

Ratings in Social Music Cataloging Services

Miika Nurminen (minurmin@jyu.fi)
Ville Halonen (vihehalo@jyu.fi)

Abstract

A general survey of social software is presented with a focus in social music cataloging services. Two example systems review and rating functionality are compared in detail. We observed a case of distorted product ratings, where overactive fan base promoted a single album at the expense of others. Despite the potential for unfair ratings, social cataloging systems are a promising form of participatory media, challenging and complementing traditional, professional reviews used in print media. Some suggestions are provided to issue the reliability and fairness of individual reviews.

1. Introduction

During the last few years, different kinds of social web applications have become increasingly popular. There is nothing essentially new here: various community-driven services, such as message forums, newsgroups and even social networking sites have been available and used for years – even before WWW. However, recently the user-generated content and metadata [6] combined with open application programming interfaces, intuitive user interfaces often powered with Ajax (Asynchronous JavaScript and XML), and explicit social network information have opened new possibilities and expanded user base for innovative web applications.

In this assignment, we focus on a specific class of social software, namely music cataloging services. Many studies about the usage patterns of social web applications have focused on earlier communities, such as the bookmarking service del.icio.us¹ [8], photo sharing service Flickr² [14], or news aggregator Digg³ [11], but music (and other) cataloging services have some distinctive characteristics, especially focus on reviews and star-ratings. Keyword tagging [8] can also be used, but because of a restricted domain, it is expected that tags are used in a more limited way like musical genre classification compared to domain-independent folksonomies [15] that general-purpose sites (e.g. del.icio.us) tend to generate.

Our goal is to assess how reviews and ratings are used in different cataloging services, and evaluate their general quality and reliability compared to traditional, professional reviews. The evaluation and discussion of the services is based on general observations during 2007 by first author and side-by side comparison using multiple frameworks by second author.

This assignment is structured as follows: existing social software applications are reviewed in Section 2. A detailed evaluation of selected social music cataloging services is provided in Section 3. The results are discussed in Section 4. Section 5 concludes the assignment.

¹ <http://del.icio.us/>

² <http://www.flickr.com/>

³ <http://digg.com/>

2. Background

We review some existing types of social software applications, discuss their general characteristics and distinguish social cataloging from others. We also provide example links to some prominent applications. Exhaustive list would be impossible to make because new kinds of social applications emerge constantly, not to mention the numerous clones of existing ones.

2.1 Brief Review of Social Software

It is complicated to define exactly what constitutes a social software application, especially because terms that are used to describe the applications itself are loose at best, starting from the elusive “web 2.0” that is used to describe the modern web environment in general. We do not attempt to give an exhaustive definition, but note that most social applications utilize a form of public, distributed metadata that is created and maintained by user community. The nature of this metadata, combined with the actual content disseminated in the application determines how the application is used in a profound way. Blogs and wikis can be regarded as special cases of social applications because the content and metadata are deeply related (e.g. wiki category tags [28] are edited in the same structure as the actual content), linked (e.g. blog posts linking to each other), and in some cases, almost indistinguishable (blog posts vs. comment discussion).

Examples of content in social software are text (wikis, forums, blogs, microblogs, messaging), pictures (e.g. Flickr, Picasa⁴), video clips (e.g. YouTube⁵, blip.tv⁶), content aggregation (Digg, reddit⁷) bookmarks (del.icio.us, Clipmarks⁸), bibliographical information (Bibsonomy⁹, CiteULike¹⁰), representations of abstract (43 Things¹¹) or concrete items (e-commerce sites with user-generated reviews, social cataloging), or personalized profiles or web pages provided by the social application environment (FaceBook¹², Ning¹³) containing any or all of above. However, it is essentially the public, distributed metadata that enables the social aspect regardless of the particular application or content. This may include e.g. time-, location-, or other structured metadata, tags or other semi/unstructured [3] metadata, discussions or ratings related to the actual content, or social relations between users. Tags are of specific interest, because they allow users to annotate content items with multiple keywords in ad-hoc way, contrasted to traditional classification (cf. e-mail client or FAT-like filesystem) where items are categorized or “filed” to a single category.

⁴ <http://picasa.google.com/>

⁵ <http://www.youtube.com/>

⁶ <http://www.blip.tv/>

⁷ <http://reddit.com/>

⁸ <http://clipmarks.com/>

⁹ <http://www.bibsonomy.org/>

¹⁰ <http://www.citeulike.org/>

¹¹ <http://www.43things.com/>

¹² <http://www.facebook.com/>

¹³ <http://www.ning.com/>

When tags from different users are aggregated together, a folksonomy [15] emerges, reflecting collective view of the most relevant concepts related to the resource. Tags can be used for various purposes, such as topic information, content type, owner information, category refinement, characteristics, self references or task organization. Even though tags are based on personal preferences, it turns out that the aggregated proportions of each tag stabilize over time [8]. Depending on the application interfaces, tags from different applications can be aggregated together (e.g. Technorati¹⁴), providing collaborative content filtering or ad-hoc semantic-like search functionality without an explicit, formal ontology [13]. Nonetheless, there has also been critique [18] whether this aggregation works in practice large-scale because of the unforeseeable diversity [14] in tag usage.

