
 soundsoftware.ac.uk

SoDaMaT

Sound Data Management Training

Steve Welburn
Centre for Digital Music
Queen Mary, University of London

The SoDaMaT Project

- » Discipline-specific data management training for digital audio and music researchers
- » So far:
 - Presentation at Sound Software 2012
 - 30 minute tutorial section at DAFx 2012 in York
 - 30 minute tutorial section at ISMIR 2012 in Porto

Outcomes from tutorials (so far)

- » DAFx tutorial was open to all attendees
 - ca. 100 audio researchers in audience
 - Reached people that did not expect it to be relevant
- » ISMIR tutorial was paid for by attendees:
 - ca. 40 in audience
 - Self selecting audience with existing interests
 - Added evidence of cases where RDM would have helped to try and engage researchers
- » Bristol Online Surveys questionnaire is in progress
- » The DropBox question...

Recovery of Overwritten Hard Disk Data

Hi, a friend of mine just overwrote two months of her PhD thesis with an older version. I know recovery of overwritten data is possible, but wonder if I'd need special hardware to do it. Does anyone know something about this ?

Thank You.

5 October 2005 Linux Forums - <http://tinyurl.com/8t7uaop>

Stolen laptop had PhD research

Thirty-five minutes spent in Langley's Willowbrook Shopping Centre cost a Surrey woman much more than she had anticipated.

Langley RCMP say that while she was shopping from 1-1:35 p.m. last Monday, someone broke into her vehicle and stole a number of items, including a Mac iBook laptop containing the research she had compiled as she worked towards her PhD.

"All that information was on that computer and she has no back-up file," said Langley RCMP spokesman Cpl. Brenda Marshall.

19 March 2008 Surrey Leader - <http://tinyurl.com/9hmtlv4>

Happiness is the return of a stolen computer, with data intact

Never has a man been so happy to see a computer full of data spreadsheets.

Claudio De Sassi's world fell apart when a car containing almost three years work towards his PhD was stolen two weeks ago.

De Sassi, a Canterbury University academic, could not hide his joy yesterday as police reunited him with his stolen laptop and backpack.

27 May 2010 The Press, NZ - <http://tinyurl.com/38sznnh>

The Lost Laptop Problem

- » 2010 Ponemon Institute report for Intel re. US laptops
 - On average, 2.3% of laptops assigned to employees are lost each year
 - In education & research that rises to 3.7%, with 10.8% of laptops being lost before the end of their useful life (~3 years i.e. within 1 PhD of allocation!)
 - 75% lost outside the workplace

- » Very similar results from 2011 European report!

Intel 2010 - <http://tinyurl.com/8c9m4bn>

Laptop Reliability

- » 2011 PC World Laptop Reliability Survey from 63,000 readers:
 - 22.6% had significant problems during the product's lifetime
 - Of which...
 - 19% had OS problems ~1 in 25 of all laptops
 - 18% had HDD problems ~1 in 25 of all laptops
 - 10% PSU problems ~1 in 50 of all laptops

PC World 2011 - <http://tinyurl.com/876qza5>

Hard Disk Failures

- » *Failure Trends In A Large Disk Drive Population*
 - Usenix conference on File and Storage Technologies 2007 (FAST '07)
 - Eduardo Pinheiro & Wolf-Dietrich Weber, Google Inc.
- » Data collected from over 100,000 disk drives at Google
- » As part of repairs procedures:
 - ~13% of disk drives replaced over 3 years
 - ~20% of disk drives replaced over 4 years

Article: <http://tinyurl.com/octz6b>

Thugs steal Christmas, doctoral dreams

A tiny television sits where a big screen used to, and a Christmas tree stands with little underneath it...

Even worse than the gifts, the crooks stole a MacBook Pro laptop and a LaCie hard drive.

The hard drive had ... her dissertation and nearly seven years of research for her doctoral degree she was set to finish in a few weeks.

