

PodCred: A Framework for Analyzing Podcast Preference

Manos Tsagkias Martha Larson Wouter Weerkamp Maarten de Rijke
e.tsagkias@uva.nl m.a.larson@uva.nl w.weerkamp@uva.nl mdr@science.uva.nl

ISLA, University of Amsterdam
Kruislaan 403, 1098 SJ Amsterdam

ABSTRACT

The PodCred framework is a framework for assessing the credibility and quality of podcasts published on the internet. It consists of a series of indicators designed to support prediction of listener preference of one podcast over another, given that both carry comparable informational content. The indicators are grouped into four categories pertaining to the *Podcast Content*, the *Podcaster*, the *Podcast Context* or the *Technical Execution* of the podcast. We adopt the term “cred” as a designation encompassing both credibility (comprising trustworthiness and expertise) and qualitative acceptability to listeners. Our podcast analysis framework is inspired by work on credibility in blogs, another medium dominated by user generated content. The PodCred framework is derived from a review of the literature on credibility for other media, a survey of prescriptive standards for podcasting, and a detailed data analysis of award winning podcasts. The paper concludes with a discussion of future work in which the framework will be applied.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

General Terms

Measurement, Experimentation, Human Factors

Keywords

Podcasts, Credibility, Quality, User preferences, Assessment

1. INTRODUCTION

Podcasts are a series of audio files distributed over the internet using feed-based syndication [17, 21]. This paper proposes an analysis framework, called PodCred, which is especially designed to support prediction of user preferences regarding podcasts. Some definitions liken podcasts to radio programs [8, 11]. However, there are three main differences between podcasting on the internet and radio broadcasting. First, podcasting is targeted at a specific listener group with a focused interest. For this reason podcasting is

also referred to as narrowcasting [10]. Second, podcasts are created with the intention that they will be available for download in the longer term and potentially reused [10]. Third, it is possible to produce and publish podcasts without specialized equipment [7]. As a result, the podosphere, the totality of all podcasts on the internet, contains a high proportion of unscripted, unedited, user generated content alongside professional content. Podcast listeners either subscribe to a podcast or download episodes individually for listening. A podcast feed is published to the internet conforming to a feed standard (an XML standard). When a listener subscribes to a podcast feed, new episodes are downloaded automatically by a feed aggregator. They can be directly loaded onto the listener’s mp3 player, but can also be listened to on a computer.

The podosphere is large,¹ growing and its growth is projected to continue [2, 11]. This prognosis implies an ever more pressing need for methods that provide searchers with intelligent access to the podosphere and help them find podcasts that they would like to listen to. Listeners need to be able to locate podcasts that treat topics that interest them and this issue has attracted recent research interest [6, 15]. However, narrowing the podosphere by topic is only part of the challenge. Within a particular topic, some podcasts will be more appealing to listeners than others.

We are interested in the question of how to characterize the inherent properties of podcasts that signal credibility and quality to listeners. The motivation driving our research is to lay the groundwork for the development of a statistic model of user podcast preference. As a first step, reported in this paper, we investigate the podosphere without anticipating considerations of the technological feasibility of automatically extracting indicators from podcasts. Such automatic extraction is necessary in order to train a model of podcast preference on large amounts of internet data. We chose an initial approach that does not take automatic indicator extraction into account, since extraction technology evolves quickly and methods not currently available might be developed in the near future. Further, we believe that the PodCred framework might serve to guide priority setting for the development of new recognition, detection and classification technologies for spoken audio.

The final goal of this research, the statistical model that will be developed in future work, will allow us to integrate information about listener preference into algorithms for podcast retrieval. Such information will make it possible to improve the ranking of a list of podcasts all related to a user query in such a way that it provides better support to the user during the podcast selection process. The model will also allow us to mine the podosphere for podcasts that have potential to appeal to listeners, but which have not yet attracted enough notice online to achieve high popularity scores.

¹Apple iTunes, one of the most extensive Podcast directories, upholds the claim for an inventory of 100.000 podcast episodes.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WICOW’08, October 30, 2008, Napa Valley, California, USA.
Copyright 2008 ACM 978-1-60558-259-7/08/10 ...\$5.00.

Since we wish to predict popularity, we do not include indicators reflecting listener scores, recommendations, reviews or download rates in our PodCred framework.

Perceptions of credibility are widely considered to have two primary components: expertise and trustworthiness. [12, 13, 19, 20]. We make the basic assumption that information seekers searching for podcasts on the internet also are sensitive to these factors, in other words, that they prefer podcasts published by podcasters with expertise, i.e., who are knowledgeable about the subject and who are trustworthy, i.e., reliable information sources and lacking in motivation to deceive listeners. We believe however, that this bipartite understanding of credibility needs to be extended in order to capture the process by which listeners select podcasts. We would like our analysis to accommodate an additional factor, parallel to the attractiveness or appeal mentioned as a secondary component of credibility in [12, 13].

