

눈으로 보는 정보시대를 열어가자

지도 정보를 이용한 정보검색

웹상에서 콘텐츠는 원래 다른사람의 콘텐츠에 새로운 정보를 추가할 수가 없다 물론 방법상으로 가능하다고 해도 그것은 소유권의 문제 때문에 쉽사리 할수 없는 일이다. 하나 열린 콘텐츠를 이용해 기존의 콘텐츠에 새로운 정보를 추가하는 방법들이 사용되어야만 하고 정보검색과 정보추가 등 사용자 입장에서 다양하고 효율적인 검색이 이뤄져야한다. 이에 지도 정보를 이용한 정보의시각화와 새로운 검색의 방법을 알아보자. 더불어 지도정보는 다양한 콘텐츠와의 연계를 통해서 정보이용의 새로운 길을 열어줄 것으로 기대되고 있다. <편집자>

정보의 비주얼화

오늘날의 정보 시스템은 컴퓨터 처리능력과 기억용량의 증대와 더불어 현저한 진보를 거듭해 왔다. 1980년대 후반부터 폭발적으로 보급되기 시작한 정보통신과 인터넷 환경은 진보를 더욱 가속시키고 있다. 초창기 정보 시스템은 처리용량의 제한으로 대개 문자나 수치로 되는 텍스트 위주의 데이터만을 처리할 수 있었기 때문에 이전의 정보 검색은 텍스트 정보를 대상으로 했다. 그러나 최근에는 정보 시스템들이 비주얼 데이터 등의 비텍스트 형식의 데이터를 취급하고 있어서 비주얼한 정보 검색이 필요해졌다. 비텍스트 정보는 종래의 정보 시스템과 비교해 볼 때, 검색 기술에 한정되지 않고 이용환경, 사용자 인터페이스 환경, 그리고 어플리케이션에 이르기까지 아주 많이 다르다. 따라서 본 원고에서는 근래 실용화된 비텍스트 정보의 하나이며, 그래픽의 대표적인 용도로서 전자지도 미디어를 이용한 정보 시스템과 거기에 적합한 새로운 비주얼한 정보 검색기술의 모습에 대해서 알아보기로 한다.

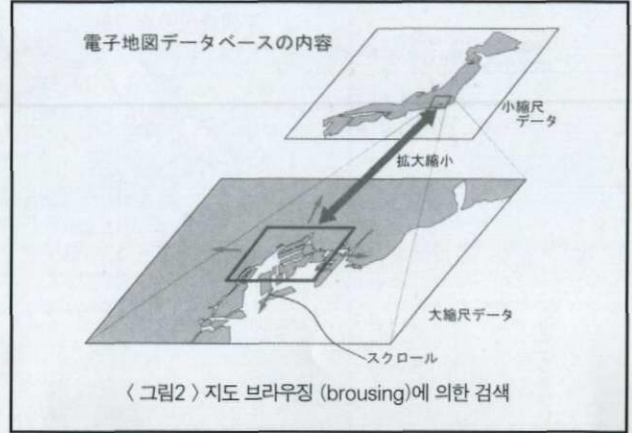
전자 지도

지도는 예로부터 종이 위에 선이나 면, 문자, 기호 등의 도형을 사용해서 지형이나 거리의 집 등을 표현해온 수치 정보이다. 전자 지도는 이를 컴퓨터에 축적해서 그래픽 기능을 이용하여 표시하는 정보 미디어로 빈번하게 활용되고 있다. 컴퓨터 표시 기능이 진보함에 따라 전자지도는 2차원의 스크린 영상에서 3차원의 형상으로도 표현할 수 있게 되었다.

전자 지도의 데이터 형식은 2차원이나 3차원에 의존하지 않고

크게 2가지의 형식으로 나누어볼 수 있다. 하나는 인터넷의 웹 브라우저에서 많이 이용하는 GIF나 JPEG 형식 등으로 그래픽을 점의 집합으로 표현하는 것이다. 비트맵 또는 래스터 형식이라 불리는데, 항공사진이나 위성화상 지도도를 이미지 스캐너를 통해 전자화 한 것 등이 그 구체적인 예이다. 반면에 건설 도면이나 전자회로 도면 등을 작성할 때 주로 이용하는 CAD(Computer Aided Design) 시스템이나 일러스트에서 주로 이용하는, 그래픽을 선이나 폴리곤이란 도형으로 표현하는 벡터 데이터를 이용한 전자지도가 있다. 벡터 데이터 형식의 전자 지도는 래스터 데이터에 비해 작성하기가 곤란하기 때문에, 일반적으로 CAD 시스템에 의해서 작업자가 수작업으로 도형을 입력해서 데이터를 작성한다. 반면에 데이터를 이용하는 면에서는 벡터 쪽이 훨씬 쉽고 응용성도 높으며, 필요한 처리능력도 상대적으로 낮아서 좋다.

