

컴포넌트 정보의 재사용을 위한 XML 기반 에이전트 구축 방법

김용식, 문창주, 송치양, 백두권

고려대학교 컴퓨터학과 소프트웨어 시스템 연구실

e-mail: {yskim, mcj, cysong, baik}@swsys2.korea.ac.kr

An agent construction Method for Component Information reuse based on the XML

Yong-sik Kim^o, Chang-Joo-Moon, Chi-yang Song, Doo-Kwon Baik
Software System Lab. Dept. of Computer Science & Engineering, Korea University

요 약

컴포넌트의 재사용성은 기존 소프트웨어의 문제점인 소프트웨어의 수요충족, 신뢰성요구 및 유지보수 용이 등을 해결할 수 있는 방법중의 하나이다. 현재 재사용 컴포넌트를 위한 웹 상에서의 정보서비스는 데이터요소의 불일치와 이미 검색된 정보를 다시 찾아야 하는 비효율적인 검색방법을 사용하고 있다. 이를 해결하기 위해, 본 논문에서는 이전에 찾은 컴포넌트 정보를 재사용 할 수 있도록 에이전트를 사용해 컴포넌트 정보를 수집하고 XML 형태로 정보를 저장, 관리하는 컴포넌트 정보의 재사용 방법을 제시한다. 즉, 이용자별로 이미 검색된 정보를 저장, 갱신관리하여, 이후 컴포넌트 검색에 있어서 이런 정보를 재활용함으로써, 검색시간의 단축, 정확한 정보검색 그리고 이용자별 차별적인 정보제공과 컴포넌트 평가를 지원할 수 있다.

1. 서론

21세기 지식기반 사회 정착화에 따른 글로벌 인터넷과 정보화 추진으로 소프트웨어의 수요는 폭발적이다. 이를 해결하기 위한 방법으로 컴포넌트의 재사용이 대두되고 있다. 현재, 통용되는 인터넷상의 컴포넌트 정보서비스는 통용 데이터 요소용어의 불일치와 이미 검색된 컴포넌트 정보를 재활용하고 있지 못하는 실정으로 효과적인 검색을 제공하지 못한다. 즉, 기존의 검색엔진이 찾아낸 정보를 활용하지 못하고, 새로운 컴포넌트 A가 필요할때 A를 찾아내지 못하고 A의 기능을 가진 C 컴포넌트를 개발을 해야만 한다면 컴포넌트는 재사용이 되지 않고 있다고 말할 수밖에 없을 것이다. 따라서 사용자가 직접 검색엔진을 이용해 관련 컴포넌트를 찾아내던가, 개발회사가 자사의 컴포넌트에 대한 정보를 자사 홈페이지에 기술해 컴포넌트를 구매자로 하여금 구매를 하도록 유도해왔다.

이보다, 진보된 방법으로 컴포넌트 데이터 레지스트리 시스템을 구축하여 컴포넌트에 대한 정보를 보다 표준화되고 일관된 형식을 갖추고자 시도되고 있다[1]. 컴포넌트 정보에 대한 키워드나 디렉토리 검색 수준의 정보서비스를 벗어나 이미 검색된 정보를 재활용하여, 검색시간과 정확성을 향상시켜주는 정보 재사용 에이전트가 필요하다[2]. 본 논문에서는 사용자는 컴포넌트 정보공유 에이전트를 사용해, 웹 기반으로 제공되는 컴포넌트 정보를 검색하고 검색 결과를 저장, 갱신하도록 하는 에이전트 시스템의 구축방법을 제시한다. 정보공유 에이전트는 사용자별 차별화된 서비스를 제공하여 타인이 검색에이전트를 사용하여 얻은 결과인

Seed URL과 컴포넌트 키워드, 번더정보를 공유할 수 있도록 한다. 사용자는 자신이 에이전트를 사용해 얻은결과를 XML로 저장함으로써, 저장된 결과에 대해서 다시 검색을 할 필요 없이 기존 검색된 정보를 재사용하여, 컴포넌트에 대한 정보를 얻는다. 이러한 방식의 에이전트는 사용자들이 에이전트를 사용함에 따라 사용자들이 구축한 정보가 축적되게 된다. 따라서 사용자가 저장한 컴포넌트 관련 URL과 공유정보는 다른 사용자에게 의해서 재사용 됨으로 컴포넌트 정보검색시간을 줄이고 정확한 검색을 한다. 또한 사용자별 필수정보를 제공하는 동시에, 이미 등록된 컴포넌트 정보에 대해서 갱신, 평가할 수 있다는 이점이 있다.

