

FREESOUND 2: AN IMPROVED PLATFORM FOR SHARING AUDIO CLIPS

Vincent Akkermans¹, Frederic Font¹, Jordi Funollet¹, Bram de Jong²,
Gerard Roma¹, Stelios Togias¹, Xavier Serra¹

¹Music Technology Group, Universitat Pompeu Fabra {name.surname@upf.edu}

²SampleSumo {bram.dejong@samplesumo.com}

1. INTRODUCTION

Freesound.org is an online collaborative sound database where people from different disciplines share recorded sound clips under Creative Commons licenses. It was started in 2005 and it is being further developed by the Music Technology Group (MTG) of the Universitat Pompeu Fabra.

Freesound's initial goal was to give support to sound researchers, who often have trouble finding large royalty-free sound databases to test their algorithms, and to sound artists, who use pre-recorded sounds in their pieces. After six years since its inception, Freesound has become one of the most popular sites for sharing sound snippets. It serves around 30,000 unique visits per day and has more than two million registered users accessing the more than 120,000 uploaded sounds. Perhaps more importantly, there is a highly engaged community of users continuously contributing to the site, not only uploading sounds but also commenting, rating and discussing in the forums about relevant topics for the community.

This remarkable growth surpassed our expectations and forced us to redesign and rebuild Freesound from scratch. This demo shows the new version of Freesound (Figure 1) which brings a number of improvements and new functionalities. Some of these improvements facilitate the access to the content and provide a new framework for developers and researchers that are interested in working with Freesound. What follows is a summarized description of the system and the types of content that can be found.

2. FREESOUND 2.0

The new version has been built with high load and scalability in mind. Figure 2 shows the block diagram of the new architecture. Retrieval of sounds can be performed us-

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.

© 2011 International Society for Music Information Retrieval.

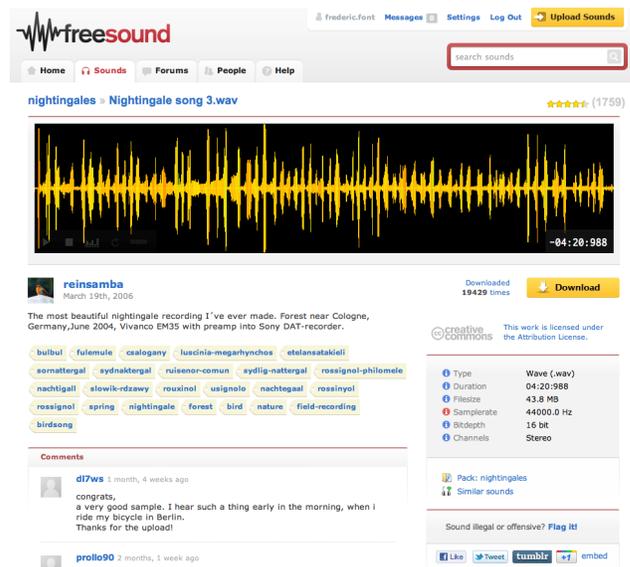


Figure 1. Sound page of Freesound 2.

ing text queries, content-based similarity search (Query by Example) or by browsing tags or geotags (Figure 3). The front-end is a Django [1] application which includes basic social interaction features (forum, sound comments, sound ratings, private messaging...). Text indexing is supported by an Apache Solr [2] server including text descriptions and tags, which allows for sophisticated text queries using the Solr query syntax. A distributed architecture is used for processing incoming sounds, producing compressed previews and waveform/spectrogram images, as well as for feature extraction. Frame-level and clip-level descriptors are available for each sound. Similarity search and feature extraction services are supported with Gaia and Essentia [3], both technologies developed at the MTG.

A new feature of Freesound 2.0 is a web API which is available through the Django application and based on the RESTful [4] principles. The API allows accessing a basic set of resources (sounds, users, sound packs) from http clients. All of the mentioned methods for searching sounds

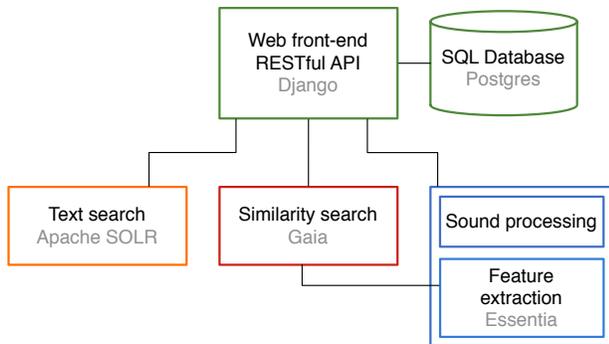


Figure 2. Freesound 2.0 architecture.



Figure 3. Interface for browsing by geotags.

are supported. Responses from the API can be obtained as JSON or XML markup. Content descriptors can be obtained through the API as JSON (frame level) or YAML (clip level). A number of clients have already been implemented for different programming languages. At the time of this writing, Python, Javascript, Actionscript and SuperCollider clients are available.

The new version of Freesound allows users to choose between three Creative Commons licenses for the uploaded sounds. These options are: CC0 (public domain), Attribution, and Attribution-NonCommercial [5]. The reason to offer these licenses is to ensure that all the content uploaded to Freesound can be reused by anyone, while giving contributing users the option to require the attribution of their work or restricting the use of their sounds to non-commercial activities. The source code of the web application is available as open source under the GNU AGPL license [6].

