

REINVENTING BUSINESS WITH CONTENT AND COLLABORATION TECHNOLOGIES

TRANSFORM™

MAGAZINE

www.transformmag.com

Vol.13, No.3 March 2004

Blueprint for Success:

**Build Nimble Content Applications
With Service-Oriented Architectures**

Cut the Cost of Paying Invoices
Automate the Last Bastion of Manual Data Entry

RIIM 2004 New Product Preview
A Sneak Peak at the Biggest Content Expo

ECM Suite Taps 'HDD WORM'
What Once Was a Paradox Now Is a Reality

콘텐츠 전달의 미래

유연하고 반복 가능한 표준 기반 웹서비스 및 웹 지향구조가 콘텐츠의 관리, 통합, 전달 방법을 변화시킬 것이다.

글 / 빌 트리페

소프트웨어 개발 접근법이 뜨고 진다. 오늘날 가장 유망한 접근법은 한창 급부상하고 있는 추세가 꺾인 후 그 주기가 반복되고 있다. 먼지가 가라앉자, 기업들은 시스템과 애플리케이션을 생산성 있고 효율적으로 만들기 위해 전력을 다해 일하고 있다.

웹서비스, 보다 광범위하게 서비스 지향구조(SOA)가 빠르게 유력한 소프트웨어 개발 및 통합 접근법이 되고 있다. SOA는 표준화된 XML 기반 메커니즘을 통해 사용할 수 있도록 만들어져 네트워크화된 소프트웨어 및 콘텐츠와 소프트웨어를 통합하는 방법으로 가장 잘 이해된다. 이 접근법에서 소프트웨어 서비스는 다른 서비스와 쉽고 공개적으로 통합될 수 있다. 일례로 환전과 텍스트 번역, 그리고 문서 형식 변환을 제공하는 소프트웨어 서비스를 유연하게 활용할 수 있다.

SOA의 매력은 애플리케이션과 데이터의 '느슨한 결합'을 향한 움직임과 밀접한 연관이 있다. 더 많은 조직에서 콘텐츠와 애플리케이션을 웹으로 내보냄에 따라, 소프트웨어 개발자들은 브라우저에 의해 제공되는 '씬 클라이언트(thin client)'로 콘텐츠와 애플리케이션을 유도할 빠르고 효율적인 수단을 모색해야 하는 부담을 안게 됐다. XML이 개발자들에게 개방된 애플리케이션간의 데이터 전송수단을 제공했듯이 SOA 역시 소프트웨어 커뮤니케이션과 인터넷을 통한 다른 소프트웨어 제어를 위한 개방된 수단을 제공한다.

본격적인 연구: 유연성의 필요

콘텐츠 관리 측면에서 보면 SOA야말로 변화에 대처해 기업에게 더욱 많은 유연성을 제공한다. 점차 증가하는 사업 운영 효율성을 지켜본 퍼시케어(PaciCare, 캘리포니아주 산타 아나 소재)는 담당 직원을 보강하거나 기술 투자를 늘려야 한다는 사실을 깨달았다. 미국 최대의 보건의료 관리업체로서, 이미 매년 수백 건의 클레임과 등록서류를 처리하고 있는 터였다.

이 업체는 서류작업의 수동처리를 생략하려는 목적으로 SOA 기반 솔루션을 채택기로 최종 결정했다. 캘리포니아주 어바인에

본사를 둔 파일넷(FileNet)의 P8 플랫폼 상에 구축된 이 솔루션 덕분에 웹서비스를 사용해 기존 시스템의 통합과 속도배치를 용이하게 해줬다. 이 시스템은 9주 이내에 본격 가동됐으며, 결과적으로 클레임 및 등록절차가 몇시간에서 몇분 단위로 빨라졌고 운영 생산성도 전반적으로 20% 향상됐다고 한다.

파일넷의 톰 도이치 제품 관리자는 "SOA 접근법의 빠른 실행 능력 덕분에 퍼시케어는 고전적인 딜레마에 대항할 수 있었다"고 말했다. 사업을 성장시키는 동시에 고객을 잘 관리하고 운영 비용을 고정 내지 절감시킬 방법은 무엇인가? "SOA는 인수합병 등의 기업 이벤트에 보다 민첩하게 대처할 수 있는 동적인 방법"이라는 게 그의 설명이다.

사실 SOA 기반 접근법을 통해 조직은 신뢰성 있는 기존 시스템의 유지 및 통합 능력을 포함한 각기 다른 동종 최고의 솔루션을 자유롭게 선택할 수 있다. 이 같은 유연성은 조직이 이질적인 데이터 소스와 각기 다른 의사소통 채널을 운영하는 데 필요한 요건에 직면하는 콘텐츠 관리에 있어 특히 중요하다. 비즈니스 커뮤니케이션 분석기관인 CAP 벤처의 존 프랭크 조사 분석관은 "다채널 의사소통에 관한 연구결과 기업의 실무자들은 소프트웨어 스위트에 걸쳐 4:1의 마진으로 동종 최고의 기술을 선호하는 것으로 나타났다. 이는 유연성뿐 아니라 확장성과 보안성에 있어서의 우위를 바탕으로 서비스 구조가 주류 문서 관리 및 콘텐츠 관리 프로젝트의 관심을 끄는 척도가 될 수 있다"고 말했다.

거의 모든 주요 기업 콘텐츠관리(ECM) 공급업체들이 제품에 이미 SOA 지원을 추가했거나 추가하는 중이라고 업계 전문가인 프랭크 질베인은 전했다. 블루벨 어드바이저의 CEO 겸 질베인 보고서 편집자이기도 한 그가 앞으로 12~18개월 안에 웹서비스가 콘텐츠 전달에서 중요한 역할을 할 것이라고 단언하는 데는 두 가지 이유가 있다.

우선 기업 정보와 콘텐츠 통합 능력의 부족으로 인해 많은 기업 업무 솔루션이 제약을 받고 있기 때문에 기업들은 복잡성을 줄이고 스케일을 제공하기 위해 웹 서비스를 찾고 있다. 둘째,



코드 호환성에 덜 의존하는 콘텐츠 중심 서비스는 완전히 성장한 웹서비스에 비해 실행이 덜 복잡하기 때문에 조만간 배치될 것이다. 게다가 콘텐츠 서비스는 보안이 문제가 되는 부분에서 덜 위험하다는 것이다.

