

마이닝기법을 활용한 사서커뮤니티 내 채용정보 및 인력수급현황 분석

Analysis of trends in library job ads using a text-mining algorithm

최윤희, 연세대학교 문헌정보학과, yh_choi@yonsei.ac.kr

송민, 연세대학교 문헌정보학과, min.song@yonsei.ac.kr

김기영, 연세대학교 문헌정보학과, gkim@yonsei.ac.kr

Yoonhee Choi, Dept. of LIS, Yonsei University

Min Song, Dept. of LIS, Yonsei University

Giyeong Kim, Dept. of LIS, Yonsei University

본 연구는 인터넷 망을 활용한 문헌정보학 전공자 채용공고를 대상으로, 마이닝기법을 적용하여 구직 실태의 특성을 파악하기 위해 수행되었다. 그 결과, 관종과 공고시기별로 유의한 차이가 나타났으며, 구직자들의 관심정도로 표현되는 공고 게시글의 조회수 또한 고용형태에 따라 관종, 분기간 유의한 차이를 보였다. 사후검증결과, 구직자의 관심정도는 여타 관종보다 전문도서관과 대학도서관의 공고에 비슷한 추세를 보이며 두 관종의 업무와 고용조건을 유사하게 인식하는 것으로 분석되었다. 이러한 결과는 도서관의 인적자원관리의 개선에 기여할 수 있는 기초자료로써 가치가 있으리라 판단된다.

1. 서론

도서관 경영에서 인적자원은 조직의 전략적 목표를 달성하기 위해 수반되는 필수 요소이다. 이러한 인적자원을 관리하는데 있어 최초 단계인 '채용'은 우수한 인재 확보 및 적재적소 배치와 관련된 경영활동으로, 경영의 성패를 좌우할 정도로 중요하게 판단된다(Wright & McMahan 1992). 채용단계에서 조직은 조직 외부에서 이루어지는 노동시장의 여건변화를 파악하고 조직내부에서 필요로 하는 자질을 분석한다. 이를 통해 조직내외부의 이해를 바탕으로 한 공고를 게재하며, 해당 조직의 인재상에 부합하는 능력을 갖춘 구직자를 모집·선발한다.

채용공고는 업무를 수행하기 위한 직원의 자격요건을 의미하며 업무수행을 통해 달성하

고자 하는 조직의 목표와도 연관된다. 또한 새로이 요구되는 자격요건은 기존 사서가 갖춘 역량범위 밖의 업무로, 조직적인 필요에 의해 특정 능력을 발휘할 수 있는 신규직원을 채용함으로써 내부적 어려움을 타개하고자 하는 노력이라 인식될 수도 있다. 문헌정보학 분야에서 채용공고를 대상으로 수행된 연구는 대부분 내용분석을 중심으로 특정 직무의 자격요건을 탐색하여 사서의 역할과 기능을 고찰하였다(Choi & Rasmussen, 2009; Han & Hswe, 2009; 김동완, 2010).

기존 연구에서 채용공고를 통한 내부의 특정 기능 이해에 초점을 두었다면, 본 연구는 도서관 관종과 고용형태, 분기를 고려한 전반적 채용현황분석 및 구직자들의 관심정도를 통한 인력수급동향을 파악하여 인터넷망을 활용한 문헌정보학 전공자 채용의 특성을 규명하고자 한다.

2. 연구의 설계

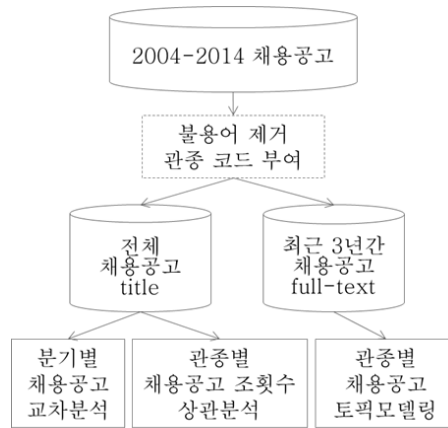
본 연구에서 알아보고자 하는 연구문제와 수행절차는 다음과 같다<그림 1 참조>.