Osuna had everything backed up on a separate hard drive in a safe, but burglars made off with that too.

"All I could think about is that all that time is gone, all that effort, everything is gone," Osuna said.

22 December 2010 KRQE - <http://tinyurl.com/9a5j56f>

Where To Keep Your Data

- » Commercial remote storage solutions (e.g. DropBox)
 - Check the T&Cs / SLA
 - Cost money
 - Not openly accessible on the web
 - No control over how data is stored
 - No control over physical location of data
 - Risk of lock-in
 - Bandwidth restrictions

JISC/DCC Curation In The Cloud : <http://tinyurl.com/8nogtmv>

Cloud Storage - Google

When you upload or otherwise submit content to our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. This license continues even if you stop using our Services (for example, for a business listing you have added to Google Maps).

1 March 2012 Google Terms of Service : <http://tinyurl.com/89dc9fa>

Cloud Storage - Microsoft

When you upload your content to the services, you agree that it may be used, modified, adapted, saved, reproduced, distributed, and displayed to the extent necessary to protect you and to provide, protect and improve Microsoft products and services. For example, we may occasionally use automated means to isolate information from email, chats, or photos in order to help detect and protect against spam and malware, or to improve the services with new features that makes them easier to use. When processing your content, Microsoft takes steps to help preserve your privacy.

19 October 19 2012 Microsoft services agreement : <http://tinyurl.com/8e4kucy>

Laptop Stolen From OSU Doctoral Student

...her car was broken into and her chrome Mac book pro was stolen.

She has a back-up for all but the last six months of research, but the most important part of the research had happened recently.

NBC4i January 06 2011 - <http://tinyurl.com/bmybv9x>

Archiving Data

- » BBC Domesday Project (1986)
 - Project to do a modern-day Domesday book
 - Used “BBC Master” computers with data on laserdisc
 - Collected 147,819 pages of text and 23,225 photos
 - Media expiring and obsolete technology put the data at risk!

Archiving Data

- » Domesday Reloaded (2011)
 - Required emulation of software
 - Images restored from original masters
 - <http://www.bbc.co.uk/history/domesday>

Lessons We Can Learn...

- » To allow long-term access to data
 - Don't use obscure formats!
 - Don't use obscure media!
 - Don't rely on technology being available!
 - Do keep original source material!

Where to publish data

- » Institutional repository – if one exists!
- » Project or research group web-sites
- » Journal Supplementary Materials
 - e.g. JASA, JNMR, CMJ
 - Check T&Cs – e.g. JASA ask for copyright to supplementary materials to be transferred to them.
- » Web archives – e.g. archive.org for audio files
- » Research data sites e.g. figshare.com

- » Talk to a librarian!