Motivation for applying an extended view of credibility to analyzing podcasts derives from recent work in our group suggesting that the need that prompts searchers to seek podcasts comprises an informational *and* an entertainment component [3]. If podcasts are a pastime, concerns in addition to strict information reliability must enter into listener podcast choices. In order to capture additional desiderata users impose, it is useful to emphasize *acceptability* aspects of podcasts. With acceptability we mean the desirability or listener-appeal of a podcast arising from factors other than the believability of its propositional or declarative content.

The connection between acceptability and credibility manifests itself in the vernacular use of the term “credibility” in the expression *street credibility*, or *street cred*. In this context, “credibility” is used with a strong connotation of acceptance and approbation. In this paper, we develop a framework for analyzing podcasts in terms of their inherent, query-independent listener appeal. We use a short form, “cred,” in the name of our framework as a reminder that we adopt an extended understanding of credibility where next to trustworthiness and expertise, acceptability and appeal play a role. We conceptualize this extended understanding to also be consistent with frameworks in which credibility is studied as a major, but subordinate aspect involved in people’s perceptions of information quality, such as mentioned in [9].

It is important to understand how the domain that is meant to be analyzed using the PodCred analysis framework is delineated. The domain we cover with our framework is podcasts, which we define to be a series of audio files published to the web at regular intervals using a feed and intended for download and playback on a portable player. Podcasts can be further divided into music-based podcasts and spoken word podcasts. Our work concentrates on podcasts containing spoken word content. Although there is no formal genre classification framework, podcasts tend to fall into genre categories, as has been noted, for example, by [8]. Two major genres of spoken word podcasts that can be easily identified are talk show podcasts, which are sometimes redistributions of shows that have run on the radio, and how-to podcasts, which give commentary or advice on particular subjects. Podcasts must be differentiated from other forms of internet multimedia, such as single audio or video files published to the web. In particular, note that we exclude from our domain of investigation question answers (cf. jabbits²), livestreams, internet radio, such as Live365,³ audio books, spoken Wikipedia articles,⁴ and sites that use speech synthesis to create feeds of audio material from content originally created

²<http://jabbits.com> – accessed 14 August 2008

³<http://www.live365.com/> – accessed 14 August 2008

⁴http://en.wikipedia.org/wiki/Wikipedia:Spoken_articles – accessed 14 August 2008

as text, such as Speakapedia⁵ and Dixero.⁶

Although we feel that it is important to clearly delineate the scope of the domain we conceived the PodCred framework to cover, we have also aimed to build in flexibility so that the framework can keep pace as the concept of podcasting develops. The most widely accepted etymology for the term *podcast* is that it is audio similar to broadcast audio created to be played on an iPod or similar player. An alternate derivation is that the *pod* of podcast stands for Play On Demand.⁷ We adopt the perspective of the latter, since we believe that in the future podcasting is sure to expand with respect to end device (for example, become more oriented to mobile phones) and/or shift medium (include increasing amounts of video content). Bloggers delight in declaring podcasting dead,^{8,9} often citing the rise of video. In any case, the basic phenomenon of a syndicated multimedia series that can be generated without professional equipment and which is targeted towards a specific audience is an enduring one. The framework we propose here provides a solid basis on which analysis of this phenomenon can build.

The PodCred analysis framework presented here is designed to be appropriate to serve as a basis for our future work. As mentioned above, our aim is to understand the factors inherent in podcasts that contribute to listener perception of the appeal of podcasts in order to predict popularity. The information search scenario is the following: a user makes use of a podcast search engine to search for a podcast on a particular topic with the goal of subscribing to that podcast. In response to the user query, the search engine returns a list of podcasts. The user reviews these podcasts by reading the feed-level metadata (i.e., podcast title and description) scanning the list of episodes and listening to, or briefly auditioning, a couple of the episodes. We are interested in understanding on the basis of a relatively quick review of a podcast, what motivates a user to chose to subscribe to one podcast over another. This information will help us to refine the results list presented to the user by the search engine by improving the ranking of more desirable podcasts.

In the next section, we discuss the motivation for proposing an analysis framework dedicated to podcasts. Then we introduce our proposed analysis framework, which takes the form of a list of indicators for podcast preference. The indicators fall into four categories: *Podcast Content*, the *Podcaster*, the *Podcast Context* and the *Technical Execution*. The next three sections discuss in turn the three sources that contributed to the development of the framework: the credibility literature, the prescriptive literature on podcasting and a human analysis of 19 popular podcasts. In the final section of the paper, we summarize and present an outlook on the role that the PodCred analysis framework will play in our future work.

2. MOTIVATION

Analyzing the credibility of internet content is a challenging task due to the lack of professional gatekeepers on the internet, who moderate or filter content as is the case in conventional media such as newspapers or television [12, 13]. Our motivation to set up an analysis framework for podcasts distinct from that applied to conventional internet content is parallel to the motivation for de-

⁵<http://shinydevelopment.com/speakapedia> – accessed 14 August 2008

⁶<http://www.dixero.com/> – accessed 14 August 2008

⁷<http://www.netcastcentral.com/podcast.php> – accessed 14 August 2008

⁸<http://althouse.blogspot.com/2007/08/podcasting-is-dead.html> – accessed 14 August 2008

⁹http://www.informationweek.com/blog/main/archives/2008/01/is_podcasting_d.html – accessed 14 August 2008

veloping credibility frameworks for blogs. Blogs are temporally structured, with new posts being added over time. Blogs are often user-generated and contain primary source descriptions of people's lives and surroundings; bloggers build a tightly knit social network structure [14]. Bloggers are individualistic and develop their own voices [22]. These characteristics hold of podcasts and podcasters as well. The literature has applied a dedicated credibility analysis framework for blogs since users approach blogs differently than other forms of web content [19, 23]. In particular, in the blogosphere the process of building credibility is a dynamic one, involving exchange between bloggers and readers, with bloggers revealing their real-world identity and personal details as part of the process to establish trust [19]. Bloggers wear their bias on their sleeve, so to say, establishing trust not by being objective, but by being open about how they are not objective, cf. [19].

