

# 텍스트마이닝을 활용한 웹사이트 FAQ 개선방안: S대학교 사례를 중심으로

## Improvement Plan of Web Site FAQ using Text Mining : Focused on the S University Case

안 수 현, 조 정 현\*, 이 상 준\*\*

세명대학교, 세명대학교\*, 세명대학교\*\*

Ahn su-hyun, Jo jeong-hyun\*, Lee sang-jun\*\*

Semyung Univ., Semyung Univ.\*, Semyung Univ.\*\*

### 요약

본 연구는 대학 웹사이트의 Q&A(질의응답) 게시판에 게재된 비정형화 된 데이터를 수집한 후 텍스트마이닝과 네트워크 분석을 활용하여 자주 등장하는 키워드 간 연관 패턴을 파악하고자 한다. 분석결과를 바탕으로 FAQ(자주하는 질문) 게시판을 구성한다면 반복적인 질문에 대한 민원을 간소화함으로써 수요자의 편의성과 행정의 효율성 향상에 기여하고 나아가 원활한 양방향 소통이 가능할 것으로 기대한다.

## I. 서론

최근 대학교에서는 대학 구성원과의 소통에 주안점을 두고 그들의 목소리에 매우 적극적인 대처를 하고 있다. 특히 교육 수요자인 학생이 고객이라는 인식을 가지고 그들의 편의를 증대시키기 위해서 Q&A(질의응답) 게시판에 올라오는 글들을 실시간 응답해주고 불편을 최소화하고자 적극적인 노력을 한다. 이 노력 중 하나로 각 대학에서는 Q&A 게시판을 운영함과 동시에 학생들이 자주하는 질문에 대해서 FAQ(자주하는 질문) 게시판을 구성하여 질문을 작성하기 이전에 미리 그들의 궁금증을 해소해주고자 한다.

FAQ 서비스는 고객이 빈번히 질의하는 민원에 관련된 정보를 미리 웹페이지에 게재하여 고객과 행정 담당자 모두 민원해소를 위해 발생하는 불필요한 시간, 노력 등의 절감효과를 줄 수 있다. 일반적으로 웹사이트의 FAQ 구성은 학생들을 직접 대하는 행정 담당자들의 주관적인 판단으로 구성되지만 이러한 판단에는 주관성이 개입될 여지가 있다.

이에 본 연구는 S대학교 사례를 바탕으로 Q&A 게시판의 비정형화된 데이터를 텍스트마이닝을 활용하여 객관적인 FAQ 구성을 제안하고자 한다. 이렇게 구성된 FAQ 게시판을 학생들에게 보다 객관적인 정보를 줄 수 있고, 행정 담당자들에게는 분야별로 구체적인 핵심 키워드 및 연관 키워드를 제시함으로써 FAQ 구성의 논리성, 체계성 향상 및 관련부서와 유기적 업무협조의 기초 자료로 활용될 수 있다.

## II. 연구설계

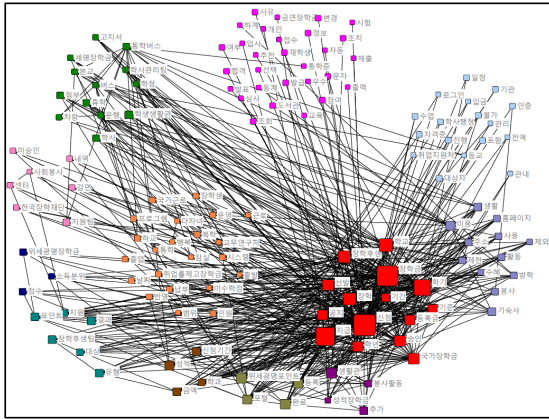
본 연구를 진행하기 위해 S대학교 웹페이지를 담당하고 있는 디지털정보원으로부터 최근 3년간 Q&A 비정형 데이터를 제공받았다. Q&A의 카테고리는 기획, 학사, 학생복지, 시설, 취업, 국제교류, 인터넷, 기타로 세분화되어 있는데, 데이터는 분류, 질문제목, 질문날짜, 질문내용, 답변제목, 답변날짜, 답변내용으로 구성되어 있다.

본 연구방법으로 채택한 텍스트마이닝은 비정형 텍스트 데이터로부터 특정한 목적에 대해 의미 있는 정보를 추출하고 가공 및 분석하는 것을 목적으로 하는 일련의 텍스트 분석 및 처리과정을 의미한다. 이를 위해 빅데이터 분석도구인 R 프로그램을 이용하여 데이터 전처리 및 단어별 출현빈도를 산출하고, 네트워크 분석도구인 Ucinet 프로그램을 이용하여 키워드 간 연결구조 파악 및 연결정도 중심성을 분석하며, 네트워크 시각화 도구인 NetDraw 프로그램을 이용하여 키워드 간 네트워크 시각화를 진행하였다.