2.2 Social Cataloging and Rating Services

Social music cataloging has emerged along with other content-varying cataloging and rating services, such as LibraryThing¹⁵ (books) and Flixter¹⁶ (movies). People can – and will – catalog or rate virtually anything. Spalding [24] has observed cataloging sites for games, comics, programs, wine, beer, and recipes among the 40 “direct” competitors to LibraryThing. Also intangibles, such as goals (43 Things), words (Wordie¹⁷) or quotes (Wikiquote¹⁸) can be cataloged. Perhaps the earliest form of online rating is related to personal characteristics – for example, RateMyFace.com¹⁹, Rate My Professors²⁰ and multiple variations of the like have been online since 1999, but they have originally focused merely on independent aggregated ratings, rather than the social aspects of user community.

In addition to independent, focused rating services, cataloging functionality can be embedded as part of other social software applications. For example, Facebook contains many applications (e.g. Flixter, Rate Your Music²¹, Shelfari²²) that provide integration to third-party rating sites. It is likely that this embedding boosts the popularity of rating sites even further, because users don't have to login to multiple applications. Even Amazon and other e-commerce websites have a kind of social cataloging functionality, because users can submit product reviews and lists. Compared to “open” cataloging services, this affects the way ratings and especially tags are assigned to products [23]. It seems that people are more eager to assign metadata to content items they regard as their “own” (i.e. filed in a collaborative cataloging service with personalized content) compared to augmenting readily available commercial product descriptions.

¹⁴ <http://technorati.com/>

¹⁵ <http://www.librarything.com/>

¹⁶ <http://www.flixster.com/>

¹⁷ <http://wordie.org/>

¹⁸ <http://wikiquote.org/>

¹⁹ <http://ratemyface.com/>

²⁰ <http://www.ratemyprofessors.com/>

²¹ <http://rateyourmusic.com/>

²² <http://www.shelfari.com/>

Social cataloging allows keeping track of personal collections – essentially a public bibliography [9] (which has traditionally been done offline if at all, perhaps with a spreadsheet), connecting to people with similar taste, and pointing out new, previously unheard material or missed “classics”. With LibraryThing, social cataloging has even been used by public organizations as a way to publish their collections [2]. To our knowledge, this hasn’t yet been done with music cataloging services, but advertisements related to the cataloged items (e.g. links to record stores or auctions with more albums from the band) have found their place along with user-generated metadata. Recommendations may emerge as a useful side-functionality (e.g. lists like “best albums of 2007”, following popular tags, or aggregated ratings [11]), but compared to services focused specifically on recommendations such as MovieLens²³ or commercially used recommendation systems such as Amazon’s “Who Bought This Item Also Bought...”, social cataloging sites focus more on general ratings, reviews, and other metadata, as well as discussions and other social interaction around the content items.

3. A Comparison of Social Music Cataloging Services

There are numerous music-related social cataloging services available, from which we have selected two – namely Rate Your Music (RYM) and Prog Archives²⁴ (PA) – for detailed evaluation. Both of them offer their users an opportunity to rate and review albums. However, there are differences between the two, like the focus on the usage of the site and its approach to reviewing. Other related but excluded services include Discogs²⁵ - a community-built music database focusing on electronic music with much similar functionality with RYM. Amazon contains also some social functionality, such as user-submitted reviews, tags and favorite lists. Metacritic²⁶ aggregates existing reviews for music, movies and other media, but is focused on professional critiques instead of user community. Many open music databases most probably owe some of their content to AllMusic²⁷, a commercial music database with little community-driven functionality.

The sites are compared using the following eight attributes: *library2.0*'s [12] general categories of look and feel, features, ease of use, community and social networking, Senko's [20] rating systems classification (stars, popularity, dual rating, value rating), Porter's [17] recommendation functionality evaluation (newness, time-sensitivity, popularity, personal and social network relevance, authority-based, collaborative), Marlow's [14] framework for tagging support (tagging rights and support, aggregation, type of object, source of material, resource and social connectivity), and finally evaluate how the service encourages users to participate to overcome Nielsen's [16] and others' [21] considerations of participation inequality. We will also give our general impression about the sites. We have evaluated the sites independently, but since they share some commonalities the Prog Archives evaluation focuses more on the differences of the two rather than stating the already discussed matters.

²³ <http://movielens.umn.edu/>

²⁴ <http://www.progarchives.com/>

²⁵ <http://www.discogs.com/>

²⁶ <http://www.metacritic.com/>

²⁷ <http://www.allmusic.com/>

3.1 Rate Your Music

Rate Your Music is a stylistically non-defined website which offers its users the ability to catalog, tag and review musical albums. Its focus is mainly on the cataloging and tagging, but there are some people who write more analytical reviews of the albums they own, so you could say that RYM is aimed more at the average user than a professional reviewer. Rate Your Music was founded in 2000 by Hossein Sharifi.

Look and feel

It is easy to see the driving principles behind Rate Your Music as soon as you arrive at the site. Design of the site is simple and inviting and the social context is present from the beginning in the form of ratings and reviews.

As you dwell deeper into the site you'll find a community which actively presents new content and renews/rejuvenates the old. There is also a quite active message board where people discuss topics related to the site, music, and other matters.

Features

As a casual observer you can e.g. see other people's lists of albums, ratings, other album related statistics and read the discussion board. The option for interaction in a social context is limited to a many to one outbound direction. This is of course natural since the community doesn't allow anonymous participation in the site.

