Final report of CATCH project WITCHCRAFT

1 General information

1.1 Title research project

Title: WITCHCRAFT - What Is Topical in Cultural Heritage: Content-based Retrieval Among Folksong Tunes

Projectnumber: 640.003.501

1.2 Project leaders (university/cultural heritage)

Name: Dr. F. Wiering
Affiliation: Universiteit Utrecht

Name co-project leader(s): Prof. dr. L.P. Grijp
Affiliation: Meertens Instituut

Name co-project leader(s): Prof. dr. R.C. Veltkamp
Affiliation(s): Universiteit Utrecht

1.3 Project participants

Name: Dr. J. Garbers
Affiliation: Universiteit Utrecht
Position: Wetenschappelijk progammeur
Start date: 01-05-2006
End date: 01-05-2010

Name: Drs. ir. P. van Kranenburg
Affiliation: Universiteit Utrecht
Position: Aio
Start date: 01-05-2006
End date: 01-05-2010

Name: Dr. A. Volk-Fleischer
Affiliation: Universiteit Utrecht
Position: Postdoc
Start date: 01-05-2006
End date: 01-02-2010

2 Research plan

2.1 Aims and objectives of the research project

(as described in the original grant application)

The WITCHCRAFT project sets as its objective to develop a fully-functional content-based retrieval system for folksong melodies stored as audio and/or notation, building on the best practices of Music Information Retrieval (MIR) research. More precisely, the project aim is the design, implementation and evaluation of a melody search engine that:

- is capable of handling large amounts of audio and notation data;
- matches relevant low-level *and* high-level musical features using similarity measure(s) that are based on music cognition and perception and reflect the musical characteristics of the folksong repertoire, including oral transmission;
- orders search output (melodies or melody fragments) by musical similarity;
- is usable for both specialists and the general public.

This system's potential will be demonstrated by **integrating** it in the *Nederlandse Liederenbank* of the Meertens Instituut, and adding a selection of early 20th century entertainment/ popular songs from the collection of the Theater Instituut Nederland to test the strengths of the melody search engine in cross-repertoire research.

2.2 Changes with respect to the original research plan

(if applicable)

After initial experiments with audio transcription on the folksong data it appeared that the quality of the output was too low to be used directly in the melody search engine. Therefore we decided to concentrate on searching encoded music notation, and to link audio files to the encodings.

Tools for large-scale data entry and management were created to support the parallel Senter Novem-funded project, Dutch Songs as Musical Content, which in turn provided an unexpectedly large corpus of encoded melodies to WITCHCRAFT. As a consequence, it was unnecessary to incorporate materials from the collection of the Theater Instituut Nederland, which moreover proved to be problematic for organisational reasons.

3 Results

3.1 Scientific results and research highlights (max. 750 words)

Include, in non-scientific language, the most important research highlights of the project.

The WITCHCRAFT project brought together specialists from two areas and two institutions: music information retrieval (MIR), represented by Utrecht University's Department of Information and Computing Sciences, and folk song research, represented by the DOC Lied of the Meertens Institute. The project's objective, the creation of a fully functional content-based retrieval system for folksong melodies, has benefited greatly from the availability of two important resources at the Meertens Institute. The first, *Onder de Groene Linde*, is a collection of more than 7000 field recordings of Dutch folk songs, mostly ballads. Additionally, there are c. 5000 hand-written expert transcriptions of these songs in music notation and very rich metadata. These metadata are stored in a much larger second resource, the *Nederlandse Liederenbank* (Dutch Song Database), containing information about 125.000 songs in the Dutch language. Parallel to the WITCHCRAFT project, the Liederenbank was made available online (http://www.liederenbank.nl) and the folksong materials were integrated into it. At the end of the WITCHCRAFT project, the melody search engine was integrated into the Liederenbank as well (http://www.liederenbank.nl/index.php?wc=true).

As this search engine operates on encoded music notation, a workflow for the production of such encodings was set up. Data entry is done through the WITCHCRAFT editor, which outputs the melody in several formats (Humdrum, MIDI and Lilypond) that are suitable for display, playback and analysis. Initially, only a small test corpus was encoded, but at the end of the project around 5800 encoded melodies had become available through various means.

An important concept in folk song research is the 'tune family', a group of melodies with a presumed common historical origin. An intrinsic difficulty in applying this concept to folk songs is that there is very often no documentary evidence to reason from, so in practice melody norm assignment is based on the assessment of musical (and textual) *similarity*. Tune families can be used as a ground truth for the evaluation of retrieval methods as follows. If one melody is taken from the tune family and used as a query, the most successful retrieval method is the one that gives, in a Google-like list of ranked output, the highest ranks to the other family members. The scientific task then becomes to create the most suitable retrieval method for musical similarity.

Tune family assignation is a complex task performed by experts. In order to understand this task better, an annotation method was developed in close collaboration with the domain experts, by which they could record the features on which they based their decisions, and to what extent. Such annotations were created for 360 melodies and have been used in the design a number of alignment methods. The Annotated Corpus is available to other researchers as well.

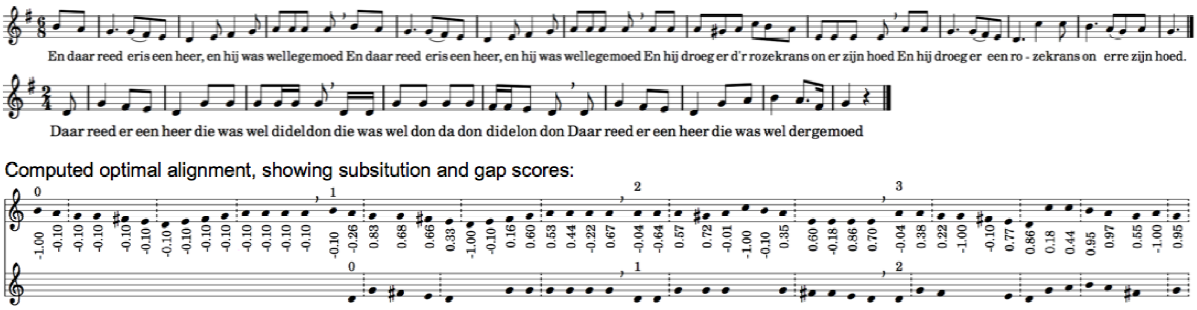
Various approaches were designed and tested in the creation of a suitable similarity measure. One that did not work was classification using global features (for example the frequency of an interval in a melody). Other methods, such as the Earth Mover's Distance, Graph Matching and Inner Metric Analysis (IMA), met with only limited success.