- 연구문제1) 신규인력을 필요로 하는 도서관종과 그 시기는 서로 독립적인가?
- 연구문제2) 관종별 채용공고에 대한 구직자들의 관심정도는 도서관종, 고용형태, 시기와 따라 다른가?
- 연구문제3) 관종별 채용공고에서 나타나는 토픽들은 무엇인가?

국내 문헌정보학 전공자 채용공고를 분석하기 위해 사서직원에 대한 대표적 채용공고사이트인 사서e마을¹⁾에 2004년부터 2014년 6월 까지 게시된 정규직과 계약직 채용정보를 수집하였다. 각 공고에 관종이 따로 기입되지 않거나 특정 관종이 디폴트값으로 설정되어 관종별로 정확하게 구분되지 않는 경우에는 타이틀을 보고 관종코드를 새로 부여하였으며, 관종코드는 Kim & Choi(2012)에서 제시된 구분을

기준으로 하였다. 타이틀만으로 관종을 확인할 수 없는 케이스는 제외하여 총 13654개 채용공고 타이틀이 수집되었다<표 1 참조>.

텍스트마이닝분석을 위한 데이터셋은 타이틀과 본문을 모두 포함하고 있으며, 최근 3년치 데이터인 채용공고 565건이 수집되었다. 생성된 문서집합은 꼬꼬마 형태소 분석기²⁾를 활용하여 체언을 추출하고, Mallet Package를 통해 불용어를 제거한 뒤 토픽모델링을 수행하였다.



<그림 1> 연구수행절차

<표 1> 2004-2014 관종별 채용공고 타이틀 수집현황(2014년 6월 현재) (단위: 건)

	공공			학교			전문			업체			대학			합계
	합계	정규	비정규	합계	정규	비정규	합계	정규	비정규	합계	정규	비정규	합계	정규	비정규	
2004	121	38	83	89	12	77	198	24	174	167	64	10	126	21	105	701
2005	264	73	191	329	17	312	327	35	292	272	106	166	219	33	186	1,411
2006	371	94	277	329	37	292	443	40	403	314	120	194	266	42	224	1,723
2007	401	85	316	466	57	409	377	42	335	289	112	177	260	55	205	1,793
2008	544	97	447	287	14	273	331	38	293	305	128	177	218	46	172	1,685
2009	460	35	425	263	15	248	193	13	180	273	117	156	169	4	165	1,358
2010	197	26	171	241	19	222	158	27	131	114	96	18	161	10	151	871
2011	349	84	265	466	9	457	166	37	129	179	133	46	187	28	159	1,347
2012	406	93	313	263	8	255	144	24	120	150	102	48	168	21	147	1,131
2013	438	84	354	225	15	210	154	25	129	129	80	49	167	18	149	1,113
2014	225	42	183	83	0	83	66	9	57	56	35	21	91	9	82	521
합계	3,776			3,041			2,557			2,248			2,032			13,654

1) 사서e마을 홈페이지(www.librarian.co.kr)는 2014년 6월에 포털 커뮤니티(http://cafe.naver.com/lisleader)로 이전됨
 2) 서울대학교 IDS연구실에서 개발한 JAVA기반의 형태소 분석기(http://kkma.snu.ac.kr/)