개발자들이 SOA를 선호하는 이유

SOA를 지지하는 여러 가지 이유가 있겠지만, 개발자들은 보다 실용적이며 선호하는 동기도 즉각적이다. 말하자면 이들은 솔루션을 실행하는 가장 빠른 경로인데, 웹 기반 애플리케이션의 경우가 특히 그러하다.

보스톤에 본사를 둔 애플리케이션 개발 및 통합업체 퀴인(Quoin)의 브래드 케인 사장은 주로 SOA 기반 접근법의 실용적인 가치를 본다. 케인은 17가지의 각기 다른 계정 정의를 갖고 있는 한 회사를 예로 들며, 대규모 조직은 본질적으로 다른 이종 데이터 소스를 보유하고 있다고 지적한다. 전통적인 소프트웨어 설계 및 개발 접근법의 경우, 조직은 데이터의 합리화 및 일반화로부터 시작하는데, 복잡하고 부담스러운 작업이라는 것이다.

“원스타나 워싱턴 포스트를 비롯한 클라이언트와의 콘텐츠 관리 관련 작업에서, 퀴인은 데이터 모델과 애플리케이션에 관한 즉각적이고 완벽한 결정을 내릴 수 있도록 해주기 때문에 SOA 접근법을 옹호한다”고 케인은 말했다. 보다 전통적인 설계 접근법에서는 사업 분석가와 소프트웨어 개발자들이 현재의 애플리케이션에 의해 제기된 것보다 장기적이고 추상적인 의문을 고려해야

한다는 설명이다.

또 신속한 개발과 민첩성이 핵심이며, 가용한 서비스를 보유하고 있으면 개발자들은 신속하게 애플리케이션들을 결집시킬 수 있다. 콘텐츠 자체를 하나의 서비스로써 활용할 필요가 있다는 말은 가치가 있다. 소프트웨어 개발자들은 콘텐츠 객체를 쉽게 식별하고 액세스하며 조작할 수 있는 방식으로 콘텐츠를 노출시키기 위해 콘텐츠 관리기술을 필요로 한다.

통상 콘텐츠를 서비스로 활용하는 데는 두 가지 애플리케이션 중 하나가 개입된다. 콘텐츠 자체는 액세스와 조작이 용이한 하나의 서비스로 쉽게 사용토록 만들 수 있으며, 구성 소프트웨어 조각을 별도로 이끌어낼 수도 있다. 이렇게 되기까지 파일넷을 비롯한 공급업체들은 콘텐츠를 다른 애플리케이션 및 프로세스와 연결할 수 있는 통일 콘텐츠 객체 모델을 강조해왔다. 인터워벤을 비롯한 또다른 쪽에서는 오픈디플로이(OpenDeploy) 기술을 사용해 이미 만들어진 구성 요소로 활용할 수 있고 ‘합성 애플리케이션’으로 결합할 수 있는 다양한 콘텐츠 관리 기능을 만들었다.

메사추세츠주 콩코드에 본사를 두고 둔, 브레드스트리트, 휴톤 미플린, 월터스 클루어 등의 고객업체에 콘텐츠관리 솔루션을 제공하는 SI(시스템통합)업체 칸다소프트(Kandasoft)의 알렉스 칼포브스키 사장은 “웹서비스는 앞으로 기업 콘텐츠 관리 시스템 어디에나 사용될 것이다. 일례로, 유연한 포매팅, 업무 및 워크플로우 규칙에 따라 특정 저장소에서 다른 저장소로 콘텐츠를 제공하는 웹 서비스가 생길 것”이라고 말했다.

‘SOA 접근법을 통해 데이터 모델 및 애플리케이션에 대한 즉각적이고 완벽한 결정을 내릴 수 있다’



퀴린의 브래드 케인 사장

웹 서비스를 통한 발행·배포

비교적 간단한 콘텐츠 지향 웹서비스의 예로 문서 발행 및 배포를 들 수 있다. 렉시스넥시스(LexisNexis) UK는 기업의 잘 알려진 법적 공지 및 세금 공지를 유연하게 발행, 배포하기 위해 Xy엔터프라이즈(메사추세츠주 소재)의 문서 합성 엔진 ‘XPP from Reading’의 웹서비스 인터페이스를 개발하고 있다. 웹서비스를 통해 사용자는 위치에 구애받지 않고 마이크로소프트 워드 파일을 업로드한 후 자동으로 최상의 고품질 조판 문서로 변환할 수 있다. 이 서비스를 통해 다음과 같은 일련의 자동 공정을 생략할 수도 있다.

- 워드 파일을 XML로 변환
- XML을 DTD에 대해 분석
- XML을 XPP로 전달
- 타이포그래픽 스타일링 및 페이지 매기기 적용
- PDF 파일 자동 생성
- PDF 파일을 웹 서버로 전송

렉시스넥시스 UK의 필 케이슬리 편집 시스템 매니저는 “웹 서비스 솔루션이 XML 기반 퍼블리싱 프로세스에 만능성을 부여하고 있는데, 원격 위치에서 작업하는 편집자들의 경우 특히 그렇다”고 말했다. 이 시스템을 통해 복잡한 인쇄를 적시에 비용 효율적인 방법으로 생성할 수 있다고 한다. 비용상의 이점 외에도, 케이슬리는 이 솔루션을 통해 공급자 기반을 합리화시키고 운영을 간소화하며 턴어라운드 시간을 단축하는 데 도움을 줄 것으로 기대했다.

콘텐츠 서비스를 갖춘 스페닝 애플리케이션

단순한 콘텐츠 전달의 예에서 본 이점이 인상적이고 분명한 반면, 진짜 본전을 뽑을 수 있을만한 SOA의 가치는 업무공정의 통일 및 자동화다. 호텔 체인인 힐튼인터내셔널은 인터워벤의 콘텐츠관리 시스템을 실행해 개별 호텔 소유주 및 관리자들이 콘텐츠를 편집, 관리할 수 있도록 하는 한편, 전반적인 힐튼의 브랜드와 사이트 콘텐츠는 중앙에서 관리했다.

