

인트라넷의 구조화는 곧 비즈니스의 구조화

정보기술 분야 전문가들은 인트라넷의 문제점들을 해소하기 위한 가장 유력한 정보기술을 'Corporate Portal'로 보고 있다. Yahoo, AOL, Microsoft, Netscape 등의 인터넷 포털 사이트들이 인기를 누리는 이유는 인터넷상의 정보들을 구조화하여 이용자들이 쉽게 정보를 찾을 수 있기 때문이다. 마찬가지로 인트라넷이 구축되어 있는 기업 역시 Corporate Portal을 구축함으로써, 기 구축된 인트라넷의 활용도를 극대화할 수 있게 된다

■ 윤진섭/ 쓰리소프트 부장



기

업에서 인트라넷은 정보네트워크의 표준으로 자리 잡아 가고있는데 포레스터 리서치(Forrester Research, Inc.)에 따르면 회사들이 인트라넷을 도입하는 첫째 이유는 전체 사업장 어디에서든지 쉽게 기업의 지식(Corporate Knowledge)을 공유할 수 있기 때문이라고 한다. 하지만 이제까지 구축된 인트라넷은 기업의 지식을 공유하기에는 몇가지 문제점들을 가지고 있다.

첫째, 오늘날의 인트라넷은 수많은 응용프로그램들과 웹서버, 파일서버, 데이터베이스와 연동되어 있지만 각각의 시스템들은 구축 당시 서로 독립적으로 개발되었기 때문에 부서별 자원이나 문서에 대한 공유가 어려운 반면 인터넷상의 야후(Yahoo)나 익사이트(Excite)같은 사이트에서는 인터넷상에 흩어져있는 여러 홈페이지들을 카탈로그와 같은 디렉토리 형식으로 쉽게 정보를 검색하고 탐색할 수 있다. 즉, 기업의 인트라넷은 해당 기업이 보유하고 있는 정보를 구조화하기 위한 디렉토리나 마스터사이트를 갖지 못하고 있다.

둘째, 메타그룹에 따르면 인트라넷상의 문서들의 양은 매년 두배씩 증가하고 있다고 한다. 기업 내부에서는 개인이나 부서에서 많은 양의 문서들이 기계적으로 양산되고 있고, 지속적으로 팽창하고 있는 인터넷으로부터 매우 가치있는 정보들이 실시간으로 유입되고 있다. 이러한 문서들을 효과적으로 구조화하지 못할 경우 각 문서들의 가치는 사라지고 만다.

셋째, 대부분의 인트라넷 이용자들은 검색방법이나 검색을 위한 질의어 선정에 익숙하지 않다. 이러한 이유로 대부분 하나의 단어를 질의어로 선택하게 되는데, 이는 너무 많은 검색결과를 초래함

으로써 가치있는 정보를 찾을 수 없게 한다.

넷째, 포레스터 리서치사의 조사에 따르면 인트라넷 사용자의 62%가 자신에게 친숙한 사이트만 이용하고 있으며, 50% 미만의 사용자만이 검색기능을 사용하고 있다고 한다. 또한 20% 미만의 사용자만이 기업전체 인트라넷을 대상으로 검색을 하고 있다고 보고하고 있는데 이는 직관적인 도구의 부족으로 사용자들이 검색에 어려움을 느끼고 있기 때문이다.

Corporate Portal이란?

IT분야 전문가들은 인트라넷의 이러한 문제점들을 해소하기 위한 가장 유력한 정보기술을 'Corporate Portal'로 보고 있다. Yahoo, AOL, Microsoft, Netscape 등의 인터넷 Portal Site들이 인기를 누리는 이유는 인터넷상의 정보들을 구조화하여 이용자들이 쉽게 정보를 찾을 수 있기 때문이다. 마찬가지로 인트라넷이 구축되어 있는 기업 역시 Corporate Portal을 구축함으로써, 기 구축된 인트라넷의 활용도를 극대화할 수 있게 된다.

델파이 그룹이 제시하는 Corporate Portal은 Content와 Community의 입장에서 보았을 때 소수의 Community를 위한 방대한 분야의 Content 서비스를 실현하여야 한다는 것이다. 즉, 기업이라고 하는 한정된 Community를 위한 서비스가 제공되긴 하지만 매우 다양한 분야의 정보들이 수집되어야 한다는 의미이다. 실제로 인터넷 Portal 사이트인 Yahoo, AOL 등에서 제공하는 정보는 전체 인터넷상에 있는 정보의 10% 수준에 불과하다고 한다. 그러나 Corporate Portal에서는 기업 내부의 정보를 포함하여 인터넷상의 전문DB, 신문기사 등 방대한 분야의 정보를 수집 서비스하여야 한다.