Although research on credibility in the blogosphere can inform an understanding of the podosphere, it is not possible to adopt a blog credibility analysis framework, such as the one presented in [19], for use in podcast analysis without modification and extension. The most fundamental difference between blogs and podcasts that motivates a dedicated podcast analysis framework is self-evident: podcasts are audio-based. Although podcasts involve text and textual descriptions in their metadata, the core of a podcast is constituted by audio and for this reason audio characteristics and also characteristics of human speech must be taken into account. A single podcast often contains rapid crossfire conversation: such exchanges are not characteristic of blogs. Other differences between podcasts and blogs are more subtle. As mentioned above, we have reason to believe that information seekers would like not only informational but at the same time entertaining podcasts [3]. Although blogs are without doubt also entertaining, the two cases have important differences. Reading a blog is a dedicated intellectual activity: the reader sits at the computer and focuses on the blog content. Multi-tasking during blog reading is effectively limited to activities such as listening to music. Listeners often search for podcasts, however, for consumption during other activities, such as housework or exercise. Consequently, an understanding of the acceptability/appeal dimension of podcasts needs to encompass aspects designed to capture the extent to which the listener can follow the content while carrying out other activities. An additional difference between blogs and podcasts is that the number of podcasts a given user can consume is much more limited than the number of blogs. Podcasts compete directly with each other for the listener's attention: subscribing to a new podcast quite possibly means dropping an old podcast [7]. A podcast analysis framework has to be able to accommodate very stringent criteria of users. We anticipate that a decision to stop subscribing to a podcast will often involve overriding clear preference indicators and be made solely on the basis of the fact that listening time is limited.

During development of our analysis framework, we are careful to maintain the perspective that podcasting has its roots in broadcasting. For this reason, credibility and quality indicators for radio often apply to podcasts. In particular, we anticipate that characteristics reflecting well crafted audio production will be important. Such a parallel has also been exploited in work on blog credibility, where credible blogs have been assumed to have the same indicators as credible newspapers [23]. Podcasts, however, cannot be analyzed as simply internet radio. As stated above, podcasts are narrowcasts targeted to a certain group and the podosphere contains a high proportion of user generated content. A proper understanding of listener preferences necessitates the development of a new assessment framework, a task to which the remainder of this paper is dedicated.

3. PODCRED ANALYSIS FRAMEWORK

This section introduces the PodCred podcast analysis framework, which consists of a list of indicators to be taken into account when assessing a podcast for credibility and appeal. In order to generate the framework, we carried out three studies: a review of the credibility literature, a survey of the prescriptive literature on how to create podcasts and a data analysis of popular podcasts. These three studies are detailed in the next section.

The PodCred framework is illustrated in Table 1 and consists of four top-level categories of indicators. The first category, *Podcast Content*, captures the quality and consistency of the informational content of the podcast. This category is designed to reflect the ability of the podcast to satisfy a particular, but yet unspecified, information need of the user. Podcast Content indicators require the podcast to be consistently about a topic. Such topical focus is necessary for a podcast to reflect a particular interest. Also included in the Podcast Content category are indicators reflecting types of content. These indicators are included to capture effects specific to user generated content, namely that information seekers place value on personal content [3]. Opinions, testimonial and recommendations are personal since they arise from the experience and convictions of an individual and not via the consensus of experts or social convention. The second category of indicator relates to the *Podcaster*. The explicit representation of the creative agent of the podcast in the framework is important. The two main components of credibility, expertise and trustworthiness, reflect that the information source is regarded as capable of intelligence and volition, both characteristics of a human agent. Moreover, specifically in the case of user generated content, credibility is built by public disclosure of personal identity and personal detail [19]. Particularly important in the Podcaster category are elements relating to the speech of the podcaster, these aim to help capture not only the potential appeal of the podcaster's persona, but also the basic ease-of-listening of the podcast. The third category of indicator is *Podcast context*. This category reflects the effect that the podcast builds its reputation by involving a whole range of other human intelligences in a network of credible sources. User generated content has been described as involving a process of information exchange [19]. When a podcast is tightly integrated with its information sources and with its listener group, not only is the source of the podcaster's content clear, but also the impact of the podcaster's content. Additionally, user generated content derives credibility by avoiding covert bias [19]. We include sponsors/stores/advertisers since they make transparent the source of podcast funding, shedding light on potential podcast bias. The final category of indicators is *Technical Execution*. These indicators are specific to podcasts and reflect the level of effort and time that have been invested in the production of the podcast.