이러한 분산 콘텐츠 생성 및 관리 접근법은 인터워벤의 SOA

기반 콘텐츠 네트워크가 있어 가능했다. 팀 데이비스 힐튼 유통 및 전자상거래 수석부사장은 “SOA 가용 콘텐츠관리 덕분에 확보된 유연성은 90년대 중반 ERP가 그랬던 것처럼 우리의 사업에서 중요한 자리를 차지할 것이다. 당장 고객의 요구를 충족시키기 위해서는 ECM이 필수적이라고 생각하며, 콘텐츠 네트워크는 경쟁 우위를 제공할 뿐 아니라 앞으로 고객의 특정한 개별 요구를 충족시킬 수 있는 발판을 마련해 줄 것이라고 믿는다”고 말했다.

캐나다의 유무선 미디어 서비스업체인 로저스커뮤니케이션 역시 인터워벤의 고객으로, SOA 중심 접근법을 개발에 적용했다. 론 맥컬리 e비즈니스 부사장은 여러 가지 실용적인 혜택을 봤다고 말한다. “우선 SOA를 통해 파일럿을 실행한 후 상당한 추가 부담을 발생시키지 않고 추가 애플리케이션으로까지 초기 실행을 확대할 수 있다. 보다 중요한 것은 SOA 기반 솔루션 덕분에 로저스는 특정업체가 독점하고 있는 API 대신 SOA 기반 개발을 전후해 기술 세트를 결집시킬 수 있었다. 앞으로는 절감된 개발 비용과 노력을 보안이나 스토리지 등의 핵심 요건으로 돌릴 계획”이라고 그는 설명했다.

디지털 자산 관리

SOA 기반 ECM 접근법의 범위는 모든 종류의 관리 대상 콘텐츠에 걸쳐 있다. 일례로 IBM은 디지털 자산 관리(DAM) 제품에 SOA 지원을 추가했고, 뉴욕주 푸킵시에 있는 마리스트 칼리지에서는 흥미로운 프로젝트가 진행되고 있다. IBM은 마리스트, 시스코와 공동으로 교사와 학생이 교육자료를 온라인으로 액세스, 관리, 배포할 수 있는 SOA 기반 시스템을 실행했다. 마리스트는 지난 12년 동안 FDR 라이브러리 컬렉션과 에미상 수상 소식 컬렉션, 다큐멘터리 및 스포츠 쇼를 비롯한 중요한 디지털 매체 아카이브를 보유하고 있다.

마리스트의 시스템은 디지털 아카이브 및 교과 자료에 대한 연결과 액세스 개선, 팀 룬용 웹 기반 포털, 계시판, 인스턴트 메시징, 학점인정코스 액세스 기능 등을 포함할 예정이다. 마리스트 칼리지의 데니스 J. 메레이 교장은 “콘텐츠 액세스 및 호환 능력은 강의 과정에서 필수적인 요소다. 이번 공동작업을 통해 얻어진 결과는 고등 교육계에 도움이 될 뿐 아니라 기업과 정부가 직원들을 교육시키는 방법까지 향상시킬 것”이라고 말했다.

마리스트의 환경은 IBM과 리얼 네트워크, 마이크로소프트, 애플을 비롯한 기타 업체들의 시스템을 비롯한 독점 디지털 미디어 스트리밍 시스템으로 가득했다. IT학과의 해롤드 앤더슨 교수는 “신규 서비스 기반 시스템을 통해 마리스트(Marist)는 단순한 추

상화 계층을 적용, 각 스트리밍 매체 시스템을 감출 수 있게 됐다”고 설명했다. 이 추상화 계층은 애플리케이션 프로그래밍을 간소화시키고 이중 시스템이 함께 작동할 수 있도록 한다.

SOA 도입 증가

SOA 기반 접근법은 분석가들과 소프트웨어 개발자 및 공급업체의 지지를 받아왔으며, 초기 파일럿 애플리케이션이 성공을 거두면서 최종 사용자 집단 내에서도 호응을 얻고 있는 것으로 보인다. 사용자들이 다른 제품의 실행 및 통합 프로젝트와 SOA 기반 프로젝트의 비용을 가늠하면서 도입이 폭발적으로 증가할 것으로 보인다.

칸다소프트의 ‘칼포브스키’는 실행이 간단하기 때문에 사람들의 마음을 끌어들이는다고 여긴다. “웹서비스 개발의 복잡성과 관련 비용은 이전의 COBRA나 DCOM에 비해 상당히 낮다. 수십 년간 각기 다르고 서로 얽힌 기업 애플리케이션 통합 접근법을 거친 끝에, 업계는 이제야 모두에게 효과가 있고 불특정 다수에게 만족스럽게 설명할 수 있는 수단을 얻게 된 것”이라고 그는 말한다.



빌 트리페
(btrippie@nmpub.com)는 보스턴에 소재한 전자출판, 콘텐츠관리, SGML 및 XML 전문 컨설턴트업체인 뉴 밀레니엄 퍼블리싱(www.nmpub.com)의 대표다.

출처

파일넷	www.filenet.com
IBM	www.ibm.com
인터워븐	www.interwovwn.com
칸다소프트	www.kandasoft.com
퀴온	www.quion.com
SAP	www.sap.com

웹 서비스 도전: 보안성과 관리성 확보

ERP 대기업인 SAP의 카즈 반 데 루 기업용 아키텍트 팀장은 소프트웨어 통합 프로젝트 중 콘텐츠 구성 요소가 주로 가장 흥미로우면서도 까다롭다고 말했다. 그 중에서도 가장 큰 난관이 바로 보안성이라고 한다. 주문이나 이행 등의 고객 중심 프로세스는 종종 개방적으로 자유롭게 배포할 수 있는 콘텐츠(예를 들어 카탈로그)를 승인된 사용자와만 공유해야 하는 데이터 및 콘텐츠와 혼합한다. 그는 “핵심적인 보안 구성 요소는 포털 자체에서의 인증이다. 소프트웨어가 이것을 제공하지 못하면 스스로 이것을 구축해야 한다”고 주장했다.