If you register you are presented with a page where you can submit your personal information and the artists/albums you want to deal with. After registration you are free to participate in the community and have privileges to alter the content of the site.

RYM has a built-in system for searching albums and artists linked to them and vice versa. It is utilized in the most part of the site's social interaction (excluding the discussion board); you can for example search users with the same albums and/or artists to your "compatibility list".

Ease of use

The site is easy to use and adopt, its structure is logical and the main functions are displayed the way that they can easily be recognized by a casual user. Usability adds to the social relevance of the site because it enables broader scope of users to participate in the community, thus diversifying the underlying social structure.

Community and social networking

From the two sites evaluated here RYM is the one leaning more to social interaction and networking than Prog Archives. This is evident in the sheer number of functions geared towards the interaction of users and the whole concept (features, appearance, usability etc.) of the site. Non-professional (meant as average user not a music-o-phile) nature of the site appears more in the content of the site than the site itself.

Another interesting aspect of the site is its connection to the basic concept of lifelogging introduced in the metaverse roadmap. The basic concept of lifelogging is to capture, storage and distribute everyday experiences and information for objects and people [22]. This can easily be adapted to RYM where people share their thoughts and feelings towards music in a specified moment of time.

Rating system

RYM employs a simple 5-star rating system (allowing half-stars) that can be applied to each issue (vinyl, cd, remastered releases etc) of a given album. Ratings can be accompanied with reviews, which vary both in length and quality from single-sentence reviews to more extensive analyses. The ratings are aggregated with an undisclosed algorithm that accounts weights to reviewers. According to the documentation of the site, the weights are affected by the number of albums rated, the diversity of the ratings, the number of reviews written, and the activity on the site. The system also gives lower weight to bursts of ratings for a particular release to counteract fraud. The ratings are further aggregated to yearly top-lists, representing a popularity rating system in Senko's [20] classification.

Users are "rated" in the discussion forums using a simple value-rating system. User's profile information contains the number of message board posts, as well as the month user has joined the community. Users are further divided to administrators, moderators and basic users with varying authority to modify contents. More extensive modifications (e.g. adding a completely new band or release, or adjusting band's genre information) may require voting or approval by a moderator.

Recommendation functionality

RYM provides a user-compatibility list based on ratings with other users, how other people have rated the common items, and the relative sizes of the respective collections. The system can also give automatic recommendations and members can manually recommend albums to each other. Yearly top-lists, as well as continuing lists of most recent ratings provide implicit recommendations. In Porter's classification [17] these represent recommendations by collaboration, social network relevance, popularity, and newness.

Tagging support

RYM supports a limited form of tagging where anyone can see specific user's tags, but tags from different users are not combined. In this way, tags in RYM are strictly for personal information organization and – based on ad-hoc browsing – not generally used very often. In Marlow's [14] terminology, RYM provides blind self-tagging, no aggregation, objects being tagged are musical releases, source of material is provided by the participants in RYM. Resource connectivity is provided in lists (both automatically generated and user-specified). Social connectivity is achieved with public friend relations and user lists. RYM allows also users to assign multiple genres to bands and subsequently vote about "correct" genres, resulting essentially to alternate tagging system. Because of the openness of the genre assignment, popular bands may receive a multitude of genres. For example, the search page of Metallica contains the following genres which may of may not represent the majority of their work:

thrash metal, speed metal, heavy metal, metal, hard rock, rock, heavy rock, arena rock, thrash, progressive thrash, proto thrash, melodic thrash, nu-metal, melodic rock, guitar metal, alternative metal, metal guitar, southern metal, bay area thrash, heavy/thrash metal, southern rock, rock & roll, grind metal, grunge, classic rock, symphonic metal, death metal, hair metal, punk, black metal, hardcore punk, novelty, bachata, emo metal

Participation

RYM provides detailed statistics of user's actions, including dates of all albums rated, lists created, number of approved artists, releases, profiles, labels, votes, uploaded covers, release corrections, and message board posts. Every album entry contains correction and image upload history, as well as names of the members that have contributed to the site. Compared to Wikipedia's anonymity, RYM keeps all contributions fairly public. This publicity has both positive and negative aspects, but probably encourages responsibility and accuracy with posted metadata, as well as "invites" to contribute more, since even small actions are numerically accounted.

3.2 Prog Archives

Prog Archives' main focus is on reviewing and defining categories for the albums specializing in progressive music (progressive rock, metal, electronic etc). Another goal of the site is to function as a source of information, and towards these ends its social and technical context is different from that of the Rate Your Music. Prog Archives was founded in 2002 by Ronald Couture and Maxime Roy.

Look and feel

Prog Archives site is aimed to people who are interested in progressive music so it is more committed to certain subculture than RYM. This shows in the appearance of the site and how it conveys its purpose to the user. There are ratings & reviews in the homepage just like in RYM but the overall feel of the site is a little more confusing. If you are interested in the subject this isn't a problem but for a random observer it can be a reason to leave.

After getting used to the structure of the site you will find out the vast resources of data concerning progressive music and its subgenres, which is the purpose of the site.

