# On the Quality of Information for Web 2.0 Services

Most Web 2.0 applications let users associate textual information with multimedia content. Despite each application's lack of editorial control, these textual features are still the primary source of information for many relevant services such as search. Previous efforts in assessing the quality of these features primarily target single applications and mainly focus on tags, thus neglecting the potential of other features. The current study assesses and compares the quality of four textual features (title, tags, description, and comments) for supporting information services using data from YouTube, YahooVideo, and LastFM.

O ne key characteristic of Web 2.0 applications, such as YouTube, is the primary role users play in creating and sharing content. Although a significant amount of the content in these applications is multimedia, users are commonly encouraged to associate pieces of textual information — *textual features* — with the multimedia objects. Common examples are title, tags, description, and user comments.

Because these textual features are user generated, however, the respective applications have no editorial control and thus can't guarantee quality, neither in terms of syntactic correctness nor of the text's semantic relationship with the object. This poses a challenge to services such as

search and advertising that primarily rely on textual features as sources of information about the objects' contents. This happens because the use of multimedia information retrieval mechanisms in Web 2.0 is still limited, possibly because state-of-the-art techniques are often ineffective under the low quality of most content and don't scale well to the size of several applications.[1]

Previous efforts toward assessing the quality of textual features primarily focused on tags, investigating how to use them to support search, recommendations, and object classification.[2-4] However, researchers haven't reached a consensus regarding their quality.[5-7] Moreover, they

**Jussara M. Almeida,**
**Marcos André Gonçalves,**
**Flavio Figueiredo,**
**Henrique Pinto,**
**and Fabiano Belém**
*Universidade Federal de Minas Gerais, Brazil*

mostly neglect the potential use of other textual features and, by typically targeting a single application, ignore potentially significant inter-application differences.

We advocate that it is necessary to investigate the relative quality of multiple textual features, across multiple applications, addressing several questions: How do different textual features compare in terms of quality for information services? And are there significant differences across applications? In other words, it is necessary to search for supporting evidence that any given feature is indeed the most promising one to be exploited for effective information services.

With this article, we take a step toward addressing these questions by assessing and comparing the quality of different textual features. We start by discussing challenges involved in assessing information quality. Next, we analyze object collections crawled from three applications: YouTube (www.youtube.com), YahooVideo (http://video.yahoo.com), and LastFM (http://last.fm). We selected these sites because of their current popularity and common set of features. YouTube and YahooVideo contain mainly videos, whereas LastFM is an online radio site, targeting audio content.

We used heuristics and user experiments to assess three relevant quality aspects — usage, descriptive, and discriminative powers — across four features: title, tags, description, and comments. Compared with our previous work,[8] this article analyzes feature quality at the granularity of individual objects, proposes a new metric for capturing discriminative power for object-classification tasks, discusses the agreement among different quality aspects, and reports results from a user experiment.

In sum, our findings can help designers decide, for example, which features are more useful and attractive to users, exhibit higher quality (being thus more important to index for supporting effective services), and have quality problems (for example, because of a lack of content or large presence of noninformative terms). Moreover, our results can motivate the design of new techniques such as new interfaces, incentive mechanisms, editorial collaboration, and content recommendation to enhance feature use and quality. Our findings provide valuable knowledge to drive the design of future Web 2.0 services and applications.

## Challenges in Assessing Information Quality

Although people intuitively know what "information quality" means, explicitly defining it depends on the characteristics of the application or domain. Considering our focus on the quality of textual features for designing effective information services, we claim that a *high-quality feature*

- has enough *amount of content* to be useful;
- provides a good description of the object's content (*descriptive power*), which is important for services that exploit objects' semantics; and
- can distinguish the object from others (*discriminative power*), even under information overload, for tasks such as separating the objects into semantic classes or into levels of relevance regarding a query.

Each of these aspects can shed light on a more precise picture of feature quality, although we aren't claiming any sufficiency property.

### Impact on Services and the Influence of Application Characteristics

We start by arguing that the amount of content, descriptive power, and discriminative power are not equally important to all services, nor must they all be present for a feature to have high quality. For example, discriminative power is important for automatic classification; if a feature contains a few highly discriminative terms, the amount of content might not be as important. However, classifiers usually combine the discriminative power of many terms (with different weights) to make decisions. Thus, more content might be beneficial, providing more evidence to support decisions. However, if many terms are ambiguous or used indiscriminately across classes, more content might be detrimental.

