

메타데이터를 활용한 그룹형 커뮤니티의 자동생성

*윤선정[○], *주우석, *윤태수, **김기홍

동서대학교 *디지털컨텐츠학부[○], **영상매스컴학부

ysj0827, savrang, tsyun, khkim@gdsu.dongseo.ac.kr

Automatic Generation of Group-type Community making efficient use of Metadata

*Sun-Jung Yoon[○], *Woo-Suk Joo, *Tae-Soo Yoon, **Gi-Hong Kim

Division of *Digital Contents, **E-Visual Communication, Dongseo University

요 약

본 논문에서는 근래에 폭발적인 성장을 하고 있는 1인 미디어의 대량의 데이터 가운데서 양질의 정보를 집중적으로 관리하고 효과적인 검색 기능을 지원하는 그룹형 커뮤니티 시스템을 구축하기 위하여 메타데이터를 활용하는 것을 제안한다. 이를 위해 특별히 교육정보만을 대상으로 하여 여기에 사용될 메타데이터 기술 요소를 개발하고 교육용 데이터에 적용 가능한 적정 카테고리를 개발하였으며 이를 검증하기 위하여 그룹형 교육 커뮤니티 EduLOG(Educational blog) 서비스를 구축하였다.

이 시스템은 새로운 교육용 커뮤니티를 개설하는 것이 아니라 기존의 많은 사용자층을 가지고 있는 1인 미디어를 활용하여 유용한 정보를 생성해 내고 공급하는 것이 가능하다는 것과 나아가 메타데이터 요소의 활용에 의해 인터넷 상에서 정확성과 신속성을 지원하는 검색 시스템 구축이 가능하다는 것을 보여준다.

I. 서론

인터넷 상에서 자신들의 관심 분야나 특정 주제를 이용한 정보 공유와 인간관계 형성을 위해 이용자들 스스로 만들어 가는 공간으로서의 카페나 클럽과 같은 그룹형 커뮤니티와 최근에 새롭게 등장한 1인 미디어인 블로그나 미니홈피는 사용의 편의성을 제공하고 관심사를 자유롭게 표현이 가능해 폭발적인 성장을 하였다[1]. 그러나 이런 서비스들은 정보 축적의 특성상 제공되는 정보의 양에 비해 양질의 정보를 검색하는데 많은 한계가 있다. 예를 들어 동일한 주제를 가진 수많은 카페의 존재, ‘뽀’, ‘스크랩’ 이라는 1인 미디어들의 특별한 복사 기능 등으로 수많은 중복 자료가 존재하여 검색은 신속성과 정확성이 떨어질 수밖에 없다.

이에 따라 본 논문에서는 새로운 개념의 시스템을 제안하면서 메타데이터 요소들에 의한 체계적인 데이터 관리를 통해서 이 시스템이 신속하고 정확한 검색 기능을 제공할 수 있는지를 살펴보고자 한다.

이를 검증하기 위해 특별히 교육용 정보만을 선별하여 EduLOG(Educational blog)라고 명명된 서비스를 구축하였다. EduLOG 서비스는 카테고리과 메타데이터에 의해 자동으로 구성되는 체인형 블로그로서 유용하고 건전한 교육 정보의 공유로 신속하며 효율적인 검색이 가능하면서도 동시에 개인의 욕구 충족 공간을 제공하는 1인 미디어의 장점과 그룹형 커뮤니티의 장점을 모두 제공하는 새로운 서비스이다.

II. 관련 연구

1. 커뮤니티의 유형

커뮤니티 서비스의 접근은 매우 편리하고 개방적이어서 다양한 이용자들이 존재하므로 많은 사람들과 함께 많은 정보들이 모이게 된다[2]. 웹 커뮤니티는 < 표 1 > 에 나타낸 것과 같이 공동의 관심사나 주제를 이용하여 이용자 그룹을 만들고 정보

공유 및 축적 서비스를 지원하는 그룹형 커뮤니티와 지극히 개인적인 관심사나 자료들을 관리하거나 자기만의 인간관계를 만들어 가면서 개성을 표출하는 1인 미디어가 있다.