Several promising new avenues were explored. One is 'group querying', a method for relevance feedback in which several related melodies that have already been found are aligned and used to find other members of the same tune family. Another is based on musical audio: after (approximate) transcription and segmentation, the most representative segments for a tune family are identified. Query melodies can be classified by comparison to these segments with reasonable success.

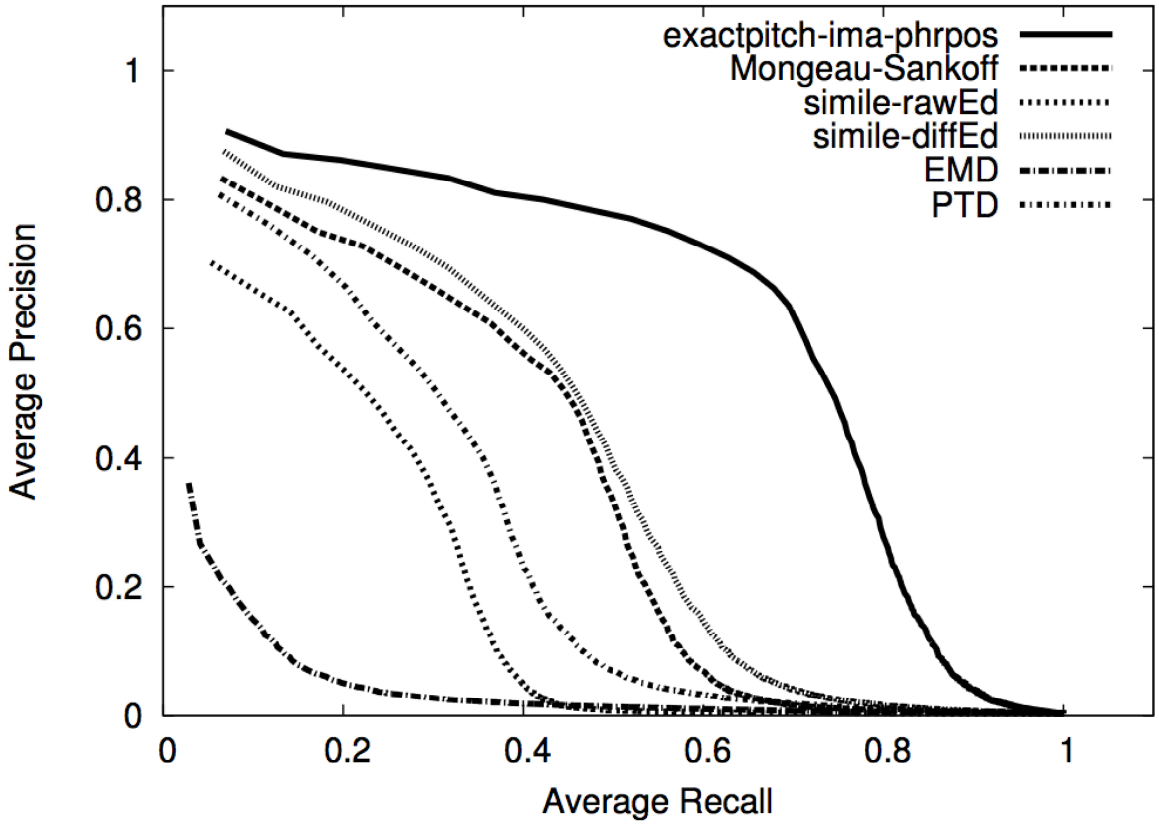
In the end, sequence alignment methods for encoded notation proved to be most successful. Such methods compare two strings of symbols by calculating the cost of inserting, deleting or substituting symbols (here: notes). When applied to folk songs, the aim is to choose these costs in such a way that the more likely a change is to occur within a tune family, the smaller the cost is. The most effective solution that emerged from the research uses a combination of three musical properties, namely pitch, metric weight using IMA, and phrase position. This method was evaluated for musicological usefulness by using it to classify 111 'hard' melodies that so far experts had been unable to assign to a tune family. About one third of these could be classified using the search engine; for one third some relationship to other melodies could be identified. This shows that an important requirement for interdisciplinary collaboration has been fulfilled in this project, namely that results attained in one domain are indeed meaningful in the other domain.

3.2 Pictures of highlights

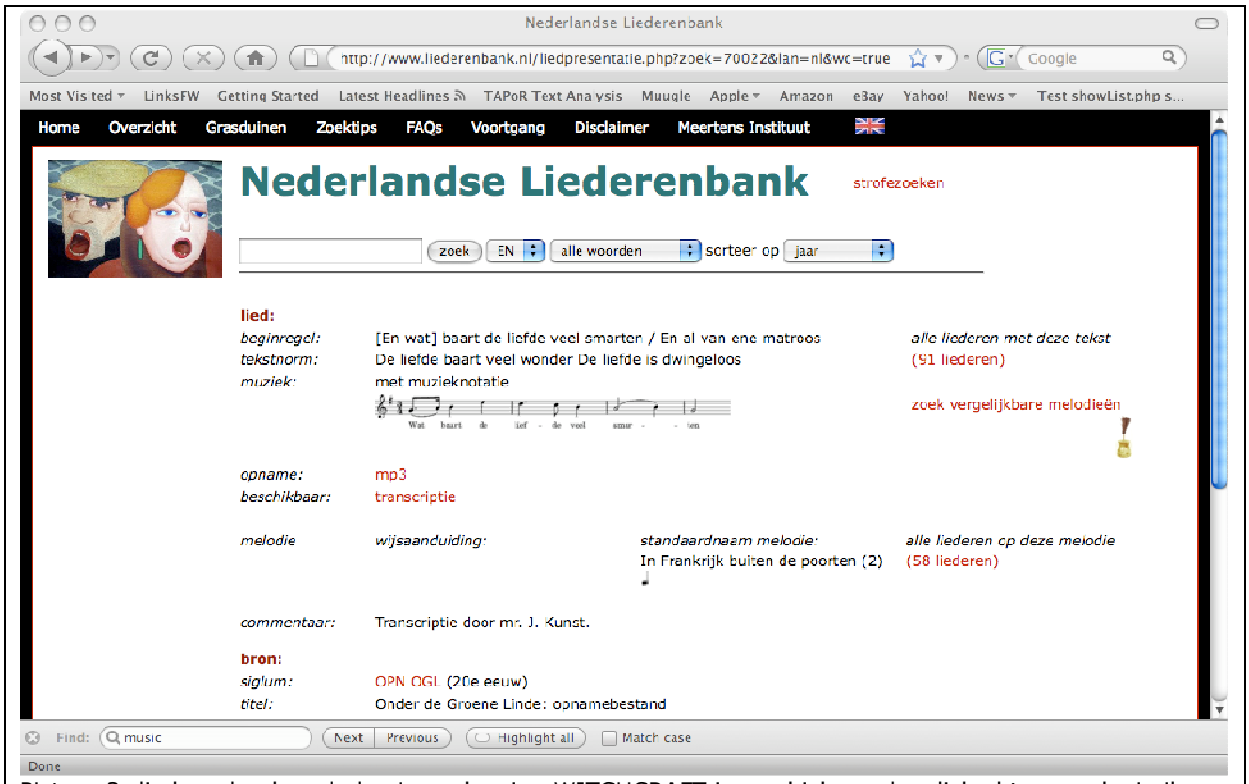
Attach at least two representative pictures of your research highlights and give a short description of these pictures.



