

Der C-3 Periodikaserver im GBV

Neue Perspektiven in der Zeitschrifteninhaltserschließung

Bericht zum Stand eines Pilotprojekts an der
SUB Göttingen in Kooperation mit der
Verbundzentrale des GBV



Inhalt

- Ausgangssituation
- Organisatorische Bedingungen
- Anforderungen an Ablösesystem
- Ergebnisse der C3-Testreihe
- Workflow
- Retro-Projekt
- Weitere Planung

Ausgangssituation

SUB Göttingen erschließt ca. 1.000 Zeitschriften

(= 8% des Bestandes) per Scanning der Inhaltsverzeichnisse

- ca. 2.300 Hefte/Jahr
- ca. 30.000 Artikel/Jahr

- aus den Fachgebieten:

- Geologie (GeoLEO)
- Finno-Ugristik (ViFaNord etc)
- Anglistik/Amerikanistik
- u.a.m.



Ausgangssituation

- überwiegend fremdsprachige Titel
- beschränkt auf lateinische Schrift
- bisher durch sog. CC-Tool (Eigenentwicklung):
 - Probleme mit Diakritika/Sonderzeichen (ASCII-Code)
 - technische Probleme
 - kein Support
 - nicht mehr auf Stand der Technik
 - keine technische Weiterentwicklung mehr

Organisatorische Bedingungen

- Akzession und Inhaltserschließung
 - organisatorisch eng verzahnt
 - jedoch räumlich getrennt
- Personal für Inhaltserschließung:
2 - 3 studentische Hilfskräfte
und 1/2 Dipl.Bibl.

Anforderungen an Ablösesystem

- Unicode-Fähigkeit (Multilingualität)
- Komfortable Diakritika-Behandlung
- Mehrplatzfähigkeit
- Anbindung an Pica-ACQ
(automatische Auftragsgenerierung bei Heftakzession)
- Prozesssteuerung und Auftragsverwaltung:
 - Heftakzession, Scannen, Indexieren, Export, Konversion, Qualitätssicherung, Katalog-Import
 - Eingangs- und Umlaufkontrolle mit Statistikfunktion
- Sehr gute Texterkennung
- Bedienerfreundliche Oberfläche
- Beschleunigung der Arbeitsabläufe
- Sowohl laufende als auch retrospektive Erfassung

C-3 - Testreihe



Testpartner:



SUB Göttingen





Phase 1: Ergebnisse



- Optimierung der C-3-Oberfläche
- Automatische Seitenintervall-Berechnung
- Maschinelle Unterstützung bei Kennzeichnung v.
 - mehreren Autoren
 - Nachnamen
- Nachbearbeitungskennzeichen auf Titel- und Heftebene
- Wörterbücher
- Multilingualität

Phase 2: Ergebnisse



- Anbindung an Pica-ACQ
 - Definition der aus Pica an den Periodikaserver zu übergebenden Daten
 - automatische Auftragsgenerierung („Scan-Auftrag“)
 - automatischer Auftragszetteldruck mit Barcode
- Prozesssteuerung, Auftragsverwaltung und Statusfortschreibung in MyBib eDoc
- Konversion der XML-Dateien in Pica3-Format
- Qualitätssicherung per Autopsie
- Demnächst: Import in Pica/OLC-Datenbank

Hefteingang - Akzession in Pica/ACQ

Titel-Ident.Nr.

C3-Ident.Nr.

Scan-Merkmal

WinBW 2.000 - [ACQ SUB GOETTINGEN Check In]

Abv Explorations : a literary and research jo / Government Coll College 1973
 0014 4075 FMAG:ZA 89730

Swets: 99001384

PPN 333888693 Bestellnummer 99006595 Eingabedatum 11-12-1998
 EPN 312696728 Zusätzl. Nr. Institutscode 2 / 0
 Lieferant 05122 Al-Abbas Budgetcode 8ck
 Aufstellung mag Typ/Stad/Freq b / k / k
 Letzte Lieferung 2006/17 Zuordnung/Ablage 89730 Bände/Hefte 1 / 0