< 표 1 > 그룹형 커뮤니티와 1인 미디어의 특징 비교

구분	그룹형 커뮤니티	1인 미디어
형태	다 대 다	일 대 다
운영 주체	회원 공동	본인
관리 형태	관리자(시삽) 따로 있음	본인
자료 유형	생성 시 주제(분류) 영역 지정	본인의 모든 관심 영역
사용 권한	회원별로 차등	본인에게만 편집 권한 있음
서비스명	카페(다음), 클럽(네이버) 등	미니홈피(싸이월드), 블로그(네이버) 등

특히 1인 미디어 중 ‘미니홈피’가 새로운 사이버 인맥을 구성하고 자기발전의 수단으로 뽐내기의 욕구 충족을 제공하는 독특한 형태라면 ‘블로그’는 ‘웹에 쓰는 일기’라는 ‘웹 로그(Web Log)’에서 유래한 것으로 ‘미니홈피’와 유사한 특성을 가지지만 텍스트 위주의 정보 제공의 매체적(Publishing) 기능이 더 부각된 미디어이다.

2. 커뮤니티 서비스의 문제점

‘카페’나 ‘클럽’과 같은 그룹형 커뮤니티의 생성은 접근이 용이하고 다양한 분류를 지원하는 등 진입 장벽이 낮은 편이지만 단순 이용자들에게 폐쇄적인 경향이 많은 편이다. 이용자가 해당 커뮤니티를 이용하려면 회원 가입, 가입 후 일정 기간의 활동을 통한 인증, 관리자의 등급 부여 등이 필요하다. 또한 회원 개인만의 욕구 충족도 자유롭게 지원해 주지는 못하므로 지극히 개인적인 용도의 자료나 글들을 관리하는 것은 쉽지 않다[7].

1인 미디어 중 ‘미니홈피’는 특성상 유용한 정보의 공유

기능이 약하며 늘어나는 사용자 관리를 위해 실명제 채택으로 개인정보 유출이 매우 심각한 상태이며 '블로그'는 '미니홈피'에 비하여 컬럼 성격을 가지고 있고 정보의 매체적(publishing) 역할이 더 강하지만 다른 사람들에게 그 정보가 노출되는 확률은 낮다.

그러나 커뮤니티 서비스들이 무엇보다 문제가 되는 것은 질적으로 유용하지 못한 글들을 효과적으로 차단하기가 어렵다는 것과 특정 키워드로 검색하면 그 내용을 주제로 한 유사한 커뮤니티가 너무나 많이 존재하고 있다는 점이다. 또한 '블로그'의 '팝', '스크랩' 등의 복사 기능은 자료의 중복을 야기 시켜 원하는 정보를 신속하게 찾는 데 걸림돌이 되기도 한다. 예를 들어 '다음'에서 'java'를 주제로 하는 '카페'를 검색한 결과 681개가 검색되었지만 이 가운데 상당수의 카페는 비활성 상태의 휴면카페로 유용한 정보가 거의 없다. 또 '네이버'에서 'java'를 키워드로 하여 블로그를 검색하면 동일한 주제를 가진 글이 3,7606 개의 서로 다른 블로그로 검색되는데 이 중 상당수가 동일한 글임을 금방 알 수 있다.

3. 메타데이터 기술요소

자동 검색엔진을 이용한 검색시스템은 많은 장점에도 불구하고 정확성과 신속성이 떨어진다는 것을 인정해야 한다. 이 문제를 해결하기 위한 시도들 중에 XML 문서를 이용한 표준화된 문서 제작이 매우 흥미롭지만 정보의 팽창 앞에서는 완전한 해결책이라고 할 수 없다.[8] 따라서 웹 커뮤니티와 같은 대량의 데이터를 필터링하는 도구로서 메타데이터 시스템이 필요한데 메타데이터는 네트워크상의 정보의 소재나 내용을 기술하고 이를 탐색하기 위한 요소들의 집합이란 의미로 통용된다. 그러므로 메타데이터는 원하는 정보를 체계적으로 선별하여 신속하고 정확한 검색이 가능하게 하며 자원의 공유와 이용을 돕는 효과적 장치라 할 수 있다[9]. 메타데이터 요소를 도출하는 과정은 다음과 같다.