전자지도는 이미 여러 곳에서 이용되고 있다. PC용 지도장 소프트웨어를 곧잘 이용하는데, 최근에는 종이 위에 인쇄된 일반 지도이다. 예전의 종이 지도는 컴퓨터를 이용하지 않고 제판했었는데, 요즘엔 대부분의 지도를 전자 지도인 CAD 시스템을 많이 이용하여 작성한다. 그 때문에 손으로 표현하기가 곤란했던 지형의 형상을 음영을 첨가하거나 위성화상을 기초로 해서 만드는 등, 매우 다양한 형태의 지도가 출시되고 있다. 네트워크상의 공동체인 네티즌이 생각해 낸, 이 웹의 자유로운 정보발신 환경은 인터넷상에서 누구라도 이용할 수 있는 정보 발신 서버(웹 서버)와, 정보를 쉽게 표현할 수 있는 정보 미디어 포맷(HTML), 즉 정보 리터라시 (* 역주 : 읽고 쓸줄 아는 능력, 즉 정보를 해독하고 작성하는 능



력)를 쉽게 획득할 수 있는 것, 그리고 정보끼리의 제휴를 가능하게 만든 커뮤니티 조성 하이퍼링크로 구성되어 있다.

한편, 웹 정보 미디어로서 전자 지도를 생각했을 경우, 정보를 발신하는 입장에서 정보 자유로이 사용할 수 있어야 한다는 점이 매우 중요하다. 그런데 전자 지도에 의한 정보발신은 일반적으로 정보 발신 측과 수신 측이 명확하게 분단되어 있는데, 이것은 웹 상에서도 마찬가지이다. 그런 의미에서 본다면 유감스럽게도 현재 상태의 전자 지도는 웹 정보 미디어로서는 아직 효과적이지 못하며 개선의 여지가 남아 있는 미디어라고 말할 수 있다.

웹 상의 전자 지도 검색

일단 전자지도가 누구에게든 발신 가능한 웹 정보 미디어로서 성립된 경우에는 현재의 웹 콘텐츠와 같은 몇 가지 문제점이 발생된다. 그 하나가 정보 검색의 곤란함을 들 수 있다. 웹 정보는 하이퍼링크를 활용해서 정보를 연결시킴으로서 검색이 가능하지만, 정보 발신측에서 반드시 관련성을 만들어야 하고 실제로 콘텐츠끼리의 연결은 현재의 웹 상에선 그다지 보편화되어 있지 않다. 즉 정보 작성은 용이하지만 그대로는 사용하기 곤란한 정보 시스템이 되어 버린다. 그래서 정보끼리 중개하는 인덱스 역할을 하는 Yahoo나 Infoseek와 같은 검색 사이트가 출현해 중요한 역할을 담당한다. 이들 검색 사이트는 사용자가 제공한 검색 조건에 일치하는 정보에 관해서 하이퍼텍스트에 의한 정보 인덱스를 제공한다.

그러나 웹상의 전자 지도 미디어를 생각할 경우, 웹 콘텐츠끼리 연결시키는 검색 엔진에는 조금은 다른 역할이 추가될 필요가 있다. 앞서 말한 것처럼 지도는 똑같은 도로를 표시한다고 해도 영역이나 축척 등이 서로 다른 경우가 존재한다. 그리고 그 지도들의 검색이나 변환은 브라우징에 의해서 비주얼하게 행해지는 것이 자

연스럽다. 이것을 웹 구성에 실제로 장착하면 [그림 4]와 같은 구성이 예상된다. 인덱스 지도와 실제의 지도에는 기본적으로 큰 구별은 없고 용도나 축척, 지역이 다른 전자지도 콘텐츠가 계층적으로 분산되어 링크된 시스템이다.

