The Application of the IODA Document Architecture to Music Data

Adam L. Kaczmarek
Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics, ul. G. Narutowicza 11/12, 80-233 Gdansk, Poland

Keywords: Document Architecture, Search Methods, Query by Humming, Query Expansion.

Abstract: This paper is concerned with storing music data with the use of document architecture called Interactive Open Document Architecture (IODA). This architecture makes it possible to create documents which are executable, mobile, interactive and intelligent. Such documents consist of many files that are semantically related to each other. Semantic links are defined in XML files which are a part of a document. IODA documents with music data have been called IODA Music Documents. Such documents consist of a file with sound, a file with lyrics and many other files with data related to the document. It is easier to search for music files in a collection of music stored in the form of IODA Music Documents. Users can search for songs on the basis of a part of their lyrics or they can perform the search with the use of humming queries. In this kind of a search users record a part of a melody that they remember and the searching system retrieves music files that match the recorded melody.

1 INTRODUCTION

This paper presents the method of storing music data with the use of a new design of documents architecture called Interactive Open Document Architecture (IODA) (Siciarek, 2011). The main feature of an IODA document is that it is not regarded as a sequence of characters and images. The document has also embedded functionality. For example, an equation in a document created in the IODA architecture, is not only presented in a graphical form, but there is also a script which makes it possible to perform calculations with the use of this equation. Readers can input their own data and verify results. Moreover, data included in IODA documents are semantically linked. There is a link between the equation and the script which implements it. This kind of document architecture was applied to music data. As a result, documents called IODA Music Documents were created. These documents are introduced in this paper.

There are many kinds of files concerned with music. Songs can by available in the form of MP3 files. There are also web pages with lyrics of songs. A file with a song and web pages with lyrics of this song are semantically related. A user can match these data, but this relation is not specified by hyperlinks or any other kind of links. This paper addresses this problem. An IODA Music Document merges sound, lyrics of a song, its musical notation and other related data.

Using IODA Music Documents extends possibilities of performing the search for music data. Introducing this kind of documents makes it possible to develop a new type of applications. The application which takes advantage of IODA Music Documents can search for songs on the basis of a fragment of lyrics. In such an application it is also possible to perform the search in a locally stored collection of music files with the use of a query by humming (QoH). This kind of queries consists of fragments of melodies recorded by a user who is willing to find a song in which the melody occurs.

2 IODA DOCUMENTS

A digital document is often perceived as an equivalent of a paper document. However, a digital document has more functionality than a paper one. A reader of a digital document can easily perform a search for words and phrases which occur in documents. There can also be tables of content with links which automatically redirect the reader to a certain page. Nevertheless, the functionality of digital documents can be much further extended

L. Kaczmarek A..
The Application of the IODA Document Architecture to Music Data.
DOI: 10.5220/0005130302680273
In *Proceedings of the International Conference on Knowledge Management and Information Sharing* (KMIS-2014), pages 268-273
ISBN: 978-989-758-050-5

with the use of IODA architecture. Documents created in this architecture are mobile, executable, intelligent and interactive.

For example, IODA documents can be used in the process of reviewing scientific papers. A review of a paper would be a document, which is a computer agent with the purpose of being filled up by reviews. First, the document creates copies of itself. Different copies are intended for different reviewers. Then, documents send themselves to reviewers with the request to input information concerned with the review. Documents may use e-mail boxes to approach reviewers.

Functionality supporting the reviewer can be also embedded in the reviewed paper. For example, equations integrated with scripts which perform calculations. Reviewers can verify themselves results presented in reviewed papers with the use of such scripts. The review in the form of an IODA document, after being completed be reviewers, sends itself back to the editor. Then, reviews from different reviewers merges into one IODA document, which is a complete set of reviews. This is only a sample scenario of using IODA documents. The architecture of IODA documents is general and it applies to multiple processes that are concerned with using digital documents.

2.1 Layers in IODA Documents

An IDOA document consists of three layers. It has a data layer, information layer and knowledge layer. An IODA document does not have a specific format. In fact, such a document can consist of many files in different formats. However, every document has a spine. A spine binds together layers of document and files which comprise of an IODA document. The spine is implemented as an XML file.

