

시멘틱 웹 기반의 정보관리 시스템 개발

박홍석(울산대 기계자동차공학부), 박성범*(울산대 기계자동차공학부 대학원),
여승훈(울산대 정보통신대학원), 이규봉(한국생산기술연구원)

주제어 : Ontology, Metadata, Semantic web, XML(Extensible markup language).

기업은 다수의 부서들이 다양한 문서를 가지고 각종 업무를 수행하게 된다. 각 부서에서 보유하고 있는 문서의 양이 많으면 많을수록 업무에 도움을 줄 수 있는 정보가 많다고 여겼던 시대는 이제 지났다. 막연하게 수집된 자료들은 결국 그 담당자조차도 기억을 할 수 없는 경우가 발생하며 불필요하게 문서가 중복되거나 보관 및 검색을 위한 공간과 시간의 낭비로 인해 유용한 정보에 대한 손실을 초래할 수도 있다. 또한 각 부서마다 문서들을 별도로 관리하기 때문에 다른 부서에서는 해당 문서에 대한 정보를 공유할 수 없게 된다.

한편 동일한 부서 내에서도 그 특성에 따라 각각 다른 포맷의 문서들이 사용되고 있다. 이러한 경우 문서가 가지는 정보의 범위는 상당히 협소하게 되어 부서간에 필요한 정보의 교환이 어렵게 된다. 따라서 현재 기업은 자원을 얼마나 효율적으로 관리하여 실제 업무에 효과적으로 적용시킬 수 있는가에 대한 해결책을 모색하고 있다.

현재 기업에서 사용하고 있는 대부분의 문서 관리형태는 전자우편이나 네트워크 환경, 또는 기업 자체 문서관리 시스템의 경로를 통해 이루어지고 있다. 이런 경우에 있어 각각의 문서를 독립적으로 활용할 수 있을 뿐 문서끼리의 상호 연관성이 결여되어 있다. 따라서 연도별 업무 프로세스의 진행 결과 및 사례에 대한 정보가 포함되어 있지 않아 설계자가 과거 차 업무에 대한 참조나 파악 등의 작업이 힘들다.

시멘틱 웹이라는 개념은 컴퓨터가 정보의 의미를 이해하고 의미를 조작할 수 있는 웹의 환경을 의미한다. 만약 문서의 각 부분을 컴퓨터가 이해할 수 있는 형식으로 기술하고 표현할 수만 있다면 복잡하게 얽혀져 있는 정보의 리소스들 사이의 의미적 연관성으로 인해 다양한 정보를 보다 효과적으로 활용할 수 있을 것이다. 이를 웹 환경 하에서 구현하면 시멘틱 웹의 기반을 갖는 시스템이 된다.

현재의 웹 문서가 가지는 내용은 기본적으로 인간만이 이해할 수 있는 수준이라 할 수 있다. 이러한 한계를 극복하기 위해서는 객체에 대한 단순한 정의가 아니라 객체들의 속성으로서 객체의 정의가 표현되도록 해야한다. 이러한 객체의 속성을 기술하기 위한 자료가 메타데이터이다. 이를 이용하여 온톨로지 기반의 데이터 연관성을 표현할 수 있는 구조를 추가하게 되면 데이터는 정보로서의 의미 뿐 아니라 지식으로서의 활용 가치 또한 가지게 된다.

이에 본 연구에서는 서로 다른 포맷의 문서 요소들을 데이터베이스에 통합하고 XML 문서 표준을 이용하여 지식기반의 새로운 웹 문서관리 시스템을 구현하고자 한다. 이로써 문서간의 불필요한 정보를 제거할 수 있을 뿐만 아니라 웹 환경의 장점인 시간이나 장소의 제약을 받지 않고 업무에 참여할 수 있어 여러 기대 효과를 기대할 수 있다.

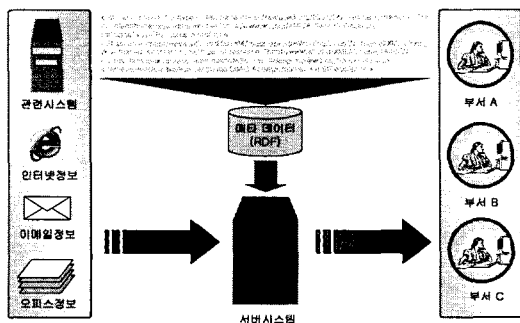


Fig. 1 Metadata를 이용한 정보통합 시스템

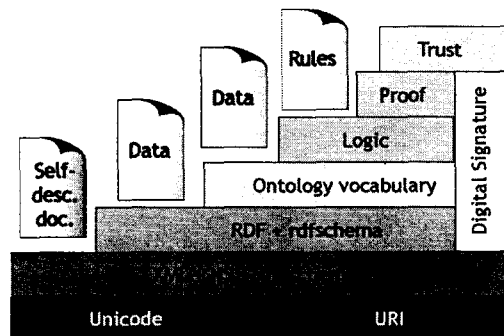


Fig. 2 Semantic web architectural dependencies