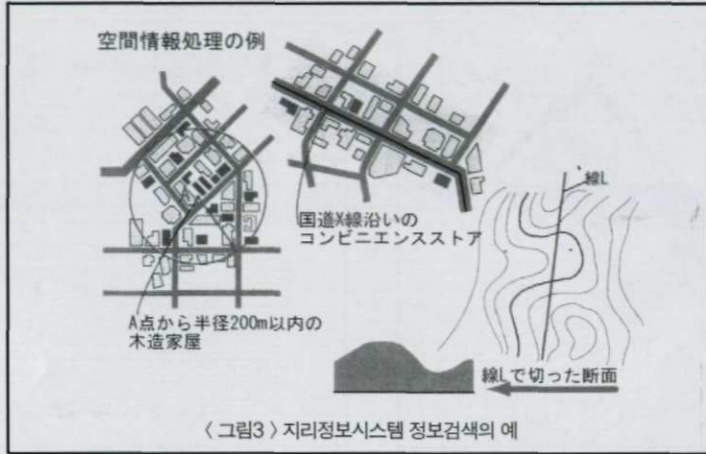
중·대 축척용 지도 검색 사이트에는 콘텐츠 관련성이 불충분한 것이 많아 일반적으로 소축척의 인덱스 지도를 제공하는 검색 사이트가 이용된다. 예를 들면 미국 연방지리 데이터위원회(FGDC)에서는 어떠한 지리정보가 어디에 있는지를 기술한 메타 데이터를 Clearinghouse 라고 불리는 데이터베이스로 관리하려고 한다. 일본에서도 일본 상공 데이터 기반 추진협의회(NSDIPA)가 데이터 정비 추진을 관계부처에서 하고 있다.

전자 지도의 리터러시와 새로운 정보 유통 스타일

웹 상에서의 전자 지도 발신의 모델은 앞에서 설명한 바와 같이 생각해 볼 수 있는데, 아직까지 누구나 발신할 수 있는 미디어로서 확립되지 못한 근본적인 문제가 남아 있다. 본 장에서는 그 문제를 해결하기 위해 개발하고 있는 JaMaPS라 불리는 새로운 정보유통 스타일을 가진 웹 플랫폼에 대해서 말하기로 한다.

전자 지도를 이용한 정보 발신에서 대두되는 커다란 문제는 지도제작이 쉽지 않다는 점이다. 지도를 개인적으로 이용할 경우, 시판된 지도를 구입해서 그 지도상에 각자가 추가하고 싶은 정보를 추가하는 방법이 가장 흔하다. 극히 일부이지만, 기본이 되는 지도상에 각자가 독자적인 지도 콘텐츠를 작성해 추가하는 것이 개인들이 만드는 지도 작성의 예이다. 이렇게 하면 쉽게 지도를 작성할 수 있다. 여기에서 설명하는 시스템은 이 개념을 웹에 적용시킨 것이다.

웹에서는 보통 다른 사람의 콘텐츠에 정보를 첨가할 수 없다. 비



록 가능하다고 해도 소유권 문제가 크게 발생한다. 그런데 방금 설명한 방법으로 지도에 정보를 첨가하는 작업이 업무상 널리 행해지고 있다. 본래의 지도 위에 투명한 필름을 덮고 그 필름 위에 지도를 그려 가는 방법이다. CAD나 도로관련 소프트웨어에서는 이러한 방법을 레이어 처리라고 한다. 본래의 도면을 변경하는 것이 아니라 새로운 정보를 추가로 표시하는 것이다. JaMaPS에서는 이 레이어처리를 분산된 웹서버에 저장된 각 레이어에 해당하는 지도 콘텐츠와 그들에 액세스해서 화면에 각각의 서버안 지도를 하나의 레이어로 처리해서 거듭 표시하는 새로운 웹 브라우저를 사용해 재현했다(그림 5)

이 때 문제가 되는 것은 각 레이어가 같은 장소에서 똑같이 맞춰져야 한다는 점이다. 보통 웹상의 전자지도는 GIF나 JPEG라는 그래픽으로 표현되는데, 거기에는 그림을 명확하게 나타내는 정보가 실려있지 않다. 그래픽으로 표현된 지도를 똑같은 장소에 똑같이 맞추려면 각각의 그래픽에, 예를 들면 위도나 경도의 좌표처럼, 일반적인 글로벌 좌표계를 사용하면 간단하다.

