

Preserving the Spirit of the Epoch: Digital Conversion of Nordic Music Magazines

Amalie Ørum Hansen
Development Consultant
Gentofte Centralbibliotek

Sergey Borovoy
CEO
ATAPY Software

Quick facts about Gentofte Central Library



Gentofte Central Library (Gentofte Centralbibliotek)

- One of six Central libraries in Denmark
- Responsibility (services) area: the Capital Region of Denmark (28 municipalities and their public libraries)
- Mission and task:
 - providing support to public libraries on buying and obtaining materials
 - providing consultancy services (technical & developmental)
 - servicing all the employees of the region's public libraries with relevant education (i.e. methodological and skills)

Quick facts about ATAPY Software



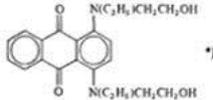
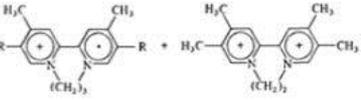
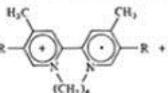
- Offices: Russia (Novosibirsk), Germany (Munich)
- Main focus areas:
 - Digitisation bureau
 - On-demand software development services in the fields of OCR & document imaging
- Completed projects in the field: 100+, track record including work for the Royal Danish Library, the National Library of Sweden, Springer, contributions to the METAe and IMPACT projects, etc.
- Strategic partnership with ABBYY and Microsoft
- Ability to handle challenging sources that yield poorly to OCR technology (due to the technology edge)

Microsoft Partner

Gold Application Development
Silver Data Platform

Scientific articles

Library cards

Reaction	Radical generation Method	Solvent	T [K]	Rate data	Ref./add. ref.
 $M^{\cdot+} + \text{Ni(C}_2\text{H}_5\text{)(CH}_2\text{CH}_2\text{OH)}$	Pulse rad. reduct. with e_{aq}^- or $(\text{CH}_3)_2\dot{\text{C}}\text{OH}$	H_2O , pH=7 (5 mM phosphate) +0.1 M 2-propanol	298	$k_f = 8.6 \cdot 10^7 \text{ M}^{-1} \text{ s}^{-1}$ $k_b = 6.8 \cdot 10^7 \text{ M}^{-1} \text{ s}^{-1}$	81Sv1
10.1.7 Bipyridinium dication + bipyridinium dication $\text{Bipy}^{\cdot+} + \text{Bipy}^{2+} \rightleftharpoons \text{Bipy}^{2+} + \text{Bipy}^{\cdot+}$					
 $\text{R} + \text{H}_3\text{C}$	Pulse rad. reduct. with e_{aq}^- or $(\text{CH}_3)_2\dot{\text{C}}\text{OH}$	H_2O , pH=7 (2 mM phosphate) +0.2 M 2-propanol	295	$k_f = 2.4 \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$ $k_b = 1.5 \cdot 10^7 \text{ M}^{-1} \text{ s}^{-1}$	84And2
$\text{O}^{\cdot+} + \text{O}^{2+}$	Pulse rad. reduct. with e_{aq}^- or $(\text{CH}_3)_2\dot{\text{C}}\text{OH}$	H_2O , pH=7 (2 mM phosphate) +0.2 M 2-propanol	295	$k_f = k_b = 5 \cdot 10^7 \text{ M}^{-1} \text{ s}^{-1}$ ^{b)}	84And2
 $\text{R} + \text{O}^{2+}$	Pulse rad. reduct. with e_{aq}^- or $(\text{CH}_3)_2\dot{\text{C}}\text{OH}$	H_2O , pH=7 (2 mM phosphate) +0.2 M 2-propanol	295	$k_f = 2.8 \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$ $k_b = 1.6 \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$	84And2

^{a)} For M, see p. 165.

^{b)} Estimated from a Marcus treatment of rate data for electron exchange with other 2,2'-bipyridinium compounds.

AACH, H.G., 1960.

Vergleich des Eiweißanteiles des Wildstammes vom Tabakmosaikvirus mit dem einer Temperaturmutante. Vererbungsl. 91:312-316, 1960.

AACH, H. G., 1957.

A 72

Serologische Untersuchungen an Mutanten des Tabakmosaikvirus. Zeitschrift für Naturforschung 12b-10:614-622, 1957.