2. 현황 및 기술

2.1 컴포넌트 정보 서비스

컴포넌트에 대한 기대와 예측에 비해 아직 컴포넌트의 정보를 공유하기 위한 노력은 아직 정립되어 있지 않은 단계이다. 컴포넌트를 개발한 회사가 자사의 컴포넌트를 홍보하기 위해서 홈페이지에 간략하게 컴포넌트 정보를 기술하거나 데모버전을 배포하고 있는 실정이다[3]. 다음의 사이트 <http://www.componentsource.com>, <http://www.imagiccom.com>에서 보면, 컴포넌트의 정보와 관련 회사를 소개를 HTML 형태로 소개하고 있는 것을 볼 수 있다. 보다 정형화된 형태를 제공하는 (<http://www.flashline.com>)의 Flashline 사의 홈페이지에서는 컴포넌트 데이터요소의 메타데이터를 담고있는 컴포넌

트 레지스트리를 가지며, 벤더로 하여금 컴포넌트 DTD 를 작성하도록 요구하는 정적인 서비스를 제공 하고 있다[4].

2.2 XML

XML 은 이미 존재하는 문서를 다른 포맷으로 변환하는 번거로움을 피하고 기존의 HTML 의 한계를 극복하기 위해 서 1996 년 W3C 의 후원으로 형성된 XML Working Group 에 의해 등장했다. XML 은 인터넷상에서 곧바로 사용될 수 있으며, 다양한 어플리케이션들을 지원하기에 문서처리 프 로그램의 구현이 용이하다. 또한, 설계하기 쉽고 정확한 형 식을 가지기 때문에 데이터베이스가 없는 경우에도 데이터 베이스 형태의 스키마를 제공한다. 이런 장점을 가진 XML 의 등장은 정보공유에 있어서 인터넷의 부각과 함께 중요한 쟁점이 되고 있다. 기존의 HTML 이 의미전달과 자연의 처 리의 문제점으로 인해 검색엔진을 가지고 얻어낸 검색 정보 를 한번 더 필터링 해야 하는 단점을 가지고 있다. 그러나 XML 은 문서 생성이 용이 할뿐 아니라 의미와 형식의 분리 하여 좀더 정확한 정보공유를 한다.

2.3 XML 과 컴포넌트

XML 은 정확하고 신뢰성 있는 정보를 제공하여, 다양하 고 복잡한 인터페이스 등을 표현해야 하는 컴포넌트 정보 명세에 효과적이다. 기존의 HTML 기반의 검색엔진이 모든 자연어를 대상으로 키워드 검색서비스를 제공하지만 좀더 세분화된 영역에서 검색서비스는 제공하지 못했다. 컴포넌 트의 검색은 이런 영역중의 하나다. XML 은 여러 문서와 용어 통일에 적합하기 때문에, XML 형태의 검색결과를 컴포넌 트 레지스트리와 같은 저장소에서 관리하기가 용이하다 [5].