Features

The features of the site are somewhat limited compared to RYM. You can rate & review the album of your choice by registering to the site's discussion board when you gain access to the editorial functions of the site. Individual profiles show user's reviews and ratings distributed by sub-genre, but for instance one can not mark album ownership or create lists as in Amazon or RYM.

The focus of Prog Archives dictates its features as they are aimed more at the direction of music and related tasks as a main focal point of the site. There is a nice feature with which you can listen to streaming song samples from the website and as a side bonus, many of the reviewed artists have a song or two that you can listen while reading their record reviews.

Ease of use

Compared to Rate My Music Prog Archives has a slightly less intuitive structure. This is not to say that it is difficult to use the site but its appearance is more cluttered than RYM's. Furthermore there is some "musical" terminology which can be confusing if one is not familiar with it.

In defense of PA must be said that it is a website for the fans of progressive music and it doesn't even try to be anything more. This commitment induces a reaction in a user where the person can relate to the site better if he/she belongs to the same subculture. This is a powerful motivator and behavioral model and quite common in the world of various music related websites.

Community and social networking

Prog Archives takes completely different approach to social networking compared to Rate Your Music where it is based more on features than the content of the site. PA encourages user to take active role in the process of content creation and participate in the discussions. These attributes are of course part of the Rate Your Music as well but more in the role of secondary activity.

Since PA only has member profiles that function as identities in discussion forum it has to create the sense of community in some other way than a list of people who like the same bands. It does this by engaging users e.g. through forum, providing information and introducing new bands via reviews and the site's own "jukebox". Cataloging of the albums and artists is one of the prominent aspects in the PA community. It also differs from the RYM approach as individual PA user does not primarily catalog his albums but defines the niche for an individual artist/album.

Another aspect to set PA apart from RYM is their relation to submitted reviews. In RYM reviews are secondary compared to rating and tagging of objects wherein the situation in PA is just the opposite. In PA reviews play a major role in the contents infrastructure of the site. They are written by "professionals" (site indicated) and normal users, this adds up to very diverse and multi-faceted examination of the material at hand.

Rating system

Like RYM, PA features both star- and popularity based rating system, as well as reviews. 5-star ratings without half-stars are used for individual albums. The results are represented as a chart with illustrative textual labels underlining the focus of the site: "Essential: a masterpiece of progressive music", "Excellent addition to any prog music collection", "Good, but non-essential", "Collectors/fans only", "Poor. Only for completionists". It's interesting to note that despite the labels, some reviews may explicitly point out the overall quality of the music in a positive way, such as "this is definitely not progressive, but it's still good music".

The length of the reviews (and generally the amount of reviews in relation to merely star-ratings) is generally longer compared to RYM. The ratings are aggregated to yearly lists, but also to sub-genre (e.g. *Symphonic Prog*, *Progressive Metal*) specific lists. According to documentation of the site, the classic calculation of the average but more weight is affected to the rating of progarchives.com collaborators and to rating with reviews is used.

Prog Archives employs an elaborate value-based rating system for users, featuring the number of posts and join date like in RYM, but also a 5-level categorization that includes newbie, groupie, senior member, collaborator/reviewer, and senior collaborator/admin. Member's status is denoted in reviews, but any registered member can post a review.

Recommendation functionality

PA provides limited recommendation functionality compared to RYM, since there are no compatibility lists or explicit recommendations. However, the genre-based top-lists provide a kind of personal preference (assuming the user is interested in specific sub-genre of progressive music) recommendations that is not present in RYM. Other popularity and newness recommendation aspects are comparable to those in RYM.

Tagging support

Prog Archives does not support user-specified tagging. All bands are assigned to the classification of progressive music sub-genres (one band can belong only to one category). The assignments to sub-genres (and the genres themselves) are not entirely stable. For example, in the end of 2007 Prog Archives underwent a revision, where *Art Rock* sub-genre was split to new genres and a few new, such as *Tech/Extreme Prog Metal* appeared. As for individual bands, Porcupine Tree was moved from *Psychedelic/Space Rock* to *Heavy Prog* and King Crimson was moved from *Art Rock* to *Eclectic Prog* among others. Genres are most likely determined by top-level participators (genre teams) – the exact mechanism is not clearly denoted in documentation, but the suggestions to both genre categorizations²⁸ and individual artist genres²⁹ are frequently discussed in the forums. Resource connectivity is limited to automatically generated lists. Social connectivity is limited to “buddy lists” in discussion forums, but these are not shown in public profiles.

Participation

Prog Archives encourages users to participate with 5-level user categorization – the higher level, more esteemed member of the community. However, to advance in levels may require considerable activity and even dedication to the community. Because of the rather specific musical orientation of the site, the initial threshold to register to the site is probably higher compared to “all-encompassing” RYM.

4. Discussion

It is indisputable that social cataloging services benefit both producers and consumers of music. Artists get direct feedback about the popularity of their work. Band profiles and

²⁸ http://www.progarchives.com/forum/forum_posts.asp?TID=43011

²⁹ http://www.progarchives.com/forum/forum_posts.asp?TID=42683

albums can be linked to advertisements in a non-intrusive way, and marketing can be directed to people that are likely to be interested (and probably buy a lot of records anyway – why else catalog them in the first place). One might argue that even without explicit advertising at all social cataloging systems boost sales of the albums and artists that are rated high in the community.