이러한 기법은 그래픽으로 표현된 전자 지도에서 일반적으로 이용 가능한 방법이며, 종래에 하나의 웹 서버에서 획득한 콘텐츠만을 표시하던 것에서 여러 개의 이른바 모듈화된 콘텐츠를 동시에 이용하는, 예전엔 볼 수 없었던 새로운 정보유통 스타일을 제공한다. 게다가 이들 콘텐츠는 비유일한 편성에 의한 상승효과로 가치가 높은 정보를 만들어 낸다.

본 장의 서두에서 말한 것처럼 레이어 처리를 이용한 이와 같은 방법은 원래의 도면 위에 고쳐 쓴 정보만을 별도의 서버로부터 독립한 정보로서 발신가능하기 때문에 그것만으로도 정보작성이 용이하다. 물론 원래의 데이터에 반복 사용할 수

있는 합의가 네트워크 상의 커뮤니티로 형성되어야 하는데, 원래 데이터를 복제해서 조금 더 정보를 고치고 발신하는 것에 비교하면 쉽게 합의를 얻을 수 있고, 또한 이러한 방법으로 인해 원래 데이터가 복제되어 사용될 위험성도 감소시킬 수 있다.

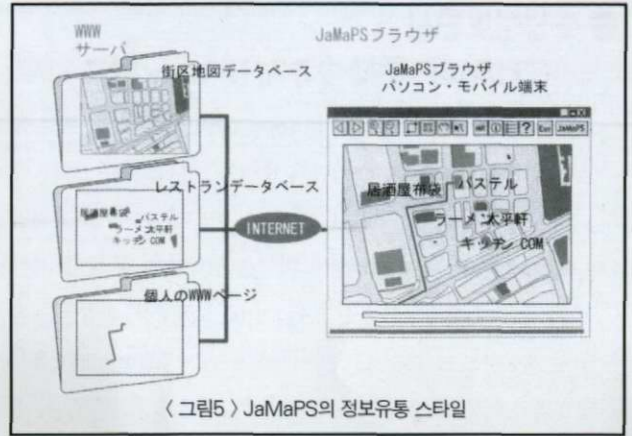
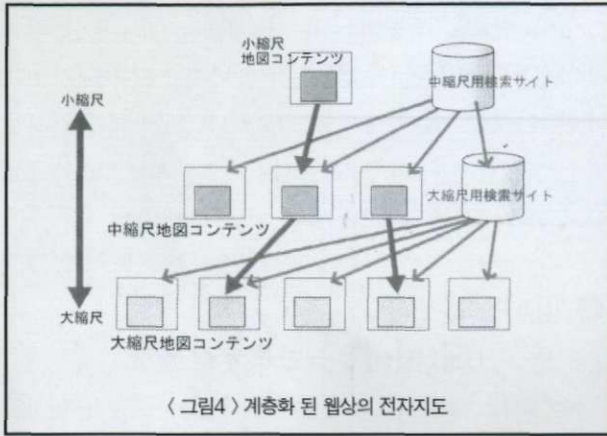
새로운 정보 검색 스타일

이상과 같은 정보유통 스타일은 지도 정보 검색에 더 새로운 기능을 필요로 한다. 이전의 지도 정보는 각각 모듈화되어 있기 때문에 개별로는 그다지 쓸모가 없고, 하나의 콘텐츠만을 검색할 수 있는 종래의 검색 서버로도 사용하기 힘들다.

그런데 모듈화된 정보를 비유일로 거듭 짜맞춤으로서 사용자의 요구를 충족시킬 수 있는 정보 조직을 검색할 수 있는 서버가 필요하다. 예를 들면 신주쿠역 주변에 주차할만한 공간을 찾을 경우, 역 주변 도로의 지도 사이트와 주차장 정보를 제공하는 지도 사이트를 찾아서 이런 것들을 다시 사용하도록 사용자에게 안내해줄 필요가 있다. 이를 위해서는 먼저 각 모듈화된 정보의 특성을 인덱스로 축적해야 한다.

한나는 Yahoo처럼 그 정보가 어떤 정보인가 하는 인덱스 정보를 사전에 정보 발신자에게 등록하도록 하는 방법이다. 앞의 Clearinghouse에서는 콘텐츠 작성자가 메타 데이터를 작성해서 그것을 스스로 등록하게 한다.

