



## Informatica Helps EMI Music Publishing Improve Copyright Compliance

### CHALLENGE

- Search and match huge database of songs titles and musical artists names to comply with copyright laws

### SOLUTION

- Informatica's identity resolution technology

### RESULTS

- Accurately identified copyright compliance and infringement issues
- Contributed to fairer distribution of royalties and increased revenue
- Increased operational efficiency in reducing time-consuming manual matches

“After consulting with specialists from Informatica, EMI experienced a 16 to 20 percent improvement in automatic match quality over its existing methods. To say the least, we were very pleased. We can match files from infringing companies and claim our rights more accurately and faster, and this results in increased revenue.”

— Alec Malyon, IT Director, EMI

### Music and Technology Pioneer Looks to Identity Resolution to Comply with Copyright Requirements

Music publishers are in the business of acquiring and marketing the rights to musical compositions by entering into agreements with composers and writers for the use of their copyrighted words and music. Their business is to ensure that the songs are heard and enjoyed, typically by licensing them for performance, broadcast and inclusion in television programs, advertisements and motion pictures. They also make certain that all royalties due on such licensing agreements are collected and distributed.

In terms of copyrights owned, controlled, or administered, EMI Music Publishing, with offices in 30 countries and rights to more than one million musical compositions, is the world's largest music publisher.

EMI has long been a pioneer in using technology to advance its business. It was the first major music publisher to establish a Web site, the first to use an advanced lyric search engine, and the first to process licensing applications online.

EMI relies on the robust identity data search-and-match capabilities of Informatica's identity resolution technology (formerly known as Data Clustering Engine)\* to identify copyright compliance and infringement issues.

00000117	emi	0	AFTER THE LOVE HAS GONE	FOSTER, DAVID/GRAYDON, JAY/CHA
txnx	100		After The Love Has Gone	J. Graydon D. Foster B. Champlin
00000205	emi	0	AL OTRO LADO DEL SOL (THE OTHER SIDE OF THE SUN)	IAN, JANIS/HAMMOND, ALBERT/GOM
txnx	92		The Other Side Of The Sun	Albert Hammond Janis Ian
00000459	emi	0	FRIENDS & CO.	SAFKA, MELANIE
txnx	92		Friends And Company	Melanie
00003011	emi	0	AFRO	SPENCE, ALEXANDER LEE
txnx	97		Grey - Afro	Alexander Lee Spence
00003026	emi	0	AFRICANO	WHITE, MAURICE/DUNN, LORENZO R
txnx	97		Africano - Power	M. White L. Dunn
00003547	emi	0	ARMAGIDEON TIME N.O.P. AS OF 10/13/84	J. MITTIO, W. WILLIAMS
txnx	92		Armagideon Time	
00003715	emi	0	BABY CAN I HOLD YOU	CHAPMAN, TRACY
txnx	95		Baby Can I Hold You	Tracy Chapman
00003747	emi	0	BABY GRAND	JOEL, BILLY
txnx	92		Baby Grand - Duet With Ray Charles	B. Joel
00003883	emi	0	BABY, I'M YOURS	MCCOY, VAN
txnx	95		Baby I'm Yours	Van McCoy
00004166	emi	0	BACK IN THE U.S.S.R.	LENNON, JOHN/MC CARTNEY, PAUL
txnx	92		Back In The USSR	John Lennon Paul McCartney
00009172	emi	0	I COULDN'T BE ME WITHOUT YOU	BILLY JOE SHAVER JOHNNY RODRIGUES
txnx	92		I Couldn't Be Me Without You	Johnny Rodriguez
00010188	emi	0	DUDE (LOOKS LIKE A LADY)	CHILD, DESMOND/TYLER, STEVEN/P
txnx	92		Dude - Looks Like A Lady	Tyler
00010358	emi	0	DEVIL GAME	HOPE, DAVID C/WALSH, STEVE
txnx	92		The Devil Game	S Walsh D Hope
00011259	emi	0	DAYBREAK	NILSSON, HARRY
txnx	100		It Is He Who Will Be King - Outro - Daybreak	
00014250	emi	0	FAME (NEW ARRANGEMENT)	LENNON, JOHN/BOWIE, DAVID/ALOM
txnx	92		Fame	John Lennon Carlos Alomar
00031683	emi	0	OLD FIVE AND DIMERS LIKE ME	BILLY JOE SHAVER
txnx	92		Old 5 And Dimers Like Me	Billy Joe Shaver

EMI uses Informatica's identity resolution technology to improve the reliability of searching and matching song titles and musical artist names to its database and to reduce the number of required manual matches.

## Massive Music Database Must Be Searched for Copyright Compliance

EMI performs data matching to determine copyrights. For example, when a performing society sends it a batch of performances or a retail organization sends EMI information about CDs, tapes, or DVDs sold for which they are not sure who holds the copyright, they must be matched against EMI's copyright database to determine if EMI is the custodian of a particular song and, if so, to whom the royalties for that performance should be distributed.

Another important function of data matching is to check for copyright compliance and infringement. There are many Web sites that offer songs for downloading and, because of recent legislation, they must now ask the music publishers for permission to use this copyrighted material. Accurate data matching means a fairer distribution of royalties and increased revenue.

Because there is no reliable universal song numbering system, EMI must use a combination of strategies to match song titles and writer names to its database. Inevitably, there are always some unmatched songs at the end of the process that must be matched manually (e.g. the song "LOVE" by the writer "SMITH"). The greater the number of songs that need to be manually matched, the greater the drain on productivity.

## Robust Identity Resolution Capabilities Prove Essential

The nature and variability of EMI's data and the potential return on investment EMI could achieve, especially in the copyright infringement area, demanded a robust identity resolution solution.

Seeking to improve the reliability of its data matching and reduce the number of required manual matches, EMI turned to Informatica's identity resolution technology. EMI was

particularly interested in the technology's ability to perform batch data matching and clustering.

Informatica's identity resolution technology can operate in a number of modes, including grouping the data in one or many files, screening an external transaction file against internal reference data, and looking for the "non-matches" between files—an important process in compliance work. The technology runs efficiently on Unix and Windows platforms and can source data from either flat files or directly from relational databases.

## Accurate Music Searches Increase Efficiency, Enforce Copyrights, and Increase Revenue

"Processing new files is low effort," says EMI IT Director Alec Maylon. "When we receive third-party files we convert to a standard format before sending it to the server (where clustering is performed), and thus need minimal changes to the rules for this job. The actual runs are well within our processing time thresholds."

"To say the least, we were very pleased. We can match files from infringing companies and claim our rights more accurately and faster, and this results in increased revenue."

### LEARN MORE

For more information on EMI Music Publishing, visit [www.emimusicpub.com](http://www.emimusicpub.com). Learn more about Informatica's identity resolution technology and the entire Informatica product platform. Visit us at [www.informatica.com](http://www.informatica.com) or call 800.653.3871.

**INFORMATICA**<sup>®</sup>  
The Data Integration Company™

Worldwide Headquarters, 100 Cardinal Way, Redwood City, CA 94063, USA  
phone: 650.385.5000 fax: 650.385.5500 toll-free in the US: 1.800.653.3871 [www.informatica.com](http://www.informatica.com)

Informatica Offices Around The Globe: Australia · Belgium · Canada · China · France · Germany · Ireland · Japan · Korea · the Netherlands · Singapore · Switzerland · United Kingdom · USA

© 2008 Informatica Corporation. All rights reserved. Printed in the U.S.A. Informatica, the Informatica logo, and Informatica's Identity Resolution Technology are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

First Published: 2008

6875 (07/30/2008)