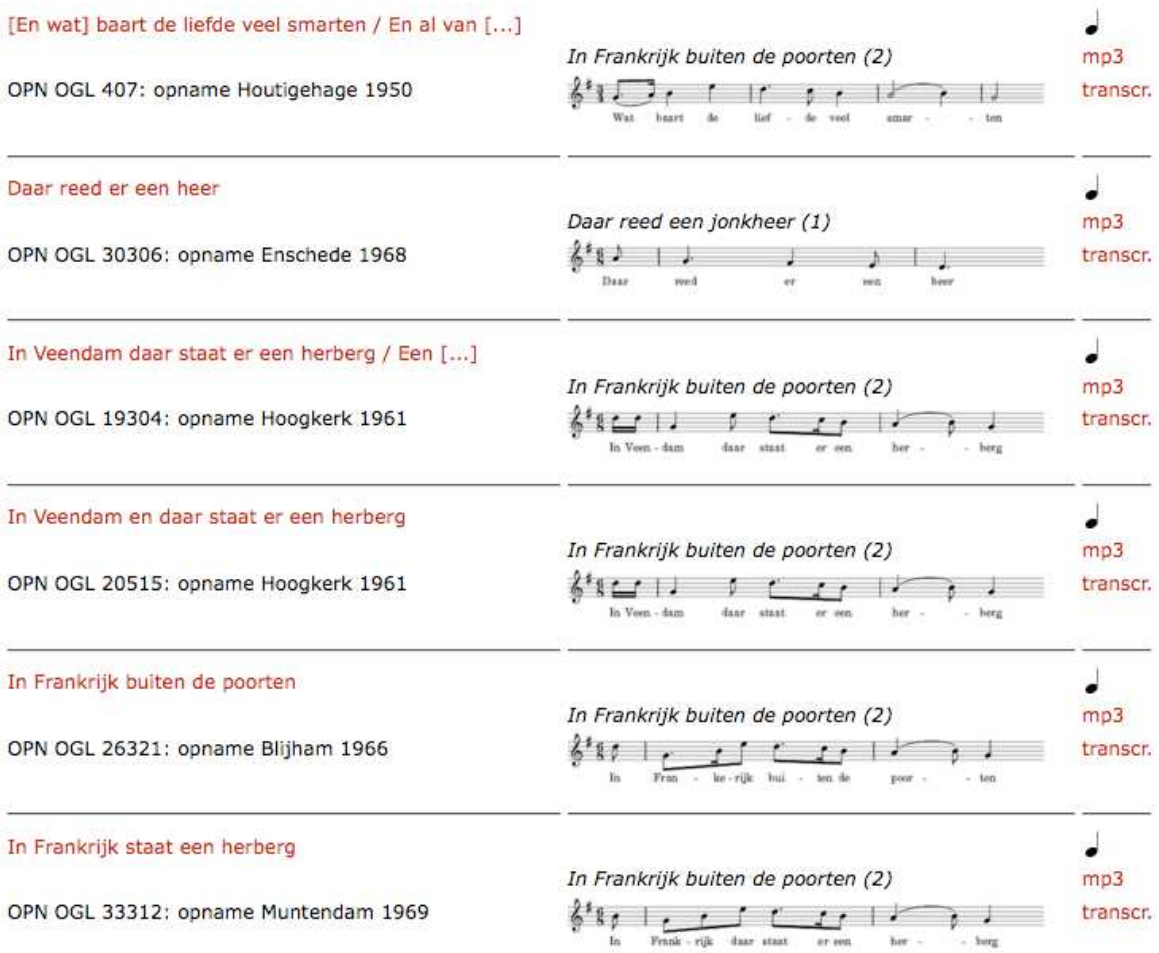
Picture 1: example alignment of two melodies



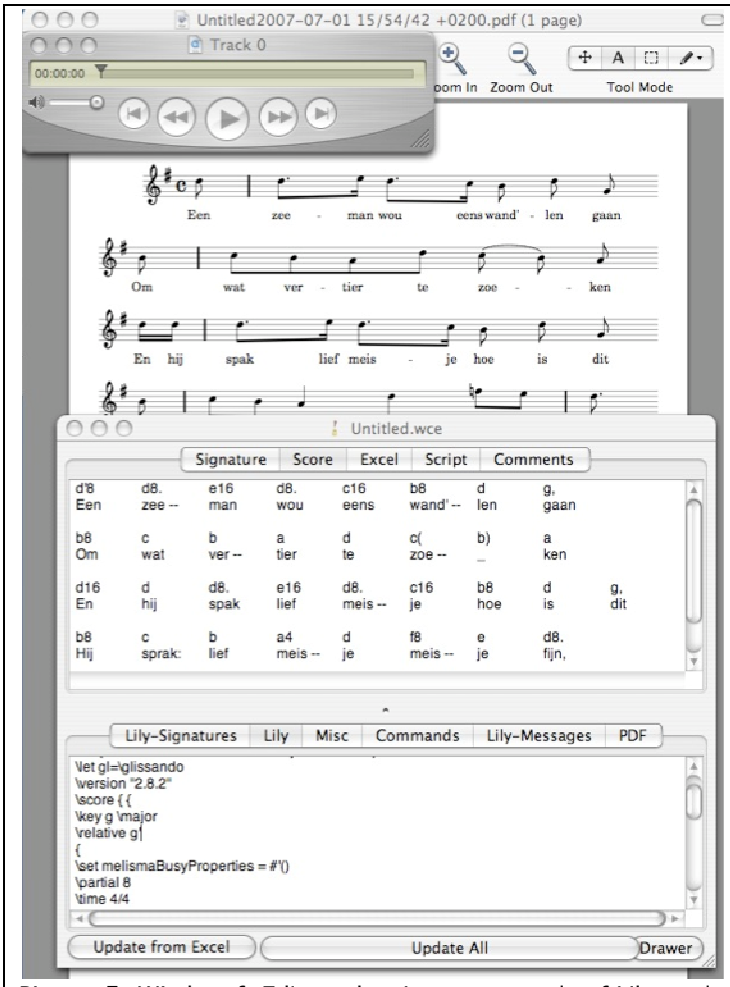
Picture 2: precision-recall graph comparing different melody retrieval methods. The best performing one, exactpitch-ima-phrpos, is the one selected for the Liederenbank.