포털에서의 인증을 통해 승인을 거친 사용자만이 여러 가지 데이터 및 소스에 접근하도록 할 수 있다. 반면 여러 웹서비스의 보안을 지나치게 강화하는 것도 옳지 않다. 카탈로그나 고객 지원 데이터베이스 등의 일부 서비스는 투명해야 한다.

반 데 루 팀장은 복수의 서비스에 걸친 인증과 관련한 성능 문제를 강조하면서, “모든 사람들에게 인증을 거치도록 하는 것은 옳지 않다”고 말했다.

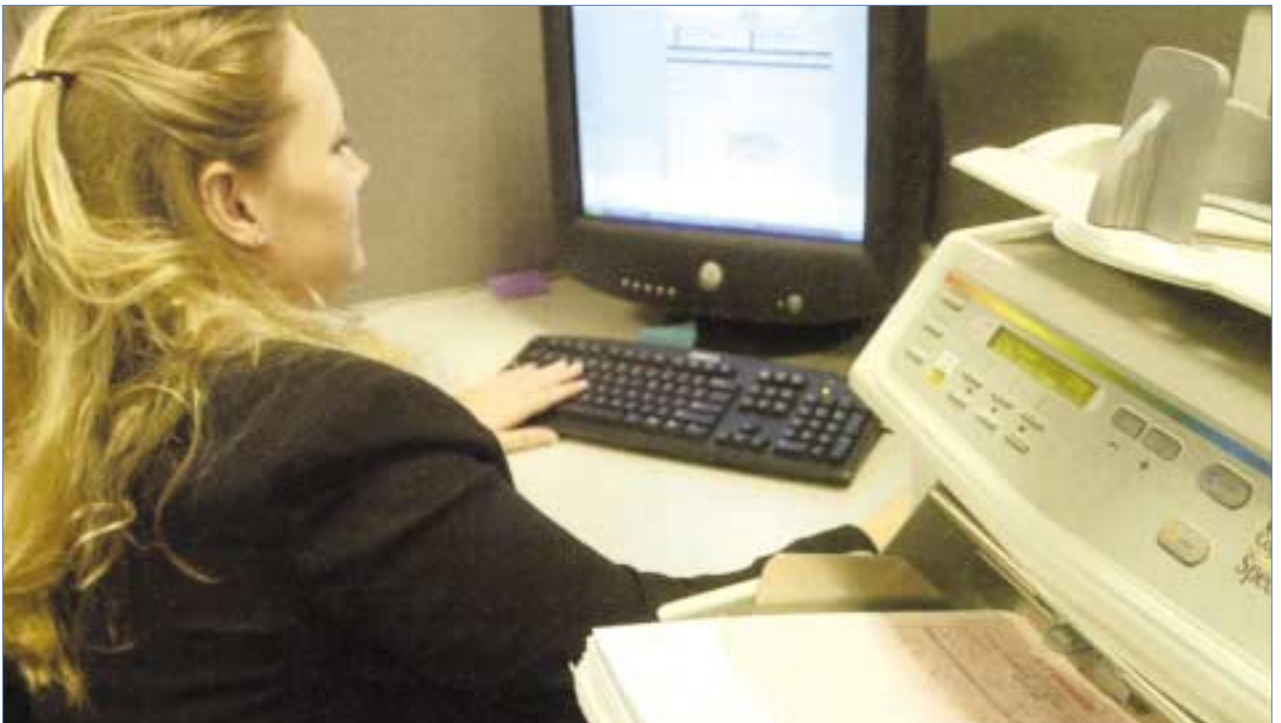
메사추세츠주 콩코드에 본사를 둔 SI업체 칸다소프트의 알렉스 칼포브스키 사장의 설명에 따르면, 웹 서비스가 웹서비스의 보안 문제를 해결하는 데 도움이 된다는 사실은 의심의 여지가 없다고 한다. 그는 “복수의 애플리케이션에 걸쳐 통일된 인증 및 검증을 위한 웹 서비스가 출현하게 되면 적법한 사용자는 허용된 시간 내지 버전 숫자 내에서 적법한 콘텐츠를 볼 수 있을 것”이라고 설명하면서, “이것을 통해 개발자들은 보안성을 애플리케이션에 적용할 준비 메커니즘을 확보할 수 있을 것”이라고 덧붙였다.

반 데 루 팀장은 SAP를 통해 보안성과 더불어 SOA 기반 접근법의 성공에 중요한 또다른 문제로 관리성을 확인했다고 말했다. “개발자들이 상상하고 배치할 수 있는 포털 서비스는 끝이 없다”고 그는 설명했다. 그러면서 SAP의 넷위버(NetWeaver) 실행 및 통합 플랫폼에 이미 관리 장치가 포함돼 조직이 웹서비스를 만들고 배치하는 이들을 관리하는데 도움을 주고 있다고 덧붙였다.

송장 결제 비용을 줄여라!

자동화된 캡처 도구로 손수 작업이 필요 없게 된 송장(送狀)은 이제 AP 부서들이 지능형 문서인식 기술을 활용해 데이터 입력 비용을 줄이고 턴어라운드 시간을 단축하려 애쓰는 공정한 게임이 됐다.

글 / 랄프 게이몬



두 회사가 송장의 관리 및 결제를 같은 방식으로 처리하는 경우는 없다. 다른 두 가지 형식 기반 프로세스가 형식 프로세스 소프트웨어를 사용해 온라인화 되거나 스캔 또는 자동화 된 반면 송장은 여전히 수동 데이터 입력이라는 문제를 안고 있다. 대부분의 기업들은 광범위한 공급업체들을 상대하며, 송장 절차는 관리 대상이 아니다. 레이아웃이 서류별로 다르므로 송장은 자동화가 거의 불가능했다.

최근 몇년 동안 서류 및 데이터 캡처 공급업체는 이미지별로 정보가 다른 위치에 나타나는 경우에도 특정 데이터의 인식 및 수집을 터득할 수 있는 정교한 인식 기술을 개발해왔다. 이른바

‘지능형문서인식(IDR)’ 기술이라고 불리는 이 기술은 마침내 성공을 앞두고 있다.