AACH, H.G., 1961.

A 112

Zur Konstanz der Aminosäurezusammensetzung im Eiweißanteil des Tabakmosaikvirus. III. Elektrophoretische Untersuchungen an TMV-Präparaten verschiedener Herkünfte. Archiv f. Microbiologie 39:253-256, 1961.

Fiction of various nature: novel, poetry, theatrical plays and travelogues

hang interessante Facts, der saavel ere vigtige for den philosophiske Menneskegrundfer, som for den Naturkyndige, for Lagen og Politikeren." Stulde dette endog nu, da Africa er ulige mere bekendt, end 1788, ikke længere findes gjeldende, og skulde endog ikke alle den i evrigt uden Tvivl sandhedselskende Herses Efterretninger være paalidelige (hans Stridemaade gier ham i det mindste mistaant for at have udsmykket et og andet, hvad han fortæller) har hans Skrift dog lige til den seneste Tid været en Hovedkilde for dem, der søgte Kundskab om Guineakysten.

Reget mindre bekendte, end noget af de her berettede Skrifter, ere de "Bidrag til Kundskab om den danske Strækning paa Guinea-Kysten" som ere indrykkede i Hr. Etatsraad Thaarups Archiv for Statistik, Politic &c. 2 Bd. S. 309-11 og 3 Bd. S. 161-268, og dog er disse ikke uden Interesse, saavel i geographisk, som i historisk og statistisk Henseende. Deriblandt findes en Beretning fra Gouverneuren og Raader paa Christiansborg af 1774, saavel om Landets Beskaffenhed og Negerenes Levemaade, som om Handelen daværende Tilstand (S. 161-92); en udførlig Beretning fra den nylig i Kiebenhavn afskede Gouverneur Wiern om de danske

Forter og Negerier, i den Tilstand hvori de vare 1788; og endelig de mærkelige Actstykker, der oplyse Historien af de Colonie-Anlag i Aquapim, som Regieringen 1788 havde overdraget Isert, der med megen Troer begyndte at grunde disse Anlag"). Af hans Indberetning om dette Forrettagendes Fremme (bateret Frederiksnoel i Guinea d. 16 Jan. 1789 — Navnet, han vilde give den nye Stad) erforer man, at han virkelig har afstiebt Negernationen i Aquapim en betydelig Landstrækning, hvorom en formelig Tractat i Kongen af Danmarks Navn blev sluttet med Nationens Overhoved; at han paa en heilidelig Maade har taget det ham af denne overleverede Land i Besiddelse, og der begyndt at opføre Bygninger; ligesom og, at han ved Negerenes Hielp i tre Uger lod bygge en Vej ned

*) Geol. omtaler disse Anlag S. 216. 175 men efter Hets Beretning synes det, som Colonien Landstrækning, i en Længde af 20 Mill og fra 5 til 10 Miles Brede, med 5 til 6 Miles Afstand fra Kysten, ikke var paa hin Side af Negerne; og Banstrækket hermed ved Transporten synes, efter dens Omkæring, indrettede meget betydelige. Hets Indberetningen ved at bygge den omtalte Vej, hvorpaa 100 til 200 Neger arbejdede i tre Uger, var omtrent 400 Mile, og Hets Kiebskammer for Taabet, næsten af alle udgjere 5 Olovers Næsti eller omtrent 600 Mile.

Och jag blev ensam kastad på ett skär!
Se'n blev det ställt. — Of vilka långa dygn!
Blott molnfri himmel öfver mig
Och ändlöst, djupblått hav omkring mig!
Ej ljud av livfylld varelse!
Ej månen väckte mig med sina skrån!
Ej ens en klare kom den lätta våg
Att skvalpa mot en sten.
Det föreföll mig som jag själv var död,
Jag högljutt talade och skrek,
Men rösten skrämde mig
Och torkan band min tunga.
Blott hjärtans jämna slag i bröstet
Mig minde att jag lemnat!
Men när en stund jag lyssnat till dess ljud
Jag snart ej mer det hörde.
Då for jag upp av bådan
Och så var våren tillrädd i synmakt föll.