Picture 3: liederenbank melody view, showing WITCHCRAFT icon which can be clicked to search similar melodies.



Picture 4: result list, after clicking WITCHCRAFT icon



Picture 5: Witchcraft Editor, showing an example of Lilypond encoding just below the middle.

3.3 (Potential) applicability of research results

Mention (potential) areas of application of the results of your research project.

The collaboration model designed for this project is not particular to folksong research but can be generalised to any computer science-humanities collaboration. It is aimed at 'deep' interdisciplinary collaboration, beyond a mere exchange of data and tools, in which research results attained in one disciplinary context contribute to insight in the other discipline.

The data entry workflow tools and melody retrieval methods designed in WITCHCRAFT can be applied to other collections of musical heritage as well. New folk song projects may benefit from both; but existing projects usually focus on digitisation and encoding, and may be particularly interested in content-based retrieval methods. In historical musicology, several source databases and online scholarly edition projects exist to which these methods are applicable as well (e.g. RISM, CMME).

In the area of Music Information Retrieval, this research is part of a trend to move away from low-level features to ones that are closer to the listener's experience of music. This trend is now also reaching the music industry. Popular applications such as Shazam work with identification methods and metadata only, and are consequently still unable to deal with similar but non-identical music. This is especially true for music that is in the 'long tail': if the industry wishes to deal satisfactory with this materials, content-based retrieval methods similar to the ones designed in the WITCHCRAFT project will be needed.

3.4 Publications and other output

List the publications, theses, conference abstracts, invited lectures and patents (applications) below.

Independent publications

Peter van Kranenburg. *A Computational Approach to Content-Based Retrieval of Folk Song Melodies.* PhD thesis, Utrecht University, Department of Information and Computer Sciences, October 2010. ISBN 978-90-393-5393-6.

J. Stephen Downie, Remco C. Veltkamp (eds.). *Proceedings of the 11th International Society for Music Information Retrieval Conference* (ISMIR2010). ISBN 978-90-393-53813

Frans Wiering (Ed). *Music and the Sciences*. *Interdisciplinary Science Reviews* 35(2), 2010.

E. Selfridge-Field, F. Wiering, G. A. Wiggins (Eds.). *Dagstuhl Seminar Proceedings 09051 Knowledge representation for intelligent music processing*, 2009. ISSN 1862 - 4405.

Peer reviewed articles, conference papers and book chapters

Anja Volk, Computermodellierungen in der Musikforschung – ein interdisziplinäres Feld. In W. Auhagen, V. Busch and J. Hemming (eds), *Kompendium Systematische Musikwissenschaft*, Laaber: Laaber-Verlag. Forthcoming.

Frans Wiering. Music and the Sciences: Introduction. *Interdisciplinary Science Reviews* 35(2), 2010, 103–105.

W. Bas de Haas, Matthias Robine, Pierre Hanna, Remco C. Veltkamp and Frans Wiering. Comparing Harmonic Similarity Measures. *Proceedings of the 7th International Symposium on Computer Music Modeling and Retrieval* (CMMR), Malaga, 21-24 June 2010. 299-315.

Peter van Kranenburg, George Tzanetakis. A Computational Approach to the Modeling and Employment of Cognitive Units of Folk Song Melodies Using Audio Recordings. In: *Proceedings of the 11th International Conference on Music Perception and Cognition* (ICMPC), Seattle, 23-27 August 2010. 794-797

Louis P. Grijp. Selling Songs: Dutch results and Experiments. In: *From Wunderhorn to the Internet: Perspectives on Conceptions of Folk Song and the Editing of Traditional songs*, edited by Eckhard John and Tobias Widmaier. Trier: Wissenschaftlichter Verlag Trier, 2010, 234-245.

P. van Kranenburg, J. Garbers, A. Volk, F. Wiering, L.P. Grijp, R.C. Veltkamp. Collaboration perspectives for folk Song research and music information retrieval: The indispensable role of computational musicology. *Journal of Interdisciplinary Music Studies*. (2009), doi: 10.4407/jims.2009.12.030.

Aline Honingh, Anja Volk. Mathematische muziektheorie: Nieuwe mogelijkheden voor muziekgerelateerd onderzoek. *Dutch Journal of Music Theory* 14(3), 2009, 181-193.

Frans Wiering, Justin de Nooijer, Anja Volk, Hermi J.M. Tabachneck-Schijf. Cognition-based Segmentation for Music Information Retrieval Systems. *Journal of New Music Research* 38(2), 2009, 139–154.

Peter van Kranenburg, Anja Volk, Frans Wiering, Remco C. Veltkamp. Musical Models for Folk-Song Melody Alignment. In: *Proceedings of the International Society on Music Information Retrieval (ISMIR 2009) conference*, 507-512.

W. Bas de Haas, Martin Rohrmeier, Remco C. Veltkamp, Frans Wiering. Modeling Harmonic Similarity Using a Generative Grammar of Tonal Harmony. In: *Proceedings of the International Society on Music Information Retrieval (ISMIR 2009) conference*, 549-554.

Meinard Müller, Peter Grosche, Frans Wiering. Robust Segmentation and Annotation of Folk Song Recordings. In: *Proceedings of the International Society on Music Information Retrieval (ISMIR 2009) conference*, 735-740.

Korinna Bade, Andreas Nürnberger, Sebastian Stober, Jörg Garbers, Frans Wiering. Supporting Folk-Song Research By Automatic Metric Learning and Ranking. In: *Proceedings of the International Society on Music Information Retrieval (ISMIR 2009) conference*, 741-746.

