all of the elements of information quality in our model: validity, accuracy, reliability, comprehensiveness, transparency, and bias.

4. NEXT STEPS AND SUMMARY

We have collected initial feedback through think aloud sessions and questionnaires from two local political bloggers with journalism backgrounds. The annotation mechanism was initially implemented as a cascading menu, however, based on user comments we are iterating on the interface to make annotation choices more immediately visible. Our next step will be to finish implementing the functionality of the application and to do more extensive usability testing with other political bloggers. We will study how bloggers use the interface and how and if the tool corresponds to their existing community argumentation routines [4] by doing a content analysis of the annotations collected using the tool.

We are also interested in studying the analytic impact of suggesting information quality issues either through automatic means or through community sharing. And as a caveat to this approach, we would like to investigate the potentially *negative* role that collaboration around information quality could engender in terms of information cascades and social proof [19]. The potential danger is in objectively falsifiable claims gaining community consensus despite their invalidity. Finally, we are interested in different ways to represent and visualize aggregate descriptions of quality from the video's annotations as a mechanism to warn viewers of poor quality information.

In summary, we have presented a simplified information quality model operationalized as a set of annotations suitable for application to a video and its transcript. We have also discussed how the design of our system, its semi-structured flexibility, and the interactions afforded by the ontology and application are expected to enhance information quality and meet our goals of facilitating sense-making, analysis, and sharing of criticisms of information quality in online political videos.

5. REFERENCES

[1] Buckingham Shum, S., Sensemaking on the Pragmatic Web: A Hypermedia Discourse Perspective. in *Proceedings of the International Conference on the Pragmatic Web*, (2006).

[2] Chklovski, T., Ratnakar, V. and Gil, Y., User Interfaces with Semi-Formal Representation: a Study of Designing Argumentation Structures. in *Intelligent User Interfaces (IUI)*, (2005).

[3] Civan, A. and Pratt, W., Supporting Consumers by Characterizing the Quality of Online Health Information: A Multidimensional Framework. in *Procedings of HICSS*, (2006).

[4] de Moor, A. and Aakhus, M. Argumentation Support: From Technologies to Tools. *Communications of the ACM*, *49* (3). 93-98.

[5] Evans, J. *Bias in Human Reasoning: Causes and Consequences*. Lawrence, 1989.

[6] Fogg, B.J. and Tseng, H., The elements of computer credibility. in *Proceedings of CHI*, (1999), 80-87.

[7] Gil, Y. and Ratnakar, V., Trusting Information Sources One Citizen at a Time. in *International Semantic Web Conference (ISWC)*, (2002).

[8] Gillmor, D. *We the Media: Grassroots Journalism by the people, for the people*. O'Reilly, 2006.

[9] Goldstein, T. *Journalism and Truth: Strange Bedfellows*. Northwestern University Press, 2007.

[10] Hoffmann, M.H.G., Logical Argument Mapping: A cognitive-change-based method for building common ground. in *Conference on the Pragmatic Web*, (2007), 41-47.

[11] Kim, P., Eng, T., Deering, M.J. and Maxfield, A. Published criteria for evaluating health related web sites: review. *BMJ*, *318* (March). 647-649.

[12] Kovach, B. and Rosenstiel, T. *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect*. Three Rivers Press, 2007.

[13] Lawrence, G. and Grey, D. Subjective Inaccuracies in Local News Reporting. *Journalism Quarterly*, *46*. 753-757.

[14] Meyer, P. Defining and Measuring Credibility of Newspapers: Developing an Index. *Journalism Quarterly*, *65*. 567-574.

[15] Rieh, S. Judgment of Information Quality and Cognitive Authority in the Web. *Journal of the American Society for Information Science and Technology*, *53* (2). 145-161.

[16] Serano, B., Buckingham Shum, S. and Motta, E., ClaimSpotter: An Environment to Support Sensemaking with Knowledge Triples. in *Intelligent User Interfaces (IUI)*, (2005).

[17] Strong, D., Lee, Y. and Wang, R. Data Quality in Context. *Communications of the ACM*, *40* (5). 103-110.

[18] Stvilia, B., Twidale, M., Smith, L. and Gasser, L. Information Quality Work Organization in Wikipedia. *Journal of the American Society for Information Science and Technology*, *59* (6). 983-1001.

[19] Sunstein, C. *Infotopia: How many minds produce knowledge*. Oxford University Press, 2006.

[20] Taylor, R. *Value-Added Processes in Information Systems*. Ablex, 1986.

[21] Yetim, F., DISCOURSIUM for Cooperative Examination of Information in the Context of the Pragmatic Web. in *Conference on the Pragmatic Web*, (2007).

[22] Yin, R. *Case Study Research : Design and Methods (Applied Social Research Methods)*. SAGE Publications, 2002.