3. TYPES OF CONTENT

In this section we describe the types of content that can be retrieved from Freesound grouped into information about “sounds”, “users” and “communities”. By “types of content” we mean the different kinds of information that are available and can be used by researchers and developers working with Freesound.

3.1 Sounds

The most relevant type of content available in Freesound are the sampled sounds and their descriptions. Since Freesound was started with research usages in mind, from the very beginning the quality of sounds and their descriptions has been prioritized over the quantity. When sounds are uploaded to Freesound, they have to be described by adding tags and a textual description (currently the average is 6.6 tags per sound with almost 30,000 unique tags). Users can also add a geotag to indicate the location where the sound was recorded (there are more than 7,000 geotagged sounds). Moreover, sounds can be grouped in packs for a better organization (almost 75% of uploaded sounds belong to packs, having around 7,000 packs in total). After the description stage done by a contributing user, all sounds are manually moderated by a group of volunteer users in order to check that descriptions are correct and that sounds are legally uploaded (can be released with a CC license). Table 1 shows the kind of information that is available from sounds and sound packs.

3.2 Users

The users and their activity result into another relevant type of information content available in Freesound, providing a different perspective than that of the sounds. Although there are more than two million registered users, only 1% of them have contributed by uploading, commenting or rating sounds, or by writing in the forums. Indeed, only ~5,300 of the registered users have contributed with at least one sound. This means that the community that is uploading to Freesound is very small compared to the number of people that are downloading¹ (Freesound has already served more than 27 million downloads since its inception). From the users that have not actively contributed to Freesound we have information about the sounds they have downloaded (which sounds have been downloaded, by who and when). From the group of contributing users we have more useful information, as shown in table 2.

¹ To download a sound it is required to be a registered user. Non registered users can still play a preview of sounds online but can not download them.

Sound
<p>Basic metadata: filename, format, samplerate, bitrate, user that uploaded the sound...</p> <p>Description: textual description, tags*, geotag*, wave form/spectrogram visualizations, audio content descriptors (clip level and frame level)**</p> <p>User annotations: comments*, ratings*, downloads* (including number of downloads and list of users that downloaded the sound)</p> <p>Related sounds: similar sounds (content-based similarity), remixes of the sound (other sounds in Freesound created with the current sound), sources of the sound (sounds that have been used to create the current sound)</p>
Sound Pack
<ul style="list-style-type: none"> - Textual description of the pack - List of sounds in the pack - Downloads* (including number of downloads and list of users that downloaded the pack)

Table 1. Available information about sound and packs.

* This information is available with the corresponding timestamps since 2005.

** For more information about the available content-based descriptors check the Freesound API documentation at <http://www.freesound.org/docs/api>

3.3 Communities

The data related to the interaction among users gives very useful information about the community aspects in Freesound. The most clear community information available in Freesound is in the forums. Freesound forums currently have more than 6,000 threads and almost 30,000 posts. They are used for purposes like requesting sound samples, discussing about recording gear and techniques or for talking about work done using Freesound samples. However, we believe that the potential to learn community information does not end here. It has been observed that there are communities that can be implicitly identified in Freesound and that are related with the interest of groups of users for particular kinds of sounds [7]. We plan to explore further these user groups, adding specific functionality to support them and developing tools that can take advantage of the resulting community data to improve current sound description techniques [8].

User
<p>Profile: username, name/surname (if provided), date of registration, about text (if provided)</p> <p>Uploads: list of uploaded sounds* and uploaded packs*</p> <p>Downloads: list of downloaded sounds* and downloaded packs*</p> <p>Annotations: comments on sounds*, ratings*, added tags*, added geotags*, forum posts*</p>

Table 2. Available user information.

* This information is available with the corresponding timestamps since 2005.

4. CONCLUSIONS

Freesound has become a consolidated audio clip sharing site with a lot of content and a strong user community behind it. With the release of the new version of the site and the inclusion of an API, the content of Freesound has become easily accessible for the development of third-party applications and for research purposes. Freesound is clearly a valuable resource for any person interested in sounds but also for the scientific community working on sound and music computing. We believe that with this new version we are facilitating the use of Freesound as a research platform.

5. REFERENCES

- [1] Django Project, <https://www.djangoproject.com>
- [2] Apache Solr, <http://lucene.apache.org/solr>
- [3] Gaia and Essentia, <http://mtg.upf.edu/technologies/essentia>
- [4] Roy T. Fielding and Richard N. Taylor, "Principled design of the modern Web architecture", *Proceedings of the International Conference on Software Engineering*, pp. 407-416, 2000.
- [5] Creative Commons licenses, <http://creativecommons.org/licenses>
- [6] GNU AGPL, <http://www.gnu.org/licenses/agpl.html>
- [7] Gerard Roma and Perfecto Herrera, "Community Structure in Audio Clip Sharing", *Proceedings of the International Conference on Intelligent Networking and Collaborative Systems*, pp. 200-205, 2010.
- [8] Frederic Font and Xavier Serra, "Extending Sound Sample Descriptions through the Extraction of Community Knowledge", *Proceedings of the International Conference on User Modeling, Adaptation and Personalization*, pp. 418-421, 2011.