The PodCred framework can be considered to belong to a class of credibility assessment approaches that has been called *Checklist Approaches* by [12]. Such approaches aim to list the factors that contribute to judgments of credibility rather than building a cognitive model of the process by which credibility is assessed. It is important to mention one way in which our analysis framework diverges from a strict checklist approach. In a strict checklist approach, the presence of all checklist factors would indicate maximum credibility. In our framework, we would like to leave open the question of whether some of the indicators are actually either/or indicators. For example, the breakdown of the podosphere into genres discussed in the introduction suggests that listeners do not require podcasts to contain both encyclopedic/factual information and discussion/opinions in equal measures, but that fact-intensive or spirited exchange of opinions could contribute to podcast appeal. Also, we would like to leave open whether or not the indicators are

| PodCred Analysis Framework | |
|--------------------------------|--|
| Podcast Content | |
| Spoken Content | Podcast has a strong topical focus Appearance of on-topic guests Use of field reports Contains encyclopedic/factual information Contains discussion/opinions Contains commentary/testimonial Contains recommendations/suggestions Podcaster cites sources |
| Content Consistency | Podcast maintains its topical focus across episodes Consistency of episode structure Presence/reliability of inter-episode references Episodes are published regularly |
| Podcaster | |
| Podcaster Speech | Fluency/lack of hesitations Speech rate Articulation/Diction Accent |
| Podcaster Style | Use of conversational style Use of complex sentence structure Podcaster shares personal details Use of broad, creative vocabulary Use of simile Presence of affect Use of invective Use of humor Episodes are succinct |
| Podcaster Profile | Podcaster eponymous Podcaster credentials Podcaster affiliation Podcaster widely known outside the podosphere |
| Podcast Context | |
| Podcaster/listener interaction | Podcaster addresses listeners directly Podcast episodes receive many comments Podcaster responds to comments and requests Podcast page or metadata contains links to related material Podcast has a forum |
| Real world context | Podcast is a republished radio broadcast Makes reference to current events Podcast has a store Presence of advertisements Podcast has a sponsor Podcast displays prizes or endorsements |
| Technical Execution | |
| Production | Signature intro/opening jingle Background music (bed) Atmospheric sound/Sound effects Editing effects (fades, transitions) Studio quality recording/no unintended background noise |
| Packaging | Feed-level metadata complete (e.g., title, description) Episode-level metadata complete (e.g., title, date) ID3 tags used Audio available in high quality / multiple qualities Feed has a logo Episodes presented with images |
| Distribution | Simple domain name Distributed via distribution platform Podcast portal Reliable downloading |

Table 1: PodCred Podcast Analysis Framework

necessarily all positive indicators. Rate of podcaster speech, for example, could contribute to listener preference if it is fast (implies mastery of the material) or if it slow (facilitating ease of information uptake). Insight into these areas will be achieved only when we start to apply the framework in our future research.

Although the PodCred framework is designed to apply to the relatively brief assessment process of selecting a podcast for subscription, it has potential to provide a basis for iterative models of credibility assessment. Typical for an iterative model of credibility assessment, users judge the credibility of content in stages (cf. the description in [12]). Users look at different aspects of the material and its presentation, investigate the background of the creative agent, look into cited sources and finally return to modify the original judgment of the content in the light of newly discovered material. We anticipate that the PodCred framework will provide a valuable reference point both in the first iteration and later ones.

4. DERIVING THE FRAMEWORK

This section describes the development of the PodCred framework in Table 1. Three sources make contributions to the framework and are described here in turn: the credibility literature, the prescriptive literature on podcasting and a human data analysis.

4.1 Credibility in the literature

The PodCred framework derives its basic skeleton and many of its indicators from work in the area of credibility assessment. The first two categories of the framework, Podcast Content and Podcaster contain indicators for *Message Credibility* and *Source Credibility*, two important streams from early research on credibility as detailed by Metzger et al. [13]. Following the account of Metzger et al. [13], investigation of message credibility has traditionally concerned itself with the impact of characteristics such as message structure, message content and language intensity including use of opinionated language. Message source credibility research concerns assessments of the person or organization who generates the message. These aspects of credibility have application not only in the area of traditional media, but also for internet content. Source and Content are the first two facets of judgment of information quality on the World Wide Web in the framework laid out by Rieh and Belkin [18]. Message credibility and source credibility have a certain overlap, and it will be noticed in the proposed framework that certain Podcast Content indicators could be argued to also be important Podcaster credibility indicators.

Hillgoss and Rieh [9] lay out a framework of credibility that can be applied across resources and across tasks. Based on a diary study using 24 participants the authors collect 12 credibility assessment types, divided into three levels, construct, heuristics, and interaction. Here, we are interested in leveraging their findings on types of credibility assessment at the heuristic and at the interaction level, since the framework we are designing aims to capture information that will shed light on an assessment process which is superficial and of relatively short duration, i.e., the subscribe/not subscribe decision. At the heuristics level, assessment types are media-related and source-related, corresponding to the classical components of credibility mentioned above. Additionally, the heuristics level contains endorsement-based assessments. In the podcast world, a podcast enjoys endorsement when that podcast finds resonance and acceptance among its listeners. Endorsement based criteria can be found in the Podcast Context category of the PodCred framework. Finally, the heuristics level contains aesthetics-based assessments. The corresponding characteristic of podcasts is simply how they sound. We add a subcategory on podcaster speech and a subcategory on podcast production to capture the sound of podcast. These

elements are designed to be the counterparts of design elements in websites, argued by [13] to contribute to website dynamism and in this way to impact credibility.