- (1) 국내의 메타데이터 요소 셀의 조사
- (2) 관련 작업에 필요한 메타데이터 요소 셀의 선별
- (3) 표준 메타데이터 셀 기준으로 선별된 셀의 매핑과 공통 요소 추출
- (4) 공통요소를 필수와 선택 요소로 분류

III. 새로운 커뮤니티 서비스와 메타데이터 기술요소 도출

1. 서비스의 배경

본 연구는 위에서 언급한 웹 커뮤니티들의 문제점들을 개선하고 사용자가 원하는 유익한 정보가 집중적이며 효과적으로 관리되어 손쉽게 접근이 가능한 시스템의 구축이 필요하다는 관점에서 출발하였고 이를 위해 메타데이터 시스템의 활용이 효과적인지 검증하는 차원에서 EduLOG(Educational blog)서비스 시스템을 구축하였다.

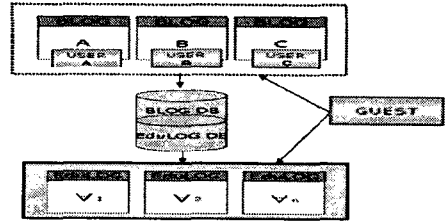
2. EduLOG 서비스 시스템

EduLOG은 'Education'과 'Blog'의 합성어로 많은 블로그 이용자들이 자신이 올리는 정보 중 양질의 교육 정보를 저장할 때 아주 간단한 조작만 추가하면 카테고리별로 통합되어 서비스되는 교육용 커뮤니티 시스템이다. 이 서비스는 기존의 블로그에 간단한 기능을 추가한 것으로 특별히 교육용으로 유익한 정보만을 대상으로 선별한 후에 교육용 정보 카테고리를 제시하고 메타데이터에 의한 체계적인 관리로 양질의 정보가 손쉽게 검색될 수 있도록 하는 서비스이다.

EduLOG 서비스 페이지는 전형적인 게시판 모습을 띄고 있는데 여기는 이용자들이 올린 블로그들이 자동으로 분류되어 관리되는 목록정보 뿐이며 각 블로그 정보는 각 개인의 블로그

에서 관리된다. [그림 1]은 EduLOG 서비스의 개념도이다.

EduLOG 서비스의 프로그램은 ASP로 구현하였으며 MS-SQL 2000 SERVER에 DATABASE를 구축하였다.



[그림 1] EduLOG 서비스 개념도

3. 메타데이터 기술요소 도출

이 시스템을 위하여 활용한 메타데이터의 선정은 세계적인 메타데이터 개발의 기반이 되고 있는 더블린코어(Dublin Core)의 메타데이터 요소를 사용하였는데 여기에는 Title, Creator, Subject, Description, Publisher, Contributor, Type, Date, Format, Identifier, Source, Language, Relation, Coverage, Rights 과 같은 요소들이 포함되어 있다 [8]. 이 중 EduLOG 서비스에 사용될 메타데이터 요소는 웹 자원의 특성상 불명확한 정보를 제거하고 검색 기준 항목으로 직접적인 키워드로 사용할 수 있으며 이용자의 추가적인 작업을 최소화 하는 것을 기준으로 선별하였는데 위에서 언급한 더블린코어 요소 중 제목, 식별자, 주제, 저자, 날짜, 설명을 선별하여 매핑하였다.

4. 교육용 카테고리 요소

교육 정보는 웹에서 다양한 형태와 정형화 되지 않은 분류 기준으로 인해 너무나 많은 종류로 나누어진다. 그러므로 본 시스템을 위해서 검색의 효율성을 높이기 위해 카테고리의 구성이 필요하였다. EduLOG에 블로그를 제공하기 위해 이용자에게 제공되는 카테고리는 교육인적자원부고시 제2004-5호(2004.6.9) 내용과 교육방송(EBS), 한국 교육학술정보원, 한국 산업인력관리공단 등에서 채택하고 있는 카테고리를 참조하여 구성하였다.[10,11,12,13] 이것은 유사한 영역으로 분류가 모호한 부분은 합하고 정보의 양이 상대적으로 적은 부분들은 제하는 등 간략화 시켜 교육영역 32가지로 구성하였다.

