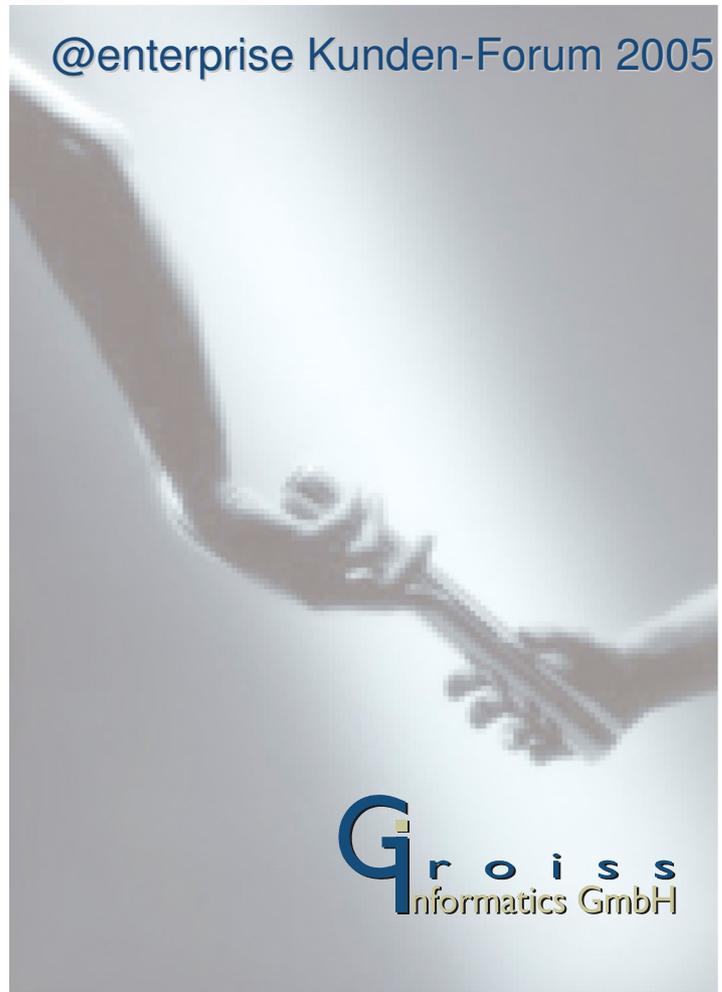




Workflow-gestützte Produktion von Rechtsinformationen bei LexisNexis

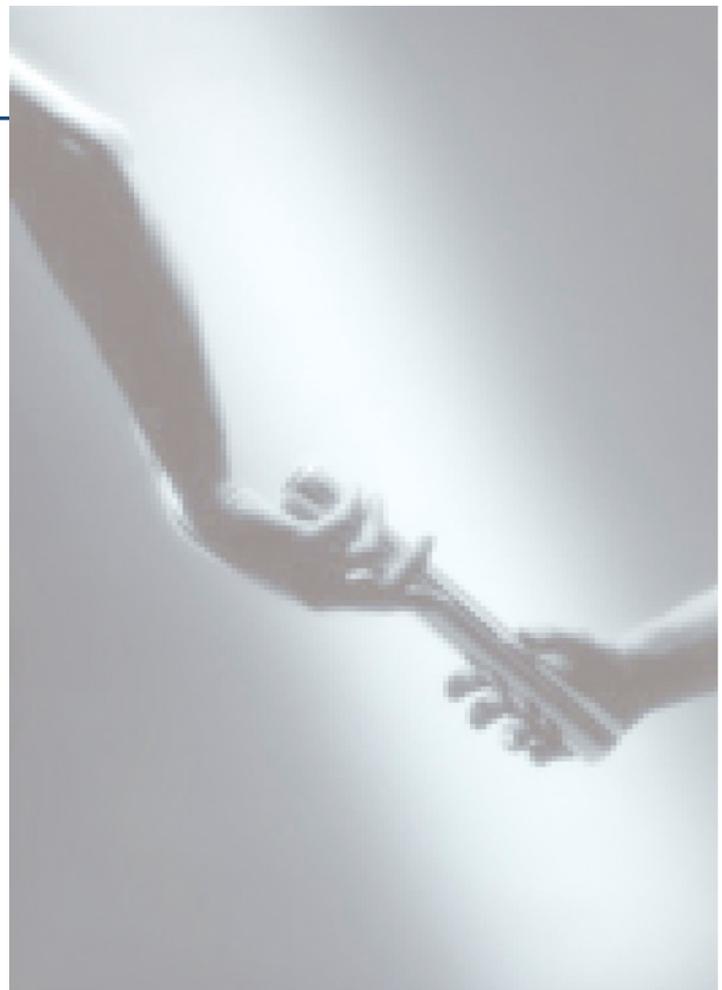
Michael Dobrovnik

@enterprise Kunden-Forum 2005



Inhalt

- ◆ Projektumfeld
- ◆ XML-Dokumentenstore
- ◆ Prozesse
- ◆ Integrationselemente
- ◆ Sicherheitsarchitektur
- ◆ Daten zur Installation



