Integration of a DAU-System in Workflow Management Systems

HEIKO MAUS
German Research Center for Artificial Intelligence (DFKI GmbH)
Overview

• DAU (Document Analysis and Understanding)
• Project VirtualOffice
• Synergetic effects
• Application scenario and example
• Outlook
Document Analysis

- Extracting information from paper sources
- Translation into a structured representation

Steps:
- Structure analysis: e.g. filtering, page segmentation
- Text recognition: OCR (Optical Character Recognition)
Understanding

Gaining contained information and semantic

Processes:
• Message type classification
• Parsing (e.g. extraction of sender)
• Pattern matching (e.g. recognising typical text phrases)
DAU: Benefits

Support in filtering, sorting, information extraction, data capturing

Applications:
• Payment document processing (e.g. bank drafts, checks)
• Forms processing (e.g. tax forms, insurance)
• Mail sorting (by address field extraction)
VirtualOffice

- DAU with flexible specialists
- Planning and learning concept
- Integration in commercial WFMS
- Use of the current context of the WFMS
- Generic approach
- Business letters as domain
DAU-system

- Specialists: each performing a specific DAU process
- Reactive Planner:
  - plans effective use of specialists
  - invokes specialists

invoked by:
- in-tray-server (incoming document)
- workflow activity (DAU-task request)
DAU-system

use of:

• Expectations stated by the workflow activities waiting for a document
• Document knowledge, e.g.
  – structure of an invoice
  – structure of an invoice of a specific company
• Corporate knowledge (addresses, product lists, ...
DAU/WFMS BrokerService

- Mediates between DAU and WFMS
- Implementation with CORBA (Common Object Request Broker Architecture)
DAU/WFMS-Broker Architecture

DAU-system

Reactive Planner Specialists

DAU/WFMS BrokerService

Expectations

WFMS A

Workflow Engine

BrokerClients

WfMC IF 3

Interface

WFMS A

WfMC IF 2

BrokerModules

API A

Workflow Instances

CORBA Communication via ORB
Synergetic Effects

- Integration of the paper-based part in the WFMS
- DAU support by knowledge from the WFMS
- Assignment of documents to the respective business process
- Support of the data capturing
- Flexible and enhanced use of DAU in the WFMS
Scenario

- Purchasing process
- Documents: invoice, delivery note, offer, ...
- Media: paper, fax, e-mail
- WFMS

Tasks:
- Assign incoming documents to waiting workflows
- Data capturing and transfer to workflow
Example

Document knowledge

Expectations

<table>
<thead>
<tr>
<th>EID</th>
<th>DAUtask</th>
<th>processID</th>
<th>activityID</th>
<th>sender</th>
<th>message type</th>
<th>orderDate</th>
</tr>
</thead>
<tbody>
<tr>
<td>456</td>
<td>GetInvoiceNo</td>
<td>3011</td>
<td>10</td>
<td>VAT</td>
<td>invoice</td>
<td>1998-10-10</td>
</tr>
</tbody>
</table>

Corporate knowledge

- product lists
- addresses

Expectations

- type: invoice
- sender: VAT
- reference: order, 1998-10-10
- DAU-task (GetInvoiceNo): 4315/67

Reactive Planner

PurchaseTecEquipment

ProcessID: 3011

9. Fax order
10. Waiting for invoice
11. Verify data
DAU-task Request

DAU-task (GetProductList):
1. Computer VAT3 3415.67
2. Monitor 21" 2700.00
Outlook

Implementation
• Staffware 97 Staffware plc, UK
• COSA Workflow, Software Ley, Germany

Goal
• Implementation of WfMC Interfaces
• CORBA Workflow Management Facility
• CORBA Vertical Domain Facility