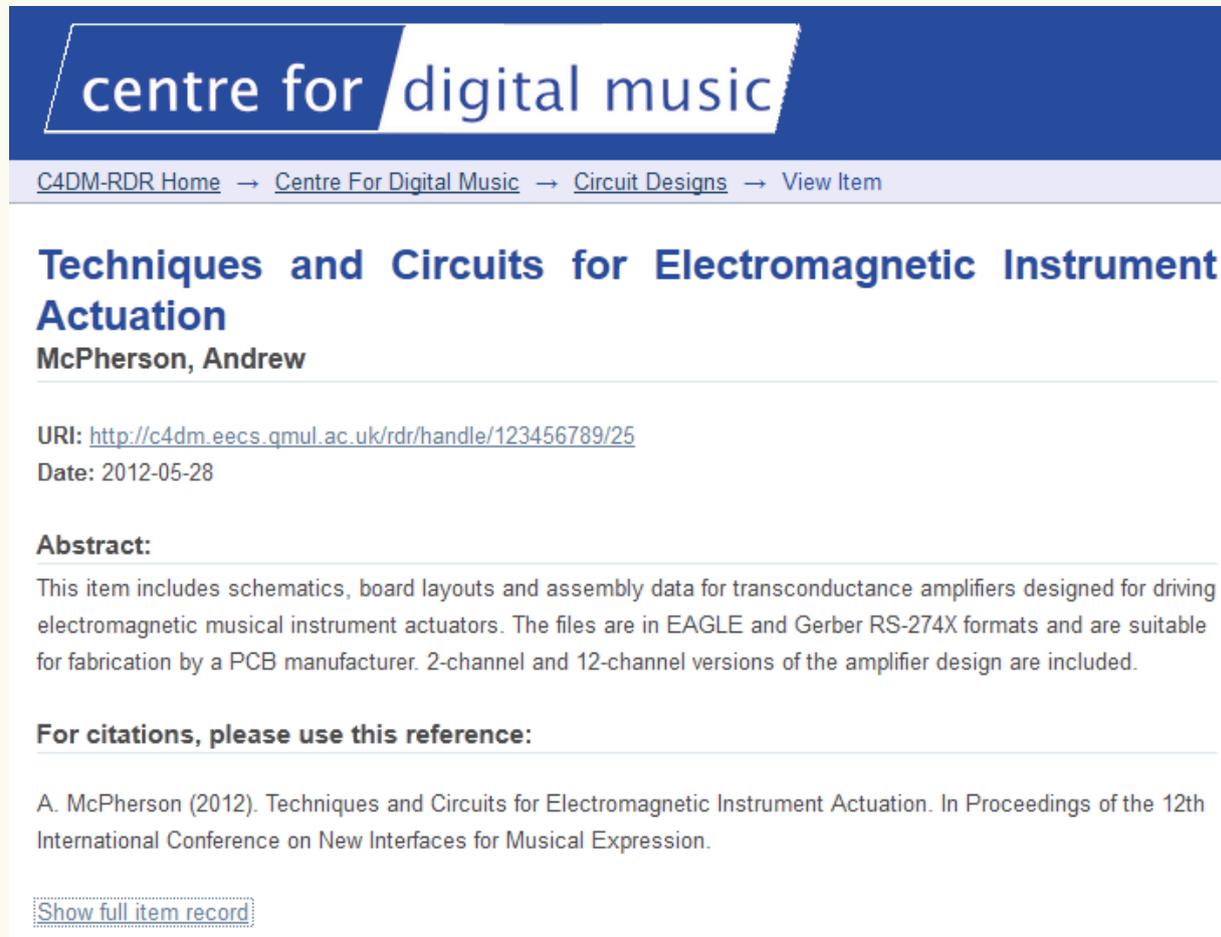
C4DM Research Data Repository

The screenshot shows the homepage of the C4DM Research Data Repository. The header features the 'centre for digital music' logo and navigation links for 'Login' and 'Help'. Below the header, the page is divided into several sections:

- Centre for Digital Music - Research Data Repository**: A welcome message and a search box with a 'Go' button. Below the search box is a link for 'Advanced Search'.
- Communities in the repository**: A section for selecting a community to browse, with a link to 'Centre For Digital Music'.
- Search the repository**: A search input field with a 'Go' button.
- Recently Added**: A list of recently added datasets, including 'The TRIOS Score-aligned Multitrack Recordings Dataset' and 'High Quality Musical Audio Source Separation'.
- Search DSpace**: A search box with a 'Go' button and a link for 'Advanced Search'.
- Browse**: A list of browse options including 'All of DSpace', 'Communities & Collections', 'Authors', 'Titles', 'Keywords', and 'By Issue Date'.
- My Account**: Links for 'Login' and 'Register'.
- Discover**: A list of discovered items, including 'Author' (Fritsch, Joachim (3), Ganseman, Joachim (1), McPherson, Andrew (1), Plumbley, Mark D. (1)) and 'Keyword' (architectural acoustics (3), reverberation (3), Room impulse response (3), Automatic music transcription (2), NMF (2), Note tracking (2)).