렌탈업체인 렌트에이센터와 발전회사인 USI에너지, 회계업체인 PRB솔즈와 시간제 컴퓨터 대여업체 HMLS 등이 IDR의 사용을 가장 먼저 시작했으며, 모두 송장을 각기 다른 형식으로 처리하고 있다. 이들의 활용 범위는 4~5가지 데이터 포인트의 입력 자동화에서부터 수백줄의 항목을 추출해 공급계약 및 할인에 관한 데이터베이스와 대조하는 데까지 이르고 있다. 이 모든 업체들이 동의하는 한 가지가 있다면 모두 투자금 회수를 원활하게 진행하면서 IDR의 시대가 도래하고 있다는 것이다.

렌트에이센터, 주당 1만5,000건 송장 처리

미국과 푸에르토리코에 2,600개의 점포를 두고 있는 렌트에이센터는 북미 최대의 렌트를 소유 체인이다. 여러 해 동안 TX 기반 업체인 플라노는 매주 처리하는 1만5,000건의 송장에 대한 접근을 향상시키기 위해 문서 이미징 시스템을 고려해왔다. 이들 송장은 유틸리티, 오프사이트 스토리지 및 기타 항목의 유지와 같은 비용을 비롯해 렌트에이센터에 제공되는 서비스를 다루고 있다. 송장은 매장 차원에서 승인을 거친 후 기업의 미수금 수취 계정(AP) 부서로 전달돼 결제된다.

렌트에이센터의 크리스티 투펠 회계 정리 팀장은 “그동안 송장을 일자 별로 분류해 박스에 넣어 다른 장소에 보관해왔다. 공급업체의 감사관이 송장 사본을 요구하면 박스를 회수해 사진을 찍어야 했다. 만약 감사관이 여러 해 동안 송부한 일련의 송장을 요구할 경우 그야말로 악몽이 펼쳐지게 된다”고 말했다.

첫 번째 설명에 대해, 투펠은 스캐닝이라는 문서 이미징 단계와 수동으로 이미지에 색인을 붙여야 하는 수고 때문에 통상적인 문서보관 및 검색을 통해 확보되는 효율성이 차감될 것이라고 추측했다. 투펠은 “이 상황을 연구한 후, 색인 작업에 사용하려는 필드가 송장을 결제하기 위해 회계 시스템에 입력해야 하는 필드와 같다는 사실을 깨달았다. 이미징의 사용을 극대화하여 회계 시스템에 대한 데이터 입력을 자동화하면 설치비용을 합리화시킬 수 있을 것”이라고 말했다.

투펠은 “전통적인 템플릿 기반의 형식 처리 시스템에 익숙하며, 처리하는 송장의 다양성과 가변성 때문이 렌트에이센터에서는 이것이 효과를 거둘 수 없음을 알고 있었다. 실제 계약업체가 9만곳이 넘고 매주 150곳을 새로 추가하고 있다”고 말했다.

투펠은 미수금 계정 컨퍼런스에서 처음으로 IDR을 사용한 송장 처리 소프트웨어를 소개했다. “이 시장의 모든 거래 업체들을 대상으로 광범위한 조사를 하기 시작했고 소수의 가격 경쟁력이 있는 애플리케이션으로 필드를 줄였다”고 설명했다. 지난해 4월 투펠은 시스템 시연에서 목격한 정확도를 기준으로 캘리포니아주 어바인에 본사를 둔 코팩스 이미지 프로덕트(Kofax Image Product)의 ‘어센트 포 페이어블(Ascent for Payable)’을 채택했다. 이 시스템은 9월까지 실무에 투입되지 못했지만, 투펠은 “다른 IT 프로젝트를 진행하는 중이었다”고 설명했다.

렌트에이센터는 현재 어센트 포 페이어블을 사용해 계약업체 명이나 주소, 송장 번호, 금액, 송장 일자 및 송장 제출 점포 수를 비롯한 송장 정보를 캡처하고 있다.

오클라호마 시티에 본사를 둔 소매업체 비즈니스 이미징 시스템(www.bisok.com)의 단 로텔리 사장은 “어센트 포 페이어블



송장이 같은 경우는 없다. 그럼에도 불구하고, 지능형 문서 인식 시스템은 공통 데이터 필드를 찾아내고 전형적으로 노동 집약적인 데이터 입력 작업에 자동화를 도입한다.

은 문서 전체에 옵티컬 문자 인식을 적용한 후 그 패턴을 이전 송장의 패턴과 비교한다. 이 과정을 통해 소프트웨어는 올바른 정보를 찾는 데 사용할 수 있는 구문론을 구축할 수 있다. 우리는 일정한 규칙을 생성하여 이것을 지원할 수 있다”고 설명했다.

투펠은 현재 렌트 에이 센터의 송장 필드 중 85%가 검증 및 교정 과정이 실행되기 전에 정확하게 캡처되고 있는 것으로 추정했다. “우리는 회계 시스템의 정보와 AP 직원들의 지식을 최대한 활용해 소프트웨어가 확신하지 못하는 데이터를 검증하고 필드를 작성한다. 모든 필드는 한두 가지 방법으로 검증된다. 일부는 기본적으로 최종 확정되지만 다른 필드는 룩업 과정이 포함될 수 있다”고 그는 말했다.

로텔리 사장에 의하면 렌트에이센터 애플리케이션의 가장 독자적인 요소 중 하나는 스캐닝과 검증 작업이 데이터 입력 전담 요원이 아닌 정규 AP 직원에 의해 이뤄지는 것이라고 한다.

투펠은 “종이 송장 정보를 입력하던 직원들을 어센트 포 페이어블을 사용한 스캐닝 및 검증 작업 부문으로 옮겼다. 처음에는 관리자들이 한주 정도의 테스트 기간을 거쳤다. 그런 다음 2주에 걸쳐 롤 아웃을 진행하면서 나머지 AP 직원들을 교육시켰다”고 설명했다.