- ◆ MBO Verlag GmbH Deutschland (Marketing, Beratung, Organisation)
 - Mittelständischer Spezialverlag in Münster
 - ca. 100 Mitarbeiter (viele disloziert) ; 70000 Kunden
 - Dutzende Produkte in den Bereichen Recht, Wirtschaft, Immobilien
 - Produkt = speziell zusammengestellter, annotierter Ausschnitt des Bundes- und Landesrechts, zusätzlich annotiert und verlinkt + Spezialpublikationen
 - Eingebunden in Such- und Navigationstool (LEXsoft)

- ◆ LexisNexis Group Deutschland
 - Teil der LexisNexis Group
 - 13.000 Mitarbeiter weltweit
 - Jahresumsatz von mehr als rund 2 Milliarden Euro
 - mehr als 30 Jahre Erfahrung in der Bereitstellung von Online-Informationen
 - Tochter von Reed Elsevier, weltweit fünftgrößter Medienkonzern, umsatzstärkster Fachverlag
 - LNG Deutschland
 - 350+ feste Mitarbeiter
 - 400+ externe Autoren
 - 400+ freie Mitarbeiter / Bearbeiter

Projekt LexSoft 2000

- ◆ Damalige Situation
 - LexSoft war unter Windows und als CGI für UNIX verfügbar
 - Die Dokumente lagen in MS-Word und proprietären Formaten vor
 - 13000 Gesetze
 - 90000 Urteile
 - 25000 Stichworte
 - weitere Dokumententypen
 - tausende Aktualisierungen pro Jahr
 - Keine Automatisierung der Ablauforganisation

- ◆ Ziel
 - Modernisierung des LexSoft Tools (Java)
 - Umstellung der Dokumentenformate (XML auf der Basis eigener DTDs)
 - **Durchgängige elektronische Unterstützung der Abläufe für**
 - Formatumstellung
 - Aktualisierung und Wartung der Dokumente
 - *Produktzusammenstellung*

Elemente der Problemlösung

- ◆ XML-Dokumentenstore

- ◆ Prozesse

- ◆ Integrationselemente

- ◆ Sicherheitsarchitektur

XML-Dokumentenstore

◆ Grundanforderungen

- Speicherung der Daten im originalen XML-Format im Filesystem (vorgegeben durch bestehende Werkzeuge)
- Suchmöglichkeiten über relevante "Metadaten"
- "reservieren" von Dokumenten (Check-Out/Check-In)
- konsequente Versionierung (fachlich (i.d.Fassung vom) und technisch)
- Verwendung in Prozessen
- Berechtigungsvergabe pro Mitarbeiter und Dokumenttyp

XML-Dokumentenstore

◆ Suchfunktionalität

- Suchmöglichkeit nach allen Metadatenelementen
- Je nach Dokumenttyp (Gesetz, Urteil, ...)
- Flexible Zusammenstellung der Queries zur Laufzeit
- Abspeichern der Queries
- Anzeige der Files und der Metadaten