— jag härde
fara,
som ej var min,
själ.

sikte vilja;
kliga,
og dem fylld.

smeka
det

ten!

ides?
tarke,
sin?
fjer älskog.

I.

Bredvid sin åttioåriga dotter stod Don Ignacin de Andrade e Silva på ballongen af sin Quinta, och blickade åt den väg, som ledde åt halvstaden.

Jag vet väl, sade han, att de ännu knappast äro att förvänta: ännu står salen alltför högt på himmelen, ännu har halvstaden knapt börjat uppstå, och mycket åter väl icke Drottningen vilja aflytta af sin vanliga siesta. Dock hade de lofvat att komma tidigt. — Men du här mig icke, lilla Clementina; du tänker aldrig annat än på din grönsa fågel!

Utän att vara gif Clemenina sin fader blott en vänlig blick; sedan fortfor hon ämnevis att smeka, ämnevis att mata den lilla paradisfågeln, som satt på hennes vänstra hand.

Fadrens blick hvilade öfver den underskåns talan, som lig utbredd under hans fötter. Efter en stund fortfor han: Du plågar vara förständigare än man annars vanligen är i din ålder, mitt barn. Men ha då dina unga ögon endast sista

Digital Conversion of Nordic Music Magazines

Project duration: 2007-2008

Three popular music journals issued in Denmark in the second half of the XX century:

Nordic Sounds - the magazine of NOMUS, the NORDIC MUSIC COMMITTEE published 1982-2006

GAFFA - a free Danish magazine, published 1983-present

MM - a Danish magazine devoted to Jazz and Rock issued 1968-1989



Digital Conversion of Nordic Music Magazines: Contributors

- **Funded by the Danish Agency for Culture, Libraries**
- **Gaffa A/S, Gaffa Nordic**
CEO Robert Borges
- **Musikbibliotek.dk, Gentofte Centralbibliotek**
Former Editor-in-Chief: Amalie Ørum Hansen (now Bibzoom.dk, State and University Library)
Editor-in-Chief: Niels Mark Pedersen



Digital Conversion of Nordic Music Magazines: Contributors

- **Det Virtuelle Musikbibliotek, State and University Library**

Former Editor: Søren Svane Hansen



THE ROYAL LIBRARY

NATIONAL LIBRARY OF DENMARK AND COPENHAGEN UNIVERSITY LIBRARY

- **The Royal Library,
National Library of Denmark and
Copenhagen University Library,
Dept. of Documentation & Digitisation**
- **ATAPY Software (digitisation contractor)
CEO Sergey Borovoy**

Digital Conversion of Nordic Music Magazines



Project goals:

- Preserving full collections of the three magazines in digital form
- Preserving and conveying the cultural “flavor” of the late XXth century in the Nordic countries, including
 - information about local music acts
 - the attitude of Nordic music community to worldwide sensations
 - snapshots of associated subculture movements, etc.
- Making this important knowledge available online

Digital Conversion of Nordic Music Magazines

Project requirements:

- Full-text recognition

- High recognition accuracy (to ensure excellent searchability of the collection)

- Excluding part of material from recognition (commercials, etc.)

- Results: industry-standard XML format

- In some parts of scope – illustrations were extracted & saved in a separate location, hyperlink placed in the XML file



Digital Conversion of Nordic Music Magazines

Complications:

- The specifics of periodicals as input material:
 - Wide format and multi-column layout (challenge for automatic segmentation)
Solution: manual after-correction of automatic segmentation
 - Colored and textured backgrounds (challenge for OCR)
Solution: semi-automated image preprocessing in graphic packages (increasing contrast, etc.)
 - A variety of fonts and font colors within one page
 - Designer fonts used
 - Skewed fragments + fragments with normal text orientation present in one page
Solution: KFI of poorly recognized or unrecognized occurrences
 - Inverted and «normal» text present on one page
Solution: multi-attempt OCR with varied settings (inverted/normal text)

Digital Conversion of Nordic Music Magazines

2

Nordic Sounds
NOMUS by the
Nordic Cultur
under the No
Ministers, Sna
DK-1205 Cope
(tel. 01-11471)

Articles in *N*
not necessari
views of NOM

Material from
can be quote
reproduced w
sion of the
Nordic Cultur
The source r
clearly ackno