2.1.1 Data Layer

The data layer of IDOA document consists of binary files. In particular, it contains files in JPG, PDF, DOC and other popular formats. Data layer also contain scripts, executable files or source code. Data layer makes an IODA document executable. Apart from the data in graphic or textural form, it provides files which can be executed by a user of IODA document. The IODA architecture also accepts running external applications or services available on the internet.

2.1.2 Information Layer

The information layer associates data available in files which consist of data layer. In the information layer, there are markings which combine fragments of graphic and textural files with executable content. For example, an equation in a PDF file can be matched with a script which calculates this equation. All these associates are implemented in XML format in the spine of a document. Thus, there is no need to modify files such as PDF or JPG to associate an executable content with some part of the file.

2.1.3 Knowledge Layer

The knowledge layer binds different parts of files containing text, graphic and the other data, which is intended to be used by a user of a document. The knowledge layer contains semantic links between fragments of document files included in an IODA document. A link can also refer to another IODA document. Moreover, in the knowledge layer, a user may store personal information such as annotations and notes. The knowledge layer can also log user's interaction with the document. In general, the knowledge layer stores additional information related to a user without modifying the basic structure of the document defined in the data layer and the information layer.

2.2 The IODA Functionality

There are four features of documents in IODA architecture. These documents are executable, mobile, interactive and intelligent. The main feature of an IODA document is that it is executable. Excitability of a document lays in its ability to run scripts and binary files included in the data layer of a document. Users can execute scripts and binary files in an IODA document by selecting some words, a fragment of an image or other visible part of a document. IODA documents contain data and information in two different formats. A document has a graphical form which can be for example printed or viewed by a user. However, parts of documents such as equations, charts, tables, algorithms and other data are also provided in the form of scripts and binary files included in data layer of an IODA document.

The mobility of an IODA document is its ability to migrate through the internet. IODA documents implement MIND architecture for the document migration (Godlewska and Wiszniewski, 2010), (Godlewska, 2010). A document has a workflow, which specifies what users should receive the document. There can also be an alternative workflow in case for example, some users who are

receivers of a document are not available. All these data are organized in XML files.

IODA documents are also interactive. Documents can support users in adding information to them and modifying data which are included in the document. For instance, if a user is filling up a form, an IODA document can present suggested words and phrases for various fields of a form. Apart from that, users of IODA documents can insert to documents personal notes and comments without modifying the main content of a document.

The intelligence of an IODA document is based on the fact that an IODA document may have a functionality of an intelligent computer agent. The data layer of a document can contain full implementation of an agent. The agent has the aim which it attempts to realize.

3 IODA MUSIC DOCUMENTS

A music file, in e.g. MP3 format, can be perceived as a kind of a document. Such files contain not only the sound of a song, but they also have ID3 tags. ID3 tags include data such as the title and the author of a song. There are also other music data stored in different files. Apart from the sound, a user can have lyrics of a song in the format of a text file or a copy of a web page. These files can be combined in IODA music document.

A document in the IODA architecture consists of many files and it sets semantic relations between parts of these files. The data layer of an IODA document is a container for different files in different formats. In case of an IODA music document, the data layer may contain the following files:

- the sound of the song (in MP3 or other format)
- lyrics of the song (in TXT or HTML file)
- the version of the song in MIDI format
- the metadata of the song (ID3TAGS)
- video
- other files (e.g. a musical notation of the song)

The data in all these files are connected by links in the information layer of an IODA document. If a song has lyrics, then the IODA document connects words in lyrics with a point of time in a song when those words occur. This connection makes it possible to verify what words are sung in different points of time in a song.

The link is neither defined in the file with a sound nor in the file with lyrics. It is a semantic link specified in the XML file, which is included in the information layer of a document. Therefore, files containing data, such as lyrics, are not modified. Similarly, a semantic link can connect points of time in a music file with some point in another file, such as for instance a musical notation of a song.