While developing the PodCred framework we paid special attention to Rubin and Liddy [19] and van House [22], work on credibility in blogs. As mentioned above, blogs and podcasts share commonalities due to their user generated and social nature. They also share the characteristic that they are temporally structured, meaning that they are published in a series that unfolds over time. The Rubin and Liddy [19] framework involves several indicators that we directly translate from the blogosphere to the podosphere and import into our PodCred system. Perhaps most importantly, we adopt from this framework the factor *blogger's expertise and offline identity disclosure*, which we turned into the subcategory Podcast Profile for application to the podcast domain. Next, we consider indicators related to the temporal nature of blogs, these are listed in the Rubin and Liddy [19] framework as timeliness and organization and we make these specific to the podosphere by introducing an indicator indicating that podcasts track recent events and one indicating their level of consistency and structure. Finally, a critical factor from the the Rubin and Liddy [19] framework is *appeals and triggers of a personal nature*. Under this factor they include literary appeal and personal connection. These are three indicators that we adopt for podcasts, integrating them as elements of the category Podcaster Style. Work by van House [22] in particular mentions the importance of the connection of on-line and off-line blogger identities and the enhancing effect of personal voice. These indicators can be found in the PodCred framework as "Podcaster eponymous" and "Podcaster shares personal details."

The literature makes clear that in addition to inherent factors included in the PodCred framework, both query-dependent and user-dependent factors have an important influence on credibility assessment. Since we hope to eventually extend the model to incorporate such factors, it is fitting that we mention them here. It is essential to recognize the relationship between how users assess content and the informational problem they are facing [18]. Because of this relationship dependency of credibility on query assessment arises. For example, medical information will be assessed in a different way from information about the personal lives of movie stars. The difference is that in the former case the information informs a potentially important decision and in the later case the user does not take any particular action as a result of the information. Metzger et al. [13] observed that factual information is more rigorously checked than entertainment information. We designed PodCred to be applicable to the whole spectrum of podcasts types, from how-to podcasts through to pure entertainment podcasts. Capturing the rigor of the user's assessment as a function of the informational need is a factor not yet built into the model. It is equally essential to recognize the relationship between the user and the assessment process. Users having a high level of information literacy assess credibility with different strategies. User aspects are also topic dependent. Users assess content in a different manner if it treats a topic that they are knowledgeable about [13]. Further, as previously mentioned, new podcasts are measured in appeal with respect to the podcasts that a user already listens to. A user replacing an old podcast with a new, more appealing podcast on the same subject will deploy a different assessment strategy than a user trying to increase topical coverage by introducing new podcasts on novel subjects to the subscription list.

4.2 Prescriptive rules for podcasting

The development of the PodCred assessment framework is informed by the prescriptive literature on podcasting. We assume that

Podcasters have come to understand what makes podcasts popular and what kind of podcasts people generally want to listen to. For this reason, we take podcaster recommendations and prescriptive guidelines into account when formulating the framework.

Our results are based on information found at websites focusing on helping podcasters produce better shows. A better podcast is considered to be a podcast that promotes the popularity of the podcaster and creates a community around the show with the ultimate goal of reaching more listeners. We look into three sources. First, *Podcast Academy*,¹⁰ a podcast containing material ranging from keynotes of podcasting conferences to interviews with guests from the podcasting domain. Second, *Podcast Underground*,¹¹ a podcast portal making information available about how to improve and enhance the content and the exposure of a podcast, including an article¹² which is composed of comments from individual podcasters who report their personal experiences, successes and failures while experimenting with the new medium. Third, *How to Podcast*,¹³ a website providing a step-by-step guide describing the steps involved in podcast production, the required key elements that should be present to make it worthwhile listening, and rough guidelines for success in terms of number of subscribers.

The study of prescriptive podcasting guidelines reinforces the inclusion of the indicators deriving from the credibility literature in our podcast analysis framework. In the rest of this section, we look at what our prescriptive sources have to say about each of the indicator categories in turn.

First, the indicators in the Podcast Content category found considerable support in the prescriptive podcast guidelines. Heavy emphasis is placed on keeping the podcast focused on one topic. The fact that one of the most prevalent suggestions found in the sources we consulted draws the attention of podcasters on treating only one topic in their podcast is hardly surprising. As mentioned above, an important difference between podcast content and radio content is the narrow target of the podcast. Podcasts are topical or of specific genre creating a pole for people with interest in that topic or genre. Listeners should be introduced to the structure within the episode, resulting in a natural flow and along with a steady pace will enhance progression. Podcasters can achieve structure and focus by doing background research before recording and even preparing transcripts that will guide them through during the show. Another suggestion is the introduction of segments in each episode that are constant throughout the podcast. Segments remind the listeners of broadcast shows and give them something to anticipate. Regularity of episode releases is of primary importance and it is cited by all our sources, who advise maintaining regularity as a method to accustom listeners to the release schedule and increase their loyalty to the program. Interviews with popular and well-known people in the domain are highly recommended.