3. 분석결과

3.1 분기별 채용광고 현황

채용광고가 특정 시기와 관계가 있는지 알아보기 위하여 교차분석의 카이제곱검증을 실시하였다. 광고 등록시기는 분기에 따라 1월부터 3월은 '1분기', 4월부터 6월은 '2분기', 7월부터 9월은 '3분기', 10월부터 12월은 '4분기'로 구분하였다. 분석결과, 분기는 관종별 채용광고와 관련이 있는 것으로 나타났다 ($\chi^2(12, N=13654)=808.931, p<.01$) <표 2 참조>. 1분기에는 학교가 35.8%로 가장 많은 채용공고를 등록하였고, 2,3,4분기에는 공공이 각각 30.5%, 27.1%, 32.9%로 다른 관종보다 많은 수의 채용공고를 등록하였다. 공개채용과 같은 경우 학교와 공공의 관종별 채용공고의 분기차이는 학교와 지방자치단체의 회계연도 간 차이가 반영된 것으로 판단된다. 「공무원임용시험령」 제47조에 따라 공개채용시험공고는 시험기일의 90일 전에 발표되어야 하는데, 3월 1일을 기준으로 회계연도가 개시되는 학교회계는 1분기(1~3월)에 신규직원 채용공고를 등록하며, 매년 1월 1일에 회계가 시작되는 지방자치단체는 주로 4분기(10~12월)에 신규직원 채용공고를 등록하여 각 기관에서 필요한 인력을 회계연도에 따라 모집한다. 이처럼 관종과 분기는 서로 유의한 관계가 있었으나 효과크기를 나타내는 분할계수가 .236으로 나타나 두 변수의 관계는 작다고 할 수 있다.

3.2 채용공고에 대한 구직자들의 관심정도

고용형태에 따른 분기별 각 관종의 채용광고 조회수 일원배치 분산분석결과, 고용형태(정규직, 비정규직)와 각 분기(1~4분기)를 고려한 8회의 분석 모두 이분산으로 나타났는데, 이는 다른 관종에 비해 학교도서관의 채용광고가 현격히 적은 수를 보이기 때문이거나 학교도서관의 채용광고에서 각 분기 내 조회수의 차이가 비슷하여 분산이 작기 때문이라는 가능성이 제기된다. 학교도서관의 채용광고에서 이와 같은 특성을 보이는 까닭은 이 광고가 주로 「초·중등교육법」에 명시된 사서교사의 자격이 충족된 구직자들에게 필요한 정보이며, 사서교사 임용시험을 주관하는 한국교육과정평가원에서 일괄적으로 공고를 내기 때문에, 등록된 공고의 수와 조회대상자의 수가 제한적이므로 조회수의 분산이 작게 나타난 것으로 보인다.

분기에 따른 채용광고 조회수는 관종별로 유의한 차이를 보였으며, 관종별 채용광고 조회수의 평균은 전문, 대학, 공공의 순으로 높게 나타났다 <표 3 참조>. 구직자들은 학교와 업체보다는 전문, 대학, 공공의 채용정보를 선호하고 있었으며 주로 졸업시점인 4분기와 1분기에 활발하게 구직활동을 수행하는 것으로 보인다.

또한 사후검증에서 전문도서관과 대학도서관은 여타 관종과의 차이에서 유사한 패턴을 보였다. 고용형태와 분기를 고려하지 않은 전체 케이스에 대한 Duncan 사후분석을 실시한

<표 2> 분기별 채용공고의 카이제곱검증 분석결과 및 분기별, 관종별 채용광고 현황

관종 분기	공공	학교	전문	업체	대학	계	N(%)
1분기	1,057(22.9)	1,653(35.8)	679(14.7)	570(12.4)	653(14.2)	4,612(100)	808.931** (.000)
2분기	974(30.5)	519(16.3)	633(19.8)	627(19.7)	437(13.7)	3,190(100)	
3분기	833(27.1)	481(15.6)	658(21.4)	585(19.0)	522(17.0)	3,079(100)	
4분기	912(32.9)	388(14.0)	587(21.2)	466(16.8)	420(15.1)	2,773(100)	