3.1 Ioda Links

It is problematic to define links without modifying files included in the data layer of an IODA Music Document. The link binds two points in files, but there are no addresses or labels in these files which can be used in defining the link. In XML files locations of links are specified with regard to the format of data file. There are also alternative methods of specifying a link location in some types of files. The different methods of locating a link are presented in Table 1.

Table 1: Defining links in IODA Music Documents.

FILE TYPE	LINK LOCATION
Files with sound (e.g. MP3) and video files	Time which elapsed since the beginning of a file
Images	Coordinates of a point
Text files	Line number and character number
Documents (e.g. PDF)	Page number, line number and character number (in case of links to text) or Page number and coordinates of a point (in case of links to images)
Web Pages (e.g. HTML files)	The location of lyrics; line number and character number with respect to lyrics location.

All methods presented in Table 1 explicitly define a point in a file. A link in XML files connects two these kinds of points.

The most complicated is defining link location in HTML files. HTML files with lyrics most often contain also other data such as images, multimedia, hyperlinks or text. For IODA Music Documents, the subject of concern is only the lyrics in this file. Links to web pages in an IODA Music Document are constructed from two elements: the location of lyrics in the file and the location in the lyrics defined with respect to the location of lyrics. The location of lyrics can be defined in different manners accordingly to the structure of a web page. The simplest method of defining it is based on finding words which consist on the first line of lyrics. The location of lyrics is the same in all links which refer

to the web page. Thus, it can be defined only once in XML file with links. When the lyrics location is specified, the location of a point in lyrics is defined similarly as in case of text files (Table 1). The location of a link is determined by a line number and a character number.

It would be also possible to define links without firstly specifying the location of lyrics on a web page. A web page can be entirely processed as a text file. All other data can be disregarded. A link can be defined in the same way as in case of text files. However, web pages available on the Internet tend to change regularly. There are parts of web pages which change very often and there is some content which usually remains the same. Links defined in IODA Music Document for a web page downloaded at some time may not apply to this web page downloaded at a slightly different time, if a web page is processed as a text file.

3.2 Ioda Files Container

IODA Music Document is not a new file format. The document consists of many files. Each file can be perceived by any application as a regular file is a file system. IODA Music Document is based on XML files which bind together different files. It is necessary to specify the location of these files.

There are no limitations on the structure of folders contain IODA document files. For example, all MP3 files can be located in the same folder and all files with lyrics can be located in the other one. It is only required to provide to an XML file a valid location of a file. The location can be either an absolute path or a path relative to the location of an XML file. However it is recommended to use relative paths and to organize a collection of files in such a way that it is easy to relocate files to another folder or device without making file paths in XML files obsolete.

4 SEARCHING FOR MUSIC FILES

Storing music files in the form of an IODA documents significantly increase possibilities of performing the search for songs. In particular, this applies to a collection of locally stored files. When users have a collection of MP3 files and they would like to find some files, they can perform the search on the basis of file names or ID3 tags included in MP3 files. Therefore, in order to find a song they need to know the name of the file, the title of a song or at least the author. Users not always have this information. They can only remember a part of lyrics or melody of a song. Possibilities of finding songs in the collection of MP3 files on the basis of these data are very limited.

4.1 Queries based on Lyrics

In a collection of music files stored in the form of IODA documents it is possible to find music files on the basis of their lyrics or melody. In an IODA document, the sound of song is integrated with its lyrics. Searching for a song on the basis of a fragment of lyrics, means searching the text files with lyrics which are included in IODA Music Documents. When there is a match such that the text searched by a user occurs in the lyrics, the IODA document containing these lyrics is presented to the user. With lyrics a user acquires also a file with sound of a song, because this file is included in an IODA document.

IODA Music Documents make it possible to deliver this kind of functionality. Such a search can be performed without any application dedicated for IODA document. All files with lyrics can be checked for occurrences of phrases that are parts of lyrics searched for by a user. However, this kind of search is inefficient because it requires opening many files with lyrics in every search process. The search can be improved by the use of an application which has a functionality of a search engine.