Second, all sources stress factors that support the indicators in the Podcaster category of our PodCred framework. Our sources point out that if a show is to become popular, the podcaster should be knowledgeable and passionate about the topic they are talking about. Although the importance of indicators related to the source of content emerges clearly from the credibility literature examined in the previous section, prescriptive rules for podcasts go one

step further underlining the importance of sharing personal experiences and stories. In this way, listeners are able to learn from podcaster's own experiences and identify with them, creating a bond between the listener and the podcaster. Podcasters report that building two emotions in the story told makes it more appealing e.g., love and humor, humor and sadness. In short, our prescriptive podcast sources provide direct justification for the inclusion of the indicators involving personal details, affect and podcaster credentials.

Third, strong supportive for Podcast Context categories emerges from the prescriptive sources. The sources advise that the podcaster should stay current with the developments in the podosphere in terms of what topics are treated in other podcasts of the same domain, but also they should be taking into consideration the comments and suggestions from their audience. With the respect to the interaction between the podcaster and their listeners, experts advise the activation of multiple interaction channels: subscription to syndication feed (e.g., iTunes), forums, voicemails, emails, blog comments, store and donation options. Podcaster's activity and response in fora discussions and comments is fundamental, refueling the interactivity cycle. To ensure sustainable listener interaction, experts advise establishing a forum instead of supporting user comments on individual posts. Post comments establish the expectation of a response from the podcaster, whereas a forum is basically communal in nature and the podcaster can mix into the discussion, but podcaster participation is not necessarily expected and listeners respond to each others' questions.

Fourth, our prescriptive podcast sources corroborated the indicators in the Technical Execution category of our PodCred framework. Experts recommend enhancing audio quality by editing the final audio, e.g., adding sound effects, cross-fades between sections, removing sentence fillers (e.g., uhm, uh), long periods of silence. A quiet recording environment and semi-professional microphones is suggested for minimizing the background noise. Another use of transcripts is that work well also for minimizing long periods of silence, use of sentence fillers or hesitations.

4.3 Human analysis of prize winning podcasts

In order to uncover additional indicators to include in the PodCred framework and to confirm the appropriateness of the framework for real world data and to seek, a human analysis is conducted on a set of podcasts on the internet known to be appealing to listeners and widely subscribed. For the analysis we chose the prize winning podcasts as announced in Podcast Awards¹⁴ for 2007.

People's Choice Podcast Awards are an annual set of awards given to the best podcasts as voted by the people. The convention begun in 2005 and it is held by a private-held company named PodCast Connect Inc. People are welcomed to nominate their favorite podcast for each category by providing the podcast's name and URL, voter's name, email address and comments. The podcasts submitted should have more than 8 episodes published from the 1st of May of the contest's year, their feeds should contain the audio enclosures and they must be in an appropriate category. The contest offers 22 prizes, one for each topical category and two extra awards for "People Choice's" and "Best Produced". Contest's categories correspond to iTunes main categories, with few exceptions where iTunes main- and sub-categories are merged together. The categories are: People Choice's, Best Produced, Best Video Podcast, Best Mobile Podcast, Business, Comedy, Culture/Arts, Education, Entertainment, Food and Drink, Gaming, General, GLBT, Health/Fitness, Movies/Films, Podsafe Music, Political, Religion Inspiration, Sports, Technology/Science and Travel. For our anal-

¹⁰<http://www.podcastacademy.com> – accessed 14 August 2008

¹¹<http://www.podcastunderground.com> – accessed 14 August 2008

¹²<http://www.podcastunderground.com/2007tips/> – accessed: 16 July 2008

¹³<http://www.how-to-podcast-tutorial.com> – accessed: 16 July 2008

¹⁴<http://www.podcastawards.com> – accessed 14 August 2008

ysis, we omitted the Best Video Podcast since the PodCred framework does not cover video content. A total of 19 unique podcasts were awarded in 21 categories. 2 podcasts were given the award for two categories; “Keith and the girl” for People’s Choice and Mature, and “Firefly Talk” for Best Produced and Movies/Films. The analysis was carried out by looking at the podcast feed, the podcast portal (if there is one) and listening to at least one, but usually several, episodes from each podcast. This process was designed to be parallel to that used in our search scenario where a user examines a podcast to make a subscribe/not-subscribe decision.

During the analysis several important indicators emerged, which we included in the framework. Nearly all the podcasts surveyed use a standard opening jingle. A large number have associated websites (i.e., podcast portals) and include images and links. The podcasters frequently cite their sources, either by providing website URLs, quotes from people, or book/article excerpts. The vocabulary used can be classified as daily, but expanded with terminology from the podcast’s focal domain. Although the syntax observed was simple with complete sentences, variance occurs when more than one speaker was involved or conversational speech takes place.