** p<.01

<표 3> 고용형태에 따른 분기별 각 관종의 채용공고 조횟수 일원배치 분산분석결과

분기	관종	정규직 채용공고 조횟수				비정규직 채용공고 조횟수			
		N	평균	표준편차	F	N	평균	표준편차	F
1	공공(a)	297	2798	1882	58.401**	760	745	466	220.852**
	학교(b)	66	2653	845		1587	520	348	
	전문(c)	84	3338	1560		595	1084	571	
	업체(d)	366	1511	919		204	674	353	
	대학(e)	86	3316	1718		567	971	547	
2	공공(a)	157	3350	2298	41.186**	817	668	421	115.231**
	학교(b)	12	2674	959		507	599	361	
	전문(c)	80	3485	1649		553	987	491	
	업체(d)	285	1698	1199		342	533	248	
	대학(e)	55	3735	1661		382	992	565	
3	공공(a)	160	3309	2414	37.599**	673	686	363	81.841**
	학교(b)	10	2668	1219		471	645	308	
	전문(c)	64	3164	1524		594	887	412	
	업체(d)	240	1561	937		345	534	260	
	대학(e)	84	3282	1388		438	895	433	
4	공공(a)	137	3801	2342	70.359**	775	756	506	56.188**
	학교(b)	115	1333	1292		273	763	449	
	전문(c)	86	3411	1443		501	1098	676	
	업체(d)	202	1677	934		264	585	386	
	대학(e)	62	3746	1703		358	1015	611	

** p<.01

결과, 5개의 관종은 4개의 집단으로 구분되었는데, 전문도서관의 평균은 다른 도서관(학교/업체/공공)과는 차이가 있지만 대학도서관과는 유의한 차이가 없다고 나타났다. 이는 구직자들이 전문도서관과 대학도서관의 채용공고를 동시에 고려하고 있으며, 두 관종에 대하여 고용조건 및 업무를 유사하게 인식하고 있음으로 판단된다.

3.3 채용공고 토픽추출

채용공고를 통한 문헌정보학 전공자의 채용동향을 알아보기 위해 최근 3년간 사서e마을에 등록된 정규직 채용공고 565건을 대상으로 토픽모델링기법을 수행하였다. 수집된 비정형 텍스트 데이터는 타이틀과 본문(업무내용, 자격요건, 제출서류 등)으로 구성되었으며, LDA(Latent Dirichlet Allocation)알고리즘을 2000번 반복수행하여 30개의 토픽을 추출하였다. <표 4>는 추출된 토픽 가운데 의미 있는

키워드집합에 연구자가 개념을 부여한 것이다.

관종별로 채용공고에 등장한 단어를 살펴보면 다음과 같은 특성을 도출할 수 있다. 먼저, 공공도서관은 채용공고에 현재 자관이 주력하고 있는 업무의 성격(주민복지, 문화진흥)을 강조하여 해당 직무에 필요한 인재를 모집하고자 한다. 학교도서관은 정규직 사서교사채용에 있어서 자격과 시험이 명확히 규정되어 있고, 사서교사 임용시험 주관기관에서 일괄적으로 공고를 내기 때문에 내용과 형식이 거의 유사한 패턴을 보여 다양한 토픽이 추출되지 않았다. 또한 문헌정보학 전공자를 필요로 하는 업체의 모집공고에서는 도서관 정보화사업이나 도서납품과 관련한 직무의 토픽이 추출되었고, 채용시 이와 관련한 경력이 요구되었다. 전문도서관과 대학도서관 채용공고의 경우 외국어 능력과 전산능력에 대해, 특정 자격증 취득자 혹은 특정 시험의 기준 점수 이상취득자를 요구하고 있어 타 관종에 비해 상당히 구체적인 능력을 기술하고 있는 것이 특징적이다.