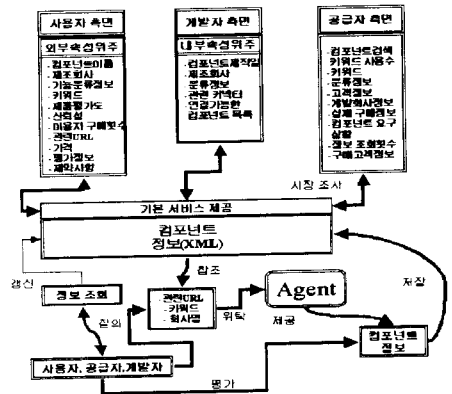
컴포넌트와 컴포넌트는 컴포지션이란 개념을 사용하여 새로운 컴포넌트를 생성함으로써 컴포넌트를 재사용 할 수 있 다. 따라서 컴포넌트의 재사용이 효과적으로 이루어 지기 위해서 는 각 컴포넌트에 대한 정보가 신속, 정확하게 제공 되어야 한다. 이런 요구사항을 만족하기 위해 XML 형태로 컴포넌트를 표현하고 저장하면, 보다 글로벌하면서도 통일 된 문서형식을 갖출 수 있어 정확한 컴포넌트 검색이 가능 하다. 실제로 이런 수요를 예측, XML 파서기를 비롯한 각종 DBMS 와 연동을 위한 유틸리티가 이미 제공되고 있는 실 정이다.

3. 컴포넌트 정보 재사용 방법 및 에이전트 설계

3.1 사용자별 컴포넌트 정보

컴포넌트를 보는 시각은 개발자, 사용자, 공급자(유통업자)에 따라 상이할 수 있다. <그림 1>은 사용자별로 요구되는 다른 컴포넌트 정보를 보여준다. 개발자는 컴포넌트의 개발에 앞서, 자신이 개발하고자 하는 컴포넌트가 이미 존재하는지 여부를 조사할 필요가 있다. 존재한다면 기존의 컴포넌트의 기능을 검사해 새로운 기능을 가진 컴포넌트와 컴포지션을 통해 새로운 컴포넌트를 개발 할 수 있다. 또한 사용자는 자신이 원하는 기능을 제공하는 컴포넌트를 정확 하고 신속하게 찾아내야 한다. 유통업자는 컴포넌트의 수요 와 공급을 분석해 시장의 동향에 맞추어 컴포넌트의 개발과 판매를 촉진하려 할 것이다. 이를 통해 이용자별 유용한 정보 를 효과적으로 제공할 수 있다.

에이전트에 의해 수집, 분류, 저장관리되는 정보는 각 컴 포넌트 각 컴포넌트 정보 서비스 사이트로부터 얻은 정보 (이름, 제작사등)에 사용자의 형상이력(키워드, 조회수, 평가 내역등)정보가 추가되어 제공된다



<그림 1> 재사용을 위한 사용자별 제공 서비스

2. 제공 서비스

1) 사용자별 서비스

기존의 검색엔진이나, 컴포넌트 정보제공 홈페이지를 보 면 컴포넌트에 대한 정보를 컴포넌트 개발자, 공급자가 스스로 갱신하거나 사용자가 컴포넌트에 대한 정보를 등록 하도록 하는 수동적인 형태를 취하고 있다. 그러나 컴포넌트 정보공유 에이전트의 각 사용자별 에이전트는 로그인 시 각 서비스 이용자(개발자, 유통업자, 사용자)별로 다른 서비스 를 제공한다. 서비스 이용자는 자신만의 ID, 패스워드를 가 지고 과거에 자신이 검색을 시도해 저장한 컴포넌트 정보검색 결과를 확인할 수 있다. 뿐만 아니라 타인이 같은 키워드 나 제품명, 회사명을 가지고 검색을 시도해 얻어낸 결과나 URL 을 참조해 새롭게 웹을 검색하는 에이전트 서비스를 이용한다[6].

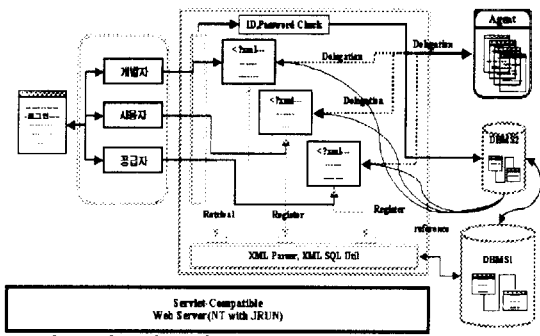
2) 컴포넌트 검색자동화

컴포넌트 검색 에이전트는 특정 URL 을 Seed 로 해서 그 문서에 링크된 홈페이지 정보 자동검색 한 후 결과를 통 하여 컴포넌트 정보가 위치한 URL 을 참조할수 있게 하는 서비스를 제공한다. 이렇게 제공된 서비스를 각 사용자는 자신이 필요한 컴포넌트 관련 정보를 저장, 평가 한다. 여 기서 Seed 가 되는 키워드나 URL 은 타인이 제공한 정보를 활용 할 수 있다. 동시에 자신이 저장한 정보는 타인이 이 용할 수 있도록 한다. 또한 기존에 타인과 자신이 검색해 놓은 컴포넌트 정보를 검색해서 삭제, 추가, 평가를 할 수 있다.