그는 어센트 포 페이어블 관련 작업을 실행하는 데 15명의 AP 직원들이 일일 10~15분을 소요하는 것으로 추산했다. 그는 “그 동안 달성한 데이터 입력 효율성 때문에 담당 인원을 25명에서 15명으로 줄일 수 있었다. 그것을 바탕으로 1년 내에 투자금을 회수할 수 있을 것으로 기대한다”고 말했다.

AP 설치의 성공을 토대로, 렌트에이센터 내의 다른 여러 부서에서도 문서 이미징 애플리케이션을 모색하고 있다고 투펠은 전했다.

‘우리가 성취한 데이터 입력의 효율성 때문에 담당 직원 수를 25명에서 15명으로 줄일 수 있었다. 또한 1년 내에 투자금을 회수할 수 있을 것으로 기대한다’



렌트에이센터의 크리스티 투펠

해외 데이터 입력 비용 타개

잉글랜드 허트포드 소재 HMSL의 이미징 서비스 부서는 일주일에 렌트에이센터만큼 많은 일일 송장을 처리하고 있으며, 보통 1만5,000건의 송장으로부터 훨씬 많은 세부 정보를 수집한다. 최근까지 HMSL은 데이터를 입력하는 데 해외의 이미지 키 입력 사업부에 크게 의존해 왔지만, 비용상승으로 인해 내부 솔루션을 모색하게 됐다.

이 업체의 시드 바실리 CEO는 “소프트웨어 솔루션을 사용하면 비용 절감에 도움이 될 뿐 아니라 턴어라운드 시간도 단축할 수 있다”고 설명했다.

업계 선두를 달리고 있는 여러 형식처리 공급업체들의 솔루션을 검토한 끝에, HMSL은 이스라엘 텔아비브에 본사를 둔 톱이미지시스템즈의 ‘이플로우(eFlow)’ 시스템을 구매기로 결정했다. 테스트 과정에서 다른 공급업체들은 우리가 정한 기준선에 못 미친 반면 TIS는 그것을 초과했다고 바실리 CEO는 말했다.

HMSL은 이플로우를 사용해 계약업체명, 송장번호 및 일자를 비롯한 헤더 데이터뿐 아니라 문자열 항목의 세부 내용까지 포착했다. HMSL는 고객으로부터 받는 구매 주문 정보 테이블과 대조하므로 문자열 항목 정보는 중요하다. 문자열 항목이 PO 정보와 일치하지 않을 경우에 대비해 HMSL는 이 사실을 고객에게 통보하는 이메일 작업 폴더 시스템을 개발했다.

이플로우 시스템은 템플릿에 의존하지 않지만 소프트웨어는 교육이 필요하다. 문서 성격에 관한 데이터베이스를 구축하려면 동일한 계약업체에서 발행한 약 3건의 송장이 필요하다. 특정 계약업체의 교육이 완료될 때까지 우리는 주로 이미지 입력 형태의 데이터 입력에 의존한다고 그는 말했다.

HMSL는 현재 이플로우를 사용해 거래 업체 8곳의 송장을 처리하고 있으며, 바실리 사장은 캡처의 디테일 수준 범위가 한줄에서 수백줄의 문자열에 이른다고 전했다. 송장 처리가 완료되면 HMSL은 XML 파일로 데이터를 클라이언트에게 보낸다.

“지금은 과거 36시간에서 48시간이 걸리던 작업을 4시간 안에 턴어라운드 할 수 있다. 하루에 처리하는 송장의 양도 두배로 늘어났고, 이제는 해외 사업부를 단계적으로 철수시키고 있다고 그는 전했다.

HMSL는 이플로우 시스템과 송장 애플리케이션을 완료하는

데 약 45만달러를 지출했으며, 이 모든 것을 실무에 투입하는 데는 약 65일이 소요됐다. 바실리 사장은 연간 약 15만달러의 비용을 절감할 수 있을 것으로 추산했다.

USI 에너지, 처리 능력 배가

조지아주 노크로스에 소재한 USI에너지에게도 송장으로부터 세부적인 문자열 항목을 추출하는 작업은 중요하다. USI는 다가구 주택이나 아파트의 공공재를 관리하며, 사용료의 처리 및 결제도 서비스에 포함된다. 이 업체는 미 전역에 약 800여건의 부동산을 소유하고 있는 14곳의 고객으로부터 월 2만5천건의 종이 송장을 처리하고 있다.

USI에너지 관리서비스사업부의 데이브 리치텔리 부사장은 “우리는 약 700곳의 공공재 업체들이 발행한 청구서를 결제한다. 물량은 계약업체에 따라 다르다. 미네소타주 스몰빌에 있는 상수도 회사의 청구서를 접수할 수도 있다. 또한 애틀란타에서 활동하는 전기회사의 청구서 수천장을 접수할 수도 있다”고 말했다.

USI는 고객으로부터의 대금 회수를 요청하는 이러한 청구서의 데이터를 전용 소프트웨어 애플리케이션에 입력한다. 그리고 고객들이 청구서를 직접 검색할 수 있도록 웹사이트에 데이터와 송장의 이미지를 게시한다.

3년여 전 USI는 처음에 OCR을 사용한 형식 처리 기술을 사용하는 방안을 고려했다. USI의 제이슨 다니엘 업무 분석관은 “그 당시 나와 있던 시스템들은 대부분 구조화된 형식에만 효과가 있었고, 그 때문에 프로젝트에 차질이 빚어졌다. 모든 공공재 사용료 청구서에는 기본적으로 동일한 정보가 들어있지만 송장 형식은 같은 것이 하나도 없다. 동일한 회사에서 발행한 청구서까지도 페이지 수와 문자열 항목이 제각기 틀리다”고 말했다.

지난 2002년, 다니엘은 또다른 공공재 송장 데이터 입력 자동화 요청을 받았다. 그는 “당시 나는 여러 가지 청구서 체계를 처리하기에 충분한 시스템이 존재한다는 사실과 우리가 필요한 맞춤 기능까지 갖춘 스크립팅 능력이 포함됐다는 사실에 고무됐다”고 전했다.