XML-Dokumentenstore

◆ Anzeigefunktionen

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
- <gesetz rang-texttyp="NULL" normgeber="Bund">
- <meta>
- <kategoriezuordnung>
  <offizielle-kategorie>2121-1-6</offizielle-kategorie>
  <mbo-kategorie>NULL</mbo-kategorie>
  <kunden-kategorie>2121-1-6</kunden-kategorie>
</kategoriezuordnung>
+ <produktzuordnung>
</meta>
- <kopf>
- <erstfassung>
  <normueberschrift>Approbationsordnung für Apotheker (AAppO)</normueberschr
```

MBO-Dokumentverwendung

Klassifikation:

Art	Name	Version	AktVers.	Akt.RS	Typ	Status	Datum	Prozess
G	AAppO.xml	1	ja	nein	Work			47
G	AAppO.xml	1	ja	nein	Work			62
G	AAppO.xml	1	ja	nein	Work		24-07-2000 18:29	37
G	AAppO.xml	1	ja	nein	Work	OUT	18-09-2000 14:53	20
G	AAppO.xml	1	ja	nein	Work		28-07-2000 00:04	60

© Groiss Informatics GmbH, 2005

11

XML-Dokumentenstore

◆ Indizierung der XML-Dokumente (XML-Shredding)

- Vorstudien zur Performance
- Festlegung relevanter Metadaten pro Dokumenttyp
 - Gesetze: DTD hat hunderte Elemente; Ca. 50 davon relevant
- Schemaentwurf als Netz von PersistentObjects (Gesetze:13 Relationen)
- Aufnahme:
 - Parsing
 - Extraktion der Metadaten
 - Konstruktion der Objektinstanzen
 - Speicherung der PersistentObjects
 - Einbringung ins Filesystem
- Verwaltungsoperationen
 - Löschen
 - Reindizierung (technisch)
 - Transformation (semantische Reindizierung)

© Groiss Informatics GmbH, 2005

12

XML-Dokumentenstore

- ◆ Verwendung von XML-Dokumenten in Prozessen
 - Eigene Klasse von Attachments
 - bloßer Verweis auf das Dokument
 - bei Bedarf "Materialisierung" des XML-Dokuments als @enterprise-Dokument
 - herkömmliche Transport- und Bearbeitungsmöglichkeiten in @enterprise
 - (versionierte) Übernahme eines Dokuments in den Dokumentenstore

XML-Dokumentenstore

- ◆ Verwendung von XML-Dokumenten in Prozessen

The screenshot shows a software interface with a task list and a detailed view of documents for a specific process. The task list is titled "Arbeitskorb von Michael Dobrovnik" and includes columns for Id, OE, Prozess, Task, Betreff, Dokumente, LNDDok, and Erhalten. The detailed view is titled "MBO-Dokumente im Prozess 55" and includes a "Klassifikation" dropdown set to "alle" and a table of documents.

