

Workshop on the Creation of Standardized Test Collections, Tasks, and Metrics for Music Information Retrieval (MIR) and Music Digital Library (MDL) Evaluation

A proposal submitted to the Second Joint Conference on Digital Libraries (JC DL 2002)

J. Stephen Downie, PhD
Assistant Professor,
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
jdownie@uiuc.edu

Executive Summary

We hope to engage the participation of all those interested in MIR and MDL research and evaluation. We plan on asking all interested parties to submit formal White Papers outlining their individual perspectives on what needs to be done to create *meaningful* MIR and MDL test collections, retrieval tasks, and evaluation metrics. We would like to hear from musicologists, music theorists, audio-retrieval experts, symbolic-retrieval experts, librarians, lawyers, industry and business representatives; in short, from all the key groups in MIR and MDL research. The compilation of these perspectives and the discussion that follows at the Workshop will form the bedrock upon which a solid foundation of future research can be built.

Problem Statement

There is a very strong feeling among the MIR and MDL research communities that we must somehow get our act together and acquire standardized collections of music (in multiple representations), standardized retrieval tasks (i.e., query sets, etc.), and standardized retrieval metrics. At the heart of all of this is very real concern that we are currently only “playing” at being objective and scientific insofar as our work is represented in our literature. One model being bandied about is the formation of a TREC-like set of collections and evaluation criteria. These collections and criteria will have to be designed specifically to address the wide variety of *music-specific* problems facing the MIR and MDL research communities. Establishing the ideal framework for developing these collections and criteria is the open question to be addressed at the Workshop.

As it stands now, notwithstanding everyone’s best efforts, MIR and MDL research is being conducted and reported upon in an *ad hoc* fashion with no mechanism for meaningful comparisons between proposed systems (i.e., each team has no standard way of comparing their results to the results of others). For example, we have absolutely no way right now to evaluate the relative merits of the various approaches when, for example, one group of researchers has a collection of MP3 rock songs, another 9354 folks in MIDI format, and yet another team uses 200 classical piano pieces in GUIDO format. Add to this problematic situation, the fact that each team reports its results in whatever metric with which they might be most familiar. Since this is the case, MIR and MDL research and development is falling short of being truly scientific in those things where science--that is, the traditional scientific method--can help advance the state-of-the-art.

Please see **Appendix A** for more complete background information concerning the need for standardized test collections, tasks, and evaluation metrics

Methods:

- 1) Communication will be grounded in the creation of "White Papers" or "Position Papers". These documents will be *scholarly* works outlining the thoughts, opinions, and arguments of their authors with regard to *their* visions of what needs to be included/considered in the creation of standardized MIR/MDL testbed collections, evaluation tasks, and evaluation metrics. I expect (and want) these documents to be *highly idiosyncratic*: a musicologist will write with respect to the needs and wants of musicologists, a lawyer with respect to legal issues, an audio engineer with regard to audio issues, and so on. It is my goal to attract such a wide range of submissions that the multi-disciplinary approach to MIR and MDL research is truly represented.
- 2) Invite some expert representative(s) from the Information Retrieval and Digital Library evaluation communities (e.g., TREC organizers (<http://trec.nist.gov>)) to provide some input on their experiences in establishing testbed collections, tasks, and metrics.

Goals:

- 1) The creation of a reference document. That is, the collection of our submitted papers would form the starting-point literature from which future arguments and work in the area of MIR evaluation could proceed.
- 2) A consensus document or "manifesto" outlining the basic framework for further development of collections and evaluation standards would be the "formal" outcome of the workshop itself.
- 3) Both documents mentioned above would be made available for all to use as foundational material for the solicitation of research funding, for the establishment of collection creation agreements, and for the possible creation of ongoing evaluation meetings and "contests."

Structure:

Part 1 (Morning): The first part of the Workshop would be formal presentations of those wishing to present their White Paper materials. Part 1 would also include the Key Note address (TBA) from an expert in evaluation issues.

Part 2 (Afternoon): The second part of the Workshop would be devoted to process of arriving at the "consensus document" mentioned above.

