

# The WITCHCRAFT Pilot: A Prototype Search Engine For Folk Songs

What is Topical in Cultural Heritage: Content-based Retrieval Among Folk Tunes

Jörg Garbers, Peter van Kranenburg, Anja Volk, Frans Wiering, Louis Peter Grijp, Remco C. Veltkamp

garbers@cs.uu.nl; petervk@cs.uu.nl; anja.volk@meertens.knaw.nl; frans.wiering@cs.uu.nl; louis.grijp@meertens.knaw.nl; remco.veltkamp@cs.uu.nl

## Current Situation

### Meertens Instituut

Nederlandse Liederbank

- Database of 130.000 Dutch Songs
- Only metadata searchable (author, title, first line, etc.)

Onder de Groene Linde

- 7000 recordings and transcriptions, compiled 1957-94
- Missing: search functionality for melodies needed for identification of songs and oral variation
- Transcriptions not yet digitized

### Department of Information & Computing Sciences Utrecht University

- Music Information Retrieval algorithms
- Muugle framework
- Usability studies

## WITCHCRAFT Goals

### Develop a content-based retrieval system for a large number of folk song melodies stored as audio and notation

- Design similarity measures, based on models of music cognition and perception
- Create a search tool that provides access to the collection of *Onder de Groene Linde* for scholars and general public
- Support testing of hypotheses about oral transmission of folk songs
- Design a music information retrieval system suitable for other large music collections

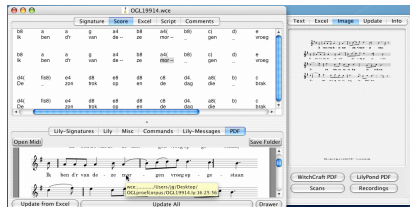


## Pilot Project

### Develop and evaluate an integrated music entry and retrieval system prototype for folk songs

- Develop a music notation editor for folk songs
- Integrate existing data (recorded audio, scanned transcriptions)
- Guide and evaluate data entry process resulting in a test corpus of 140 digitized melodies from *Onder de Groene Linde*
- Evaluate Muugle's melodic similarity search on the test corpus
- Explore ways of using the new tools for addressing research questions such as classification, identification and tracing of melodic variants

### Music editor



### Manual Compilation of Variants

Diagram illustrating the manual compilation of variants. It shows a sequence of musical staves (OGL19914m, OGL20512m, OGL21003m, OGL21818m, OGL25904m, OGL26905m, OGL38411m, OGL37303m, OGL37511m, OGL41104m, OGL49313m) with red annotations and arrows indicating relationships between different versions of a melody. The lyrics "Ik ben d'r van de - ze mor - gen vroeg op - ge - staan" are visible. Above the staves, a sequence of notes is shown: a g a b a b c/e e d c b d e d e d. The Muugle logo is visible in the top right corner of the diagram area.