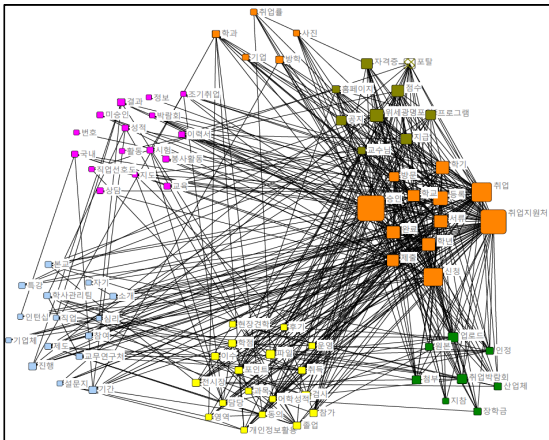
## III. 연구결과

본 연구는 텍스트마이닝을 이용하여 Q&A 비정형 데이터를 분석하였다. 웹페이지에 게재되는 질문의 유형과 내용을 분석함으로써 카테고리별 질문의 유형과 내용을 분석함으로써 각각의 특성을 조사하고, 이용자의 정보요구를 파악하고자 하였다. 연구결과 그림 1, 그림2, 그림 3, 그림 4와 같이 카테고리 별로 유사점을 가진 단어들이 군집을 형성하였다.

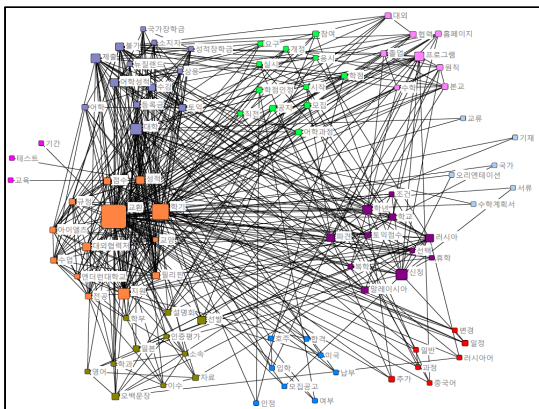
이 연구결과를 토대로 S대학교에서는 행정 담당자들이 직접 각 군집이 의미하는 내용에 대해 FAQ를 작성하였고, 주관적인 판단을 배제한 채 분류·정리된 질의응답에 대한 모범답변을 작성하였다. 내용은 간단명료하고, 이해가 쉽도록 이용자 관점의 용어를 사용하였으며, 친밀감을 주는 구어체를 사용하였다.



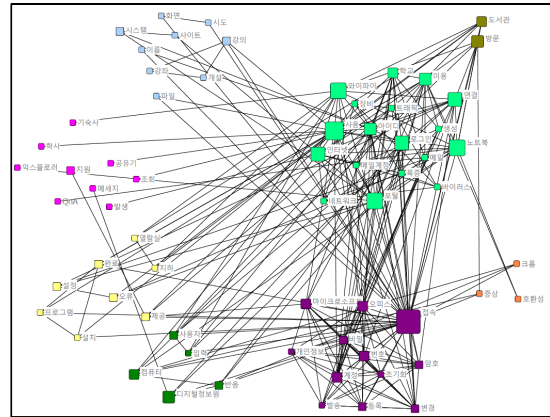
▶▶ 그림 1. 학생복지 Q&A 네트워크 분석



▶▶ 그림 2. 취업 Q&A 네트워크 분석



▶▶ 그림 3. 국제교류 Q&A 네트워크 분석



▶▶ 그림 4. 인터넷 Q&A 네트워크 분석

#### IV. 제언

본 연구결과를 토대로 S대학교의 방대한 질의응답들이 분석되어 민원서비스의 질을 향상시킬 수 있는 FAQ를 완성하였고, 이는 대학 행정업무의 효율성 향상과 학생의 알권리 충족에 기여할 수 있을 것이다. 이에 더 나아가 매년, 매 학기별 학생들의 질의응답은 변화할 것이므로 이에 대한 대비도 필요한데, FAQ의 각 분야별, 주제별 조회건수를 파악함으로써 웹페이지를 방문한 학생들의 주요 관심사항을 파악해야 한다. 그리고 대학 웹사이트의 Q&A, 개인 이메일, 오프라인 민원 등을 지속적으로 관리하여 학생의 요구에 신속하게 대응할 필요가 있으며, 재질문의 빈도와 원인, 요청 서비스에 대한 적절한 서비스를 제공하고 있는지 등의 만족도 평가도 함께 이루어져야 한다.

#### ■ 참고 문헌 ■

- [1] 광기영, 소셜 네트워크 분석, 청람, 2017.
- [2] 백영민, R을 이용한 텍스트 마이닝, 한울아카데미, 2017.
- [3] 허명희, R을 활용한 사회네트워크분석 입문, 자유아카데미, 2015.
- [4] Matloff, N., The Art of R Programming: A Tour of Statistical Software Design, No Starch Press, 2011.
- [5] Zumel, N., Mount, J., Practical Data Science with R, Manning Publications, 2014.