The revolution in music distribution and storage brought about by digital technology has fueled tremendous interest in and attention to the ways that information technology can be applied to this kind of content. The rapidly growing corpus of digitally available music data requires novel technologies that allow users to browse personal collections or discover new music on the world wide web, or to help music creators to manage and protect their rights. The general field of information retrieval (IR) is devoted to the task of organizing information and of making it accessible and useful. An information retrieval process begins when a user specifies his/her information needs by means of a query. The retrieval system should then deliver from a given data collection all documents or items that are somehow related to the query. For example, in the case of a typical web search, the query may consist of a text string of words and the search engine should deliver all text documents containing the specified words.

While ten years ago, most digital content was textual, it has now expanded to include audio, images, video, and other types of multimedia documents. This particularly holds for the music domain, where listeners enjoy ubiquitous access to huge music collections containing audio recordings, digitized images of sheet music, album covers, and an increasing number of video clips. Such huge amounts of readily available music require retrieval strategies that allow users to explore large music collections in a convenient and enjoyable way. Most of the available services for music recommendation and playlist generation rely on metadata and textual annotations of the actual audio content. For example, a music recording may be described by the name of the artist or composer, the title of the piece, or the song lyrics—editorial data that is typically created manually by domain experts. Typical query terms may be a title such as “Day Tripper” when searching for the song by The Beatles, or a composer’s name such as “Beethoven” when looking for the Fifth
Symphony. In this scenario, a user needs to have a relatively clear idea of what he or she is looking for (see Figure 7.1a). To overcome these limitations, recent retrieval systems complement editorial metadata with general and expressive annotations, which are also referred to as tags. Such tags may describe the musical style or genre of a recording, and also include information about the mood, the musical key, or the tempo. Many music recommendation systems rely on a large number of tags that have been generated by different users, automatically extracted from music blogs, and enriched by statistical information on user behavior and music consumption. Even though such tags may be quite noisy, they still express certain general trends and describe the music content in a statistical and human-centered way.

While text-based retrieval systems can be very powerful, they require the audio material to be enriched with suitable metadata—an assumption that is often not valid, in particular for less popular music or music material that is scattered in unstructured data collections. Furthermore, not all retrieval scenarios can be handled by a purely text-based approach. How should a retrieval system be designed if the user’s query consists of a short excerpt of a CD recording or a melody sung into a microphone? What can be done if only a few measures of a musical score are available? How can a user be satisfied if he or she looks for music with a specific rhythmic pattern or harmonic progression which have not been annotated? To handle such scenarios, one requires content-based retrieval systems that only make use of the raw music data, rather than relying on manually generated metadata. The term content loosely refers to any kind of information that can be directly derived from the music material to be queried, compared, and retrieved.

In this chapter, we present various content-based retrieval strategies that follow the query-by-example paradigm: given a music representation or a fragment of it (used as a query or example), the task is to automatically retrieve documents from