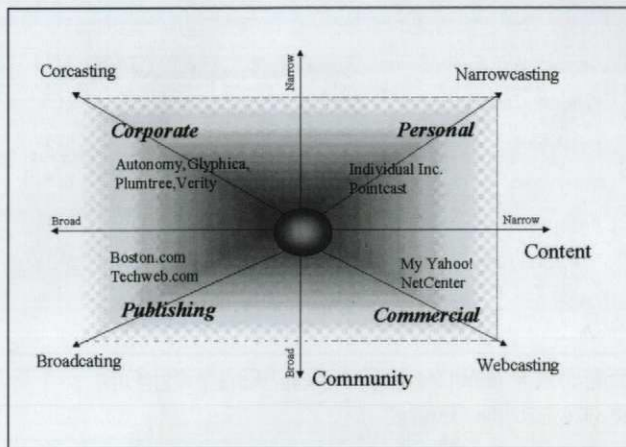
Corporate Portal의 목적

Corporate Portal 구축의 기본 목적은 기업내 조직원들이 원하는 정보를 가장 빠르게 얻게끔 하는 것이다.

포춘1000대 기업중 Corporate Portal을 도입한 기업들을 대상으로 한 설문조사결과 Corporate Portal의 도입 목적으로 32%의 기업이 지식경영시스템(KMS)을 꼽고 있는데 그 이유는 지식경영시스템이 정보의 공유에서 시작하기 때문이다.

Corporate Portal의 구조

Corporate Portal은 일반적으로 크게 7개의 모듈로 구성되어



〈그림 1〉 델파이 그룹이 제시하는 Corporate Portal
(Source: Delphi Group)

있는데 각 모듈은 유기적으로 결합되어 있다.

첫번째 Taxonomy모듈은 정보에 대한 분류체계를 구축관리하는 모듈로 사람 또는 기계학습에 의존하여 운용된다. 이 모듈은 정보원(Information Source)에 의한 분류뿐만 아니라 각종 메타데이터 및 주제어에 의한 분류를 가능하게 함으로써 모든 정보를 구조화한다.

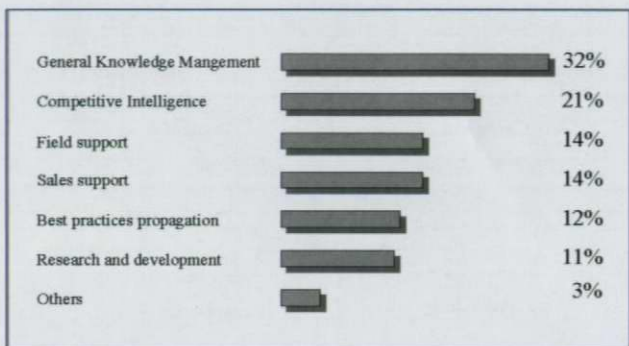
두번째 Metadata모듈은 태그(Tag)데이터를 이용하여 정보에 접근할 수 있도록 하는 모듈로 SGML/XML에서 사용되는 Element나 Attribute, 그리고 DBMS의 필드에 기록되는 각종 서지사항들로 정보를 찾아 볼 수 있도록 한다.

세번째 Search모듈은 전문검색(Fulltext Retrieval)엔진을 이용한 주제어 검색을 제공하는데 단순한 키워드 검색외에 언어학적인 특징을 고려한 유사어, 유사철자어, 동의어 등에 대한 질의어 확장을 가능하게 한다. 이 모듈은 분류방법이나 분류체계를 수정하기 위한 통계적 정보들을 피드백(Feedback)한다.

네번째 Integration모듈은 Metadata와 Search모듈의 결과를 통합하여 중복된 정보를 제거하는 기능을 한다. 이 모듈은 Learning모듈에 의해서 기동된다.

다섯번째 Publishing모듈은 Integration모듈에 의해서 통합된 결과를 출력하기 위한 특정 형식으로 데코레이션하는 기능을 한다. 일반적으로 인트라넷에서의 출력방식은 PDF 또는 HTML이므로 Publishing모듈은 이러한 양식에 맞추어 정보를 출력한다.

여섯번째 Presentation모듈은 사용자와의 인터페이스를 제공한다. 사용자의 요구사항(예를들어 검색질의어)을 받고, 사용자에게 정보를 제공(예를들어 검색결과)하는 등의 역할을 담당한다.