The editor of
accepts no re
unsolicited m
tographs and

Editor: *Knud*
Editorial cons
Jens Brincker
Translators: *M*
Cbesnutt and
Editorial secr
Hansen

Publisher (wi
sibility in a
Danish pres
Lasse Lindba
Graphic desig
marsson and
lanck, Liber
Chris Bjerrega
Film-setting by
Copenhagen/
Printed in Sv
Liber Grafiske

This number
was sent to the
1982
Date of public
1982

commodate the regional cultur
stitutions and activists. This has
turn strengthened Nordic ident

Most importantly, it has inc
human contacts among the art
their understanding of one ano
cultures. The result is what we
"Nordic culture", with at least f
ferent but related dialects. Only
kind of cooperation is maintai
developed can any common E
an identity grow up naturally.

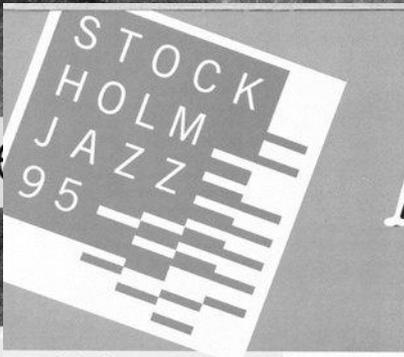
At a recent seminar in Mainz
"Oper—zwischen Tradition un
periment", it struck me how ge
Germans are at forging and ma
ing contacts with other Germa
king countries. There were rep
tatives from Switzerland and L
tenstein as well as Israel and H
and German-speaking commu
in Europe and America. The F
too, have their cultural links w
countries in Africa; the British
Commonwealth and the Britis
Council. Countries work toget
because of the common herit
age they feel language has
given them.

In our Nordic co-
operation langu-
age can be a di-
visive as
well as a uni-

elle oppgaver. Hva vil St
"Politiet foretar ident
på grensene." heter det.
vis sikre på at det snart
bordet forslag om at det
sensitive politio
ppgaver

fying factor. I remember well
ding a Young Nordic Music F
seminar where I thought that
cussions should have been hel
Scandinavian languages. The
delegation stormed out. They

An Ice



by ODD SNEEGEN

Stockholm Jazz is a new festival which has just had its premiere. What is special about this festival is that the main focus is on the composer of jazz music. It has also been the ambition of the festival to reflect the Nordic jazz scene. Last, but definitely not least, a part of the festival has been devoted to seminars in collaboration with the music colleges.

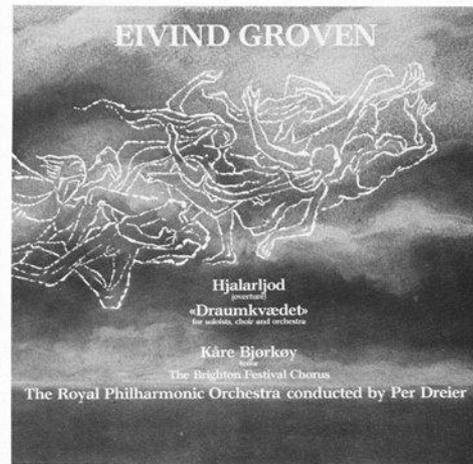
The commissioning of works from the composers who take part in the festival also means that the audience is given the chance to make the acquaintance of completely new compositions.

It is common practice when organizing festivals of western art music to focus on the composer; this gives audiences an opportunity to become more familiar with the musical language and works of a few select composers. In jazz it is a well-known fact that the composer also often plays an active role as a musician. However, there are a certain number of artists in the jazz world who have put most of their energy into composing. These composers have often had a significant influence on the stylistic development of jazz: they include Thelonious Monk, George Russell, Duke Ellington, Charlie Mingus and Bob Brookmeyer, as well as Nordic jazz composers such as Palle Mikkelborg, Jan Johansson, Lars Gullin, Jan Garbarek and Heikki Sarmanto.