Other services, such as search, explore term frequencies to discriminate among objects (such as to generate ranks). The repetition and co-occurrence of query terms in one or multiple features might help support this task. In contrast, for services that exploit content-based filtering (such as recommendation), good descriptive power might be more important. Finally, high descriptive or discriminative power might be of little use if the feature is absent in most objects.

| Quality aspects | Editorial collaboration | Editorial incentive | Clearer category semantics | Larger class sizes | Greater object popularity |
|---|---|---|---|---|---|
| Amount of content | Positive | Positive | Positive (some cases) | No clear impact | Positive (collaborative features) |
| Descriptive power | Negative | Not measured | No clear impact | No clear impact | Positive (collaborative features) |
| Discriminative power | Negative | Not measured | Positive | Positive | No clear impact |

*Table 1. Impact of application characteristics on quality aspects.\**

\* Positive and negative indicate whether the characteristic tends to favor higher or lower quality, respectively, possibly in specific scenarios. No clear impact reflects insignificant correlations.

Also, different application characteristics might impact how users edit textual features and, thus, their quality. For instance, the level of editorial collaboration the application allows is important. The same feature might carry more or less content depending on whether its editing is restricted to the object owner (a restrictive feature) or available to any user (a collaborative feature). In the applications we analyze here, title is restrictive and comments are collaborative, whereas tags are restrictive only in YouTube and description is collaborative only in LastFM.

Furthermore, various mechanisms might contribute to enhance feature quality. The use of tags for organizing personal libraries (as in LastFM) might serve as an incentive for users to provide more and higher-quality content. Similarly, an attractive and easy-to-use interface might also impact quality. For example, LastFM descriptions are typically edited in a wiki-like manner, which might encourage use. Finally, the semantic overlap across object categories might also influence the quality (discriminative power) of content associated with different objects.

Table 1 illustrates this discussion, presenting the potential impact of several characteristics on each quality aspect. Such observations are based on our assessment of feature quality, particularly on results from correlation analyses, which we discuss later on.

## Automatic Quality Assessment

Having defined a scope for information quality, its assessment becomes the next challenge. Whereas experiments with volunteers might produce good estimates of descriptive and discriminative powers, the unavoidable degree of subjectivity in such experiments might affect results. Moreover, the dependency on the aspect under evaluation (such as the relevance of an object to a query or its pertinence to a category) requires assessment in the context of a specific service.

Because manual assessment is costly, automatic means to evaluate descriptive and discriminative powers might be a better alternative. Heuristics, such as term frequency, inverse document (feature) frequency,[8] information gain, and entropy,[9] capture these powers to some extent and can be applied to larger object samples at lower cost. However, heuristics invariably have limitations, such as being focused on specific issues (and thus only partially capturing the target aspect) and having biases that impact their effectiveness in specific scenarios. Nevertheless, heuristics can still produce important, cost-effective insights. That said, we next define heuristic metrics used to assess each quality aspect. We refer to the content of a feature associated with an object as a *feature instance.*

The amount of content of a feature instance is estimated by the number of unique stemmed terms in it, thus avoiding counting variants of the same stem as different terms.

For assessing descriptive power, we used an adaptation of a metric previously proposed as part of an information retrieval model for structured Web documents.[10] We adapted this metric to textual features of Web 2.0 objects and refer to it as *feature instance spread* (FIS).[8]

We initially define the *term spread* of term $t$ in object $o$, $TS(t, o)$, as the number of feature instances $f$ associated with $o$ containing $t$:

$$TS(t,o) = \sum_{f \in o} I(t,f),$$

where

$$I(t,f) = \begin{cases} 1 & \text{if } t \in f \\ 0 & \text{otherwise} \end{cases}.$$

| Table 2. Fraction of empty feature instances. | | | | |
|---|---|---|---|---|
| **Application** | **Title (%)** | **Tags (%)** | **Description (%)** | **Comments (%)** |
| LastFM | 0 | 19 | 53 | 55 |
| YahooVideo | 0.2 | 16 | 1.2 | 97 |
| YouTube | 0 | 0.1 | 0 | 23 |

The FIS of an instance *f* is the average *TS* across all terms in *f*. We could apply various *filtering criteria*, such as disregarding popular (and perhaps too broad) terms or taking the *k* terms with the largest *TS*, to compute FIS.