Lesesaal Magazin

bw=z ul=cc e=gl

M # Erwartet Lieferung N
 Numerisch

j 0 06-03-07 2007/18 2007/18

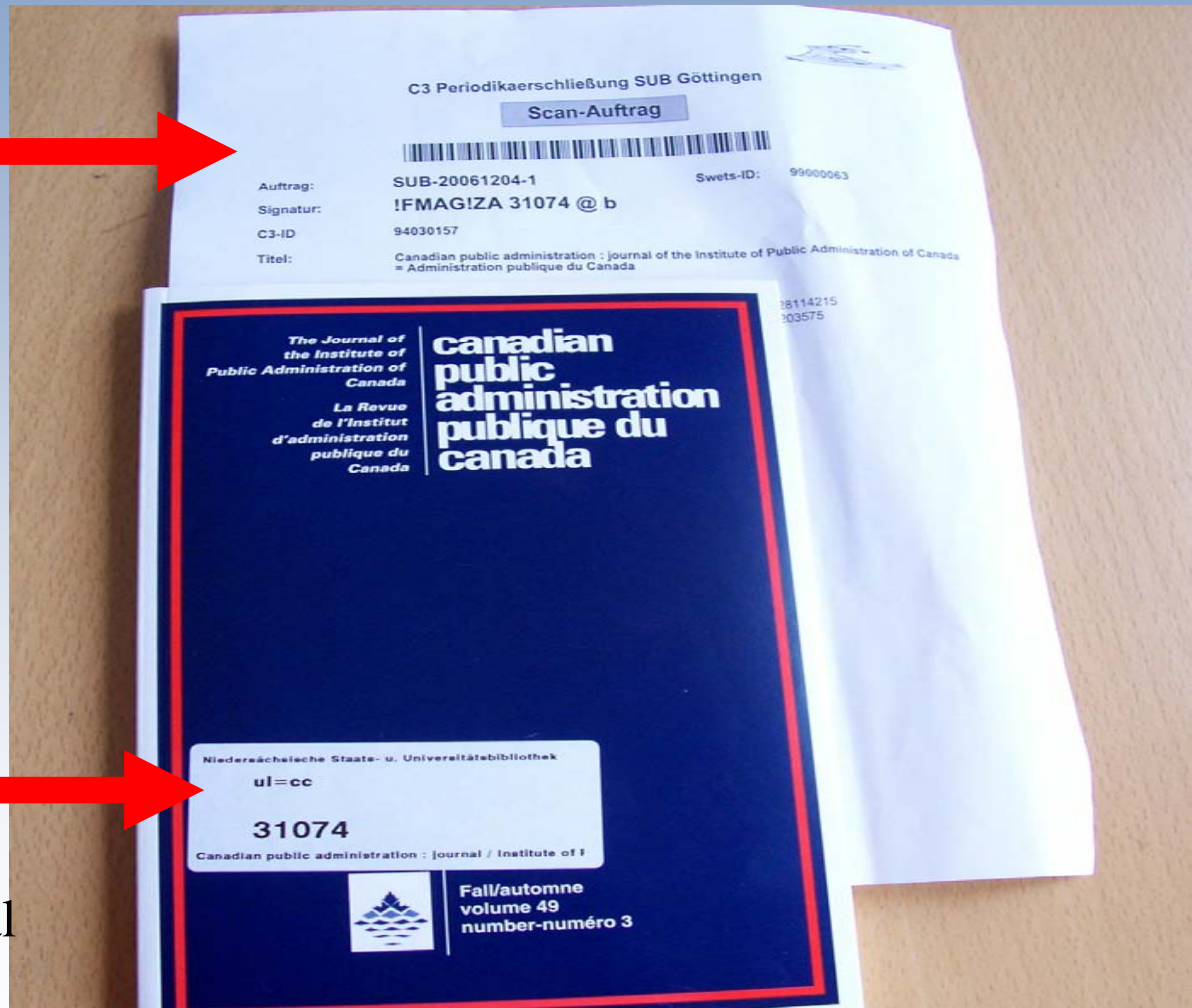
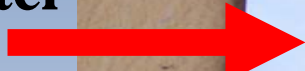
eingegang. Heftlieferung Mahndikator j=ja