However, the reliability and overall quality of the ratings and recommendations can be questioned. Gillmor [7] stated that blogs have become a form way of participatory media, complementary to traditional print media and broadcasting. Online music reviews in social applications have clearly some similarity (but not necessarily authority) to professional reviews in print media. Surowiecki [25] stated that the aggregation of information in large and independent groups yields often to more accurate results compared to decisions done by few or one “expert” decision maker. This can also be seen as a justification of folksonomies – most popular tags related to a particular resource represent the collective (and in one sense, “correct”) opinion of the user community. In general, this seems to work fairly well with ratings, as can be seen from the following examples:

The albums regarded as best or worst (see Figure 1) seem to generate a nonsymmetrical, almost power law –like distribution, reflecting a strong consensus in the community.

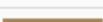
0.5		35
1.0		10
1.5		12
2.0		33
2.5		63
3.0		154
3.5		359
4.0		1077
4.5		1592
5.0		3498

Figure 1. Ratings distribution of "Abbey Road" by The Beatles (Rate Your Music, 20.3.2008)

Most releases tend to generate normal-like distributions. In Figure 2, the average is near to center as well, but obviously this doesn't have to be the case. Both in RYM and PA many albums tend to generate a normal-like distribution with most votes around 3.5-4. This small skewness towards the upper ratings could be explained by the fact that people tend to own (and subsequently rate) albums that they like.

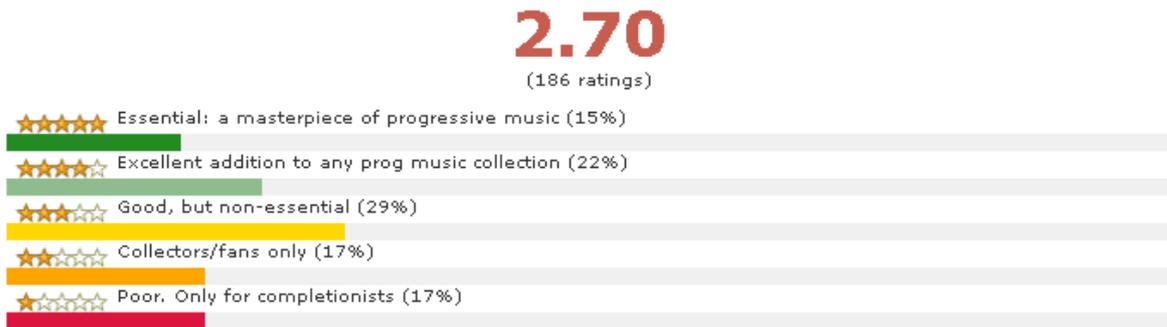


Figure 2. Ratings distribution of "90125" by Yes (Prog Archives)

When band is relatively unknown but has a few active fans, it can get high ranks with very few ratings. This can be detected with traditional statistic tests – whenever a release gets only few rates, the validity of them should be questioned. An interesting case (see Figure 3) are albums that generate an irregular distribution – no peak is clearly noticeable (although 3.5 stars seems to be one vote ahead of others in the example below, but that can be hardly called as “peak” since both 2.5 and 5 stars have received 6 votes). This might either indicate that the crowd has not “decided” yet – as the album get more votes, the results might later aggregate either to a normal or to a power law –like [1] distribution. Another possibility is that the release divides opinions – perhaps the record has received a lot of airplay due to marketing and receives both positive and negative attention. In either case, these kinds of distributions seem to be the surprisingly commonplace in RYM, especially with pop and eurodance albums.

0.5	████████	3
1.0	████████	4
1.5	████████	4
2.5	████████████████	6
3.5	██████████████████	7
4.0	████████████████	5
4.5	████	1
5.0	████████████████	6

Figure 3. Ratings distribution of "No time to chill" by Scooter (Rate Your Music, 20.3.2008). Note the non-existence of ratings with 2 or 3 stars.

As Yang [29] and others have noted, online rating systems are subject to manipulation, fraud, and in general, unfair ratings. We encountered an illustrative example of this in a “rating battle” between *In Rainbows* album by Radiohead and *Talvikuningas* by CMX. Radiohead is highly popular English alternative rock band that has a considerable fan community in RYM – to the extent that the band is explicitly noted in RYM FAQ about tips for posting new topics in the forums:

Before you create a new topic, ask yourself if it will really spark a thoughtful discussion. Your topic should be broad enough to let people voice various opinions

yet specific enough so it doesn't diminish into nonsense. [...] Just because your topic is unique doesn't mean it's new. For instance, "The Beatles vs. Motörhead" may have never been done before, but the "this vs. that" formula is overused and gets old after a while. Here are some of the most common topics that are irritatingly frequent:

- *Band X (Beatles, Nirvana, Joy Division) are overrated*
- *Band X vs. Band Y*
- *Anything related to Radiohead, Nirvana, or The Beatles*
- [...]

On the other hand, CMX is an eminent Finnish rock band that also has a notable fan base in Finland, but is virtually unknown elsewhere because of the Finnish lyrics. CMX released their 12th album, *Talvikuningas* 5.9.2007 in a special limited box set that contained the album along with high-quality lyrics and art book with only 8000 numbered copies (a standard issue was released separately a few months later). Because the release was specially directed for the fans, it got many top ratings in RYM (even though relatively few overall ratings) and soared as #2 album for 2007 by the end of September 2007.