<http://c4dm.eecs.qmul.ac.uk/rdr>

C4DM Research Data Repository



The screenshot shows a web page with a blue header containing the text 'centre for digital music'. Below the header is a breadcrumb trail: 'C4DM-RDR Home → Centre For Digital Music → Circuit Designs → View Item'. The main title is 'Techniques and Circuits for Electromagnetic Instrument Actuation' by 'McPherson, Andrew'. The URI is 'http://c4dm.eecs.qmul.ac.uk/rdr/handle/123456789/25' and the date is '2012-05-28'. An abstract section follows, describing the item's content. Below the abstract is a section for citations, and at the bottom is a 'Show full item record' link.

centre for digital music

C4DM-RDR Home → Centre For Digital Music → Circuit Designs → View Item

Techniques and Circuits for Electromagnetic Instrument Actuation

McPherson, Andrew

URI: <http://c4dm.eecs.qmul.ac.uk/rdr/handle/123456789/25>
Date: 2012-05-28

Abstract:

This item includes schematics, board layouts and assembly data for transconductance amplifiers designed for driving electromagnetic musical instrument actuators. The files are in EAGLE and Gerber RS-274X formats and are suitable for fabrication by a PCB manufacturer. 2-channel and 12-channel versions of the amplifier design are included.

For citations, please use this reference:

A. McPherson (2012). Techniques and Circuits for Electromagnetic Instrument Actuation. In Proceedings of the 12th International Conference on New Interfaces for Musical Expression.

[Show full item record](#)

<http://c4dm.eecs.qmul.ac.uk/rdr>

C4DM Research Data Repository

For citations, please use this reference:

A. McPherson (2012). Techniques and Circuits for Electromagnetic Instrument Actuation. In Proceedings of the 12th International Conference on New Interfaces for Musical Expression.

[Show full item record](#)

[Export the entire dataset as a single package](#)

Files in this item

- +-- [mrp-amp-12channel.zip](#) [12-channel amplifier circuit designs, 396.7Kb]
- +-- [mrp-amp-2channel.zip](#) [2-channel amplifier circuit designs, 141.7Kb]

The following license files are associated with this item:

- [Creative Commons](#)

This item appears in the following Collection(s)

- [Circuit Designs](#)
Schematics, board layouts, assembly data for electronic circuits



Except where otherwise noted, this item's license is described as Attribution 2.0 UK: England & Wales

<http://c4dm.eecs.qmul.ac.uk/rdr>

C4DM Research Data Repository

The image shows a Google Scholar search interface. The search bar contains 'mirex multif0' and the search button is highlighted. Below the search bar, it shows 'Scholar' and '46 results (0.11 sec)'. On the right, there are buttons for 'My Citations' and a notification icon with '0'. The search results are listed under the 'Articles' tab. The first result is 'Evaluation of multiple-F0 estimation and tracking systems' by M Bay, AF Ehmann, JS Downie, published in Proc. of ISMIR, 2009. The second result is 'Multiple-instrument polyphonic music transcription using a convolutive probabilistic model' by E Benetos, S Dixon, published in the 8th Sound and Music Computing Conference, 2011. The third result, highlighted with a blue border, is 'MIREX MultiF0 Development Set' by E Benetos, G Grindlay, published in 2012. The fourth result is 'MULTIPLE-F0 ESTIMATION AND NOTE TRACKING FOR MIREX 2012 USING A SHIFT-INVARIANT LATENT VARIABLE MODEL' by E Benetos, S Dixon, published in 2012. On the left side, there are filters for 'Legal documents', 'Any time' (with sub-options for 'Since 2012', 'Since 2011', 'Since 2008', and 'Custom range...'), 'Sort by relevance', 'Sort by date', and checkboxes for 'include patents' and 'include citations'. There is also a 'Create alert' button.