Projected Participation:

I project the enrollment of approximately 40-50 persons. I project the submission of 10-20 White Papers.

Response to initial expressions of interest in the Workshop itself has been quite strong. Please see Appendix B (i.e., the initial list of Workshop signatories). I am also heartened by the breadth of interests represented (e.g., computer science, audio engineering, library and information science, musicology, etc.) and by the wide range of nationalities represented in the list of initial signatories. Missing from the list are those who have emailed me personally, some of whom have already submitted abstracts of their White Papers. Enrollment will most assuredly be enhanced by the presence of members of the three big NSF funded MDL projects: Indiana, Michigan and Johns Hopkins.

Communication Channels:

The MIR and MDL communities have their own listserv with 300+ members subscribed (music-ir@ircam.fr). The organizers of the *International Conference on Music Information Retrieval* (ISMIR 2002, Paris, October 2002) have created a substantial collection of multi-disciplinary mailing lists through which announcements will be made. Interactive communications will be hosted *via* WWW pages and forms at <http://music-ir.org> where I am the Webmaster.

Biographic Information:

J. Stephen Downie is an assistant Professor at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. He holds a Bachelor's in Music Theory and Composition, an MLIS, and a PhD, all from the University of Western Ontario. He teaches in the areas of information retrieval, information representation, and information processing at the undergraduate, masters and doctoral levels. Stephen's research and publications focus upon music information retrieval and informetrics. He led the Exploratory Workshop on Music Information Retrieval at SIGIR '99. Dr. Downie was named a faculty fellow of the National Center for Supercomputing Application for 2000-2001. He has been very active in helping establish a vibrant music information retrieval research community through his organizational efforts to make the International Symposia on Music Information Retrieval an ongoing success: He was ISMIR 2000 programme chair, ISMIR 2002 general chair, and now chairs the ISMIR 2002 Panels and Tutorials committee. He is head of the Mellon-funded Music Information Retrieval Annotated Bibliography Project (<http://music-ir.org>).

Appendix A

Information Concerning the Need to Create Standardized MIR Test Collections, Tasks and Evaluation Metrics for MIR Research and Development

Background:

On 16 October 2001, at the Second International Symposium on Music Information Retrieval, Bloomington, IN, (ISMIR 2001: <http://ismir2001.indiana.edu>), Drs. Don Byrd and J. Stephen Downie convened an informal session to discuss issues pertaining to the creation of standardized MIR test collections, retrieval tasks, and evaluation metrics. After vigorous debate, a general consensus was reached that the MIR research community *must* move forward on the creation of such collections, tasks, and metrics. At the behest of Dr. David Huron, the following resolution was drawn up by Matthew Dovey for submission to the final plenary session of ISMIR 2001, 17 October 2001. The resolution was overwhelmingly endorsed by those present.

It was further resolved that we encourage all those who have an interest in these matters to express their support for the resolution by allowing them to electronically “sign” the resolution (<http://music-ir.org/mirbib2/resolution>). This resolution and its appended signatures will be used to solicit from relevant funding bodies and rights holders the materials necessary to make the establishment of standardized MIR test collections, tasks, and evaluation metrics a reality.

The Resolution :

There is a current need for metrics to evaluate the performance and accuracy of the various approaches and algorithms being developed within the Music Information Retrieval research community. A key component for the development of such metrics would be a corpus of test data consisting of both audio and structured music data. Such a corpus would need to be readily available to the research community with international clearance of all relevant copyright and other intellectual property rights necessary to use this data solely for the purpose of music information retrieval research.

At the International Symposium on Music Information Retrieval, 2001, hosted by Indiana University, a resolution was passed endorsing support for activities which would lead to the establishment of such a corpus and hence evaluation metrics for this community.