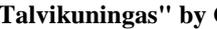
0.5		19
1.0		6
1.5		6
2.0		8
2.5		4
3.0		13
3.5		17
4.0		31
4.5		30
5.0		73

Figure 4. Ratings distribution of "Talvikuningas" by CMX (Rate Your Music, 20.3.2008)

Radiohead released their 7th album, *In Rainbows* 10.10.2007. It got immediately highly positive reviews in RYM. The album circulation was boosted by Radiohead's innovative marketing strategy of releasing album on-line³⁰ in digital, DRM-free format, allowing user to choose the price to be paid for the download – "It's really up to you" was noted in the checkout page. The next day, *In Rainbows* soared as #3 album for 2007 with hundreds of positive reviews, but *Talvikuningas* was still holding the second place with only tens of mostly 5-star reviews.

³⁰ <http://www.inrainbows.com/>

0.5	█	54
1.0	█	21
1.5	█	32
2.0	█	78
2.5	█	103
3.0	█	277
3.5	█	508
4.0	█	1051
4.5	█	1174
5.0	█	782

Figure 5. Ratings distribution of "In Rainbows" by Radiohead (Rate Your Music, 20.3.2008)

This created some confusion about the rating aggregation mechanism because the number of ratings was so small, and apparently caused discontent among the most dedicated Radiohead fans, who felt that In Rainbows should be at the top. As a result, over 10 0.5-star ratings with unfavorable reviews appeared to Talvikuningas. To make things worse, several 0.5 star reviews appeared also to In Rainbows – apparently from resentful CMX fans. Additional positive ratings were also added to Talvikuningas, e.g. one 5-star review noting “I dedicate this rating to all Radiohead's fans”. Naturally, this was felt inappropriate by vast majority of both CMX and Radiohead fans and was widely discussed in the RYM message board³¹. For example, nick *wehopethatyouchoke* wrote on 11.10.2007:

i noticed a nasty little thing going on, that i thought was worth bringing to public attention: the album Talvikuningas by obscure finnish band CMX is being kind of hacked by some dumb RH fans with the obvious purpose to help "in rainbows" reach the second spot for 2007.

they already did their job, as it was #2 this morning and now is #134...

specifically i'm referring to the users ____, ____, and ____, who rated "in rainbows" 5 stars, and, by sheer coincidence, gave the poor finnish 0.5 stars right away... as you can guess by my nickname, i'm a huge RH fan myself, but i just can't stand these kind of things.

In the subsequent discussion³² *rachel* noted that some of the negative reviews to Talvikuningas were due to multiple accounts of the same person. Further, somebody who had rated In Rainbows with 5 stars had given 0.5 stars to all other albums in the 2007 top 10 –list. Multiple accounts are explicitly prohibited in RYM user guide, but otherwise it is an extremely delicate or even impossible issue to determine if the review is “correct”, since it is a matter of taste. RYM uses weighting to handle the problem, as site founder Sharifi points out during the discussion:

³¹ http://rateyourmusic.com/board_message/message_id_is_1159146_and_board_id_is_1&show=20&start=500

³² http://rateyourmusic.com/board_message/message_id_is_1175501

There are over 10 million ratings. If we actually allowed people to submit potentially false ratings, then we would spend so much time trying to police the ratings that we would never have time to enjoy the site or do anything else music-related here.

*But more importantly, there would be tons of false accusations against people who actually have listened to the music, and it would sour their experience on RYM (this has happened **many times** in the past, and it's the reason why we don't allow "call-out" threads. Often times, the accusers are completely wrong).*

We have a system in place to filter out crap ratings, and it works amazingly well. I wish we could simply display the weight next to each rating to prove it, but of course that would allow people to more easily abuse the system.

During the following weeks, the debate subsided and In Rainbows is currently holding the second rank for best albums for 2007, followed by Talvikuningas as the third. It seems that the relatively high number of 0.5 star ratings did not severely affect the rank of Talvikuningas (especially considering the total number of ratings, which is in completely different scale compared to Radiohead). Nevertheless, the issue raises questions about the validity and even comparability of aggregated ratings.

As a side note, In Rainbows was adopted with much less drama in Prog Archives (CMX is not listed in PA). Radiohead is categorized in a sub-genre *Crossover Prog*, but In Rainbows is not listed as a top album of the genre at all (earlier Radiohead album "Kid A" was listed as the 15th best in its own category, reflecting the orientation to "progressive" music). As for the yearly charts, In Rainbows ranked 34th best progressive album for 2007. It can be concluded that active fan base, as well as the focus of the site has a huge effect on the popularity of a particular release - as Radiohead fans in RYM have shown.