Id	OE	Prozess	Task	Betreff	Dokumente	LNDDok	Erhalten
<input checked="" type="checkbox"/> 55	MBO	LNKAus	KommKommentieren	AO_0002.xml	KAForm		15-02-2005 15:07
<input type="checkbox"/> 56	MBO	LNKAus	KommAufbereiten	LNK_1_AltTZG.xml	KAForm		14-02-2005 10:32
<input type="checkbox"/> 57	MBO	LNKAus	KommAufbereiten	LNK_10_AltTZG.xml	KAForm		14-02-2005 11:09
<input type="checkbox"/> 58	MBO	LNKAus	KommAufbereiten	LNK_12_AltTZG.xml	KAForm		14-02-2005 13:11
<input type="checkbox"/> 59							
<input type="checkbox"/> 60							

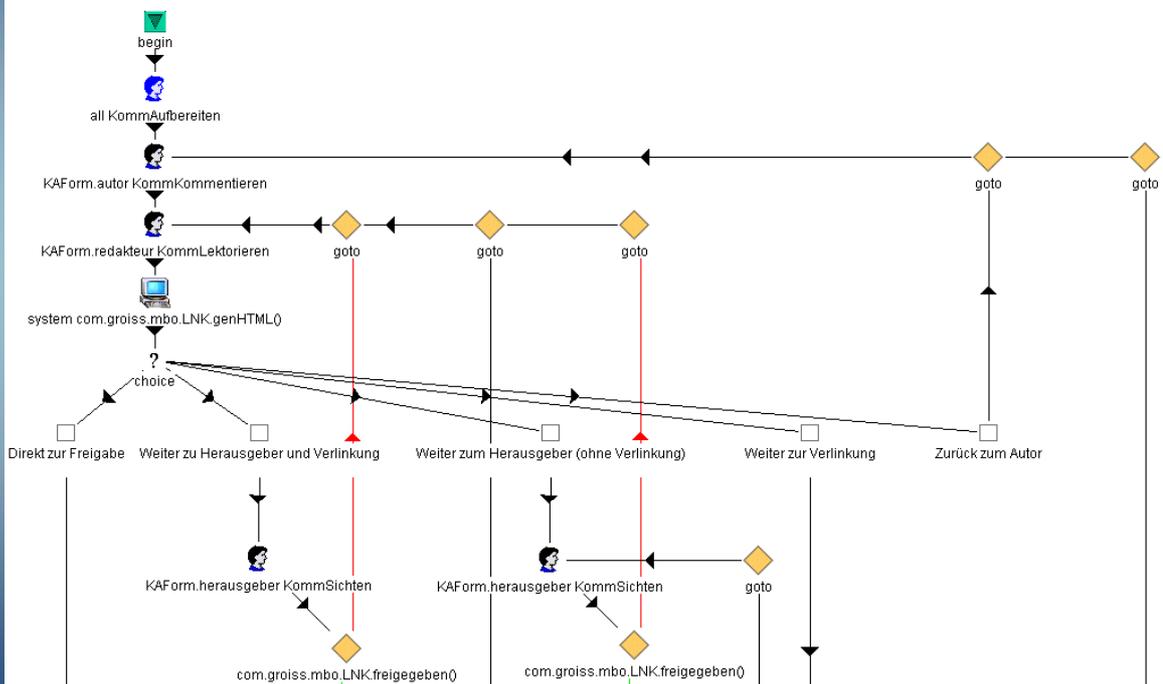
Art	Name	Version	AktVers.	Akt.RS	Typ	Status	Datum	Benutzer	Größe
<input type="checkbox"/> J	AO_0002.xml	1	ja	ja	Work	OUT	12-02-2005 20:33	55	3592
<input type="checkbox"/> G	BGB.xml	1	ja	ja	Work	IN	25-09-2005 10:19		4358043

Prozesse

- ◆ Verschiedene Abläufe je nach Dokumententyp
- ◆ Unterscheidung nach Neuaufnahme / Überarbeitung
- ◆ Wesentliche gemeinsame Elemente
 - Hierarchie von Bearbeiter / Supporter / Spezialist
 - automatische Überprüfung der XML-Texte
 - Aufruf von externen Funktionen (System / Batch-Schritte)
 - Generierung abgeleiteter Dokumente
 - Weitere Überprüfungen
 - weitestgehende Sicherstellung der Dokumentenqualität
 - "externe" Bearbeiter
- ◆ Weitere Prozesse zur Unterstützung der Produktion (CD-Pressen, ...)

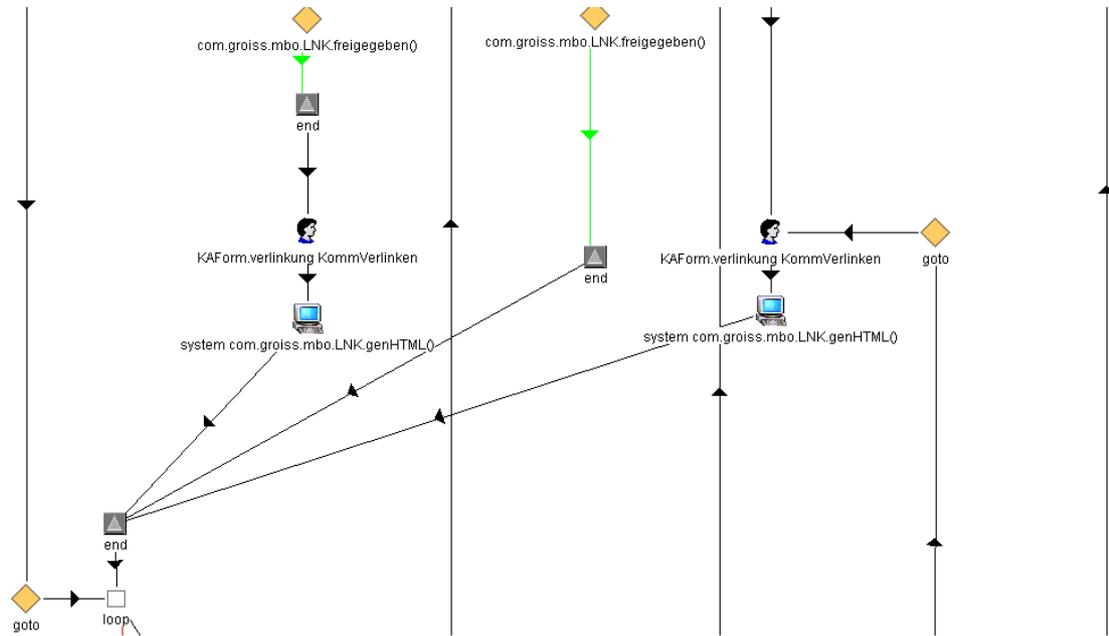
Prozesse

◆ Beispiel: Basiskommentare



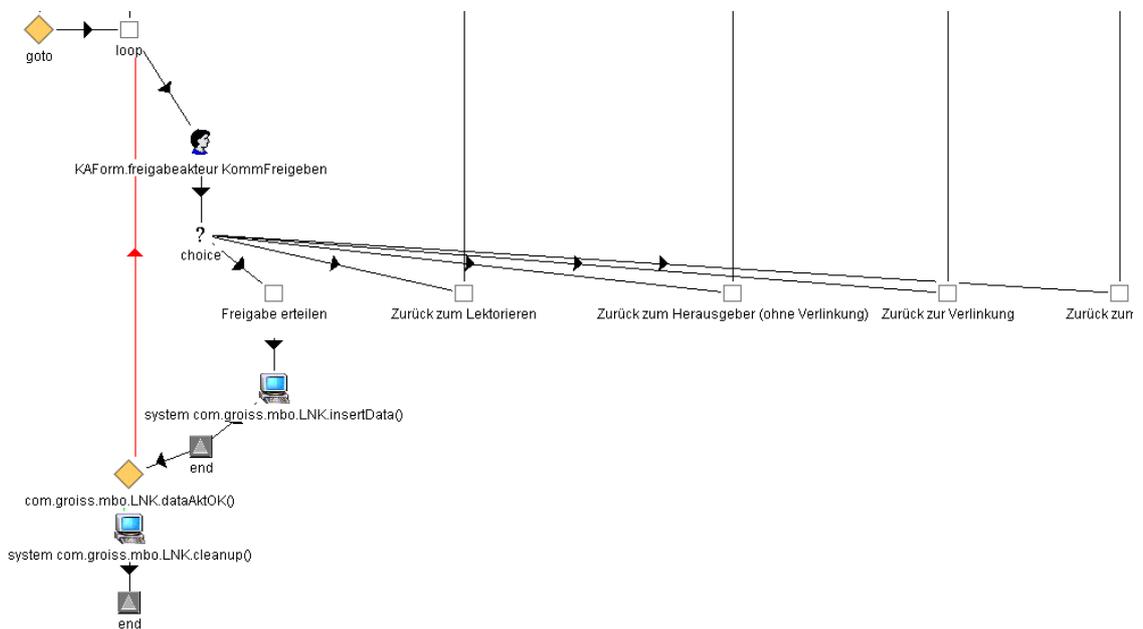
Prozesse

◆ Beispiel: Basiskommentare (II)



Prozesse

◆ Beispiel: Basiskommentare(III)



Weitere Integrationselemente

- ◆ Kommunikation über DB
 - Insert löst Prozessstart aus
 - Prozesse liefern Abrechnungsdaten
- ◆ Systemschritte / Batch-Jobs
 - externes XML-Parsing
 - XSLT
 - LND-spezifische Tools
- ◆ Zugriffe auf das System durch "Bridge"
 - Kombination aus (externem) Servlet und RMI-Client
- ◆ Mail-Schnittstelle
 - Einbringen und Löschen von bestimmten Dokumententypen
- ◆ *XML-Dokumentenstore*
 - *als RMI-API*

Sicherheitsarchitektur

Istsituation (Mitte 2004)