The intuition behind FIS is that terms appearing in several features associated with the same object have a better chance of being related to its content. For example, if the term Madonna appears in four features of an object (*TS* = 4), there is a high chance that it's related to the famous singer.

We assess discriminative power for qualifying object-classification tasks, thus assuming that objects are precategorized into semantic classes. We propose a new metric, *feature instance class concentration* (FICC), which estimates how strongly the instance's content indicates the object's preassigned category.

We initially define *term class concentration* (TCC) of term *t* occurring in instance *f* of feature *F* associated with object *o* as the fraction of instances of *F* containing *t* that are associated with objects of the same class of *o*:

$$TCC(t,F,o) = \frac{\sum_{f \in F} I(t,f) \times \text{class}(f,o)}{\sum_{f \in F} I(t,f)},$$

where

$$\text{class}(f,o) = \begin{cases} 1 & \text{if class}(o_f) = \text{class}(o) \text{ and} \\ & o_f \text{ is the object that contains } f \\ 0 & \text{otherwise} \end{cases}.$$

We compute the FICC of *f* as the average *TCC* across all terms in *f*. Again, we can apply filtering criteria to disregard, for instance, unpopular terms that might undesirably inflate FICC because they occur in few objects (classes), or very common (domain-independent) terms.

We choose to assess discriminative power targeting object classification tasks for several reasons:

- many applications let users associate classes with objects,
- classification tasks support other services, such as tag recommendation, and

- preassigned object categories allow for automatic evaluation.

Other metrics, such as inverse feature frequency[8] or some information-to-noise ratio, could be used to assess discriminative power, particularly if categories are unavailable, or other services (such as search) drive the evaluation. Combining multiple metrics could also improve quality assessment. We choose to evaluate each quality aspect separately to facilitate interpreting results. The evaluation of other (combined) metrics is left for future work.

## Assessing the Quality of Textual Features

We assessed the quality of title, tags, description, and comments by quantifying their usage, descriptive, and discriminate powers in collections of real objects and associated features. Our analyses focused on English content, covering approximately 181,000, 160,000, and 100,000 objects crawled from YouTube, YahooVideo, and LastFM, respectively.[8] For YouTube and YahooVideo objects, we also collected the category assigned by the video owner, considering a predefined list of options. For LastFM, we collected a sample of musical genres from AllMusic (www.allmusic.com), which enabled manual precategorization of 6,400 objects.

### Feature Usage

As Table 2 shows, all the features except for title have a significant fraction of empty instances in at least one application, as previous research found with Flickr data.[6] Interestingly, 16 to 19 percent of the LastFM and YahooVideo objects did not contain tags, which could significantly impact tag-based services. Comments are greatly under-explored in LastFM and YahooVideo, and many LastFM objects don't have descriptions. All features are significantly explored only in YouTube. Whereas this might be partially explained by YouTube automatically filling title and tags if no content is provided, the large presence of comments and description might reflect different usage patterns.

| Application | Title | | | Tags | | | Description | | | Comments | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Average | Maximum | CV* | Average | Maximum | CV | Average | Maximum | CV | Average | Maximum | CV |
| LastFM | 1.80 | 23 | 0.47 | 27 | 269 | 1.50 | 90 | 3,390 | 1.06 | 110 | 22,634 | 3.55 |
| YahooVideo | 6.30 | 16 | 0.39 | 13 | 52 | 0.52 | 22 | 141 | 0.71 | 52 | 4,189 | 2.51 |
| YouTube | 4.60 | 36 | 0.43 | 10 | 101 | 0.60 | 40 | 2,071 | 1.75 | 322 | 16,965 | 1.94 |
| * Coefficient of variation (CV) | | | | | | | | | | | | |

*Table 3. Amount of content in nonempty feature instances (number of stemmed terms).*

Table 3 shows the average and maximum amount of content in nonempty instances, along with the corresponding coefficients of variation (CV) across objects. In all applications, on average, the title is the smallest feature, followed by tags, description, and comments; this trend is aligned with the expected degree of user verboseness. CVs also indicate that titles exhibit the lowest variability. In particular, titles tend to be shorter in LastFM, where they usually contain artist names, with at most two terms in 89 percent of the objects.