Signatories

Yan Huang
Engineer
Panasonic Speech Technology Lab
Sat Feb 9 01:39:44 2002

Hannah Frost
Media Preservation Librarian
Stanford University
Tue Jan 29 12:34:32 2002

Michael Nelson
Associate Professor
Faculty of Information and Media Studies University of Western Ontario
Thu Jan 3 12:31:37 2002

Yazhong Feng
Ph.D student
Department of Computer Science,Zhejiang University,China
Wed Dec 19 19:40:29 2001
Yohannes Abraha
Mr.
Research and Documentation Center,Erirea
Mon Dec 17 11:03:05 2001

Jeremy Pickens
Research Assistant
Center for Intelligent Information Retrieval Department of Computer Science University of Massachusetts Amherst
Mon Dec 10 12:15:49 2001

Anna Pienimäki
Assistant Researcher
Dept. of Computer Science, University of Helsinki, Finland
Tue Dec 4 02:53:16 2001

Jyh-Shing Roger Jang
Associate Professor
Computer Science Department, National Tsing Hua University, Taiwan
Sat Dec 1 08:51:54 2001

David Meredith
Research Fellow
City University, London
Wed Nov 28 10:10:53 2001

Jiyoung Shim
graduate student
Yonsei University
Tue Nov 27 07:36:34 2001

Pierre-Yves ROLLAND
Assistant Professor of Computer Sc. and Multimedia Tech.
Universite d'Aix-Marseille 3, France
Mon Nov 19 08:54:16 2001

Ichiro Fujinaga
Professor
Peabody Conservatory of Music Johns Hopkins University
Wed Nov 14 19:59:46 2001

Sayed Choudhury
Hodson Director of the Digital Knowledge Center
Johns Hopkins University
Wed Nov 14 13:32:37 2001

Mikael Fernstrom
Mr
Interaction Design Centre, Univeristy of Limerick, IRELAND
Tue Nov 13 10:22:45 2001

Donncha S. Ó Maidín
Head of Department of Computer Scienc
University of Limerick, Limerick, Ireland
Tue Nov 13 09:21:19 2001

Dirk Van Steelant
Master Computer Science
KERMIT - Ghent University, Belgium
Tue Nov 13 02:13:08 2001

Kjell Lemström
Researcher
Dept. of Computer Science, University of Helsinki, Finland
Tue Nov 13 02:00:33 2001

John McPherson
Ph.D Candidate
University of Waikato Hamilton New Zealand
Mon Nov 12 17:12:46 2001

Jane M. Subramanian
Music Cataloger/Archivist
SUNY Potsdam Potsdam, NY
Mon Nov 12 13:37:24 2001

Fabien Gouyon
PhD candidate
Audiovisual Institute, Pompeu Fabra University, Barcelona
Mon Nov 12 08:28:41 2001

George Tzanetakis
PhD candidata
Computer Science Department, Princeton University
Sat Nov 10 12:48:43 2001

Jane Singer
PhD candidate
Hebrew University of Jerusalem (Israel)
Fri Nov 9 06:52:18 2001

Micheline Lesaffre
Musicologist, PhD. Candidate
IPEM, dept. Musicology, Ghent University, Belgium
Fri Nov 9 04:42:51 2001

Taelman Johannes
R&D Engineer
Institute for psychoacoustics and electronic music, Ghent University, Belgium
Fri Nov 9 04:37:44 2001

Dirk Moelants
Assistant
IPEM-Dept. of Musicology Ghent University, Belgium
Fri Nov 9 04:36:28 2001

Mark Sandler
Professor of Signal Processing
Queen Mary University of London
Wed Nov 7 03:09:27 2001

Min HUANG
Associate Professor
University Library, Shanghai Jiao Tong University
Tue Nov 6 19:16:05 2001

Craig Burket
Senior Software Development Engineer
Stellent, Inc. 62 Forest St., Suite 100, Marlborough, MA 01752, USA
Tue Nov 6 12:46:09 2001

Cheryl Martin
Director of Bibliographic Services
McMaster University, Hamilton, Ontario, Canada
Mon Nov 5 08:54:37 2001

Lenore Coral
Music Librarian
Sidney Cox Library of Music and Dance Cornell University
Mon Nov 5 08:07:45 2001

Dick Vestdijk
librarian
on private title, The netherlands
Mon Nov 5 07:40:47 2001

Steven Blackburn
Independant Researcher
Formerly with the University of Southampton, UK.
Sat Nov 3 11:41:37 2001