- ◆ LND betreibt einen @enterprise Workflow Server (WFS)
 - (plus 2 Testserver, hier nicht weiter relevant)
- ◆ Als DB kommt MS-MSQLServer 2000 zum Einsatz
- ◆ Der WFS greift auf Netzwerklaufwerke (Dokumentenstore / Samba-Server) zu
- ◆ Es kommen HTTP/HTTPS-Clients zum Einsatz
 - Aus dem Intranet
 - **Aus dem Internet**
- ◆ keine Trennung der Netzwerke

Sicherheitsarchitektur

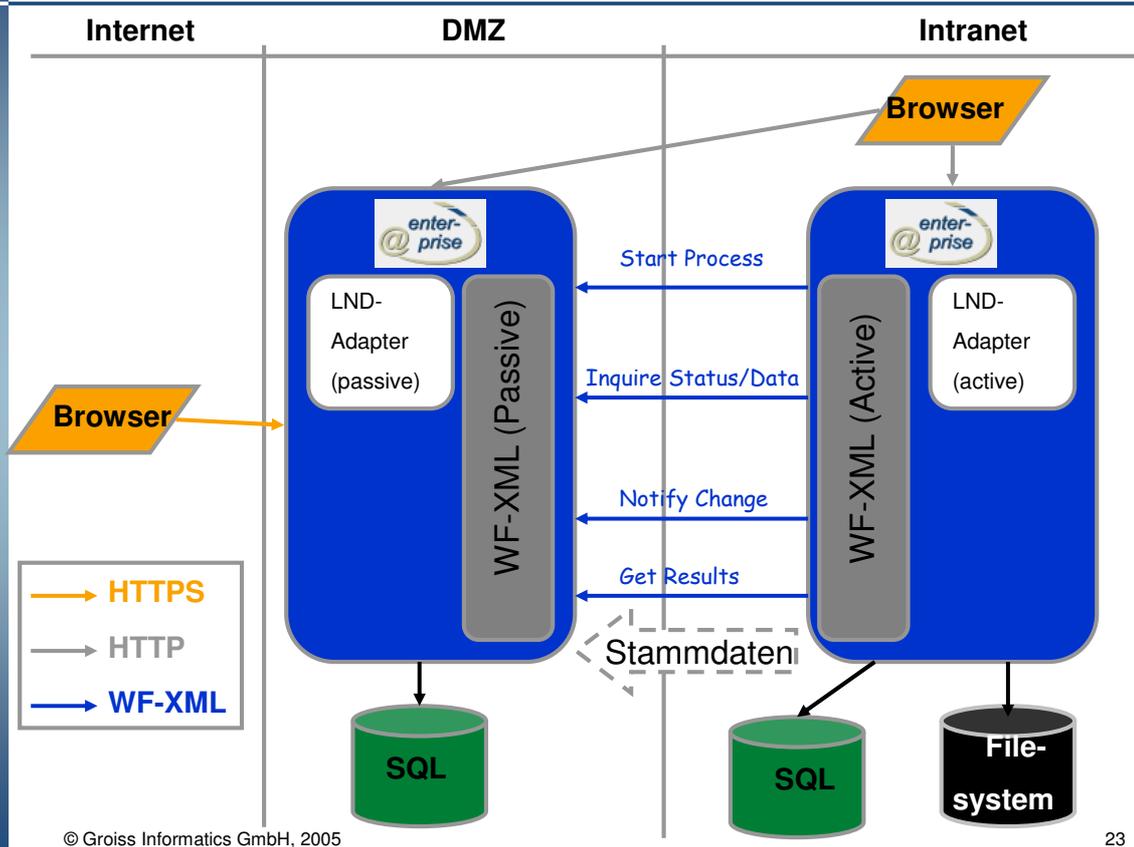
Neue verbindliche Sicherheitsrichtlinien ab 10/2004:

- ◆ Es sind konzernweite Sicherheitsrichtlinien zu beachten:
 - Kein Zugriff aus dem Internet auf Server
 - Ausnahmen: spezielle Protokolle/Port in eine DMZ können offen bleiben
 - Aber: **kein Zugriff aus der DMZ aufs Intranet**
- ◆ Ausnahmen:
 - Spezielle von LND zur Verfügung gestellte Clients (HW/SW vorkonfiguriert, mittels VPN)
 - Nur für einige wenige Mitarbeiter vorgesehen (Kostengründe/Adminaufwand)
- ◆ **Bisheriger Zugang zum WFS durch externe Mitarbeiter nicht mehr so aufrecht zu erhalten.**
- ◆ **Aber: Arbeit mit dem WFS muss trotzdem auch von externen Mitarbeitern aus dem Internet weitgehend gleich möglich sein**
- ◆ **Forderung (an uns selbst): möglichst kein Prozessredesign durch die Sicherheitsanforderungen**

Lösungsalternativen

- ◆ Außerkraftsetzung der Sicherheitsrichtlinien
 - Nicht machbar
- ◆ Externen Zugriff unterbinden
 - Nicht machbar
- ◆ Platzierung des WFS in der DMZ
 - würde zu einer Platzierung des SQLServers und des Fileservers in der DMZ führen → ist nicht machbar
 - → Einzige Möglichkeit: Platzierung des WFS im Intranet
- ◆ **Konsequenz: Logische Teilung des Systems**
 - Anstatt eines WFS werden 2 WFS eingesetzt
 - Der bisherige Server wandert ins Intranet **ohne Konnektivität von außen**
 - Ein zweiter Server wird in der DMZ platziert und ermöglicht den Zugang von außen

Skizze der Lösungsarchitektur



Beschreibung der Lösungsarchitektur

- ◆ Bisheriger WFS wird im Intranet platziert (Interner WFS, IWFS)
 - Ist von außen nicht erreichbar
 - Ist von innen erreichbar (HTTP)
 - Ist von der DMZ nicht erreichbar
 - Kann auf Server in der DMZ zugreifen (IWFS)
 - Kann auf Mailserver zugreifen (senden SMTP / lesen IMAP)