Comparing the video applications, descriptions and comments tend to be larger in YouTube, possibly due to its larger audience. In contrast, titles and tags are larger in YahooVideo. The larger tags might be due to their collaborative nature. However, compared to LastFM, where the feature is also collaborative, YahooVideo tags are only slightly larger than the restrictive YouTube tags. This is possibly because, in contrast to LastFM, where tags are used for organizing personal libraries, there is no clear incentive for YahooVideo users to explore them. We also found significantly larger descriptions in LastFM, possibly due to its wiki-like collaborative nature and to the type of the associated object — that is, users might feel inclined to write more about an artist than about a specific video.

We further analyzed the amount of content across different object categories. We measured the linear correlation coefficient $\rho$ between the average amount of content and category size (number of objects),[11] finding a range of values ($-0.44 \geq \rho \geq 0.68$) with no clear relationship to any analyzed application characteristic. In some cases, these differences might be related to category semantics. For instance, the YouTube "Gaming" category has, on average, the largest titles, tags, and descriptions, possibly because these features are commonly used to promote the (typically product-related) "Gaming" videos. We also quantified the correlations between the amount of content and object popularity (number of views). Positive correlations exist for largely adopted collaborative features, such as tags ($\rho = 0.17$), description ($\rho = 0.31$), and comments ($\rho = 0.86$) in LastFM and comments in YouTube ($\rho = 0.26$). The insignificant correlations for YahooVideo collaborative features ($-0.007 \leq \rho \leq 0.004$) might be due to its lower audience and lack of incentives for use.

Altogether, title has the best coverage in terms of feature presence, but it contributes the smallest amount of content. Tags typically contribute more content but are absent in many objects. Description and comments follow a similar trend.

In the following analyses, to reduce the impact of usage on assessing descriptive/discriminative power, we disregarded the (largely absent) YahooVideo comments, considering only objects containing nonempty instances of all analyzed features.

**Descriptive Power**

Figure 1 shows the FIS distributions across objects for each feature and application. The *feature instance rank* (*x*-axis) is the position occupied by an instance in the FIS ranking (starting from the highest values) produced considering all feature instances. The figure also shows overall averages and CVs.

For each feature, FIS values vary greatly across objects. Nevertheless, in all applications, title is the most descriptive feature, followed by tags and description, both with significantly lower FIS values. Comments are the least descriptive feature because they commonly contain a lot of noise in the form of unrelated or nonexistent terms used in discussions loosely related to the object. Filtering the most popular terms from the FIS computation has little impact because they represent a small fraction of all terms, given the heavy-tailed term popularity distributions.[8]