Generieren Laufzettel Etikett Umlaufliste

Mahnfristen 121212

Zeitschriftenheft nach Akzession

Scan-Auftragszettel



Etikett mit



- Besitzvermerk
- ACQ-Scan-Merkmal
- Standortangabe

Scan-Auftragszettel


C3 Periodikaerschließung SUB Göttingen
Scan-Auftrag



Auftrag: SUB-20070302-3 Swets-ID: 33004334

Signatur: !FMAG!ZA 89730 @ b

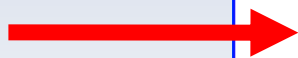
C3-ID: 99006595

Titel: Explorations : a literary and research journal of the Department of English Language and Literature, Government College Lahore

ACQ-Nr:	59005555	EPN:	312696728
PPN:	333888693	ZDB-Nr:	20573601
ISSN:	0014-4975*	ISBN:	
Bes.klassifikation:	!1064!5409!{16 03 Englische Sprache und Literatur}		
SSIS-Code:	7.25		

Heftinformation: 2005/16/0

Nachbearbeitung:
Swets-Nummer fehlt!



BCS-2 Scannen und übertragen

The screenshot shows the BCS-2 software interface. The main window displays a scanned document page with the following text:

Präefatiö.....(Waseem Anwar)

Collaboration and Difference in Page, Stage and Film:
Wilder, Williams and Wharton..... (Martha Bower) 1

The Image Pattern in Othello.....

The Vexing but Diverting We
Gulliver, and the Reader.....

Interview: Hal Ryder.....
Farhan EbadatYar Kh.....

Turkish Woman as Represent
Century Muslim Society.....

Wordsworth: The Nature Hate.....

Caddy, 'My Heart's Darling
..... (F.....

Beauty and Domestic Bli
Subversion in Toni Morrison'.....

Overlaid on this is a smaller window titled 'Scan-Aufträge: [4/27]' which contains a table of scan tasks:

Auftrag	Typ	Datum	Anzahl	Größe	Üb.	Dt.	A-Nr.	Status	Wann	Liefen-Act
407	S0001473	MYFB	25.02.07 08:11	2	63.810	-	SUB-20070126-368			
408	S0001474	MYFB	25.02.07 08:17	2	65.214	-	SUB-20070126-367			
409	S0001475	MYFB	25.02.07 08:26	2	64.532	-	SUB-20070126-366			
410	S0001476	MYFB	25.02.07 08:37	2	63.678	-	SUB-20070126-365			
411	S0001477	MYFB	25.02.07 08:39	2	64.158	-	SUB-20070126-364			
412	S0001478	MYFB	25.02.07 08:41	2	60.668	-	SUB-20070126-363			
413	S0001479	MYFB	25.02.07 08:43	2	61.360	-	SUB-20070126-362			
414	S0001480	MYFB	25.02.07 08:45	2	64.354	-	SUB-20070126-361			
415	S0001481	MYFB	25.02.07 08:47	7	266.900	-	SUB-20070126-665			
416	S0001482	MYFB	25.02.07							
417	S0001483	MYFB	25.02.07							
418	S0001484	MYFB	25.02.07							
419	S0001485	MYFB	25.02.07							
420	S0001486	MYFB	25.02.07							
421	S0001487	MYFB	25.02.07							
422	S0001488	MYFB	02.03.07							
423	S0001489	MYFB	05.03.07 10:45	2	66.444	-	SUB-20070126-477			
424	S0001491	MYFB	05.03.07 12:52	2	96.908	-	SUB-20070126-478			
425	S0001492	MYFB	05.03.07 13:23	1	26.636	-	SUB-20061213-1			
426	S0001493	MYFB	05.03.07 13:26	2	97.600	-	SUB-20061213-1			
427	S0001494	MYFB	05.03.07 13:29	2	7.262	-	SUB-20061213-1			

Below the table is a dialog box titled 'Scan-Aufträge Übertragen' with the text 'Wollen Sie die selektierten Scan-Aufträge übertragen?' and buttons 'Ja', 'Nein', and 'Abbrechen'.