USI는 뉴욕 주 테리타운에 소재한 데이터캡의 ‘태스크마스터(TaskMaster)’ 소프트웨어를 채택했는데, 부분적으로는 비주얼 베이직 스크립팅 기능이 있었기 때문이다. “당시 또다른 애플리케이션도 고려하고 있었지만 전용 스크립팅 언어를 갖추고 있어서 프로그래머들이 교육을 받는데 시간이 걸렸을 것”이라고 다니엘은 설명했다.

테스트 준비가 완료되기 전까지 개발에 4개월이 소요됐으며,

가장 큰 난관은 회사의 데이터베이스를 검증 프로세스와 통합하는 일이었다고 다니엘은 말했다. USI는 또한 스크립팅을 사용해 특정 계약업체로부터 수령한 각각의 문서에 적용할 규칙을 만들었다.

다니엘은 “애플리케이션 전체가 키워드로 작동한다. 송장의 모든 문자에 OCR을 적용한 후 거래업체명이나 총액 등의 키워드와 그 뒤의 숫자 또는 문자를 검색하기 시작한다. 전체 영역을 픽업하는 지역을 생성하는 문자열 항목을 위해 맞춤 스크립트를 생성했다. 이 애플리케이션은 발견된 키워드를 토대로 지역을 분석한다”고 설명했다.

그의 설명에 의하면 특정 계약업체 송장의 필드 정확도를 70~80%에 달하도록 조절하는 데 3~4시간 정도가 소요된다고 한다. 이러한 규칙을 사용해 송장의 내립차순 숫자를 처리하면 정확도가 90%를 넘어선다는 설명이다.

USI의 송장 시스템은 3개월의 테스트 기간을 거친 후 지난해 9월 실무에 투입됐으며, 지금은 이것을 이용, 가장 많은 양의 송장에 대해 약 30여건의 공공재 요금 청구서를 처리한다고 다니엘은 말했다. “이것은 우리의 전체 분량 중 약 절반을 차지한다. 이 애플리케이션 덕분에 하루에 처리할 수 있는 송장의 수가 일인 당 300건에서 최대 600~700건으로 늘어났다”는 것이 그의 설명이다.

다니엘은 “앞으로도 자금 회수율이 줄어드는 시점까지 자동으로 처리되는 계약업체의 청구서수를 지속적으로 늘릴 예정”이라고 말했다. 또한 “계약업체로부터의 물량이 줄어들면서, 규정 마련을 추진하려는 노력이 더 이상 합리적이지 않게 됐다”고 설명하면서, “앞으로는 전통적인 이미지 데이터의 키 입력 방식으로 돌아갈 것”이라고 덧붙였다.

EDI의 유산 감사

세계 최대의 미수금 회수 감사기관이라고 자임하는 PRG솔츠(애틀랜타 소재)는 다량의 고객 송장을 처리하는데, 이들 중에는 북미 최대 규모의 소매업체들도 포함된다. 이 고객들 중 상당수가 EDI를 통해 송장의 90% 이상을 처리하지만, 조 맛사넬리의 설명에 의하면 나머지 10~15%만 해도 상당한 수의 송장이며 상당한 비용회수 기회라고 한다.

PRG의 업무에는 올바르게 결제가 됐는지 송장을 관찰하는 일도 포함된다. 여기에는 이중 청구나 적절한 할부 수준 등의 문제에 관한 송장 검토가 포함된다. 맛사넬리 문서관리사업부 사장은 “각 송장의 문자열 항목 데이터는 우리의 업무에 중요하다. 전용 소프트웨어에 데이터를 입력하고 그것을 구매 주문이나 수

취 정보 등의 문자열 항목 세부 내역과 대조한다”고 설명했다.

맛사넬리는 “통상 송장을 시스템에 있는 기존 데이터와 합류시키는 데 필요한 헤딩 정보라면 무엇이든 종이 송장에서 캡처한다. 문자열 수준에서 품목 번호, 수량, 단가, 확대 비용 및 기타 할부 또는 할인 등을 문자열 별로 최소한으로 캡처한다”고 말했다.

PRG는 역대로 해외 서비스국을 통해 데이터 캡처를 처리해왔지만 회사측이 원하는 것보다 비용이 많이 들고 시간도 제때 지켜지지 않았다. 지난 2002년 12월, 세 가지의 대표적인 송장 처리 시스템을 도입해 4개월에 걸친 평가기간을 거쳤다. 맛사넬리 사장에 따르면 설치의 용이성, 사용자 인터페이스 및 성능을 기준으로 애플리케이션을 평가했으며, 세 가지 범주 모두에서 리드소프트(Readsoft, 로스엔젤레스 메타이리 소재)가 가장 높은 평가를 받았다고 한다.

지난해 4월, PRG는 리드소프트의 송장 소프트웨어에 최소 10만달러를 투자했다. 이 프로젝트는 아직 파일럿 단계에 있다고 한다. 세 가지 각기 다른 송장 프로젝트에서 이것을 실행하면서 문서를 지원한 결과 75%의 시간에서 성공을 거둔 것으로 밝혀졌다. “클레임까지 생성할 수 있었다는 얘기”라고 그는 설명하면서, “성공을 거두지 못한 경우에도 반드시 그 원인이 소프트웨어에 있는 것은 아니었다”고 덧붙였다.

송장 소프트웨어는 최적화 모듈을 사용해 사용자가 계약업체 프로파일을 설정하는 데 도움을 주는데, 맛사넬리의 설명에 의하면 채 5분도 걸리지 않는다고 한다. 테스트 송장 결과가 기대했던 수준에 못 미칠 경우, 그것을 조정하는 데 10~15분이 추가로 소요될 수 있다고 그는 말했다. “송장 문자열 항목의 복잡성은 변한다. 계약업체가 항상 한두 줄에서 문자열 항목 정보를 일정하게 유지한다면 간단하다. 항목당 문자열 수가 변한다면 이 과정이 다소 까다로워질 수 있다”는 것이 그의 설명이다.

맛사넬리는 “PRG는 문서의 분량과 복잡성에 의해 결정되는 특정 경우에 한해 송장 처리 소프트웨어를 적용하고 있다. 특정 계약업체가 발행한 200건 이하의 송장이라면 해외에서 데이터 입력을 처리하는 것이 더 효율적인 것이다. 하지만 우리의 상당수 고객들은 같은 계약 업체들을 상대하므로, 일단 템플릿이 구축되면 그것을 여러 고객에 걸쳐 사용할 수 있다”고 설명했다. 또한 “확고한 투자 회수율 수치가 나와 있지는 않지만, 투자금을 회수할 수 있다는 데는 의심의 여지가 없다”고 덧붙였다.