As a heuristic, FIS might be influenced by the variability of sizes of feature instances associated with an object; larger instances
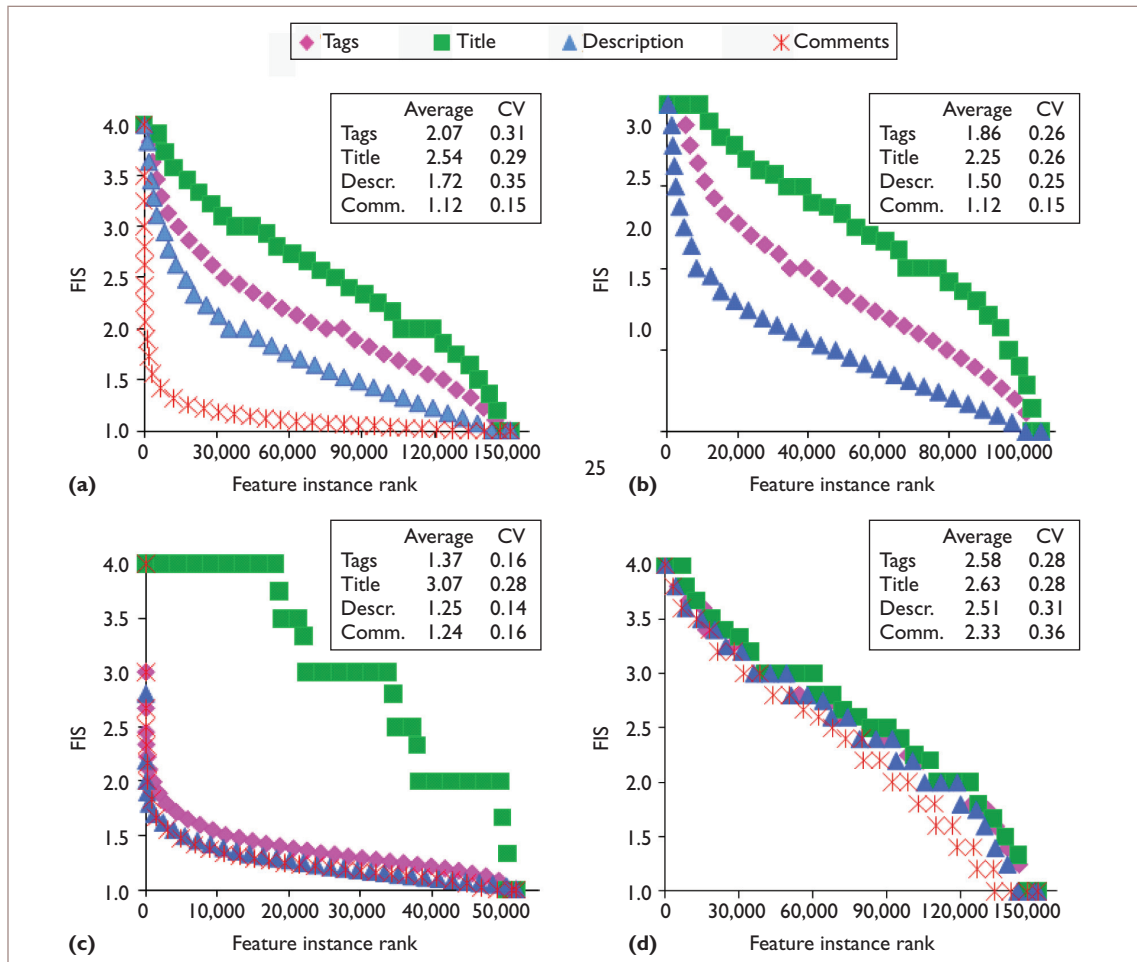
Figure 1. Assessing the descriptive power of textual features using the distribution of feature instance spread (FIS). The results from (a) YouTube (all terms), (b) YahooVideo (all terms), (c) LastFM (all terms), and (d) YouTube (top-five TS terms) vary across objects.

might have lower values simply because there is a higher chance that most of their terms don't appear in the other (smaller) features. This is aggravated when there is a large discrepancy across features, as in LastFM. That is, because titles tend to be much shorter, the other features tend to have much smaller FIS values (see Figure 1c).

To reduce the impact of instance sizes, we recomputed the FIS values considering the $k$ terms with largest $TS$. As Figure 1d illustrates, results improve for all features for YouTube and $k = 5$. As an example, consider a YouTube video entitled "Bedlam cube official Guinness world record" (http://www.youtube.com/watch?v=R1KcN72JbV4). Its set of 52 tags includes the terms *video* and *fitness*, which are only vaguely related to the video's content. By considering only the five tags with the highest $TS$ — world, record, Guinness, cube, and

bedlam — the FIS increases from 1.3 to 3.8. Regardless, the relative ordering of the features, in terms of FIS, remains the same.

Both within and across applications, restrictive features tend to have a higher average FIS, possibly because the same user (object owner) will likely repeat terms across an object's features.[8] As an exception, tags in YahooVideo, despite being collaborative, have a higher FIS than the restrictive description, possibly due to the lack of incentives for collaborative tagging in the system.