Frans Wiering. Meaningful Music Retrieval. In: *1st Workshop on the Future of Music Information Retrieval, ISMIR 2009*, 1-3. http://www.columbia.edu/~tb2332/fmir/Papers/Wiering-fmir.pdf

Frans Wiering, Remco C. Veltkamp, Jörg Garbers, Anja Volk, Peter van Kranenburg, Louis P. Grijp. Modelling Folksong Melodies. *Interdisciplinary Science Reviews*, 34(2–3), 2009, 154–171.

Garbers, J., Volk, A., Van Kranenburg, P., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). On Using Pitch and Harmony in Folk Song Variation Retrieval. *Mathematics and Computation in Music: First International Conference, MCM 2007, Berlin, Germany, May 18-20, 2007. Revised Selected Papers*. Ed. T. Klouche and T. Noll. Berlin, Heidelberg: Springer, 2009. 97-106.

Volk, A., Garbers, J., Van Kranenburg, P., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). Comparing Computational Approaches to Rhythmic and Melodic Similarity in Folksong Research. *Mathematics and Computation in Music: First International Conference, MCM 2007, Berlin, Germany, May 18-20, 2007. Revised Selected Papers*. Ed. T. Klouche and T. Noll. Berlin, Heidelberg: Springer, 2009. 78-87.

Anja Volk. The Study of Syncopation using Inner Metric Analysis, Linking Theoretical and Experimental Analysis of Metre in Music, *Journal of New Music Research*, Vol 37, No. 4, 2008, 259 – 273.

Frans Wiering. Digital Critical Editions of Music: A Multidimensional Model. In: *Modern Methods for Musicology: Prospects, Proposals and Realities*, edited by Tim Crawford and Lorna Gibson. London, Ashgate, 2009. 23-46

Casey, M., Veltkamp, R.C., Goto, M., Leman, M., Rhodes, C., Slaney, M. (2008). Content-based music information retrieval: current directions and future challenges. Proceedings of the IEEE, 96(4), 2008, 668-696.

Garbers, J., Wiering, F. (2008). Towards structural alignment of folk songs. *Proceedings of the Ninth International Conference on Music Music Information Retrieval.* Ed. J.P. Bello, E. Chew, D. Turnbull. Philadelphia. 381-386.

Haas, W.B. de, Veltkamp, R.C., Wiering, F. (2008) Tonal Pitch Step Distance: a Similarity Measure for Chord Progressions. *Proceedings of the Ninth International Conference on Music Music Information Retrieval.* Ed. J.P. Bello, E. Chew, D. Turnbull. Philadelphia. 51-56.

Kranenburg, P. van. (2008) Assessing Disputed Attributions for Organ Fugues in the J.S. Bach Catalogue. *Computing in Musicology* 15. 120-137. http://www.ccarh.org/publications/cm/15/cm15-07b-vankranenburg.pdf

Nooijer, J. De, Wiering, F., Volk, A., Tabachneck-Schijf, H.J.M. (2008). An experimental comparison of human and automatich segmentation. In: K. Miyazaki, M. Adachi, Y. Hiraga, Y. Nakajima, M. Tsuzaki (eds.), *Proceedings of theTenth International Conference on Music Perception and Cognition (ICMPC)*. CD-ROM. 399-407.

Nooijer, J. De, Wiering, F., Volk, A., Tabachneck-Schijf, H.J.M. (2008). Cognition-based segmentation for music information retrieval systems. In: C. Tsougras, R. Parncutt (eds.), *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08)*, http://web.auth.gr/cim08/.

Volk, A. (2008). Persistence and Change: Local and Global Components of Meter Induction using Inner Metric Analysis. *Journal of Mathematics and Music* 2:2. 99-115.

Volk, A., Chew, E. (2008). Reconsidering the Affinity between Metric and Tonal Structures in Brahms' Capriccio Op. 76, No. 8. *Computing in Musicology* 15. 138-171.

Volk, A., Kranenburg, P. van, Garbers, J., Wiering, F., Veltkamp, R.C., Grijp, L.P. (2008). A manual annotation method for melodic similarity and the study of melody feature sets. *Proceedings of the Ninth International Conference on Music Music Information Retrieval.* Ed. J.P. Bello, E. Chew, D. Turnbull. Philadelphia. 101-106.

Wiering, F. (2007). Can Humans Benefit from Music Information Retrieval. *Adaptive Multimedia Retrieval: User, Context, and Feedback. 4th International Workshop, AMR 2006*. Ed. S. Marchand-Maillet, E. Bruno, A. Nürnberger, M. Detyniecki. LNCS 4398. Springer. 82-94.

Volk, A., Garbers, J., Van Kranenburg, P., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). Applying Rhythmic Similarity Based on Innner Metric Analysis to Folksong Research. *Proceedings of the Eighth International Conference on Music Information Retrieval.* Ed. Simon Dixon, David Bainbridge, Rainer Typke. 293-296.

Van Kranenburg, P., Garbers, J., Volk, A., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). Towards Integration of MIR and Folk Song Research. *Proceedings of the Eighth International Conference on Music Information Retrieval.* Ed. Simon Dixon, David Bainbridge, Rainer Typke. 505-508.

Garbers, J., Van Kranenburg, P., Volk, A., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). Using Pitch Stability among a Group of Aligned Query Melodies to Retrieve Unidentified Variant Melodies. *Proceedings of the Eighth International Conference on Music Information Retrieval.* Ed. Simon Dixon, David Bainbridge, Rainer Typke. 451-456.

Pinto, A., Van Leuken, R., Demirci, M.F., Wiering, F., Veltkamp, R.C. (2007) Indexing Music Collections through Graph Spectra. *Proceedings of the Eighth International Conference on Music Information Retrieval.* Ed. Simon Dixon, David Bainbridge, Rainer Typke. 153-156.

First International Conference of the Society of Mathematics and Computation in Music, Berlin. http://www.mcm2007.info/pdf/fri2b-volk.pdf

Volk, A. (2007) Applying Inner Metric Analysis to 20th century compositions, proceedings of the conference "Mathematics and Computation in Music", Berlin.