These characteristics are incorporated in our proposed framework. Additionally, we looked for corroboration of the indicators from the credibility literature and prescriptive guidelines. We made observations for a number of selected indicators from the PodCred framework, recording the number of podcasts in which these indicators were present. In Table 2, the findings of the podcast analysis are presented. Observations are divided over the four main indicator categories of the PodCred framework. The absence of an indicator in the table implies that it was not analyzed and should not be taken to mean that it was never observed.

Several observations we made during the analysis were particularly striking. Most of the podcasts analyzed are of conversational style, usually with two speakers; one host and one guest. Frequently, a podcast has one host with more than one guest, or has a multi-speaker format, like “ShowGirls” with seven women co-hosts. Episode release regularity ranges from two episodes per day to monthly or even absence of regular releases. The majority of podcasts is published on a daily or weekly basis. All but one podcast comes with complete feed metadata, half of them limit their podcast-level metadata to one sentence descriptions. At the level of episodes, rich metadata is observed for all except two shows which do not supply such information. Finally, the analysis revealed that interactivity between the podcaster and the listeners is an important characteristic of good podcasting. Three-quarters of the podcasters address the listeners directly and the same amount receive a large volume of comments. The community building factor is quite high in our sample having 10 podcasts providing a forum for their listeners. Some podcasters respond directly to the episode comments, but podcasters also respond on fora or by giving feedback from inside the episode.

5. DISCUSSION AND OUTLOOK

This paper has presented the PodCred analysis framework that we have developed for podcasts based on a review of the credibility literature, a survey on the prescriptive literature on podcasting and human data analysis. The PodCred framework was developed to provide a basis for future work.

The first step will be to carry out additional data analysis in order to confirm that the indicators in this framework are appropriate to identify podcasts with low appeal. In order to carry out such an analysis, we will compare podcasts treating the same topic ranked by popularity. We expect that indicators we have discussed here will emerge as important in distinguishing podcasts that are popu-

| Observed indicator | No. of podcasts |
|----------------------------------|-----------------|
| Category Podcast Content | |
| Topic podcasts | 13 |
| Topic guests | 8 |
| Opinions | 14 |
| Cite sources | 15 |
| One topic per episode | 9 |
| Consistency of episode structure | 14 |
| Interepisode references | 8 |
| Category Podcaster | |
| Fluent | 17 |
| Presence of hesitations | 7 |
| Normal speech speed | 8 |
| Fast speech speed | 10 |
| Slow speech speed | 1 |
| Clear diction | 14 |
| Conversational vocabulary | 7 |
| Invective | 1 |
| Humor | 10 |
| Multiple emotions | 4 |
| Personal experiences | 15 |
| Credentials | 10 |
| Affiliation | 4 |
| Podcaster eponymous | 10 |
| Category Podcast Context | |
| Podcaster addresses listeners | 15 |
| Episodes receive many comments | 15 |
| Podcaster responds to comments | 9 |
| Links in metadata/podcast portal | 13 |
| Advertisements | 10 |
| Social network | 4 |
| Forum | 10 |
| Technical Execution | |
| Opening jingle | 16 |
| Background music | 7 |
| Sound effects | 8 |
| Editing effects | 10 |
| Studio quality recording | 13 |
| Background noise | 5 |
| Feed-level metadata | 18 |
| Episode-level metadata | 16 |
| High quality audio | 13 |
| Feed has a logo | 11 |
| Associated images | 11 |
| Simple domain name | 14 |
| Podcast portal | 16 |
| Logo links to podcast portal | 7 |

Table 2: Analysis of 19 award winning podcasts

lar from podcasts that are not popular. Further, we do not exclude the possibility that the framework presented here might be extended with indicators that signal that a podcast is not popular. This analysis must be of large enough scale in order to make an estimation of the relative importance of indicators in the PodCred framework. Indicators that turn out not to be important may be eliminated from consideration or given low priority in further investigation.

The next step is to tackle our goal of applying automatic methods to extract features that correlate with the PodCred indicators. We can then collect large amounts of data from the internet and use it to train statistical models that capture podcast preference. The methods that we anticipate applying to automatically extract indicators include text analysis, audio processing and speech recognition techniques. Some indicators will be quite easy to automatically extract. E.g., completeness of feed and feed item metadata and also regularity and recency of feed item publication can be ascertained by parsing the feed and analyzing or comparing the relevant feed elements. Others will be more challenging. Laughter and other

non-speech sounds such as applause can be automatically detected, (cf. e.g., [4]) making it possible to identify not only the likelihood or humor but also the presence of certain production effects, like music bed. Number of speakers, turn duration and alternation patterns can be automatically extracted and used to predict whether or not the podcast is a discussion or contains conversational speech (cf. e.g., [5]). Richness of vocabulary can be calculated on speech recognition transcripts, if compensation is made for potential recognizer vocabulary limitations. For indicators involving topic and type, we would like to apply classification techniques (cf. e.g., [16]) to metadata and speech recognition transcripts.