Moreover, average FIS values vary only slightly across categories for all features and applications. Thus, the object category doesn't strongly impact the feature's descriptive power (regarding neither the category's semantics nor size). In contrast, significant positive correlations exist between object popularity and FIS for largely adopted collaborative features, such
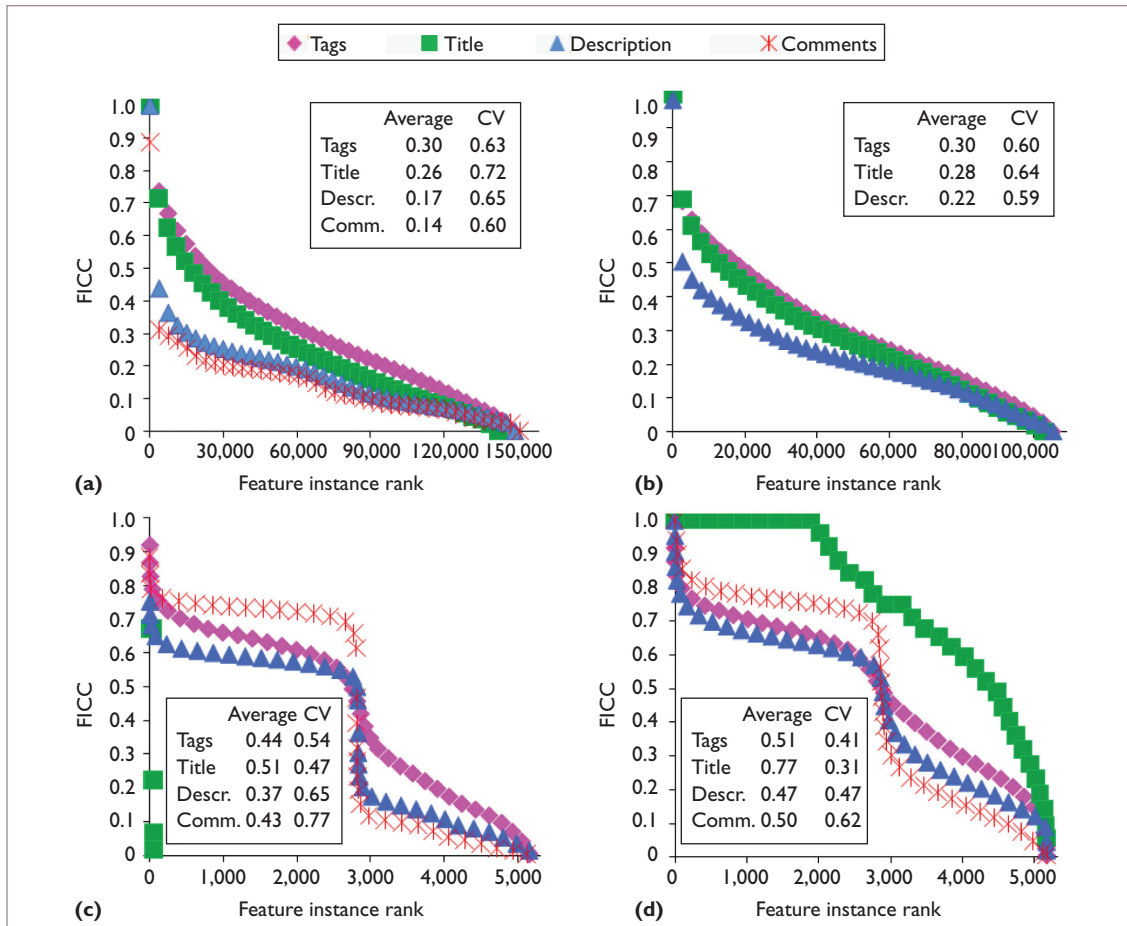
Figure 2. Assessing the discriminative power of textual features by the distribution of feature instance class concentration (FICC). The results for (a) YouTube (with filtering), (b) YahooVideo (with filtering), (c) LastFM (with filtering), and (d) LastFM (without filtering) also show variability across objects. Averages and CVs are computed only over objects with FICC values.

as tags ($\rho = 0.48$) and description ($\rho = 0.35$) in LastFM. Exceptions are comments ($-0.14 \leq \rho \leq 0.0$) and tags ($\rho = 0.008$) in YahooVideo, possibly due to the significant presence of noise (in comments) and lack of incentives for tagging.

We also ran a small experiment with 17 volunteers to investigate whether FIS captures, to a certain degree, users' perception of the features' descriptive power. The volunteers were graduate and undergraduate computer science students familiar with Web 2.0. We selected 10 popular videos from our YouTube collection and asked each volunteer to rate each associated feature according to how well it describes the video. The possible rates were 0, 1, and 2 for the following criteria: feature content is not related, partially related, or completely related to the video's content, respectively.