John Vallier
graduate student (info studies/ethnomusicology)
UCLA
Sat Nov 3 09:40:08 2001

David Dorman
Consultant
Lincoln Trail Libraries System Champaign, IL
Fri Nov 2 12:11:41 2001

Ronald Rousseau
professor
KHBO, Dept. Industrial Sciences and Technology, B-8400 Oostende Belgium
Fri Nov 2 01:02:39 2001

Joyce McGrath
Foundation member of IAMLANZ in 1968 now retired
Formerly Art & Music Librarian at the Stat Library of Victoria, Melbourne, Australia
Thu Nov 1 17:34:25 2001

Pamela Thompson
Chief Librarian
Royal College of Music Prince Consort Road London SW7 2BS UK
Thu Nov 1 16:50:54 2001

James M Turner
Professeur agrégé
École de bibliothéconomie et des sciences de l'information, Université de Montréal
Thu Nov 1 15:23:31 2001

Pamela Juengling
Music Librarian
University of Massachusetts/Amherst
Thu Nov 1 14:10:23 2001

Leslie Troutman
Associate Professor of Library Administration, Music Library
University of Illinois at Urbana-Champaign
Thu Nov 1 12:35:04 2001

Adriane Swalm Durey
PhD Candidate
School of Electrical and Computer Engineering, Georgia Institute of Technology
Wed Oct 31 11:01:05 2001

Donald Byrd
Senior Scholar
Indiana University
Wed Oct 31 08:07:15 2001

Aymeric Zils
PhD candidate
SONY Computer Science Labs, Paris, France.
Wed Oct 31 03:31:23 2001

Jean-Julien Aucouturier
Assistant Researcher
SONY Computer Science Labs, Paris, France.
Wed Oct 31 02:50:55 2001

Chaokun Wang
Ph.D. candidate
Dept. of Computer Science and Technology, HIT(Harbin Institute of Technology), P.R.C.
Tue Oct 30 19:31:33 2001

Nigel Nettheim
Honorary Adjunct Fellow
Macarthur Auditory Research Centre Sydney
Tue Oct 30 15:21:14 2001

Timothy DiLauro
Deputy Director, Digital Knowledge Center
Johns Hopkins University Milton S. Eisenhower Library
Tue Oct 30 14:52:05 2001

Allen Renear
Associate Professor, Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
Tue Oct 30 14:25:31 2001

David Bainbridge
Lecturer
Dept. of Computer Science, University of Waikato, New Zealand
Tue Oct 30 14:20:20 2001

Matija Marolt
Ph.D. candidate
Faculty of Computer and Information Science, University of Ljubljana, Slovenia
Tue Oct 30 14:11:58 2001

Tim Crawford
Project Manager, OMRAS
Department of Music King's College Strand LONDON WC2R 2LS United Kingdom
Tue Oct 30 09:56:34 2001

Beth Logan
Principal Research Scientist
Cambridge Research Laboratory, Compaq Computer Corporation
Tue Oct 30 08:01:54 2001

David Nichols
Visiting Lecturer in Digital Libraries
Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign
Mon Oct 29 14:42:27 2001

Mary Wallace Davidson
Head, William & Gayle Cook Music Library
Indiana University, Bloomington, USA
Mon Oct 29 11:03:48 2001

Elena Ferrari
professor
Università degli Studi dell'Insubria, Como (Italy)
Mon Oct 29 09:31:48 2001

Dr Stefan Rueger
Research Lecturer
Department of Computing Imperial College, London, UK
Mon Oct 29 09:00:56 2001

Margaret Cahill
Junior Lecturer
University of Limerick, Ireland
Mon Oct 29 07:51:50 2001

Seymour Shlien
research scientist
communications research centre
Mon Oct 29 06:45:28 2001

Juan Pablo Bello
Ph.D. candidate
Queen Mary College, University of London
Mon Oct 29 05:57:19 2001

Bozena Kostek
Assoc. Prof. (Ph.D., D.Sc., Eng.)
Sound & Vision Eng. Dept., Technical University of Gdansk, Poland
Mon Oct 29 04:25:30 2001