 - Kann auf SQL-Server und Samba-Shares zugreifen
- ◆ Weiterer WFS in der DMZ (Externer WFS, EWFS)
 - Ist von außen über HTTPS erreichbar
 - Ist von innen über HTTP erreichbar
 - Hat eigenen SQL-Server in der DMZ
 - Hat keinen Zugriff auf Server im Intranet (IWFS, Samba)
 - Kann Mails senden (SMTP)

Beschreibung der Lösungsarchitektur

- ◆ Kommunikation zwischen den WFS
 - Alle Aktionen werden vom IWFS ausgelöst
 - Start von Prozessen
 - Statusabfrage von Prozessen
 - Abbruch von Prozessen
 - Abholen von Ergebnissen
 - Verwendetes Protokoll
 - WF-XML
 - Erweitert um die Modi Aktiv / Passiv

- ◆ Prozessdesign:
 - Am IWFS: wie bisher / keine (kaum) Änderungen
 - Am EWFS: im wesentlichen Miniprozesse mit einem einzigen Schritt

Beschreibung der Lösungsarchitektur

- ◆ LND-Adapter Passiv:
 - Verarbeitet eingehende WF-XML-Messages
 - Reagiert nur auf Anforderungen durch den IWFS
 - Start der Miniprozesse
 - Abfrage des Status der Miniprozesse
 - Abbruch der Miniprozesse
 - Bei Fertigwerden der Prozesse:
 - Keine direkte Kontaktaufnahme zum IWFS
 - In eine Warteschlange zur Abholung durch den IWFS bereitstellen

Beschreibung der Lösungsarchitektur

◆ LND-Adapter Aktiv:

- Veranlasst Kommunikation mittels WF-XML
- Wird periodisch von einem Timer angestoßen
 - Ermittelt, ob Prozesse am EWFS zu starten sind, startet diese (Formulare / @enterprise-Dokumente / Notizen)
 - Holt die Ergebnisse aus der Warteschlange des EWFS ab, aktualisiert die Daten / Dokumente der Prozessinstanz am IWFS und leitet die Instanzen weiter
- Kann über Link aus der Prozesshistorie am IWFS ggf. den Zustand des korrespondierenden Miniprozesses am EWFS ermitteln
- Kann Prozesse am EWFS abrechnen

◆ Konfiguration

- Angabe, welche Kombinationen von Prozessen / Tasks / Rollen für den EWFS relevant sind
 - Kennzeichnung von Benutzern/Rollen als extern
 - Durch Zuordnung eines Rechts

Beschreibung der Lösungsarchitektur

◆ Stammdaten-Abgleich zwischen IWFS und EWFS

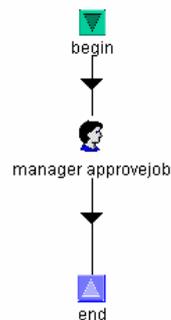
- bestimmte Stammdaten werden nur am IWFS gepflegt und zum EWFS gespiegelt