IV. 새로운 커뮤니티 서비스 시스템의 구성과 검색

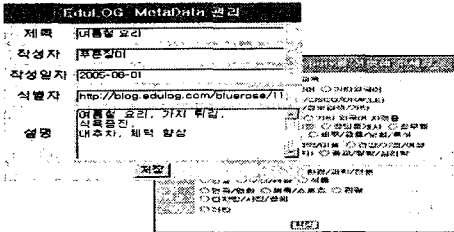
EduLOG 서비스 시스템의 구성은 크게 컨텐츠 저장을 위한 몇 가지 단계와 컨텐츠 검색 단계로 구성된다.

1. EduLOG 서비스의 저장

이용자는 자신의 블로그에 글을 저장하는 마지막 단계에서 EduLOG에 등록 여부를 선택할 수 있는데 원하지 않으면 자신의 블로그에만 글이 등록되고 EduLOG 서비스로 진행이 이어지지 않는다. 그리고 나면 EduLOG의 검색 기능과 정보의 분산 기능을 증가시키기 위해 메타데이터를 결정하는 단계이다.

[그림 2]의 항목 중 제목, 저자, 날짜 정보, 식별자(블로그의 URL)는 자동으로 출력되며 수정 가능하고 설명은 작성자가 입력하면 된다. 여기서 식별자와 제목은 블로그의 '스크랩' 기능의 문제점을 해결하는 요소로 볼 수 있다. 사용자가 다른 블로그에서 스크랩한 글을 자신의 블로그에 올린 경우에 제목, 식별자 등을 분석하여 변경 사항이 없는 경우는 EduLOG에 등록되지 않는다. 그리고 주제는 카테고리 선택 작업으로 대체되는데 이 때 적절한 카테고리를 선택하는 문제는 전적으로

이용자의 몫이지만 사이트에서 제공되는 도움말을 이용하여 더 적합한 카테고리 지정이 가능하고 때로는 '기타' 영역으로 분류해 두면 관리자에 의해 재분류될 수 있다. 여기서 글이 등록될 때 사용하는 메타데이터는 스크랩한 글이 중복되어 등록되는 것을 예방하는 기능을 제공한다. 이것은 블로그를 이용한 검색에서 가장 큰 문제점을 개선한 것이라고 볼 수 있다.

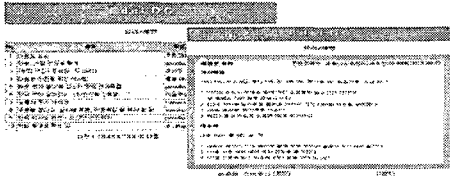


[그림 2] 메타데이터 등록과 카테고리 결정

그 다음 단계로 사용자가 EduLOG에 글을 등록하면 일정 요건을 충족하지 않는 글을 필터링하는 과정을 거치는데 본 시스템에서는 자동화 도구를 사용하는 과정을 실제 적용하지 않았다. 추가로 질적으로 우수한 자료 관리를 위해 관리자가 수동으로 시스템을 모니터링하여 주기적으로 등록된 글을 점검, 정리, 재분류할 필요가 있다.

2. EduLOG 서비스의 검색

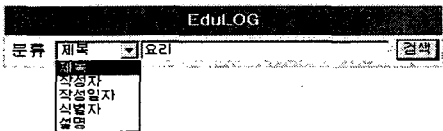
EduLOG 서비스의 검색은 키워드 검색외에 카테고리를 이용한 검색과 메타데이터를 이용한 검색이 가능하다. EduLOG 서비스 페이지에서 제시된 카테고리를 지정하면 [그림 3]의 좌측과 같은 지정된 카테고리의 목록을 보여준다.



[그림 3] 특정 카테고리의 EduLOG 서비스 페이지와 원하는 글 보기

여기서 글이 읽히면 조회 수가 증가하고 해당 글의 작성자를 클릭하여 등록자의 블로그로 이동할 수 있다. [그림 3]의 우측은 EduLOG 서비스에서 특정 글의 내용을 보여주는 페이지이다.