반면에 Infoseek처럼 정보수집 로봇이 자동적으로 인터넷상의 웹 콘텐츠의 내용을 순회해서 수집하는 구조와 인덱스 정보의 자동생성기술을 통합하는 것을 생각할 수 있다. NTT가 얻은 모바일인포서치 실험에서는 LIG라 부르는 로봇이 사용되어 위치정보를 가리키는 언어가 포함된 앵커를 검출해서 그 전에 콘텐츠를 우선적으로 수집하는 전략을 취한다. 다만 인덱스를



생성할 때쯤 되면 콘텐츠의 내용 자체를 해석할 필요가 생기기 때문에 등록방식에 비해 인덱스의 정밀도는 좋지 않다. 인덱스 생성의 성능은 자연언어처리기술, 특히 HTML 등의 콘텐츠 해석 기술을 따르는 경향이 많은 것 같다.

SYG (Scalable Vector Graphics)

전자지도를 표현하는 데는 벡터데이터와 래스터 데이터 두 가지의 형식이 존재하며 벡터데이터 쪽이 이용하기 쉽다는 것은 앞서 말한바 있다. 더욱이 앞장의 마지막에 기술한 로봇형의 검색 시스템에서는 더더욱 벡터 데이터 쪽이 이용하기 쉽다. 벡터형의 데이터에서는 그래픽이 섬세한 도형데이터(그래픽 대상)의 집합으로 구성되어, 지명이나 역명 등의 문자사례 정보도 도형정보의 일종으로 쉽게 처리할 수 있도록 기술되어 있기 때문이다. 따라서 문자를 대상으로 한 현상태의 검색 시스템에서도 쉽게 인덱스 정보가 생성가능하고, 그래픽 대상을 인덱스로한 검색 방법도 실현가능하다.

그에 비해서 래스터 데이터에서는 이용자에게 문자로서 인식할 수 있는 정보도 실제로는 점의 집합으로밖에 기술되어 있지 않기 때문에 화상인식기술과 같은 고도의 기술을 점의 상태로 검색 서버를 재구축하지 않으면 안 된다.

그런데 웹 상에서 유통시킬 수 있는 그래픽 데이터는 지금 으로서는 GIF와 JPEG 두 종류이고 이들은 모든 래스터 데이터이다. 그에 비해서 현재 W3C에서는 인터넷의 표준 벡터 데이터포맷으로 해서 SVG(Scalable Vector Graphics)를 책정하고 있다. 본 포맷은 XML에 근거하고 있기 때문에 HTML 등의 웹 콘텐츠와 매우 친화적이고, XML의 특성을 살려서 그래픽 대상에 지리정보 시스템에서 불가결한 속성 데이터를 유


연하게 추가할 수 있도록 한다.

다른 정보미디어와 전자지도의 통합

근래 다른 종류 미디어 사이의 변환기술은 근래 매우 주목받고 있다. 대 화면용 HTML과 극소 화면용 텍스트 사이의 변환기술로서 문서 요약기술이나 HTML과 음성간의 변환기술 등, 특히 모바일 컴퓨팅 서비스를 하나의 타겟으로 해서 서비스가 시작되고 있다.

전자지도는 실제 공간에 흩어진 물체나 정보 등을 시각적으로 표현하는 것에 뛰어난 표시방법을 사용자 단말기에 제공한다. 게다가 벡터 그래픽 데이터를 이용하면 확대 축소해도 화상이 깨지지 않고, 스크랩도 용이하기 때문에 특히, PDA 등의 휴대정보기기에서도 화면 사이즈나 전송속도에 유연한 서비스가 제공 가능하다.

그래서 지도로서 표현되지 않는 공간정보, 예를 들면 HTML 콘텐츠 안의 전화번호나 주소, 추상적인 표현 등으로 기술한 정보를 미디어 변환을 통해 지도상에 표시하는 것이 가능해지면, 전자지도를 이용한 검색이 효과적으로 활용될 수 있다. [모바일 서치 실험]에서는 주소가 기술된 콘텐츠에 대해서 위도와 경도를 부여하고 있다. 앞으로도 위치정보를 부여하고, 공간정보를 검색에 이용할 수 있는 정보의 양은 급속하게 증가해갈 것으로 기대된다.

이상과 같이 특히 웹상에서 정보 미디어로서의 전자지도는 아직까지는 미숙하지만 웹의 이용범위나 정보유통 스타일을 크게 바꿀 만한 잠재력을 갖고, 앞으로 크게 발전할 수 있는 정보 미디어이다. 21세기를 향한 새로운 사회기반, 정보시스템 기반으로서 전자지도가 널리 공헌할 것을 기대한다. 

- 다카키 사토루
- 마쓰모토 가즈노리