이 시스템의 장점은 컴포넌트 정보검색을 위한 시간을 여 러 사람이 투자한 시간만큼 단축시킬 수 있다는 점과 컴포넌트에 대한 정보를 통해 개발자나 공급자에 대한 정보와 컴포넌트를 사용해본 사용자의 정보도 얻을 수 있다. 또 컴포넌트 정보를 개발자와 사용자 그리고 공급자에 따라 차별 제공할 수 있으며 컴포넌트 정보가 다수 사용자에 의해서 자동 생성, 갱신 된다는 점이다.

3.3 정보 재사용을 위한 에이전트구축방법

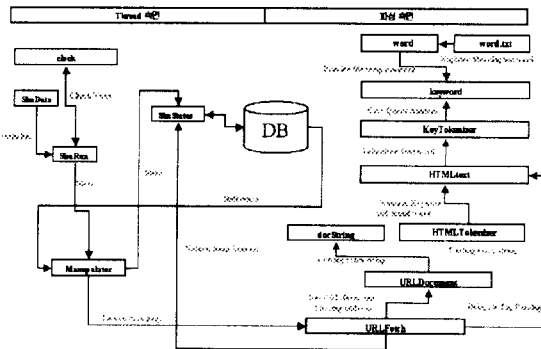
<그림 2>에서 시스템을 이용하는 이용자는 로그인과 동시 에 개발자, 사용자, 공급자 별로 구분되어 각기 상이한 정보 를 제공하는 사용자 인터페이스를 제공 받는다.



<그림 2> 컴포넌트 정보공유 에이전트의 활용과 시스템 구성도

동시에 이전에 컴포넌트 정보공유 에이전트를 실행해서 얻은 결과에 대한 저장결과를 DBMS2 로부터 가져온다. 또한 DBMS1 에 저장된 컴포넌트 정보를 가져와 새롭게 에이전트에 검색을 위탁해 새로운 컴포넌트를 찾아 낸다. 이때 DBMS1 은 모든 사용자가 등록한 컴포넌트 정보의 재사용을 위해서 개발자, 사용자, 공급자별로 나누어 제공하며 새로운 검색을 원하는 사용자에게 Seed 와 키워드, 회사명 등을 제공하여 검색 에이전트에게 검색방향을 제시한다. 일정 시간 검색을 마치고 이용자들은 검색결과에 대해서 자신의 평가를 기록하여 자신과 다른 이용자가 이용할 수 있도록 하는 컴포넌트 정보 등록이 가능하다. 이런 컴포넌트 정보 공유 에이전트의 사용과 이에 수반되는 컴포넌트 정보의 갱신은 비순차적이고 실시간으로 이루어 진다[7].

자동 문서검색을 지원하는 에이전트의 설계는 XML 형태의 문서와 HTML 형태에 대한 문서가 파싱이 가능하도록 해야 한다. 이를 위해서는 자바로 구성된 문서파싱 에이전트를 가지게 된다. <그림 3>은 HTML 문서의 파싱과 쓰레드의 사용을 보여주는 클래스 관계도 이다



<그림 3> XML, HTML 문서를 파싱하여 컴포넌트 정보를 얻어 내기 위한 에이전트의 설계

3.4 시스템 구현시 고려사항

1) XML, HTML 지원

정보검색, 공유를 위한 에이전트는 기존의 HTML 을 지원하는 것은 물론이고 XML 문서도 지원을 해야 한다.