"A Freak w

20

Eivind Groven's *Draumkvædet* is based on an old Norwegian ballad about Olav Asteson, who fell into a deep sleep which lasted all the twelve days of Christmas. When he awoke he went to church and related the vision he had experienced, in which he not only wandered through the Kingdom of the Dead but also saw the Day of Judgment. Groven (1901-77) used three of the melodies to which the ballad is still sung in his treatment of this ancient song from Telemark. The recording was made in connection with a concert performance in London, and the English choir catches the Norwegian linguistic tone and the rough and ready style of the music remarkably well. Only the opera mannerisms of the tenor soloist seem a little out of place.



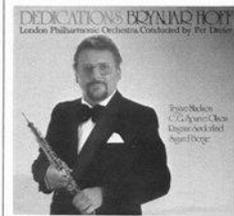
NORWEGIAN COMPOSERS



Sparre Olsen was born in 1903 and is thus a veteran of Norwegian music. The eight pieces

on this record span a period of more than half a century right up to *Metamorphosis* for solo cello from 1982. The sleeve notes rightly emphasise that Sparre Olsen has made his most notable contribution as a vocal composer, which makes it all the more odd that the anthology only includes one vocal item, *Two Edda Poems* (1946) for soprano and orchestra. The poems are sung pleasantly but with almost no text articulation by Toril Carlsen, and the lack of a word-sheet is particularly acute in the circumstances.

Here the splendid Norwegian oboist Brynjar Hoff plays four works dedicated to him personally. The sleeve shows him in a suitably Romantic get-up, and the contents correspond precisely to the packaging. Söderlind adds a taste of Arabic exoticism to his Romantic broth, while there is an Indian touch to Sigurd Berge's music. But the results are always gratifying, and it is no wonder that this record has been one of the biggest sales successes in the series.



Digital Conversion of Nordic Music Magazines

Complications:

■ Format specifics:

- Input material: TIFF, PDF - the latter sometimes containing a text layer, which was sometimes partially unrecognizable (in vector format)

Solution: OCR of PDF as image-only (sometimes), or KFI of unrecognized information

- Output requirements: XML format, different from that produced automatically by ABBYY FineReader

Solution: export to Microsoft Word, semi-automatic XML markup in required format

■ Language specifics:

- Several languages (Danish & English) on one page (additional challenge for OCR even with correct dictionaries enabled)

- Critical information in Danish

Solution: operators with linguistic background, manual verification

Digital Conversion of Nordic Music Magazines

XML format requirements:

- One XML file per article (including multi-page articles)
- Metadata block: magazine name, issue (e.g. 2007-12), article number on page
- Article level tags: type of article (interview, musical review, news piece, etc.), author, abstract
- Formatting tags : title, subtitle, text type (bold, italic, bold italic), etc.
- Special tags (occurrences of proper and geographical names, etc.)

Digital Conversion of Nordic Music Magazines

Process phases (scope of work):

1. Analysis/segmentation of pages (automatic)
2. *Segmentation cross-check & correction (manual)*
3. OCR (automatic)
4. *Verification/correction or KFI if unrecognized (manual)*
5. Export to Microsoft Word (automatic)
6. *XML markup (manual, semi-automated)*
7. *XML file aggregation (merging several files related to one article)* (manual)*
8. XML validation (automatic)

* Optional phase (is case of issues with automatic article segmentation)

Digital Conversion of Nordic Music Magazines

Tools and technologies used:

ABBYY FineReader 8.0 (the latest ABBYY FineReader version at the time):

- Segmentation

- OCR

- Verification (manual, in ABBYY FineReader interface)

- Export to Microsoft Word

Third-party XML validation software:

- XML file validation

In-house-developed macros (VB):

- XML markup of Microsoft Word files

ABBYY®



Digital Conversion of Nordic Music Magazines



Det Virtuelle Musikbibliotek



Results:

All project requirements fulfilled; results accepted and published

Nordic Sounds and MM magazines: available at the website of Online Music Research Library (www.dvm.nu/periodical)

GAFFA magazine: available online at <http://gaffa.dk/arkiv>

Thank you!

Amalie Ørum Hansen
Development Consultant
Gentofte Centralbibliotek
ahan@Gentofte.dk
Tel.: +45 39 98 58 47

Sergey Borovoy
CEO
ATAPY Software
sergeyb@atapy.com
Tel.: +7 383 363 96 99