One of search methods used in search engines is based on creating a reverse index (Langville and Meyer, 2012). Such an index contains lists of words. Each word is correlated with a list of web pages that contain this word. When a search is performed for a user's query, the search engine retrieves a list of web pages that contain words included in the query by reading data from the index. A similar method can be applied to searching for lyrics in IODA Music Documents. An application that performs the search can have a list of words with lists of documents containing them. When a user searches for a fragment of lyrics the application can limit the number of potentially relevant music files to those files which contain words included in the fragment of lyrics provided by a user.

4.2 The Query by Humming

The IODA architecture also makes it possible to search for music files on the basis of a fragment of their melodies. Applications which can perform this

kind of search are called Query by Humming (QbH) systems (Kotsifakos et al., 2012). In these applications users can sing and record a part of a melody. The search is performed on the basis of such a humming query. There are many web sites, which provide this kind of functionality, but performing such a search in a local collection of MP3 files is problematic. Formats, such as MP3, contain sound of a song which is dedicated for users. In a song, there can be many instruments and vocals. In MP3 file there is only one stream of data containing all sounds, which comprise on a song. Performing search with humming query in MP3 files is problematic as it is hard for computer algorithms to detect the main melody of the song and match it with users' humming query.

In case of IODA music files, it is possible to perform QbH search. In an IODA document, apart from lyrics and file with a sound of a song such as MP3, there can be another file with an equivalent of a song in a form of a MIDI file. MIDI file contains only precisely defined sequences of sounds played on various instruments. Every instrument taken into account in MIDI file corresponds to different stream of data in this file. It is much simpler to match humming queries with MIDI files than match these queries with MP3 files.

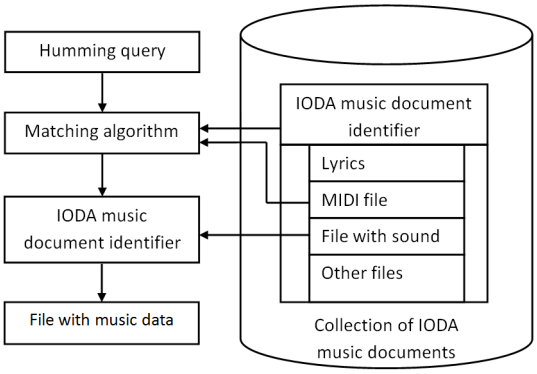

Figure 1: Retrieving music files on the basis of humming queries in the collection of IODA music.

The process of acquiring songs on the basis of humming queries in the collection of IODA Music Documents is presented in Fig. 1. The search in IODA documents, on the basis of humming query, is performed similarly as the search for fragment of lyrics. First, user's query is matched with melodies in MIDI format. Then, when a MIDI file is relevant to the query the IODA document containing this MIDI file is provided to the user. The document contains also full song in MP3 or other format, which a user can take advantage of.

4.3 Query Expansion

Users can be supported in the search with humming queries similarly as users of search engines are supported in searching for web pages. Search engines provide functionality designed to facilitate users in forming appropriate queries. For example, if a word in a query contains a spelling error, then a search engine suggests a correct form of a word. When a user types a query a search engine can present a list of popular queries, which contain letters typed by a user in a text box of a search engine. This method is called real-time query expansion (White and Marchionini, 2007). There are also interactive query expansion methods which are used after the search for a given users' query is complete (Fonseca et al., 2005). Apart from a list of found web pages for a user's query, search engines present a list of similar queries semantically related to user's query. If users are not satisfied with search results, they can modify their query and perform the search again.

The search for an IODA Music Document can be performed by a query by humming system which can also support users in creating humming queries. Users' humming query can be inaccurate and users may have some problems with correctly singing a part of a song that they are looking form. A system designed to search the collection of the IODA Music Document can provide users with suggestions concerned with improving their queries. These suggestions can be generated by Music Clustering by Directions algorithm (Kaczmarek, 2013). This algorithm is designed to present to a user fragments of melodies which are related to the humming query that a user recorded. Melodies are presented in a form of a tag cloud with musical notation (Fig. 2).

