**Matthew Woolhouse**
**McMaster University, Hamilton, ON**

**Dan Tidhar**
**King's College, London, UK**

# Querying 1.8 million music downloads from 13 countries

**Abstract:** Findings from an international study are reported in which music download data supplied by the Nokia Corporation were analysed from sociocultural perspectives. Novel methods capturing musical diversity and adventurousness were employed for each country in the study. Current research involving download data and Music Information Retrieval are also discussed.

**Résumé:** Les résultats d'une étude internationale seront discutés. Les données relatives aux téléchargements de musique fournies par Nokia Corporation sont analysées selon une perspective socioculturelle. De nouvelles méthodes de capture de la diversité musicale et du degré d'aventure ont été utilisées pour chaque pays de l'étude. Seront également discutés la recherche actuelle sur le téléchargement de musique et le repérage de l'information musicale.

## 1. Introduction and background

This paper reports the core findings of a study conducted in 2010 when the first named author was the consultant psychologist for the Nokia Corporation's global *Psychology of Music Campaign*, a research programme designed to evaluate current global music downloading trends. The data in the study were generated by Nokia's online music store system, which allows users to purchase and download music directly onto a mobile device or a home computer (Molteni & Ordanini, 2003; Lindholm & Keinonen, 2003). In sum, the metadata of 1.8 million music-tracks, downloaded from 13 geographically distinct and culturally diverse countries, were analysed from a sociological and musicological perspective. In addition, the paper also outlines current research which capitalizes on the availability of the download data as well as the audio itself, and involves music information retrieval (MIR) techniques (Downie, 2003).

## 2. Aims

Among the research aims of the study conducted in 2010 were to explore sociocultural factors such as the level of musical diversity and individual adventurousness within each country, and the level of domestic versus international music consumed. In addition, countries' musical genre preferences were correlated with music preference-personality measures (North, 2010; Rentfrow & Gosling, 2003) in order to derive a profile of national "characteristics" for each country. In the current MIR-related study we extract audio features and apply machine learning techniques with the goal of predicting download patterns based on musical properties. We enrich our analysis of the download patterns by using extracted audio features in conjunction with metadata, an approach which has proved useful in the past (e.g. Pampalk et al., 2004).

**3. Methods**
In the first study, music download metadata such as 'store country', 'user name', 'device type', 'date', 'time', 'track catalogue number', 'artist' and so on, were organized into a series of relational databases and searched using the *MySQL* implementation of *SQL* (Structured Query Language; e.g. Groff, et al. 2009). Data generated by these searches were analysed using statistical software and graphed accordingly. As there are no standardized tests for musical diversity and individual adventurousness for countries, heuristic, novel statistical methods were devised to explore these phenomena. For example, a country's musical diversity was calculated using the standard deviations of the download percentages of its top-10 genres. The rational for this approach is that high standard deviation scores produced by countries with highly skewed genre data indicate relative musical uniformity or conservativeness (they tend to like only one or two types of music, while ignoring others); low standard deviation scores produced by countries which download different genres equally indicate relative musical diversity or plurality (they tend to like many types of music equally). Individual adventurousness was examined by calculating the average number of musical genres downloaded by the individuals of each country. These data were then compared to each country's musical diversity score to infer levels of social cohesion. An example of this process is outlined in the *Results* section below.

Subsequently, a more in-depth survey of musical adventurousness calculated the standard deviation of each individual's genre distribution, and the median SD observed within each country. The lower the median SD for a particular country, the more evenly distributed were downloads across multiple genres, and thus the level of musical plurality or adventurousness was greater. This was implemented using the Perl language, accessing the data through the MySQL Perl API.

Over the past decade a number of studies have sought to link musical preference with personality (for overviews see North & Hargreaves, 2008; Clark et al. 2010). While these studies are not entirely consistent, a few broad trends have emerged to suggest that it is possible, in theory at least, to deduce something about a nation's "character" based on its musical preferences. Using tables of musical genre by personality dimension, the national character of each country in the study was calculated in relation to the following: self-esteem, creativity, attitude to work, outgoingness, gentleness, and how "at ease" a country is with itself.