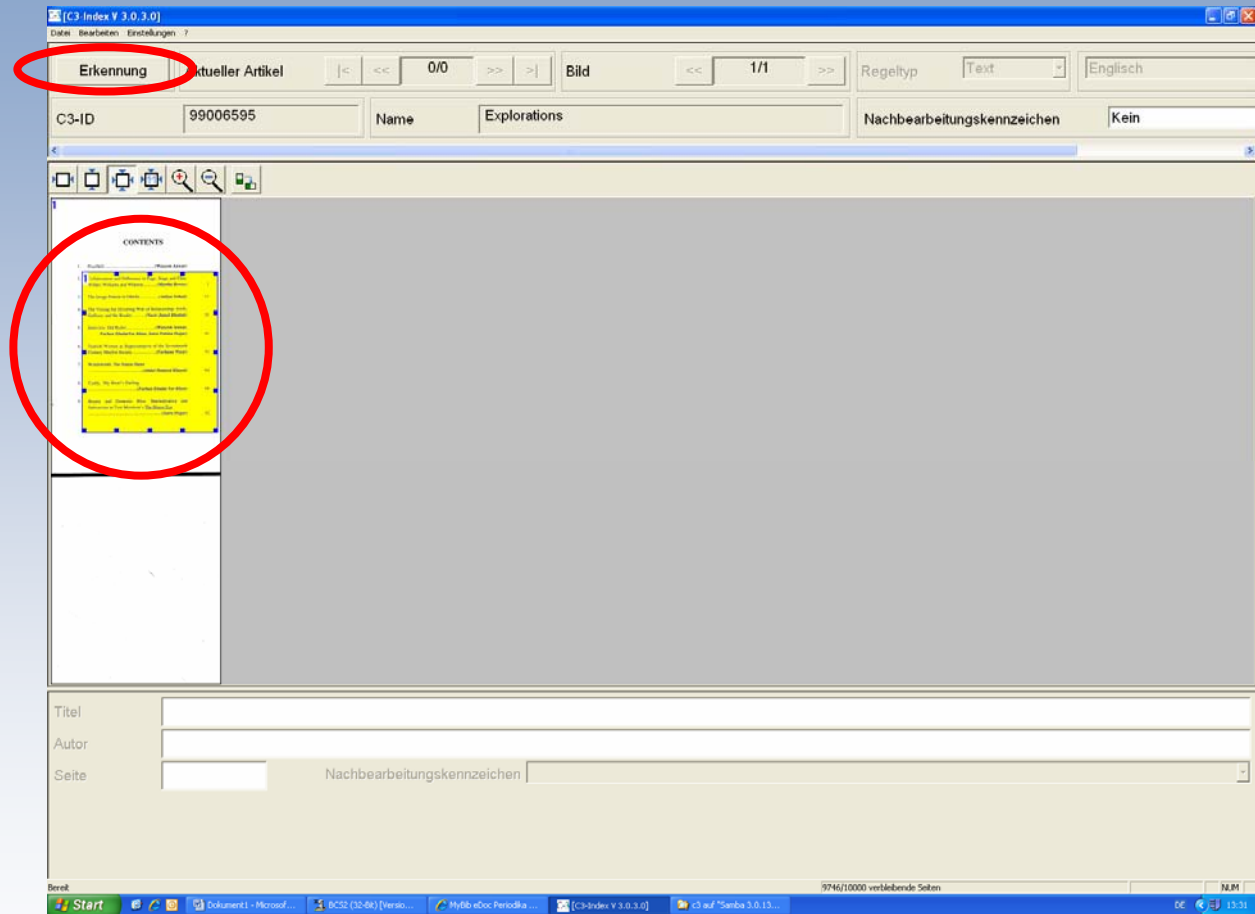


C-3 Template

The screenshot displays the 'C-3 Template V 3.0.3.0' software interface. On the left, a 'CONTENTS' table of contents is visible. The main area contains three dialog boxes for creating new separators: 'Neuer Separator für "Titel"', 'Neuer Separator für "Autor"', and 'Neuer Separator für "Seitennummer"'. On the right, a metadata form is shown with fields for 'Publikation CS-ID', 'Name des Titels', 'Sprache', 'Kommentar', 'Interpretationstyp' (circled in red), 'Reihenfolge', and 'Attribute'. A 'Sprachen auswählen' dialog is also open, showing a list of languages with 'Deutsch' selected. The taskbar at the bottom shows various open applications including 'Start', 'Dokument1 - Microsoft...', 'DC32 (32-bit) [Versio...', 'Mylib eDoc Periodika...', 'C3-Index V 3.0.3.0', 'C3 auf 'Simba 3.0.13...', and 'C3-Template V 3.0.3...'.