Google mirex multif0

Scholar 46 results (0.11 sec) My Citations 0

Articles Did you mean: mirex *multiflow*

Legal documents [PDF] Evaluation of multiple-F0 estimation and tracking systems ismir.net [PDF]
M Bay, AF Ehmann, JS Downie - Proc. of ISMIR, 2009 - ismir2009.ismir.net
... 3. RESULTS AND DISCUSSION The evaluation results of two iterations of the **MIREX multi-F0** estimation task (2007-2008) are presented here. ... Multiple F0 Estimation in Polyphonic Music, Available at http://www.music-ir.org/mirex/2008/abs/mirex08_multif0_Cao.pdf. ... Cited by 19 Related articles View as HTML All 3 versions Import into EndNote

Any time Since 2012 Since 2011 Since 2008 Custom range...

Sort by relevance Sort by date

include patents include citations

Create alert

[PDF] Multiple-instrument polyphonic music transcription using a convolutive probabilistic model qmul.ac.uk [PDF]
E Benetos, S Dixon - 8th Sound and Music Computing Conference, 2011 - eecs.qmul.ac.uk
... Finally, a hidden Markov model-based note tracking method is employed in order to provide a smooth piano-roll transcription. The system was tested on recordings from the RWC database [9], the Disklavier dataset in [2], as well as the **MIREX multi-F0** woodwind quintet [10]. ... Cited by 10 Related articles View as HTML All 4 versions Import into EndNote

MIREX MultiF0 Development Set
E Benetos, G Grindlay - 2012 - c4dm.eecs.qmul.ac.uk
... Description: This multi-track recording is used as a development set for the **MIREX multi-F0** and note tracking tasks. ... The recordings and annotations can also be accessed from <http://www.music-ir.org/evaluation/MIREX/data/2007/multif0/index.htm> (login required). ... Cached Import into EndNote

[PDF] MULTIPLE-F0 ESTIMATION AND NOTE TRACKING FOR MIREX 2012 USING A SHIFT-INVARIANT LATENT VARIABLE MODEL music-ir.org [PDF]
E Benetos, S Dixon - 2012 - music-ir.org
... The resulting piano-roll transcription matrix is given by: $P(p, t) = P(t)P(p|t)$ (8) In Fig. 2, the transcription matrix $P(p, t)$ for an excerpt of the **MIREX multi-F0** woodwind quintet recording can be seen, along with the corresponding pitch ground truth. ... Related articles View as HTML All 2 versions Import into EndNote

EPSRC Principles

- » The UK Engineering and Physical Sciences Research Council (EPSRC) states:
 - Data should be freely available with as few restrictions as possible
 - Data should remain accessible and usable for future research (10 years after last use!)
 - Metadata should be available to enable reuse
 - Results should say how to access the data
 - Users should acknowledge the sources of their data
 - Data management policies and plans should exist

<http://tinyurl.com/993p6v6>

Four Steps To RDM

Four steps that any researcher can take towards managing data:

- Preserve
- Document
- Organise
- Publish

Developing online materials

- » E-Learning materials
 - Specific tools
 - May be difficult to move between systems
- » Traditional web-site
 - Static
 - Content management
- » Wiki
 - Community involvement
 - Content easily updated and evolves
 - Well known amongst CS audiences

SoDaMaT Wiki

- » https://code.soundsoftware.ac.uk/projects/sodamat/wiki/SoDaMaT_Wiki
- » <http://tinyurl.com/bmxlu9o>
- » Currently project based
 - Centres on SoDaMaT project with sections for work packages
 - Includes review of existing data
 - Section for developing the online materials

Developing online materials

- » E-Learning materials
 - Specific tools
 - May be difficult to move between systems
- » Traditional web-site
 - Static
 - Content management
- » Wiki
 - Community involvement
 - Content easily updated and evolves
 - Well known amongst CS audiences

Coming up...

- » Wiki updates
 - Extending content further
 - Making content more user-friendly
- » Results of BOS evaluations
 - And update of materials
- » Longer tutorials targeted at:
 - Postgraduate researchers
 - Academics, Pls, supervisors