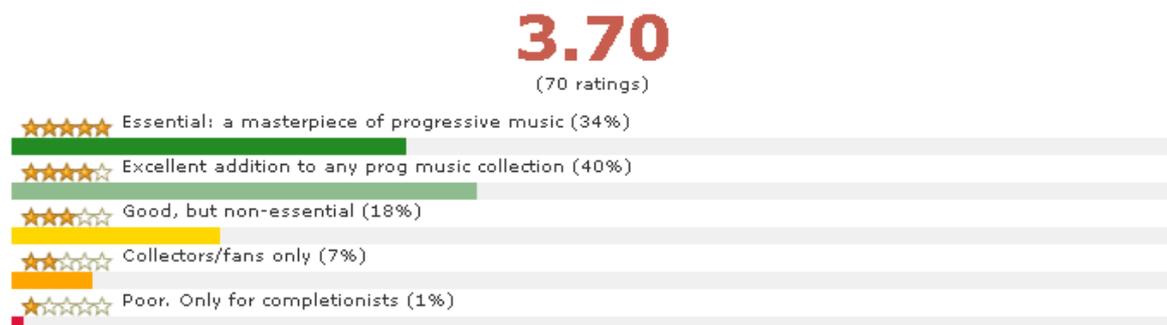


Figure 6. Ratings distribution of "In Rainbows" by Radiohead (Prog Archives, 20.3.2008)

Although the In Rainbows vs Talvikuningas -case can hardly be regarded a serious problem in itself, it indicates a fundamental problem with collaborative rating systems in general. Bustos [4] notes that review systems can be used for false – both positive and negative – advertising by companies in a similar way as fake blogs [10] are used. Various methods have been suggested to counteract manipulation of the ratings, such as the signal-based approach

by Yang [29]. Bustos [4] suggests that a Amazon- or Digg-like dual rating system (e.g. “Was this review useful to you?”) should be applied to social review systems more widely.

The rating battle can be seen as an example of a situation where “wisdom” of the crowds – a fundamental prerequisite for successful folksonomies – is questionable. Active fan base and promotion of a single album at the expense of another blurred independency and diversity of opinion of the reviewers, which are essential for successful group-based decision making, along with decentralization and aggregation mechanisms [25], which were present. Reliable ratings may also be distorted by the well-known mechanism of conformity [26]: sometimes we go as far as converting to the majority opinion, but most of the time we just choose to comply with it even if we don't believe it. For example, if 3498 RYM users have concluded that Abbey Road deserves five stars, it is easy to lean to the more positive ratings. Salganik [19] observed that this happens also with previously unknown songs: social influence contributes to inequality, and – perhaps surprisingly – to unpredictability to success of particular songs. The result has a resemblance to the popularity distribution of blogs [21] and in general, connectivity distribution in any scale-free networks [1] – the power law.

From a business perspective, Clemons [5] argues that the amount of bad reviews does not really matter, since the people giving bad reviews do not belong to target segment of the product anyway. It's the amount of good and excellent reviews that matter – in this case both albums represented in the case are undoubtedly successful. As a minor effect from marketing point of view, we note that numerous negative reviews do pull an album down in the top charts, which could potentially result to less attention from casual users. In the case of Talvikuningas, drop from #2 to #3 was of no consequence.

We conclude that social music rating systems provide a useful alternative to traditional, professional reviews. The dependencies in social networks may polarize the popularity of individual artists, but on the other hand, provide more diversity in “the long tail” side of the distribution [21]. Unfair reviews and ratings is an issue that needs to be tackled. The professional critiques in traditional media are to have lesser impact due to collective ratings, but they are unlikely to be replaced completely. Rather, professional reviewers should be recognized in community-driven systems as well (perhaps as authorities), as is partially done in Prog Archives with user level categorization. While user-generated content can be of surprisingly good quality (cf. Wikipedia [27]), professional reviews are still needed as well to provide general credibility and quality to rating systems.

5. Conclusion

Social cataloging services benefit both producers and consumers of music. Artists get direct feedback about the popularity of their work. Recommendation systems can be linked to advertisements in a non-intrusive way to a potentially receptive audience. New bands can promote themselves in an open environment. Users can keep track of their collections with a stronger sense of community compared to “isolated” catalogs and socialize with people with similar interests. Rating systems can also serve as a powerful recommendation mechanism by different ranked lists or genre classifications.

The general quality of ratings seems to depend on the community. For example, one could argue that Prog Archives contains reviews of better quality compared to those in Rate Your Music, but the reviews in PA are probably more useful only if the user is interested in progressive music in the first place. As the size and popularity of the community grows, so does the potential for abuse of the ratings. Users also shape the communication culture and values of the community. In general, the “correctness” of individual reviews is extremely hard to determine, because it is fundamentally a matter of taste. From the perspective of group-based decision making, the validity of the decision can be questioned if the independency of the reviewers is biased. Besides, one should remember that popularity does not necessarily equal quality, which is a strong argument in favor of traditional, “authoritative” reviews.

Despite the clear advantages of social cataloging, rating systems are prone to manipulation and unfair reviews. It is technically rather easy to e.g. create multiple identities to boost ratings for particular albums in a similar way to fake blogs or even spam. If the identities are not real (i.e. bots), it is theoretically straightforward to filter them out, but a more ambiguous case can be made from overactive fan base determined to vote their favorite album to the top. Several methods have been suggested to alleviate the situation, such as assigning weights to ratings, dual ratings to individual reviews, and detection of fake reviews based on users’ rating patterns, but clearly more research is needed to build trust on collaborative rating systems and to make them a credible alternative for reviews in traditional media.