With 95 percent confidence, the average rates given to title, tags, description, and com-ments were $1.62 \pm 0.09$, $1.57 \pm 0.08$, $1.44 \pm 0.1$, and $0.89 \pm 0.09$. Thus, also according to this experiment, title and tags are the most descriptive features, and comments are the least descriptive. We explain the lack of a clear consensus between title and tags as the best feature by the high degree of subjectivity in user perception, which might be influenced by other factors such as the amount or diversity of content and the features' discriminative power.

**Discriminative Power**

Figure 2 shows the distributions of FICC, computed considering only terms appearing in at least 50 instances, along with overall averages and CVs. Results for other filtering criteria are similar. Like FIS, FICC shows significant variability across objects. In both video applications, the most discriminative feature is tags — which are meant for object organization and

classification — followed by title, description, and for YouTube, comments.

LastFM results exhibit different patterns, however. In particular, title is the least discriminative feature because most title instances carry only specific and rare terms (such as artist names), which are disregarded by our filtering, thus leaving those title instances with no FICC value. In fact, those terms tend to have low generalization capabilities (rarely appearing in both training and test sets), possibly resulting in poor classification.[8] Nevertheless, they might be helpful for searching purposes — that is, for discriminating the single most relevant object related to the artist. In fact, Figure 2d shows that, without filtering, titles have excellent discriminative power because many terms tend to occur in only one (correct) class.

Surprisingly, Figure 2c also shows that comments have good discriminative power in many LastFM objects. This might be due to a skew in our collection that could reflect a bias in the application: approximately 50 percent of the objects belong to the same category ("Pop/Rock"). Thus, many terms are associated with this category, which might affect our FICC results. In fact, this bias might also explain the shape of the three top curves in Figure 2c. Nevertheless, like in the other applications, in general, tags have good discriminative power in LastFM. Moreover, like FIS, FICC tends to be higher in restrictive features, possibly because the object owners tend to use terms more related to the class they assigned to the object.

We found strong positive correlations between the average FICC and category sizes across features and applications ($0.7 \leq \rho \leq 1.0$), reflecting the aforementioned bias toward larger classes. This bias is inherent to the nature of discriminating objects into different categories and thus might affect any metric that expresses that capability. In contrast, the correlations between FICC and object popularity are typically low ($0.00 \leq \rho \leq 0.12$).

Finally, compared with YouTube and YahooVideo, LastFM has higher average FICC in all features. Whereas the more skewed object distribution might contribute to that, the semantic overlap between categories could also be a relevant factor. Although such overlap is hard to quantify, the boundaries of different categories seem more clearly defined in LastFM (for example, "Blues" and "Jazz") than in YahooVideo and YouTube (for example, "Comedy" and "Entertainment").

## Correlations among Quality Aspects

Finally, we investigated the degree of agreement among quality aspects by quantifying the correlations among the amount of content, FIS, and FICC (with filtering criterion) for each feature and application. We found most correlations to be low ($-0.25 \leq \rho \leq 0.25$), indicating no strong influence of one aspect (as captured by the corresponding heuristic) over the other. This might be because each heuristic was designed to target one aspect, thus possibly overlooking relevant factors for other aspects. For instance, whereas some popular terms might appear in many features of an object, thus having high *TS*, they might also appear in objects of many classes, thus having low TCC.

Exceptions are strong negative correlations between FIS and the amount of content for tags and description in YouTube ($\rho = -0.32$, $\rho = -0.34$) and YahooVideo ($\rho = -0.53$, $\rho = -0.46$), reflecting the impact of instance sizes on FIS. Considering only the top-five *TS* terms, the two aspects are positively correlated in LastFM tags ($\rho = 0.80$), description ($\rho = 0.36$), and comments ($\rho = 0.31$), and in YouTube comments ($\rho = 0.41$). Because these features are collaborative, the larger the number of terms, the higher the chance they contain a few good object descriptors.

We analyzed the quality of four textual features in three Web 2.0 applications with respect to feature usage, descriptive, and discriminative powers. Regarding usage, a trade-off between object coverage and the amount of content per object leads to no clear winner. Moreover, titles emerged as the most descriptive feature in the three applications, followed closely by tags. Finally, tags tend to carry the most discriminative terms in both video applications, followed somewhat closely by title. In LastFM, in contrast, title presents the best results, and though not generalizable, they might be a useful feature for searching. For other services, such as classification, tags are a better feature.