Dr. Xavier Serra
Director
Audiovisual Institute, Pompeu Fabra University
Mon Oct 29 03:55:00 2001

Inigo Barrera
Ph. D. Candidate
UPC (Polytechnical University of Catalonia), Barcelona, Spain
Mon Oct 29 02:37:49 2001

Shyamala Doraisamy
MPhil/PhD candidate
Imperial College, London, UK.
Mon Oct 29 02:30:27 2001

Thomas Sødring
Ph. D. Candidate
Dublin City University
Mon Oct 29 02:16:25 2001

Giuseppe Frazzini
Information Technologist
Teatro alla Scala, Milano
Mon Oct 29 01:25:47 2001

Gaetan Martens
Master Computer Science
IPEM + Applied Mathematics And Computer Science, Ghent University, Belgium
Sun Oct 28 03:10:11 2001

Koen Tanghe
R&D engineer
IPEM (Institute for Psychoacoustics and Electronic Music), Ghent University, Belgium
Sat Oct 27 15:35:55 2001

Michael Good
CEO
Recordare LLC
Fri Oct 26 23:41:14 2001

Cheng Yang
Ph.D. candidate
Stanford University
Fri Oct 26 18:19:50 2001

Jonathan Foote
Sr. Research Scientist
FX Palo Alto Laboratory, Inc.
Fri Oct 26 12:49:18 2001

Les Gasser
Associate Professor
Graduate School of Library and Information Science University of Illinois at Urbana/Champaign
Fri Oct 26 12:18:47 2001

P. Bryan Heidorn
Assistant Professor
University of Illinois, Graduate School of Library and Information Science
Fri Oct 26 12:14:13 2001

Sally Jo Cunningham
Senior Lecturer
University of Waikato
Fri Oct 26 11:23:48 2001

Eric Isaacson
Associate Professor of Music Theory
Indiana University, Bloomington, USA
Fri Oct 26 11:14:34 2001

Jay Kim
Ph. D student
Rutgers University
Fri Oct 26 11:13:25 2001

Matthew J. Dovey
R&D Manager, Libraries Systems
Oxford University
Fri Oct 26 10:55:41 2001

Antony Gordon
System administrator/cataloguer
British Library National Sound Archive
Fri Oct 26 10:54:02 2001

Emanuele Pollastri
Engineer - PhD Student
Università degli Studi di Milano Dipartimento di Scienze dell'Informazione
Fri Oct 26 10:53:40 2001

Alan Smeaton
Professor of Computing
Dublin City University, Ireland
Fri Oct 26 10:52:37 2001

Maurizio Longari
MS - PhD Student
Università degli Studi di Milano Dipartimento di Scienze dell'Informazione
Fri Oct 26 11:08:18 2001

Jeffrey M. Vinocur
Undergraduate
Cornell University, Ithaca, N.Y., USA
Fri Oct 26 11:45:54 2001

Perry Roland
Information Technologist
University of Virginia
Fri Oct 26 13:08:28 2001

J. Stephen Downie
Assistant Professor
Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign
Fri Oct 26 13:02:54 2001

Appendix B

Signatories as of 11 February 2002 to the Initial Call for Expressions of Interest

Taken from: http://music-ir.org/mirbib2/eval_workshop

Suzanne Lodato
Associate Program Officer, Scholarly Communications
The Andrew W. Mellon Foundation
Mon Feb 11 13:48:43 2002

Adam Berenzweig
PhD Candidate
Columbia University
Mon Feb 11 10:21:28 2002

A set of corpi (of various sizes and focus) will be a great help to the field. Some tricky licensing issues lurk here; are the record labels behind us?