또한 EduLOG 서비스에 등록된 글은 검색의 신속성과 효율성을 위해 메타데이터를 이용한 검색이 지원되는데 메타데이터와 키워드의 조합으로 원하는 정보 검색이 더 정확하고 빨라질 수 있다. [그림 4]



[그림 4] 메타데이터를 이용한 검색

메타데이터에 의한 자동 필터링 기능은 동일한 글들이 여러 블로그에서 제공되는 문제점을 개선하고 카테고리별로 집중적으로 분류되어 관리됨으로서 검색의 효율성을 높일 수 있다.

V. 결론

미니 홈페이지 블로그와 같은 1인 미디어에 의한 대량의 정보 가운데 일반 이용자들은 질적으로 우수한 정보를 찾는 데 많은 수고와 요구된다.

본 논문에서는 기존 커뮤니티 서비스의 문제점을 분석하고 이것을 개선하기 위해 EduLOG 서비스를 제안하고 구축하였다. 이 서비스는 개인은 기존의 블로그에서 활동하는 것과 크게 다르지 않지만 다른 이용자들에게 유익이 되는 글을 교육용 정보로 등록하면 자동으로 지정된 카테고리나 메타데이터에 의해 그룹형 커뮤니티에 포함되는 서비스이다.

특히 이 시스템을 구축하는데 있어서 검색의 효율성과 정확성을 높이기 위하여 더블린 코어에 의한 메타데이터 요소를 도출하고 자동으로 메타데이터에 의한 분류 및 필터링이 가능하도록 구성하였다. 또한 유용한 교육 정보를 관리하기 위한 카테고리기를 개발하고 적용하였다. 특별히 메타데이터를 이용한 관리 기능은 블로그에서의 흔히 볼 수 있는 무수한 중복된 글을 걸러내고 양질의 정보가 집중 관리될 수 있게 해주며 키워드 검색과의 조합으로 유용한 교육정보를 신속하고 정확하게 검색이 가능하게 지원해 준다. 그러므로 EduLOG 서비스는 1인 미디어의 장점과 그룹형 커뮤니티의 장점을 동시에 갖고 있으면서 유용한 교육 정보를 제공하는 지식 검색소의 역할을 담당할 수 있게 된다.

참고문헌

- [1] 김철수, 싸이월드에는 과연 다음을 넘어서었을까? : 상식을 뒤집는 웹 기획 아이디어, 길벗출판사, 2004
- [2] 최선희, 인터넷 커뮤니티의 현황과 전망, 정보통신정책, 제 12권, 11호, pp. 62-66, 2000
- [3] 이태석, 신수미, 다중 웹사이트간 상호연동을 통한 게시판 통합운영시스템에 대한 연구, 한국정보과학회 추계학술발표대회, 제 30권, 2호, pp. 472-474, 2003
- [4] Daum Communications Corp, http://tab.search.daum.net/dsa/search?nil_profile=g&nil_search=btn&w=cafe&q=java.
- [5] NHN(주), <http://cafeblogsearch.naver.com/search.naver?where=post&query=java&start=7>
- [6] SK Communications, <http://www.cyworld.co.kr>
- [7] 유진환, 김길성, 편현장, 이진아, 한영석, 커뮤니티 블로그의 설계, 한국정보과학회 추계학술발표대회, 제 31권, 2호, pp. 595-597, 2004
- [8] 한국과학기술정보연구원, 청소년 대상 과학정보 메타데이터 구축, 성균관대학교 부설 정보관리연구소, pp. 33-34, 2003
- [9] 김홍기, 보이지 않는 공간의 혁명, 월간 마이크로소프트, pp. 244-251, 2002.4
- [10] 교육인적자원부, http://www.moe.go.kr/laws/laws04.html?Select_Flag=02&pageNum=3&subNum=4
- [11] 한국교육방송공사, <http://www.ebs.co.kr>
- [12] 한국 교육학술정보원, <http://community.edunet4u.net/live.asp>
- [13] 한국산업인력관리공단, <http://www.q-net.or.kr/common/ssologin.jsp>.