C-3 Indexierung





C-3 Indexierung

The screenshot shows the C3-Index V 3.0.3.0 application window. The main content area displays a document with the following text:

Praefātiō..... (*Waseem Anwar*)

**Collaboration and Difference in Page, Stage and Film:
Wilder, Williams and Wharton..... (*Martha Bower*)** 1

The Image Pattern in Othello (*Aaliya Sohail*) 11

At the bottom, a metadata table is visible:

Titel	Collaboration and Difference in Page, Stage and Film: Wilder, Williams and Wharton		
Autor	(martha @Bower)		
Seite	1-11	Nachbearbeitungskennzeichen	Kein

The 'Seite' field is circled in red in the original image.



C-3 Indexierung

The screenshot shows the C3-Index V 3.0.3.0 interface. At the top, there are navigation buttons and a status bar showing '7/8' and '1/1'. Below that, a header section contains 'C3-ID: 99006595', 'Name: Explorations', and 'Nachbearbeitungskennzeichen: Kein'. The main content area displays a list of articles:

- Wordsworth: The Nature Hater(Abdul Hameed Khayal) 64
- Caddy, 'My Heart's Darling**(**Farhan Ebadat Yar Khan**) 68
- Beauty and Domestic Bliss: Internalisation and Subversion in Toni Morrison's The Bluest Eye

At the bottom, a metadata table is visible:

Titel	Caddy, 'My Heart's Darling
Autor	Farhan Ebadat Yar @Khan
Seite	68-82
Nachbearbeitungskennzeichen	Kein 123 (Namensansetzung) Kein

A red circle highlights the 'Nachbearbeitungskennzeichen' field in the metadata table.



C-3 Indexierung

The screenshot shows the C3-Index V 3.0.3.0 interface. At the top, there are navigation controls and a search bar. The main area displays a document snippet with two entries:

Caddy, 'My Heart's Darling(*Farhan Ebadat Yar Khan*) 68

Beauty and Domestic Bliss: Internalisation and Subversion in Toni Morrison's The Bluest Eye(*Saira Dogar*) 82

The second entry is highlighted in yellow. Below the document, a metadata table is visible:

Titel	Beauty and Domestic Bliss: Internalisation and Subversion in Toni Morrison's The Bluest Eye		
Autor	Saira @Dogar		
Seite	82-xxx	Nachbearbeitungskennzeichen	Kein

The 'Seite' field is circled in red. The status bar at the bottom indicates '9745/10000 verbleibende Seiten'.

Konversion XML → Pica3

- kann zu beliebigem Zeitpunkt gestartet werden
- läuft online ab
- umfasst Validierung
 - mehr als 19 Autoren
 - fehlende Heftangaben
 - fehlende/fehlerhafte Swets-ID
- pro Heft 1 Datei
- kann sequentiell abgearbeitet werden:
Korrekturlesen per Autopsie





Konversion XML → Pica3

```

<?xml version="1.0" encoding="UTF-16" ?>
<!-- automatically generated by C3_ConnectSingle.dll -->
- <workitem C3ID="94017463">
- <ILL_MAIL>
  - <workitem>
    <c3id>94017463</c3id>
    <name />
    <band />
    <ausgabe />
    <jahr />
    <transaction-group-qualifier>SUB-20060724-1</transaction-group-qualifier>
    <acq>94017463</acq>
    <swetsid>99001065</swetsid>
    <epn>12796827X</epn>
    <ppn>129077143</ppn>
    <zdbid>28733</zdbid>
    <title>Information bulletin // The Library of Congress</title>
    <collationstatement>2006/65/5</collationstatement>
    <publicationdate>2006</publicationdate>
    <issue>5</issue>
    <volume>65</volume>
    <placeofpublication>Washington, DC</placeofpublication>
    <publisher>Libr.</publisher>
    <callnumber>NR ZB 676 @ b</callnumber>
    <issn>0041-7904* </issn>
    <isbn />
    <classificationcode />
    <ssgcode>24,1</ssgcode>
  </workitem>
</ILL_MAIL>
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<USERFIELD NAME="" />
<NAME>Information Bulletin</NAME>
<YEAR />