- Benutzer / Rollen / Rechte / Funktionen / Tasks / Objektklassen / Formulare / Applikationen / OEs / OE-Klassen
- Ausgelöst durch Administrator und / oder Timer am IWFS
- Erstellung eines Export-Files
- Import dieses Files am EWFS

Konsequenzen der Lösungsarchitektur

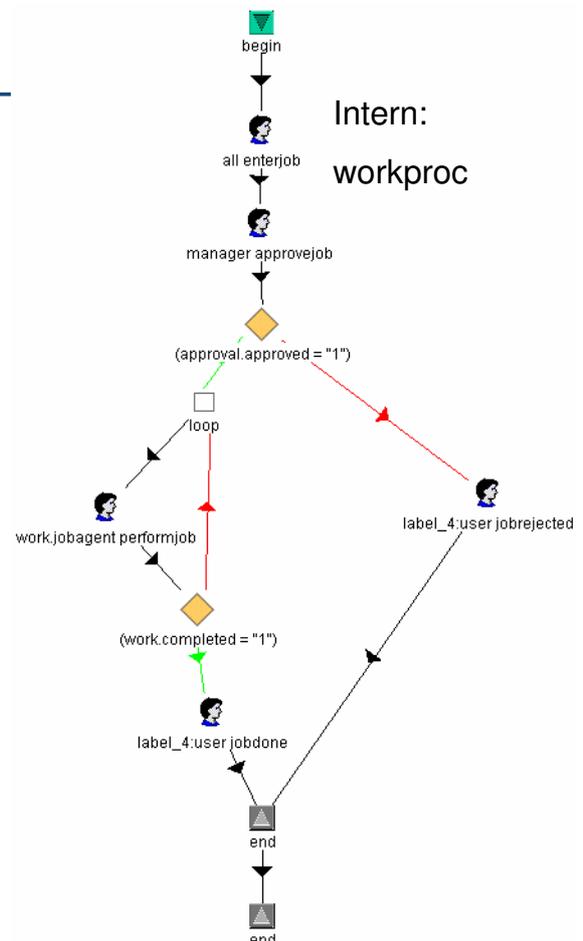
- ◆ Sicherheitsauflagen können erfüllt werden
- ◆ Bisheriger WFS / bestehende Prozessdefinitionen können im wesentlichen unverändert übernommen werden
- ◆ Zusätzlicher administrativer Aufwand
 - Betreuung von zwei Systemen
- ◆ Einschränkungen der Funktionalität
 - Zurückgehen
 - Keine aussagekräftige historische Sicht für externe User

Prozesse

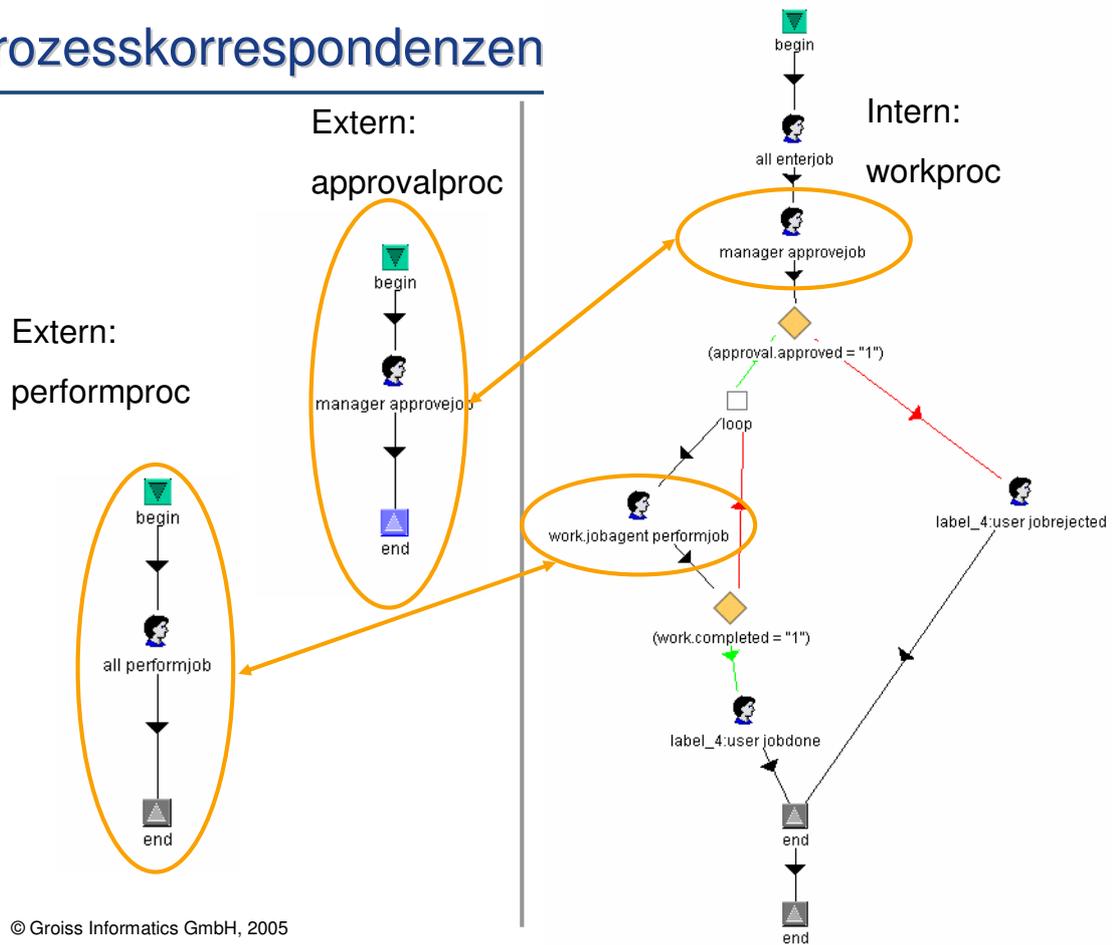
Extern:
approvalproc



Extern:
performproc



Prozesskorrespondenzen



Spezifikation der Prozesskorrespondenzen

- ◆ Auf Prozesstypenebene über XML-Mapping-File
 - Abbildung der Tasks der Internen Prozesse auf externe Prozesse
 - Möglichkeit der Angabe von Partnern
 - Einschränkungen auf Formularebene
 - Einschränkungen auf Dokumentenebene
- ◆ Auf Agentenebene durch Zuordnung eines Rechts "Extern"
 - Direkt oder über Rolle
- ◆ Prozess wird "externalisiert", wenn
 - Prozess/Taskkombination in XML-Mappingfile angegeben ist und
 - der Agent das Recht "Extern" hat

Spezifikation der Prozesskorrespondenzen

◆ <?xml version="1.0"?>

<ExternTasks>

<Process appl="wfxmldemo" id="workproc" version="any">

<Task id="performjob">

<RemotePartner id="ep62extern"/>

<RemoteProcess appl="wfxmldemo" id="performproc" version="0"/>

<Input>

<ProcessForm id="work"/>

<DMSDocuments>

<FormType id="work" version="1"/>

</DMSDocuments>

</Input>

</Task> (*)

</Process> (*)

</ExternTasks>

Interner
Prozess

Interner
Task

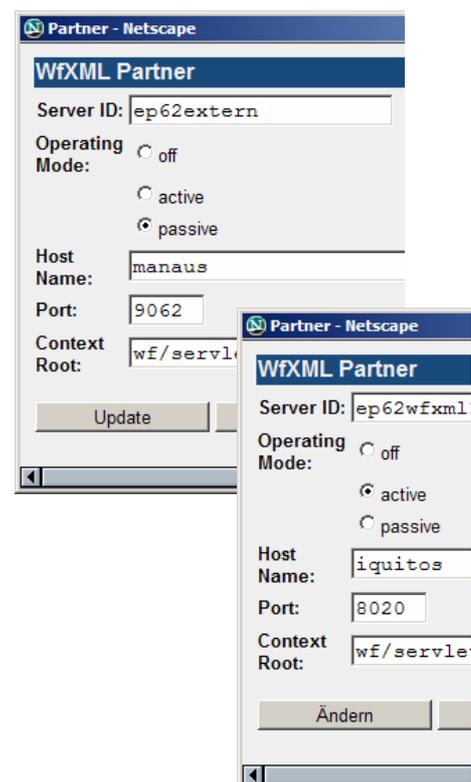
Externer
Prozess

Formular
(mehrere)

Dokument-
Typen

Konfiguration der Systeme

- ◆ WF-XML Aktivieren
 - WFXML aktivieren: aktiv/passiv/off
 - WFXMLTask Aktivieren
 - Partner (extern/intern) Eintragen
- ◆ Zusätzlich am aktiven System
 - Administration / Timer
 - wfxmlactive-timer aktivieren
 - → **LND-Adapter Aktiv**
- ◆ Mapping-File schreiben
- ◆ Stammdatenabgleich



Weitere Aspekte der Architektur