Garbers, J. (2006) An Integrated MIR Programming and Testing Environment. *Proceedings of the Seventh International Conference on Music Information Retrieval.* Ed. Roger Dannenberg, Kjell Lemström & Adam Tindale. 332-333

Kranenburg, P. van (2006) Composer attribution by quantifying compositional strategies. *Proceedings of the Seventh International Conference on Music Information Retrieval.* Ed. Roger Dannenberg, Kjell Lemström & Adam Tindale. 375-376

Bosma, M., Wiering, F., Veltkamp, R.C. (2006) Muugle: A Modular Music Information Retrieval Framework. *Proceedings of the Seventh International Conference on Music Information Retrieval.* Ed. Roger Dannenberg, Kjell Lemström & Adam Tindale. 330-331

Bosma, M., Veltkamp, R.C., Wiering, F. (2006). Muugle: A music retrieval experimentation framework. *Proceedings of the 9th International Conference on Music Perception & Cognition.* 1297-1303.

Other articles and conference papers

Frans Wiering, Peter van Kranenburg. Searching melodies using sequence alignment. Vakidioot 09-10(4), 2010. 3-7.

Peter van Kranenburg, Frans Wiering. Een melodieënzoekmachine voor de Nederlandse liederenbank. To be published in *DIXIT*, December 2010.

Meinard Müller, Peter Grosche, Frans Wiering. Towards Automated Processing of Folk Song Recordings. In: *Dagstuhl Seminar Proceedings 09051, Knowledge representation for intelligent music processing.* 2009 ISSN 1862 - 4405.

Peter van Kranenburg. Folk Song Alignment. In: *Dagstuhl Seminar Proceedings 09051, Knowledge representation for intelligent music processing.* 2009 ISSN 1862 - 4405.

Frans Wiering. Towards meaningful music (information) retrieval. In: *Dagstuhl Seminar Proceedings 09051, Knowledge representation for intelligent music processing.* 2009 ISSN 1862 - 4405.

Jörg Garbers. Software frameworks for systematic music processing. In: *Dagstuhl Seminar Proceedings 09051, Knowledge representation for intelligent music processing.* 2009 ISSN 1862 - 4405.
Remco C. Veltkamp, Frans Wiering. Seven years of music retrieval @ UU. In: *Fascination for computing. 25 jaar opleiding informatica.* Universiteit Utrecht, 2009. 161-175
Jörg Garbers. Bridging Music Information Retrieval and Folk Song Research - The Computational Setup of the WITCHCRAFT Project, *International Conference on Acoustics*, Rotterdam 23-26 maart 2009
Wiering, F. (2008). Zoeken naar melodieën. *Informatieprofessional* 2008-5. 31-33.

Technical reports

Aline Honingh, Anja Volk. *De Kracht van Wiskunde in Muziekonderzoek.* ILLC Prepublication series, PP-2009-23.
Volk, A., Kranenburg, P. van, Garbers, J., Wiering, F., Veltkamp, R.C., Grijp, L.P. (2008). *The Study of Melodic Similarity using Manual Annotation and Melody Feature Sets.* Technical Report UU-CS-2008-013.
Volk, A. (2008). *The Generation of Metric Hierarchies using Inner Metric Analysis*. Technical Report UU-CS-2008-006.
Kranenburg, P. van. (2008). *On Measuring Musical Style - The Case of Some Disputed Organ* Fugues *in the J. S. Bach (BWV) Catalogue.* Technical Report UU-CS-2008-026.
Breunese, N. *Passieve Cooperatieve Filtering voor Online Aanbevelingssystemen.* Master's thesis. INF/SCR-2008-008
Van Kranenburg, P., Garbers, J., Volk, A., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). *Towards Integration of MIR and Folk Song Research*. Technical Report UU-CS-2007-016
Nooijer, J. de. (2007). *Cognition-based Segmentation for Music Information Retrieval Systems.* Masters Thesis. INF/SCR-2006-072.
Jörg Garbers, Peter van Kranenburg, Anja Volk, Frans Wiering, Louis Peter Grijp, Remco C. Veltkamp, Martine de Bruin. *The WITCHCRAFT Baseline Measurement And Pilot Project.* Technical Report UU-CS-2006-064

Presentations (selection)

Peter van Kranenburg. On Content-Based Folk Song Melody Retrieval. *International Ballad Conference*, Terschelling, Netherlands, 5 July 2010.
Peter van Kranenburg. Introduction: Doing Computational Musicology. *Perspectives on Computational Musicology.* Amsterdam, 5 October 2010.
Anja Volk. A Moment of Opportunity for Computational Musicology. *Perspectives on Computational Musicology.* Amsterdam, 5 October 2010.
Anja Volk.Wiskundig onderzoek naar de ritmische structuur van muziek (Lezing Alumnidag Radboud Universiteit Nijmegen, Departement Wiskunde). 2 Oktober 2010.
Frans Wiering. Musical Meaning and Music Information Retrieval. *e-Science for Musicology,* Edinburgh, 1-2 juli 2009.
Frans Wiering: Musicology (Re-)mapped: Netherlands. ESF workshop *Musicology (Re-)Mapped***,** Warsaw, 18-21 november 2009.
Anja Volk, Mathematics and the groove in music, *Nationale Wiskunde Dagen*, Noordwijkerhout, 6 februari 2009
Jörg Garbers. The WITCHCRAFT project, *Digital Strategies in Heritage Conference*, Rotterdam, 9-10 december 2009
Grijp, L. Diachronic Research with the Dutch Song Database. Symposium Technical Challenges and developments in 21th century folk music archiving, Instituut voor Muziekwetenschap, Hongaarse Akademie van Wetenschappen, Boedapest, Hongarije, 12-06-2008.
Grijp, L. WITCHCRAFT & Nederlandse Liederenbank. Meeting CATCH Open House, Centraal Museum, Utrecht, 23-05-2008.
Grijp, L. Dutch Song Database & WITCHCRAFT. Meeting EthnoArc, Wissenschaftskolleg, Berlin, Duitsland, 14-03-2008.
Wiering, F. (2007) Querying Incipits. Collaborative Database Meeting (Renaissance Music) CCH, King's, London, 25 January 2007.
Volk, A. (2007): Applying inner metric analysis to rhythm perception, analysis and classification, Music Cognition Reading Group, University of Amsterdam. 26 January 2007.
Wiering, F. (2007) WITCHCRAFT: melody retrieval in Dutch folksongs, and other MIR projects at the ICS. Guest lecture Universitat Pompeu Fabra, Barcelona, 21 February, 2007.
Volk, A. (2007). The Study of Persistence and Change in Meter using Inner Metric Analysis. ISMS seminars, Goldsmiths College, University of London, March 22 2007
Van Kranenburg, P. (2007). Towards Integration of MIR and Folk Song Research. Presentation at WITCHCRAFT Workshop, Meertens Institute, Amsterdam, 24 April 2007.
Volk, A. (2007). Rhythmic Similarity in the Oral Tradition of Folksongs. Rhythm Perception and Production Workshop, Dublin, 1-5 July 2007
Wiering, F. (2007) Presentatie WITCHCRAFT als onderdeel CATCH. Werklunch NWO met Judith van Kranendonk, Directeur-generaal Ministerie OCW, 4 July 2007.