```

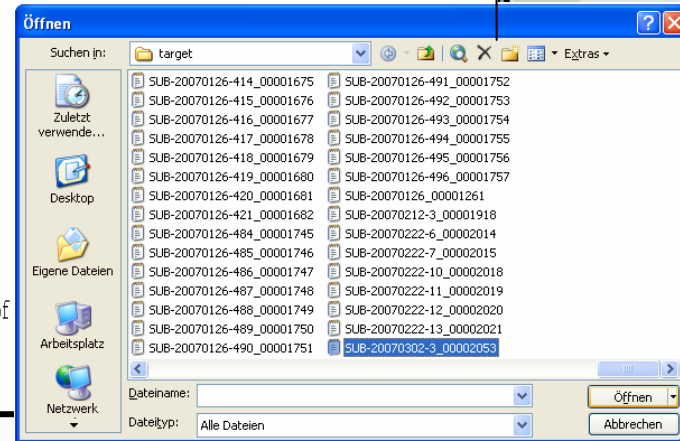
```

#####
# Es liegt mindestens ein Artikel mit Nachbearbeitungskennzeichen vor
#####
0500 Aou
1100 2005
2185 99001384
2199 SUB-20070302-3-1
3011 Martha@Bower
4000 Collaboration and Difference in Page, Stage and Film: Wilder, Williams and
Wharton
4070 /v16/j2005/a0/p1-11

0500 Aou
1100 2005
2185 99001384
2199 SUB-20070302-3-2
3011 Aaliya@Sohail
4000 The Image Pattern in Othello
4070 /v16/j2005/a0/p11-31

0500 Aou
1100 2005
2185 99001384
2199 SUB-20070302-3-3
3011 Nasir Jamal@Khattak
4000 The Vexing but Diverting Web of
Reader
4070 /v16/j2005/a0/p31-41

```



Retro-Projekt

Retrospektive Erfassung bereits gebundener
Zeitschriften im Rahmen eines DFG-Projekts

Beisp.: ViFa Nord – Projektumfang 1998 - 2006

Problem:

- Automatische Auftragsgenerierung unabhängig von aktiver Akzession
- Fehlende Informationen über
 - Anzahl Bände, Benennung der Bände
 - Anzahl Hefte pro Band, Art der Heftzählung

Retro-Projekt

Lösung: Halbtechnisches Verfahren

1. Manuelle Vorgabe:

- Titel-Ident.Nummern
- Heftangabe, ab der die Auftragsgenerierung bei allen Heften mit Stadium f = akzessioniert erfolgen soll

Titel	PPN	EPN	Best.Nr.	Swets-ID	Band ab
Emakeele Seltsi aastaraamat	271181958	128198672	94029909	99001358	1998/35
Eesti Kirjanduse Seltsi ...	171031644	319105725	99011759	99001359	1998/1997/27
Estonia	166595934	229205232	97013855	99001360a	1998/13/1
Estonia	519509110	799215414	6030119	99001360	2006/1
Estonian literary magazine usw.					

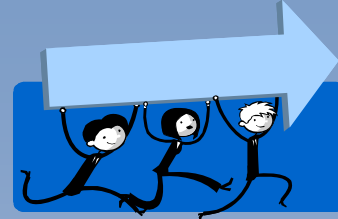
Retro-Projekt

2. Batch-Lauf:

Automatische Generierung aller in ACQ aufgeführten *und* bereits akzessionierter Hefte, beginnend ab dem Genannten



Weitere Planung



- Import in Pica/OLC-Datenbank
- Mehrplatzfähigkeit
- Update u. Neukonfiguration des Periodikaservers
- Statistische Auswertungen nach Erweiterung der Bearbeitungsstatus
- Erprobung des neuen Freiform-Modus
 - Konversion in OCR-Text
 - Indexierung durch manuelle Separierung Autor-Titel-Seite
- Vorlagen in nichtlateinischen Schriften
- Parallel-sprachige Inhaltsverzeichnisse



Weitere Planung



- Anpassung der Organisationsstrukturen
 - Qualitätssicherung ohne Heft bzw. Autopsie?
 - Scan-Merkmal ul=cca (Autopsie erforderlich)
 - Bearbeitungsmerkmal im Heft entbehrlich?
 - Verzicht auf separat geführte Excel-Tabelle





*Vielen Dank
für Ihre Aufmerksamkeit*