PRG는 업무의 결제 측면에도 송장 소프트웨어를 사용해 종이 송장을 전자 데이터를 사용할 수 있는 EDI와 유사한 거래 프로세싱으로 변환한다. 종이 송장에 EDI와 동일한 규정이 적용되지 않는 경우가 많으며, 그 결과 결제 오류에 취약할 수 있다.

스캔 이미지 파일 크기를 최소화시키는 방법

문서의 스캐닝과 디지털화 작업에서 겪게 되는 난관 중 하나는 스토리지 소요량이 엄청나다라는 것이다. 이번 호에서는 전문가들이 이러한 부담을 덜 수 있도록 조언한다.

글 / 페니 렌트 크로스먼



수잔 모이스, 인포트렌즈 리서치그룹(메사추세츠주 마쉬필드) 컨설턴트

두가지 요소가 파일 크기에 가장 큰 영향을 미친다. 비트의 심도와 해상도가 바로 그것이다. 비트 심도는 주로 2진 스캐닝의 경우 2비트, 단색 고화질은 8비트, 컬러 캡처는 24비트로 설정돼 있다. 비트 심도를 추가할 때 픽셀당 더 많은 정보를 효과적으로 수집함으로써 파일 크기가 늘어나는 것이다.

그 다음으로 파일 수를 조절할 수 있는 척도는 해상도다. 대부분의 문서 이미징 애플리케이션에서 통상 해상도 범위는 100dpi에서 300dpi 사이지만, 특별한 경우 600dpi까지 높아질 수도 있다. 해상도가 증가하면 파일 크기도 그만큼 늘어난다.

가장 중요한 문제는 사용이 가능토록 페이지 정보를 캡처하는 것이다. 2진 스캐닝이나 단색 고화질, 또는 컬러 캡처 중 필요한 것을 파악하는 일부터 시작한다. 그런 다음 가능한 한 허용할 수 있는 결과를 도출하면서도 가장 낮은 해상도를 적용한다.

개인적인 경험에 의하면 2진 스캔의 경우 통상 적절한 결과물을 얻기 위해 높은 해상도(200dpi 이상)를 필요로 하는 반면, 단색 고화질이나 컬러 스캔은 낮은 해상도에서도 좋은 화질을 생성한다(100dpi부터 시작해보라). 문서를 각기 다른 비트 심도와 해상도 설정에서 스캐닝해 화질과 파일 크기에 미치는 영향을 측정해보면 된다.



하비 스펜서, 하비 스펜서 어소시에이츠(뉴욕주 이스트 노스포트) 대표

이미지 파일의 크기를 줄이는 가장 빠르고 쉬운 방법은 낮은 해상도(인치 당 도트로 측정)로 저장 또는 스캔하는 것이다. 물론 파일의 해상도가 낮을수록 판독성도 떨어진다. 대부분의 문서관리 시스템은 TIFF 이미지의 최소 해상도를 200dpi로 예상하며, 통상 옵티컬 문자 인식 소프트웨어는 적절한 문서 판독을 위해 300dpi를 필요로 한다. 가능한 범위까지 해상도를 낮춰 TIFF 그룹 4 이미지의 크기를 줄이고 얼룩제거 기능(despeckler)을 사용해 원치 않는 배경(예를 들어 문서의 핑크색 배경)을 없애며, 이미지의 기울기를 보정하면 된다. 이 모든 수고를 통해 TIFF 그룹

4가 최적화되도록 흰색 영역의 'run length(흑백 픽셀의 라인)'를 향상시킬 수 있다.

색조 정보(단색 또는 컬러)가 들어있는 JPEG의 경우 2진톤 이미지에 비해 설정된 해상도가 낮은 것이 보통이므로 150 또는 100dpi 이미지를 생성하는 쪽을 고려해볼 수 있다. 음영의 수를 줄이면 파일 크기도 줄일 수 있다. 불필요한 음영을 없애도 된다. 마지막으로 압축률을 조절하면 된다. 75% 정도의 비율이면 적절한 JPEG 영상을 제공하면서 파일 사이즈도 작아진다.

〈Copyright CMP Media LLC〉

Transform Bookshelf



업무 공정 관리:

제3의 물결

저자: 하워드 스미스 & 피터 핑거
Meghan-Kiffer press, 39.95달러

‘제3의 물결’에서는 존경받는 업계의 베테랑 두명이 업무공정 관리를 고려하고 계획해야 되는 이유를 설명한다. 이들은 실무집단에서 ‘그러한 노력을 하지 않을 것’이라던가 ‘시도해 보기는 했는데 별로였다’라든가, 또는 시간이 없는 등의 공통적인 반대 의견을 설명한다. 마지막의 경우, 업무공정 모델링은 IT업체들이 보유하고 있었을 때에도 소프트웨어 엔지니어링 공정에서 비생산적인 추가 단계로 인식됐다. 반면 업무 공정 시스템은 그 어떤 것보다 빠르고 신속한 애플리케이션 개발 방법이다. 공정 모델을 소프트웨어로 해석할 일이 전혀 없다고 응수한다.

스미스와 핑거는 BMP 원칙 및 기술을 사용하기 시작한 보잉사 등의 기업을 예로 든다. 보잉은 CAD/CAM 설계 도구와 공정을 전세계에 걸쳐 엔지니어와 고객, 유지보수 요원, 프로젝트 관리자 및 부품 공급업체들과 전자적으로 공유하는 방법으로 777 항공기를 웹상에서 설계했다. 물리적인 모델은 없었다. 종이로 된 청사진도 없었다. 이 공정의 성격은 777은 밀집하게 결합돼 날아다니는 하나의 부품 덩어리라는 슬로건에 잘 나타나 있다. 보잉은 공정 체류를 통해 8~12개월 내에 항공기를 납품하는 것을 목표로 하고 있다고 한다.