<표 4> 관종별 채용공고의 토픽 추출 결과

관종	주제명	주제 구성 단어
공공	사서자격	정사서,자격증,도서관,소지자,경력자,준사서,소지자로서,정규직,공공도서관,사서모집
	공무원시험	지방공무원,임용시험,지방직,시행계획,자료조직개론,한국사,행정학개론,정보봉사개론,시간제,경기도
	경력우대	경력,자격,소지,사서직,도서관,근무한,불인정,근무경력,해당분야
	업무성격-주민복지	투철한,봉사정신,이해,서비스,어린이도서관,소지자,사회복지
	업무성격-문화진흥	문화복지,구민,성북문화재단,증진,문화예술진흥,성북구,설립,관심,홍보마케팅
학교	교원자격, 임용시험	사서교사,소지자,자격증,교원자격증,취득예정자,도서관,학년도,임용시험,중등교사,준사서
전문	법학전문	법원도서관,경력경쟁채용시험,사법연수원,사서서기보,시행계획,서울고등법원,정보제공
	의학전문	행정업무,최선,치료,프로그램,운영,병상,전문병원,의학통계
	기록물관리	취득,이상,기록관리학,기록물관리학,학사학위,문헌정보학,역사학,교육과정,석사학위
	영어능력	toeic,teps,toefl,기록물,기록물관리,ibt,cbt,보유자,전문요원,직원채용
	일본어능력	전문사서,library,문헌정보학과,일본어,종합전시관,사내도서관,일본어능력,석사학위자
	전산능력	성과물,프로젝트,능통자,관리업무,db구축,동종업종,문헌정보학전공자,계정관리,웹마스터,업데이트
업체	MARC입력	marc,졸업자,문헌정보학과,도서관,marc입력,dls,준사서,사서직,입력작업,유경험자
	도서납품	marc작업,준사서자격증소지자,다산서고,담당직원,제공고,도서납품,경력우대,충원시,대학졸업
	해외출판사	국내외,프레젠테이션,springer,가능한,능통자,커뮤니케이션,한국지사,product,지원자
	도서관솔루션	도서관,마케팅,kolas,뛰어나신,솔루션,rfid,표준자료관리시스템,컨설팅,유지보수,졸업예정자
대학	국립대공개채용	도서관학과,자료정리,독서활동,해당자,부합,기본계획,모집분야,정보봉사론,자료조직론
	영어능력	toeic,toefl,teps,ibt,전공자,어학시험,졸업자,기준점수,강의연계서비스,ibt기준
	전산능력	가능자,정보처리기사,oracle,과워포인트,아래한글,운영가능자,도서관전산시스템,powerbuilder
	기록물관리	학술정보관,사서팀,기록보존,기록물관리,사회학,기록물,한국과학기술원,전문가

4. 결론 및 향후 연구

본 연구에서는 도서관 조직의 경영활동 가운데 인적자원관리의 첫 단계인 채용에 초점을 두어, 사서e마을에 2004년부터 현재까지 등록된 채용공고를 분석하였다. 통계분석을 통해 고용형태와 분기, 관종, 조횟수에 대한 변수간 관계의 유의성을 확인하였으며, 수집된 텍스트데이터는 마이닝기법을 활용하여 관종별 채용공고의 특성을 살펴보았다.

본 연구는 마이닝분석에 활용된 데이터가 최근 3년간의 것으로 토픽추출에 대한 통시적인 접근에 제한이 있었다. 향후 연구로는 온라인 사서취업커뮤니티 구직게시판의 Q&A 게시글을 분석함으로써 연도에 따른 관종별 구직동향을 파악할 수 있을 것이며, 다년간 데이터의 수집·분석을 통해 구직실태의 통시적 접근을 제안할 수 있다.

참 고 문 헌

김동완. (2010). 시스템 사서의 역할 분석을 통한 미국의 도서관 정보기술 동향 연구: 2006년부터 2010년까지의 채용공고 분석을 중심으로. 한국비블리아학회지. 21(4)"149-160.

Choi, Y. & Rasmussen, E. (2009). What Qualifications and Skills are Important for Digital Librarian Positions in Academic Libraries? A Job Advertisement Analysis. The Journal of Academic Librarianship. 35(5): 457-467.

Han, M. J. & Hswe, P. (2009). The Evolving Role of the Metadata Librarian: Competencies Found in Job Descriptions. Library resources and Technical services. 54(3): 129-141.

Kim, G. & Choi, Y. (2012). A Suggestion of Criteria for Categorizing Libraries into Types: Linking between Library and Information. Journal of Korean Society for Information Management. 29(1): 395-404.

Wright, P. M. & McMahan, G. C. (1992). Theoretical Perspectives for Strategic Human Resource Management. Journal of Management. 18(2): 295-320.