〈그림 2〉 포춘1000대 기업들이 Corporate Portal을 도입한 이유
(Source: Plumtree Software)

다. 이 모듈은 Learning모듈을 기동시키거나 Learning모듈에 의해서 기동된다.

일곱번째 Learning모듈은 사용자의 응답과 시스템 프로세스, 문서간의 반복적인 피드백을 제공하는데 사용자의 요구사항에 맞도록 스스로 개선된다. 이 모듈의 구현을 위해서는 기계학습원리가 필요한데, 신경망(Neural Network) 알고리즘이나 유전자 알고리즘이 적용되고 있다.

지식베이스(Knowledge Base)의 구축

앞에서 논의한 Corporate Portal의 7개모듈 외에 각 기업은 기업 고유의 지식베이스를 구축하여야 한다. 지식베이스는 하나의 주제 단위로 구축된 구조화된 질의어(Query)사전으로, 각각의 주제는 특정 산업(Industry)분야와 관련된 것 뿐만 아니라 소비자의 관심분야에 관련된 것 등 매우 다양하다.

각 기업은 이러한 지식베이스의 구축을 통하여 조직원들간의 공감대를 형성함과 동시에 정보의 공유를 원활하게 실현할 수 있다. 일반적으로 사용자의 검색 질의어들은 대부분 사용자가 속한 산업에 기반을 두게되므로 지식베이스의 질의어들은 기업내부 또는 인터넷상의 정보들을 검색할 때 매우 유용하게 이용된다. 즉, 사용자 스스로 검색 전문가가 될 필요없이 단지 지식베이스로부터 질의어를 선택하기만 하면 전문가처럼 검색할 수 있게된다. 또한 이러한 형식의 지식은 분류체계를 결정하는데 이용하기도 한다.

지식베이스의 질의어들은 질의어 제작 전문가들과 각 분야별 전문가들이 협력하여 만들어지는 것이 일반적이나 조직원들의 검색질의어나 정보수집방식과 관련된 통계적인 자료들이 참조되어 진다.

그렇다면 지식베이스와 시소러스(Thesaurus)와는 어떻게 다른가? 시소러스는 지식베이스에 수록된 여러 종류의 지식 가운데 하나로 Broader-to-Narrow term, Synonym, Related term 등으로 단어간의 관계를 표현하고 있다. 지식베이스는 시소러스와 비교하여 다음과 같은 특징을 가진다.

- CONTEXT(Sentence/Paragraph에서의 단어의 위치와 순서)
- THE MORE EVIDENCE, THE BETTER(시소러스는 모든 단어들을 동일한 비중으로 검색)
- PROXIMITY
- LOCATION/ZONE/SOURCE/CURRENCY/STRUCTURED FIELDS
- INEXACT MATCHES (유사철자, 통계적 관계 등의 표현)
- FREQUENCY(출현빈도)

인트라넷의 구조화

현재 기업내 인트라넷의 문제점들은 앞에서 이미 언급한 바 있듯이 일관성의 결여 및 분산되어 있다는 점이다. 그러나 이렇듯 산재되어 있는 정보들을 통합하여 하나의 서버에서 운영을 하려고 한다면 그것은 새로운 인트라넷을 하나 추가 개발하는 것에 지나지 않을 것이다. 인트라넷의 구조화는 먼저 정보위치의 투명성을 확보하는데서 출발하여야 하는데, 정보위치의 투명성이 확보되면 검색이용자들은 정보의 위치에 상관없이 원하는 정보를 얻을 수 있다.

인트라넷의 구조화를 통하여 기업이 얻을 수 있는 장점은 여러 가지가 있겠지만 그 중에서 가장 특징적인 것은 생산성의 향상이라는 측면이다. 정보가 중요시되는 시대에 원하는 정보를 얼마나 빠른 시간내에 얻을 수 있는가는 곧 기업의 생산성과 직결되는 문제이다.

동일한 정보를 1시간에 얻는 기업과 10분에 얻는 기업의 경쟁에서 누가 이길 것인가? 이러한 정보 획득 시간은 인터넷 시대에 더욱 중요한 요소가 되었다.

기업이 경쟁에서 이기기 위한 요소들은 한마디로 말할 수 있는 성질의 것은 아니다. 하지만 기업이 경쟁에서 이기기 위한 좋은 솔루션 중에 하나는 바로 인트라넷의 구조화를 통하여 기업의 생산성과 마켓 인텔리전스를 향상시키는 것이라고 주저없이 말할 수 있을 것이다.