2) 확장성, 독립성

인터넷을 기반 XML, HTML 을 처리하기 위해서는 에이전트를 플랫폼에 독립적인 자바를 사용하며 에이전트가 획득한 결과를 저장, 서비스 하기위해서 여러 환경에서 호환성

과 확장성이 뛰어난 서블릿을 사용한다. 이런 서블릿과 XML 의 사용은 기존 DBMS 에서 제공하는 유틸리티등을 함께 이용할 수 있다.

3.5 기대효과

컴포넌트 정보공유검색 에이전트를 사용하면 다음과 같은 효과를 가질 수 있다. 첫째, 통합된 컴포넌트 정보를 한곳에서 단시간에 찾아서 활용할 수 있다. 이때 이용자는 사용자별로 차별화된 서비스를 제공 받아 여러 사용자가 컴포넌트 정보를 검색해놓은 결과를 재사용할 수 있는 효과를 가진다. 둘째, 사용자들이 얻어낸 정보를 바탕으로 새로운 검색 및 기존의 검색결과에 대한 평가정보를 재사용 할 수 있다.

재평가는 컴포넌트 정보에 대한 신뢰성과 정확성을 이용자들이 높여갈 수 있다는 것을 의미한다. 셋째, 컴포넌트를 이용하는 사용자, 개발자, 공급자에게 컴포넌트에 대한 정보를 XML 형태로 받아서, 이를 각각의 사용자에게 다시 복잡한 변형의 과정을 거치지 않고도 서비스를 제공한다. XML 의 적용은 어플리케이션이 용이해 보다 단순한 인터페이스로 다양한 컴포넌트 정보제공 서비스를 할 수 있다.

4. 결론

웹상에서 산재 되어 있는 컴포넌트 정보를 하나의 시스템에 통합시키고 사용자가 검색한 결과를 자신은 물론, 타인이 재활용 할 수 있는 에이전트 구축방법을 제시하였다. 본 논문에서 제안한 방법은 기존에 획득한 컴포넌트 정보를 사용자별로 차별화 된 서비스를 제공하는 동시에 에이전트를 사용해 얻어낸 결과를 XML 형태로 저장, 관리한다. 이렇게 저장된 컴포넌트 정보는 다른 사용자가 유사 컴포넌트를 검색 하려할 때, URL 이나 키워드 등의 검색정보를 제공하므로 효과적인 컴포넌트 검색의 방향을 에이전트에게 제시한다. 또한 사용자별 정보를 분리, 가공해 개발자, 사용자, 공급자에 따라 상이한 컴포넌트 정보를 제공한다. 이런 에이전트의 활용함으로 사용자는 컴포넌트 정보를 단시간에 정확히 얻을 수 있다.

추후 연구로는 지식기반에 의한 검색과 저장에 추가되고, 컴포넌트 메타정보를 XML 형태로 변환하여 체계적인 정보교환이 여러 시스템에서 가능한 시스템의 설계 및 구현이 필요하다.

참고문헌

- [1] 임성빈, 문창주, 백두권 "통합된 컴포넌트 공유환경 구축을 위한 컴포넌트 레지스트리", '99 한국 정보과학회 추계 학술대회 논문집, 26 권 2 호, p441-443, 1999
- [2] 남궁영환, 나홍석, 이정옥, 백두권 데이터 레지스터에 기반한 XML 문서 검색 시스템의 설계
- [3] 관련 컴포넌트 기술, 판매 사이트(<http://www.imagi.com.com>, <http://www.componentsource.com>)
- [4] 채진석, 나홍석, 백두권 "XML 데이터 타입을 위한 데이터 레지스트리의 설계 및 구현", 한국정보과학회 추계학술발표논문집, 25 권 2 호, pp.650-652, 1998
- [5] XM 소개, ㈜언어기술, 1999.8.7 <http://ibase.co.kr/solution/xml>
- [6] Jon Bosak, XML, Java and the future of the Web, Mar 1997, URL: <http://sunsite.unc.edu/pub/sun-info/standard/why/xmlapps.html>
- [7] InnerView XML Search System, Techno2000 Project, Inc 1999.8.13, <http://www.cs.caltech.edu/~adam/papers/xml/xmarks-the-spot.html>