Shyamala Doraisamy
MPhil/PhD candidate
Imperial College, London
Mon Feb 11 04:13:12 2002

David Bainbridge
Senior Lecturer
University of Waikato, New Zealand
Mon Feb 11 04:12:58 2002
Formulation of an MIR testbed/evaluation suite is crucial for the development of this field

Joe Futrelle
Sr. Research Programmer
National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign
Sat Feb 9 11:23:28 2002

Jon W. Dunn
Assistant Director for Technology, Digital Library Program
Indiana University
Fri Feb 8 17:57:08 2002

Matthew J. Dovey
R&D Manager, Library Systems
Oxford University
Fri Feb 8 17:28:21 2002

Jyh-Shing Roger Jang
Associate Professor
Department of Computer Science
Fri Feb 8 08:54:33 2002
I'm interested in content-based music retrieval via query by singing, humming, and tapping, etc.

Perry Roland
Information Technology Specialist
University of Virginia
Fri Feb 8 07:59:24 2002
I'm interested in music representation.

Nicola Orio
Assistant Professor
University of Padova Department of Electronics and Computer Science
Fri Feb 8 07:51:20 2002

Massimo Melucci
Associate Professor
University of Padova Department of Electronics and Computer Science
Fri Feb 8 07:50:06 2002

Micheline Lesaffre
Musicologist
IPEM (Institute for Psychoacoustics and Electronic Music), Ghent University, Belgium
Fri Feb 8 04:23:29 2002
I support and I am interested because my research for the MAMI project concerns the development of a conceptual framework for classification, description and representation of musical sound in the field of content-based musical audio mining. A main characteristic of the MAMI-project is its focus on music as audio signal.

Jia-Lien Hsu
Researcher
Computer & Communication Lab., Industrial Technology Research Institute, Taiwan.
Thu Feb 7 19:11:38 2002

Perfecto Herrera
Research Assistant
Universitat Pompeu Fabra
Thu Feb 7 11:03:57 2002

Dan Ellis
Assistant Professor, Electrical Engineering
Columbia University
Thu Feb 7 09:24:41 2002
We've struggled with defining our own artist and genre ID corpora. It's a huge boost to a field to have common tasks for meaningful comparisons between different approaches. But it's also incredibly delicate, as an imperfectly-designed task can channel lots of effort towards exploiting peculiarities when aceing the task becomes more important than the general problem...

Koen Tanghe
R&D engineer
IPEM (Institute for Psychoacoustics and Electronic Music), Ghent University, Belgium
Thu Feb 7 08:50:38 2002
You have my interest and support. Won't be at the workshop though. Some viewpoints will be sent by email. Hope this helps.

Beth Logan
Member of Research Staff
Compaq Computer Corporation
Thu Feb 7 08:42:04 2002

I have a background in speech recognition where there is a long history of standard evaluation tasks. I'd be interested in helping set up something for music (audio) evaluations. It's unlikely I can attend the workshop however but please keep me updated.

Jay Kim
Ph. D Student
Rutgers, the State University of New Jersey
Fri Feb 1 17:03:04 2002

Yaxhong Feng
Ph.D candidate
Institute of Artificial Intelligence, Zhejiang University, China
Thu Jan 31 10:09:36 2002
I am concerning what will be discussed.

Donald Byrd
Senior Scholar
School of Music, Indiana University
Thu Jan 31 05:44:20 2002

Eric Isaacson
Associate Professor of Music Theory
Indiana University
Tue Jan 29 21:44:13 2002
My interest is as a music theorist, with interests in theory, analysis, and instruction. My primary interests is in searching of score-derived data, though the synchronization of score data and audio is important to me as well.

Tim Crawford
Mr
King's College, London, Department of Music
Tue Jan 29 10:14:07 2002
I want to use MIR in musicology, so am equally interested in 'known-item' searches. I am still to be convinced that we have any understanding of musical 'relevance'.

Tae Hong Park
Grad Student
Princeton Music Dept.
Tue Jan 29 10:00:40 2002

Elizabeth Davis
Head, Music & Arts Library
Columbia University New York, NY
Tue Jan 29 09:16:22 2002

George Tzanetakis
PhD candidate
Computer Science Department, Princeton University
Tue Jan 29 09:14:27 2002
audio information retrieval evaluation, implementation issues with large audio collections

J. Stephen Downie
Assistant Professor
Graduate School of Library and Information Science University of Illinois at Urbana-Champaign
Mon Jan 28 17:09:13 2002
I look forward to seeing you all at JCDL '02!