The most promising feature for a specific service depends strongly on the relative importance of each quality aspect to service effec-

tiveness. Our findings provide valuable insights for understanding information quality on the Web 2.0 and can drive the design of future applications and services. 🕮

## References

1. S. Boll, "MultiTube – Where Web 2.0 and Multimedia Could Meet," *IEEE MultiMedia*, vol. 14, no. 1, 2007, pp. 9–13.
2. D. Ramage et al., "Clustering the Tagged Web," *Proc. 2nd ACM Int'l Conf. Web Search and Data Mining*, ACM Press, 2009, pp. 54–63.
3. R. Schenkel et al., "Efficient Top-k Querying Over Social Tagging Networks," *Proc. 31st Ann. Int'l ACM Conf. Research and Development in Information Retrieval*, ACM Press, 2008, pp. 523–530.
4. B. Sigurbjornsson and R. van Zwol, "Flickr Tag Recommendation Based on Collective Knowledge," *Proc. 17th Int'l World Wide Web Conf.*, ACM Press, 2008, pp. 327–336.
5. K. Bischoff et al., "Can All Tags Be Used for Search?" *Proc. 17th ACM Conf. Information and Knowledge Management*, ACM Press, 2008, pp. 193–202.
6. C. Marshall, "No Bull, No Spin: A Comparison of Tags with Other Forms of User Metadata," *Proc. 9th ACM/IEEE CS Joint Conf. Digital Libraries*, ACM Press, 2009, pp. 241–250.
7. P. Heymann, G. Koutrika, and H. Garcia-Molina, "Fighting Spam on Social Web Sites: A Survey of Approaches and Future Challenges," *IEEE Internet Computing*, vol. 14, no. 6, 2007, pp. 36–45.
8. F. Figueiredo et al., "Evidence of Quality of Textual Features on the Web 2.0," *Proc. 18th ACM Conf. Information and Knowledge Management*, ACM Press, 2009, pp. 909–918.
9. T. Mitchell, *Machine Learning*, McGraw-Hill, 1997.
10. D. Fernandes et al., "Computing Block Importance for Searching on Web Sites," *Proc. 16th ACM Conf. Information and Knowledge Management*, ACM Press, 2007, pp. 165–174.
11. J.L. Rodgers and W. Nicewander, "Thirteen Ways to Look at the Correlation Coefficient," *The American Statistician*, vol. 42, no. 1, 1988, pp. 59–66.

**Jussara M. Almeida** is an associate professor of computer science at the Universidade Federal de Minas Gerais, Brazil. Her research interests include performance modeling and analysis of large-scale distributed systems, workload and user-behavior characterization and modeling, as well as Web 2.0 quality of information. Almeida has a PhD in computer science from the University of Wisconsin–Madison. Contact her at jussara@dcc.ufmg.br.

**Marcos André Gonçalves** is an associate professor of computer science at the Universidade Federal de Minas Gerais, Brazil. His research interests include information retrieval, digital libraries, text classification, and text mining. Gonçalves has a PhD in computer science from Virginia Tech. He's an affiliated member of the Brazilian Academy of Sciences. Contact him at mgoncalv@dcc.ufmg.br.

**Flavio Figueiredo** is a PhD student at the Universidade Federal de Minas Gerais, Brazil. His research interests include social networks, quality of information in user-generated content, and user-behavior modeling. Figueiredo has an MSc in computer science from the Universidade Federal de Minas Gerais, Brazil. Contact him at flaviov@dcc.ufmg.br.

**Henrique Pinto** is an MSc student at the Universidade Federal de Minas Gerais, Brazil. His research interests include social networks and content recommendation. Pinto has a BSc in computer science from the Universidade Federal de Minas Gerais, Brazil. Contact him at hpinto@dcc.ufmg.br.

**Fabiano Belém** is an MSc student at the Universidade Federal de Minas Gerais, Brazil. His research interests include knowledge management, information retrieval, social networks, and digital libraries. Belém has a BSc in computer science from the Universidade Federal de Minas Gerais, Brazil. Contact him at fmuniz@dcc.ufmg.br.

**cn** *Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.*