Grijp, L.P. (2007). French-Dutch relations in theatre music, investigated with the Database of Dutch songs. Centre de Musique Baroque de Versailles, Versailles, Frankrijk, 08-07-2007

Wiering, F. (2007). The Witchcraft Project: A Progress Report. Meeting of the International Musicological Society Study Group on Musical Data and Computer Applications. Zürich, 9-10 July 2007

Grijp, L.P., De Bruin, M., Volk, A., Garbers, J. Van Kranenburg, P. (2007). Strophic Heuristics and Melodic Similarity in Song Research. Third Conference on Interdisciplinary Musicology, University of Tartu, Tallinn, Estland, 16-08-2007.

Grijp, L.P., Wiering, F., Veldkamp, R.C., Volk, A., Van Kranenburg, P., Garbers, J. (2007). Starting up WITCHCRAFT. In search of the principles of oral variation in music. 37th International Ballad Conference, University of Aberdeen, Kyle of Lochalsh, Schotland, 30-08-2007.

Grijp, L.P. (2007) WITCHCRAFT en het geheim van de mondelinge overlevering. Digitaal Erfgoed Conferentie, De Doelen, Rotterdam, 12-12-2007.

Anja Volk. Rhythm in music as a dynamic process. Presentation at the inaugural conference 'Rhythm, Time and Temporal Organisation' of the Institute for Music in Human and Social Development. 2 - 4 June 2006, University of Edinburgh, Scotland

Frans Wiering. Can Humans Benefit from Music Information Retrieval? Keynote speech at 4th International Workshop on Adaptive Multimedia Retrieval, Geneva, 2006

Posters (selection)

Steven Ness, Peter van Kranenburg. An Online Interface to Explore Audio Segments. *International Society on Music Information Retrieval (ISMIR 2010) conference.* Utrecht, Netherlands, 9-18 August 2010.

Jörg Garbers and Peter van Kranenburg. Bridging Music Information Retrieval and Folk Song Research - The Computational Setup of the WITCHCRAFT Project. *International Society on Music Information Retrieval (ISMIR 2009) conference*, Kobe 26-30 oktober 2009; SIREN 2009, Enschede, 5 november 2009.

Peter van Kranenburg, Anja Volk, Frans Wiering, Remco C. Veltkamp. Musical Models for Folk-Song Melody Alignment. *International Society on Music Information Retrieval (ISMIR 2009) conference*, Kobe 26-30 oktober 2009; SIREN 2009, Enschede, 5 november 2009: nominated for Best Poster Award.

Garbers, J., Van Liempt, M. Stash, N.V., Nussbaum, N. (2007). Unified Access and Search. Poster *CATCH Midterm Event*, Den Haag 29/30 November 2007

Garbers, J., Van Kranenburg, P., Volk, A., Wiering, F., Grijp, L.P., Veltkamp, R.C. (2007). WITCHCRAFT: Modeling Folk Song Concepts for Music Retrieval. *CATCH Midterm Event*, The Hague, 30 November 2007

Jörg Garbers, Peter van Kranenburg, Anja Volk, Frans Wiering, Louis Peter Grijp, Remco C. Veltkamp. The WITCHCRAFT Pilot: A Prototype Search Engine For Folk Songs. *SIREN 2006*, Utrecht, 12 Oktober 2006

Software and manuals

Peter van Kranenburg. Framework for Sequence Alignment. Software, 2010

Jörg Garbers. Testing Environment for MIR methods. Version 2009. Software.

Jörg Garbers. Rubato. Version 2009. Software.

Jörg Garbers. Alignment software. Version 2009. Software.

Jörg Garbers. Using C++ and Java implementations in the Melody Search Engine using Preferences and Scripting. Software, 2009

Chris de Vries, Jörg Garbers, Peter van Kranenburg. A Slider Interface for Weighted Testing Environment Configurations. Software, 2009

Peter van Kranenburg. Converter for melody encodings. Version1.44 Software, 2009

Garbers, J., WitchCraftEditor. Version 1.13. Software. 2008

Jole, S. van, Software keyboards for Yahmuugle. Software

Garbers, J., Van Kranenburg, P., Software: A Data Flow Infrastructure for the Liederenbank, November 2007

Garbers, J., de Bruin, M., Software Integration of the Melody Search Engine with the Liederenbank (www.liederenbank.nl)., December 2007

Garbers, J., Software: Adaption of the Musipedia Keyboard for the Liederenbank., December 2007

Jörg Garbers, Peter van Kranenburg, Anja Volk. Codering proefcorpus Witchcraft. User Manual for WitchCraftEditor

Events

Symposium Perspectives for Computational Musicology. Amsterdam, Meertens Institute, 5 October 2010.

11th International Society for Music Information Retrieval Conference (ISMIR 2010), Utrecht, 9-13 August 2010.

Utrecht Summer school in Music Information Retrieval (USMIR). Utrecht Univerity, 2-6 August 2010.

Knowledge representation for intelligent music processing, Dagstuhl Seminar 09051, January 2009

MIR Symposium and Workshop, Utrecht University, 20 February 2007.
CATCH Meeting, Meertens Instituut, 27 April 2007.
International Musicological Society Study Group Meeting on Musical Data and Computer Applications. Zürich, 9-10 July 2007

In the news

ISMIR 2010
- Press release Utrecht University: http://www.uu.nl/NL/Actueel/Pages/Wetenschappersbesprekennieuwstemuziektechnologie%C3%ABninUtrecht.aspx
- Kennis voor de toekomst: 75e lustrum Universiteit Utrecht. http://www.cm.uu.nl/lustrum/index.php/hoog-geeerd-publiek/173-de-nieuwste-muziektechnologieen
- 21 August 2010. Frans Wiering and Remco Veltkamp in Technisch Weekblad: "Computer wordt muzikaal". http://www.cs.uu.nl/research/publicity/computer-wordt-muzikaal.106403.lynkx.htm
- 3 August 2010. Frans Wiering and Remco Veltkamp in BNN Niewsradio program Denktank. http://www.cs.uu.nl/research/publicity/BNR_Nieuwsradio-3Aug2010.mp3

PhD thesis Peter van Kranenburg
- Press release Utrecht University: http://www.uu.nl/NL/Actueel/Pages/Slimmezoekmachinevindtvariatiesopmelodie.aspx
- 29 September 2010. Peter van Kranenburg in Mijkes Middag. http://player.omroep.nl/?aflID=11502435
- 5 October 2010 Peter van Kranenburg in Degids.fm. http://degidsfm.vara.nl/Radio-Nieuwslicht.7746.0.html

4 Results of interest to the general public

Please give an up-to-date summary of the project, in Dutch, in language accessible to the general public. EW may use this summary for publicity. Include the objective and problem definition/assignment, and the design. Then give a concise description, in Dutch, of the results achieved/new insights gained in the project.