References

- [1] A.-L. Barabási. *Linked: How Everything Is Connected to Everything Else and What It Means*. Plume, 2003.
- [2] A. Bell. *Advantages of Social Cataloging*. School 2.0, 12.11.2007
<http://school20.ning.com/profiles/blog/show?id=595650%3ABlogPost%3A8731>
- [3] P. Buneman: *Semistructured data*. Sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems.
<http://doi.acm.org/10.1145/263661.263675>
- [4] L. Bustos. *Fake Customer Reviews, Bad Product Reviews: What to Do?* Get Elastic, 16.8.2007. <http://www.getelastic.com/fake-customer-reviews-bad-product-reviews/>
- [5] E.K. Clemons, G. Gao, and L.M. Hitt. *When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry*. 39th Annual Hawaii International Conference On System Sciences, 2006. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1579524
- [6] A. J. Gilliland-Swetland: *Setting the stage*. In M. Baca (ed.): *Introduction to Metadata: Pathways to Digital Information*. Getty Research Institute, 2000.
<http://www.getty.edu/research/institute/standards/intrometadata/>

- [7] D. Gillmor. *We the Media*. 2004. <http://www.authorama.com/book/we-the-media.html>
- [8] S.A. Golder and B.A. Huberman. *Usage patterns of collaborative tagging systems*. Journal of Information Science, vol 32(2). <http://jis.sagepub.com/cgi/reprint/32/2/198>
- [9] D.G. Hendry and A. Carlyle. *Hotlist or Bibliography? A Case of Genre on the Web*. 39th Annual Hawaii International Conference on System Sciences, 2006. <http://doi.ieeecomputersociety.org/10.1109/HICSS.2006.203>
- [10] J.L. Horton. *Marketing And Blogs: What Works*. Technical report, Online Public Relations, 2003. <http://online-pr.com/Holding/MarketingandBlogs-WhatWorks.pdf>
- [11] K. Lerman. *Dynamics of collaborative document rating systems*. 9th WebKDD and 1st SNA-KDD workshop on Web mining and social network analysis. 2007. <http://doi.acm.org/10.1145/1348549.1348555>
- [12] "library2.0". *Librarything, Shelfari, and Gurulib: Social Cataloging Sites Compared*. Librarytwopointzero, 21.10.2006. <http://librarytwopointzero.blogspot.com/2006/10/librarything-shelfari-and-gurulib.html>
- [13] D.L. McGuinness. *Ontologies Come of Age*. In D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster, (eds). *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 2003
- [14] C. Marlow, M. Naaman, D. Boyd, and M. Davis. *HT06, tagging paper, taxonomy, Flickr, academic article, to read*. Seventeenth conference on Hypertext and hypermedia, 2006. <http://doi.acm.org/10.1145/1149941.1149949>
- [15] A. Mathes. *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*, 2004. <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>
- [16] J. Nielsen. *Participation Inequality: Encouraging More Users to Contribute*. Alertbox, 9.10.2006. http://www.useit.com/alertbox/participation_inequality.html
- [17] J. Porter. *Which Movie to Watch? An Overview of Recommendation Systems*. Bokardo, 14.9.2005. <http://bokardo.com/archives/quick-overview-of-recommendation-systems/>
- [18] J. Porter. *The Del.icio.us Lesson*. Bokardo, 2.5.2006. <http://bokardo.com/archives/the-delicious-lesson/>
- [19] M.J. Salganik, P.S. Dodds, and D.J. Watts. *Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market*. Science, vol 311(5762). <http://www.sciencemag.org/cgi/content/abstract/311/5762/854>

- [20] P. Senko: *Rating Systems In Social Networks*. OntoInfo, 1.10.2006
<http://www.ontoinfo.com/2006/10/01/rating-systems-in-social-networks/>
- [21] C. Shirky. *Power Laws, Weblogs, and Inequality*. Clay Shirky's Writings About the Internet, 8.2.2003. http://www.shirky.com/writings/powerlaw_weblog.html
- [22] J. Smart, J. Cascio, and J. Paffendorf. *Metaverse Roadmap Overview – Lifelogging*. 2007. <http://www.metaverseroadmap.org/overview/02.html#LL>
- [23] T. Spalding. *When tags work and when they don't: Amazon and LibraryThing*. Thingology, 20.2.2007. <http://www.librarything.com/thingology/2007/02/when-tags-works-and-when-they-dont.php>
- [24] T. Spalding. *SocialCatalogers: For people who make social cataloging applications*. Thingology, 27.2.2007. <http://www.librarything.com/thingology/2007/02/socialcatalogers-for-people-who-make.php>
- [25] J. Surowiecki. *The Wisdom of Crowds*. Little, Brown, 2004.
- [26] C. Thurlow, L. Lengel, and A. Tomic. *Computer Mediated Communication*. Sage Publications, 2004.
- [27] D. Weinberger. *Is the Web as weak as its weakest link?* Everything Is Miscellaneous, 28.10.2007. <http://www.everythingismiscellaneous.com/2007/10/28/is-the-web-as-weak-as-its-weakest-link/>
- [28] J. Voss. *Collaborative thesaurus tagging the Wikipedia way*. Technical report, arXiv:cs/0604036v2. <http://arxiv.org/abs/cs/0604036>
- [29] Y. Yang, Y.L. Sun, J. Ren, and Q. Yang. *Building Trust in Online Rating Systems Through Signal Modeling*. 27th International Conference on Distributed Computing Systems Workshops, 2007. <http://dx.doi.org/10.1109/ICDCSW.2007.27>

(web-references linked 20.3.2008)