The functionality of this kind of tag cloud is such that it plays fragments of melodies when a user selects them. The purpose of the algorithm is to present different kinds of melodies related to users' query. If there is a melody that a user considers as relevant to a song that she or he is looking for, then this melody can be added to user's query in order to improve this query and make it more accurate.

Music Clustering by Direction (CBD) algorithm derives from Clustering by Direction algorithm which was designed to support users of search engines in forming queries (Kaczmarek, 2011). It is a kind of interactive query expansion method. The results of the algorithm were presented to the user in a form of a tag cloud containing words related to user's query. These words were used to expand the query.

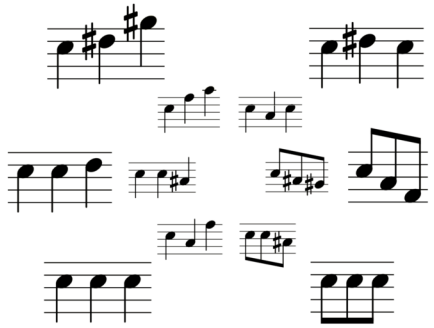

Figure 2: The interface in Music Clustering by Directions algorithm.

4.4 Other Functionalities

The IODA Music Document, apart from features described in previous sections, preserves also functionalities of a regular IODA document such as mobility and interactivity. For example, users of IODA Music Documents can make personal notes related to music files. These notes are stored in knowledge layer of an IODA document. Moreover, an IDOA Music Document can have a workflow which can be used in delivering music files. This functionality can be used by on-line music stores for specifying receivers of music files and providing them with these files.

5 CONCLUSIONS

Taking advantage of the IODA architecture can became a widely use method for storing music data. The architecture supports users in collecting, retrieving and verifying information. There is also a great potential for development of applications dedicated to this architecture. Such application can support users in creating and modifying XML files included in the layers of a document. Applications can also be used to perform different kinds of searches in IODA Music Documents.

ACKNOWLEDGEMENTS

This work was supported by the National Science Center grant no. 2011/01/B/ST6/06500.

REFERENCES

Fonseca B. M., Golgher P., Pôssas B., Ribeiro-Neto B., and Ziviani N. 2005. Concept-based interactive query expansion. In *Proceedings of 14th ACM Conference on Information and Knowledge Management, (Bremen, Germany)*, ACM. 696-703.

Godlewska, M., and Wiszniewski, B. 2010. Distributed MIND – A New Processing Model Based on Mobile Interactive Documents, *Parallel Processing and Applied Mathematics LNCS 6068*, Springer Berlin / Heidelberg, 244-249..

Godlewska, M. 2010. Agent System for Managing Distributed Mobile Interactive Documents. *Agent and Multi-Agent Systems: Technologies and Applications LNCS 6071*, Springer Berlin/Heidelberg, 390-399.

Kaczmarek, A.L. 2011. Interactive Query Expansion With the Use of Clustering-by-Directions Algorithm, *IEEE T. Ind. Electron.* 58, 8 (Aug. 2011), 3168-3173.

Kaczmarek, A.L. 2013. Information Retrieval with the Use of Music Clustering by Directions Algorithm. In *Proceedings of International Joint Conferences on Computer, Information and Systems Sciences and Engineering CISSE13* (University of Bridgeport, USA, Dec 12 - 14, 2013), 79:1-79:6.

Kotsifakos A., Papapetrou P., Hollmén J., Gunopulos D., and Athitsos V. 2012. A survey of query-by-humming similarity methods. In *Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '12)*. ACM, New York, USA. 5:1-5:4.

Langville A.N. and Meyer C.D. 2012. *Google's PageRank and Beyond: The Science of Search Engine Rankings Paperback*, Princeton University Press.

Siciarek, J., and Wiszniewski, B. 2011. IODA - an Interactive Open Document Architecture. In *Procedia Computer Science 4, ICCS 2011, Proceedings of the International Conference on Computational Science,* Elsevier, 668-677.

White R.W., and Marchionini G. 2007. Examining the effectiveness of real-time query expansion. *Inf. Process. Manage* 43, 3 (May. 2007), 685-704. ACM, New York, USA.