- ◆ Prozessstart mit Rolle
- ◆ Externer Prozessstart und weiterlaufen als interner Prozess
- ◆ Statusinfos
 - technisch für Administrator
 - Historie für Anwender

	MBO	if (PUForm.presswerk = "1") true	15.0
			13.3
V		MasterCDVersenden	15.0
H	MBO	Preprocessing: de.lexisnexis.wf.ProUpdate.setDueDate("duedate")	13.3
J		PresswerkEingang	15.0
V	MBO	Preprocessing: de.lexisnexis.wf.ProUpdate.resetAuftragsb() Postcondition: de.lexisnexis.wf.ProUpdate.auftragsbErhalten()	14.0

Remote process instance: id = 120358, role = Lager

- ◆ Berechtigungsprüfung für URL(-Präfixe) über URL-Checker
 - Kombination von URL * IP * Recht | Deny bestimmt über Zulässigkeit
 - Verbot der Administration von außen

Zur Installation

- ◆ Systemkonfiguration
 - jeweils: WindowsServer2003 – SQLServer 2000 - @enterprise 6.2
 - 2 x Intel Xeon 3 GHz (extern 2,8 GHz)
 - 3,5 GB RAM (extern 2,5 GB RAM)
- ◆ Mengengerüst
 - Größe der Datenbank: 32GB
 - Dokumentzahlen ca.:

43.000 Gesetze * 2 Versionen	3.000 Handbücher
20.000 Juristische Kommentare	3.000 Rundschreiben
41.000 Aktuelles * 2,5 Versionen	3.000 Verwaltungsanw.
70.000 Stichworte * 1,5 Versionen	37.000 Artikel
350.000 Urteile	
 - in Summe 700.000+ Dokumentversionen zu je 25 KB = 17 GB Gesamtgröße
 - Prozessinstanzen

Jahr	intern	extern
2004	21.000	19.000
2005	35.000	71.000

Zur Installation

◆ Personal:

- Administrativ
 - mehrere fachliche (Neben-) Administratoren (1st Level Support)
 - ein Hauptsystemadministrator (2nd Level Support)
 - *Groiss Informatics (3rd Level Support)*

◆ Development:

- LexisNexis Deutschland: (2-3 Personen)
 - zunehmend eigenständige Adaption des Systems
 - Entwicklung von Teilkomponenten
 - kooperative Prozessdefinition
- Groiss Informatics: (1-2 Personen)
 - Erstellung Grundsystem
 - Mitwirkung / Durchführung bei Analyse / Design / Implementierung neuer Funktionen und Aspekte

Zur Installation

◆ Weitere Aktivitäten

- Ausrollung bei LexisNexis AT (vormals : ORAC Verlag)
 - teilweise andere Basissysteme (ORACLE/Linux)
 - Refactoring:
 - Trennung allgemeine / landesspezifische Teile
- Ggf. Wechsel des DBMS bei LexisNexis DE