**4. Results**
The research revealed numerous intriguing differences and similarities between countries from the Americas, Europe, the Middle East and Asia. For example, individuals within Germany showed the greatest levels of musical adventurousness; that is, someone within Germany was more likely to download different types of music than, for example, their counterparts in Italy or Sweden. The United Arab Emirates (UAE) had the greatest musical diversity (i.e. widest musical tastes) in the study, closely followed by the United Kingdom. This indicates that musical genres were downloaded more equally in these countries than in other countries. The least musically diverse were

Turkey and Malaysia. Both these countries' top-10 genre preferences were massively dominated by pop music; 47% in the case of Malaysia, and 42% in the case of Turkey. As a consequence, in these two countries other genres were downloaded relatively infrequently.

On the face of it one might expect a country's musical diversity to match its individuals' musical adventurousness; however, this is not necessarily the case. For example, although the UAE had the highest diversity score in the study, the musical adventurousness of its inhabitants ranked only seventh. This might imply that individuals within the UAE are musically quite conservative (seventh in musical adventurousness), but that these individuals are highly different from one another (as indicated by the UAE's high musical diversity score). This mismatch, between on the one hand individual conservatism and on the other hand high diversity at the national level, may indicate that the UAE suffers from a lack cohesion in certain ways. This is in contrast to Malaysia, where the low individual musical adventurousness score matched the country's level of musical diversity; i.e. there are lots of musically conservative individuals (with low individual adventurousness scores), but happily their musical tastes are all rather similar (as indicated by Malaysia's low diversity score).

The genre preferences and national "characteristics" of each country in the study are too numerous to report here. Suffice to say, based on musical preference, Brazilians appear to be at ease, Finns, Germans and Swedes are creative, the British, Italians, Russians and South Africans are outgoing, Malays and Singaporeans brim with self esteem, while Mexicans, Turks and Emiratis work hard!

Finally, and on perhaps a more serious note, the study also examined the proportion of downloads for each country that were domestic versus international. Turkey was the strongest consumer of its own music; in fact, downloads of music by Turkish artists in Turkey accounted for well over half of all downloads (59%). Italy was the second strongest consumer of domestic music, followed by the UK, Finland and Mexico. Singapore was the country with the most international tastes in music; here, domestically produced music accounted for only 1% of all its downloads. South Africa, Malaysia and the UAE were also major musical importers. The extent to which these data correlate with foreign language acquisition, or with other social and cultural metrics, has yet to be investigated.

**5. Concluding remarks**
The contribution of these data to our understanding of the unfolding story of music's globalization is potentially very significant. To date, multinational communications companies have been reluctant to allow their commercially generated data to be scrutinized with the intention of illuminating the ways in which music is currently being used and consumed on a global scale. The authors are therefore heavily indebted to the Nokia Corporation for their generosity in providing unfettered access to the data reported in this study, and for continued access in support of further research in areas such as MIR. These data are a unique snapshot of current global music listening practices and give an insightful view of the rapidly changing patterns in music consumption presently sweeping the globe.

**References**

Clarke, E.F., Dibben, N.J. & Pitts, S.E. 2010. *Music and Mind in Everyday Life*. Oxford: Oxford University Press.

Downie, J. S. 2003. Music information retrieval. *Annual Review of Information Science and Technology,* 37: 295–340.

Molteni, L. & Ordanini, A. 2003. Consumption Patterns, Digital Technology and Music Downloading. *Long Range Planning,* 36(4): 389-406.

North, A.C. 2010. Individual Differences in Musical Taste. *American Journal of Psychology,* 123(2): 199-208 .

North, A.C. & Hargreaves, D.J. 2008. *The Social and Applied Psychology of Music*. Oxford: Oxford University Press.

Rentfrow, P.J. & Gosling, S.D. 2003. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology,* 84(6): 1236-1256.

Lindholm, C. & Keinonen, T. 2003. *Mobile Usability: How Nokia Changed the Face of the Cellular Phone.* McGraw-Hill, Inc. New York.

Groff, J.R., Weinberg, P.R., & Oppel, A.J. 2009. *SQL, The Complete Reference (3rd ed.).* McGraw-Hill.

Pampalk, E., Dixon, S. & Widmer, G. 2004. Exploring Music Collections by Browsing Different Views. *Computer Music Journal,* 28(2): 49-62.