Het Meertens Instituut herbergt onder de naam *Onder de groene linde* een waardevolle collectie van ruim 7000 opnames van Nederlandse volksliederen. Sinds juni 2007 zijn de opnames publiek toegankelijk via de website van de *Nederlandse liederenbank* (www.liederenbank.nl). Om een bepaald liedje te vinden kan in de liederenbank in metadata (zoals titel, zanger, opnamedatum, etc.) worden gezocht. Het zou nuttig zijn om ook op melodische inhoud te zoeken. Daarom is in kader van het WITCHCRAFT-project gezamenlijk door Meertens Instituut en Universiteit Utrecht een melodieënzoekmachine ontwikkeld.

Een belangrijk begrip uit de melodieleer is *tune family*. Hiermee wordt een groep melodieën aangeduid die als variant van elkaar kunnen worden beschouwd. Gegeven een melodie uit een tune family moet een effectieve melodieënzoekmachine andere leden van de tune family een hoge ranking kunnen geven in een Google-achtige resultaatlijst.

De meest succesvolle methode voor het vinden van gelijkende melodieën is gebaseerd op uitlijning van reeksen symbolen. Er bestaat een algoritme dat de optimale uitlijning van twee reeksen symbolen kan bepalen door de 'kosten' van de uitlijning te berekenen. In de gekozen methode worden de kosten berekend op basis van drie muzikale kenmerken: toonhoogte, ritme en frasestructuur.

Bij een test met 360 zoekvragen en een collectie van bijna 5000 melodieën bleek dat in 90% van de gevallen een relevant zoekresultaat bovenaan de resultaatlijst stond en dat zich in 99% van de gevallen een relevant zoekresultaat bij de eerste 10 melodieën van de resultaatlijst bevindt. Deze methode is als prototype in de Nederlandse liederenbank opgenomen.

5 Follow-up

5.1 Personnel

In what field and in what position do project participants work after the project is finished?

dr. Jörg Garbers: currently unemployed.
dr. ir. Peter van Kranenburg: researcher at the Meertens Institute, Amsterdam, active in computational folk song research.
dr. Anja Volk: has just received a VIDI grant for computational musicology; she will be employed at Department of Information and Computer Sciences, Utrecht University

5.2 Research

Will there be any follow-up of the research performed within the project? In what direction/form?

The follow-up of the WITCHCRAFT project includes:

Anja Volk's VIDI project *Modelling musical similarity over time through the variation principle* (2011-2016), which will be based at Information and Computer Science, Utrecht University. One of the case studies involves folksong.

Computational folksong research at the Meertens Institute into computational modelling of musical memory and oral transmission and into the identification of high-level meaningful patterns in melody. Funding applications are being prepared for this research.

Tools created in the WITCHCRAFT project are being further developed into reliable, robust and usable tools in the WITCHCRAFT Plus project, enabling better access to folk song collections by end users and music researchers. It will also finalize the integration of the melody search engine into the Nederlandse Liederenbank.

Other project initiatives are in preparation with various Dutch and international partners in the areas of computational folksong research, interoperability with other collections of musical heritage (e.g. broadcasting archives) and research into music similarity measures that better capture the cognition and human experience of music.

5.3 Retrospect

Mention experiences, "lessons learned" of consequence for future CATCH-projects

The availability of a generous quantity of high-quality data prepared and annotated by experts was crucial for the success of WITCHCRAFT. Generally, the importance of such data cannot be overemphasised for any form of computational research into digital cultural heritage or humanities, and means should be provided to enable the creation and/or acquisition of such data as part of, or in parallel to the research projects.

Communication between computer scientists and domain experts/musicologists was a complex issue. Even though the parties involved had discussed the topic years in advance, mutual understanding was sometimes difficult to reach. The solution we found was to make reaching this understanding part of the project. This has resulted in a model for cooperation between computer scientists and domain experts, and in insight in how domain experts conceptualise their area of expertise and reach their decisions.

Computer science research produces methods and implementations that can act as a proof of concept. To make these suitable for domain experts and end users is a complex task, involving for example provisions for adequate data management, and dealing with additional requirements of content features, usability and robustness. These issues can only be partially addressed in a scientific research context. It is an important contribution to the impact of CATCH projects that further tool development is supported by CATCH Plus, but it is unfortunate that the actual progress of CATCH Plus is hampered by so many factors that have only little to do with the actual content of the implementation projects.

It was critically important for the success of WITCHCRAFT to work with a cultural heritage institution (the Meertens Institute) that was strongly committed to supporting the project and using its outcomes. This made the practical results of the project fit in a larger context. In particularly, the ongoing work on the Nederlandse Liederenbank, especially its online version, was very beneficial to the project.

5.4 Suggestions for further research

Provide suggestions for further research

In the domain of computational musicology, folk song research in particular:

- suitable audio transcription and analysis methods
- content creation methods, in particular for digital scholarly editions
- interoperability between European collections
- visualisation of song/composition relationships, especially using geographical and temporal information
- computational modelling of musical memory and oral transmission
- quantitative study of musical style and influence

In the domain of music information retrieval:

- designing music retrieval methods that better capture the cognition and human experience of music, notably by addessing 'hooks', salient, meaningful and easily memorised patterns in music
- user studies into musically meaningful features using Web 2.0 methods
- segmentation and pattern discovery models
- integration of audio and notation retrieval approaches
- methods for searching distributed collections
- copyright issues in unlocking music collections

6 Signature

I hereby declare that the information provided on this form is true and accurate.
Completed on: 15-10-2010

Name : dr. Frans Wiering
Project leader