Initial applications of automatically extracted credibility indicators in the area of information retrieval for user generated content suggest that with the appropriate statistical modeling techniques the negative effects of noise can be overcome and useful scores can be generated. In recent work, Weerkamp and de Rijke [23] implemented several credibility indicators in a blog post retrieval system. Results show that taking credibility into account can improve retrieval performance. Other research has focused on using the notion of quality to improve access to user generated content. Weimer et al. [24] use five categories of features to classify forum posts as either good or bad (high or low quality) posts, achieving satisfactory results. Further research has addressed community question answering. Agichtein et al. [1] show that using a large set of features, both high quality questions and high quality answers can be identified with satisfactory performance.

In sum, the PodCred assessment model put forward in this paper is designed to be both detailed and flexible enough to model the podosphere as it continues to expand to include more content and broader range of phenomena, (i.e., video) and also serve as the a launching point for a larger research program on the statistical modeling of podcast preference.

6. ACKNOWLEDGMENTS

We thank our anonymous referees for their valuable comments.

This research was supported by the E.U. IST program of the 6th FP for RTD under project MultiMATCH contract IST-033104, and by the Netherlands Organization for Scientific Research (NWO) under project numbers 220-80-001, 017.001.190, 640.001.501, 640.-002.501, 612.066.512, STE-07-012, 612.061.814, 612.061.815.

7. REFERENCES

- [1] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne. Finding high-quality content in social media, with an application to community-based question answering. In *Web Search and Data Mining (WSDM)*, pages 183–194, 2008.
- [2] Arbitron/Edison. The podcast consumer revealed 2008: The arbitron/edison internet and multimedia study, 2008. URL http://www.edisonresearch.com/2008_Edison_Arbitron_Podcast_Report.pdf.
- [3] J. Besser, K. Hofmann, and M. Larson. An exploratory study of user goals and strategies in podcast search. In *Workshop on IR 2008 FGIR der Gesellschaft für Informatik*, 2008.
- [4] K. Biatov and J. Kohler. An audio stream classification and optimal segmentation for multimedia applications. In *MULTIMEDIA '03: Eleventh ACM international conference on Multimedia*, pages 211–214. ACM, 2003.
- [5] K. Biatov and M. Larson. Speaker clustering via bayesian information criterion using a global similarity constraint. In *SPECOM: 10th International Conference SPEECH and COMPUTER*, 2005.
- [6] I. Celma and Y. Raimond. Zempod: A semantic web approach to podcasting. *Web Semant.*, 6(2):162–169, 2008.
- [7] M. Geoghegan and D. Klass. *Podcast solutions: The complete guide to podcasting*. friends of ED, 2005.
- [8] V. Heffernan. The podcast as a new podium. *The New York Times*, July 2005.
- [9] B. Hilligoss and S. Y. Rieh. Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing and Management*, 44(4):1467–1484, 2007.
- [10] J. Louderback. Master radio techniques and avoid radio traps, pnme 2007: Master radio techniques, 2008. URL <http://podcastacademy.com/2008/06/18/pnme-2007-master-radio-techniques/>.
- [11] K. Matthews. Research into podcasting technology including current and possible future uses. *Electronics and Computer Science, University of Southampton*, 2006.
- [12] M. J. Metzger. Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *J. Am. Soc. Inf. Sci. Technol.*, 58(13):2078–2091, 2007.
- [13] M. J. Metzger, A. J. Flanagan, K. Eyal, D. R. Lemus, and R. McCann. *Credibility in the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment.*, pages 293–335. Mahwah, NJ: Lawrence Erlbaum, 2003.
- [14] G. Mishne. *Applied Text Analytics for Blogs*. PhD thesis, University of Amsterdam, 2007.
- [15] J. Ogata, M. Goto, and K. Eto. Automatic transcription for a web 2.0 service to search podcasts. In *Interspeech*, Antwerp, Belgium, August 2007.
- [16] G. Paass, E. Leopold, M. Larson, J. Kindermann, and S. Eickeler. SVM classification using sequences of phonemes and syllables. In *Proc. of the 6th Euro. Conf. on Principles of Data Mining and Knowledge Discovery*, pages 373–384, 2002.
- [17] L. J. Patterson. The technology underlying podcasts. *Computer*, 39(10):103–105, 2006.
- [18] S. Y. Rieh and N. J. Belkin. Understanding judgment of information quality and cognitive authority in the WWW. *J. Am. Soc. Inf. Sci. Technol.*, 35:279–289, 1998.
- [19] V. L. Rubin and E. D. Liddy. Assessing credibility of weblogs. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs (CAAW)*, 2006.
- [20] S. Tseng and B. J. Fogg. Credibility and computing technology. *Commun. ACM*, 42(5):39–44, 1999.
- [21] F. van Gils. PodVinder: Spoken Document Retrieval for Dutch Pod- and Vodcasts. Master's thesis, U. Twente, 2008.
- [22] N. van House. Weblogs: Credibility and collaboration in an online world. unpublished ms., 2002.
- [23] W. Weerkamp and M. de Rijke. Credibility improves topical blog post retrieval. In *HLT-NAACL*, pages 923–931, 2008.
- [24] M. Weimer, I. Gurevych, and M. Mühlhüser. Automatically assessing the post quality in online discussions on software. In *ACL 2007 Demo and Poster Sessions*, pages 125–128, 2007.