

# 회의 내용정보 온톨로지화에 관한 연구

최현지 · 정회경 · 김창수\*

배재대학교

## A Study on the Ontology of Conference Content Information

Hyun-ji Choi · Hoe-hyung Jung · Chang-su Kim\*

Paichai University

E-mail : hyunji930624@naver.com / hkjung@pcu.ac.kr / ddoja@pcu.ac.kr

### 요 약

최근 인터넷의 급속한 발달에 따라 정보들이 기하급수적으로 증가하고 있다. 이러한 수많은 정보들을 원활하게 커뮤니케이션하기 위해서 다양한 연구가 이루어지고 있다. 최근에는 인공지능 및 빅데이터 기술을 적용한 관련 연구가 활발하게 이루어지고 있다. 하지만 괄목할 만한 결과를 도출해내지 못하고 있다. 그 원인 중 하나는 언어 및 지식 표준의 결여라는 심각한 한계에서 찾을 수 있다.

현재는 멀티미디어적인 접근 방식의 회의에 대한 연구가 활발히 이루어지고 있는 실정이며, 점차적으로 지식기반 회의 시스템에 관해서도 관심을 가지기 시작했다. 멀티미디어적인 접근 방식의 회의의 경우는 기존 오프라인 회의의 장단점을 그대로 온라인으로 표현 한 것에 지나지 않으며 정작 중요한 실제 회의 내용 및 과정에 대한 정보 관리는 소홀하게 된다.

이에 본 논문에서는 회의 내용정보를 온톨로지화 할 수 있는 방안에 대해 연구하고 온톨로지화된 정보를 체계적으로 분석하는 방안을 제시한다.

### ABSTRACT

Recently, according to the rapid development of the Internet, information is increasing exponentially. A lot of this information

Various studies are being conducted in order to communicate smoothly. In recent years, related researches applying artificial intelligence and big data technologies have been actively conducted. However, it has not produced remarkable results. One of the causes can be found in the severe limitation of the lack of language and knowledge standards.

Currently, there is an active research on conferences using a multimedia approach, and gradually, interest in knowledge-based conference systems has begun. In the case of a meeting with a multimedia approach, the advantages and disadvantages of the existing offline meetings are expressed online as they are, and the management of information on the actual contents and process of important meetings is neglected.

Therefore, in this paper, we study a plan to convert conference content information into an ontology, and propose a method to systematically analyze the ontology-formed information.

### 키워드

conference content information, knowledge, knowledge-based conference systems, ontology

### 1. 서 론

최근 인터넷의 급속한 발달에 따라 인터넷 사용자가 급속하게 증가하고 있으며, 정보들이 기하급

수적으로 증가하고 있다. 이러한 수많은 정보들

을 원활하게 커뮤니케이션하기 위해서 다양한 연구가 이루어지고 있으나 이 분야에서 상당한 가능성을 시사했던 인공지능 또는 빅데이터 관련 연구가 괄목할 만한 결과를 도출해내지 못하고 있다. 그 원인 중 하나는 언어 및 지식 표준의 결여라는

---

\* corresponding author

심각한 한계에서 찾을 수 있다[1-3].

비 정형화된 지식들을 대상으로 한 검색과 더불어 지식간의 연관 관계에 기초한 지식 마이닝(Knowledge Mining)을 비롯하여, 논리분석을 통한 지식 발견(Discovery)에 관한 요구 사항들이 증가하고 있다[2-4].

본 논문에서는 온톨로지의 지식화 과정과 온톨로지 마이닝 기법과 함께 사용자가 직접 개입이 가능한 형태로 하여 회의 내용정보를 온톨로지화할 수 있는 방안에 대해 연구하고 온톨로지화된 정보를 체계적으로 분석하는 방안을 제시한다.

## II. 온톨로지 분류

본장에서는 자연어 처리의 가장 단순한 형태로서 특정한 카테고리 분류되는 문서들에 공통적으로 자주 나타나는 문장 형태와 키워드들을 온톨로지 패턴으로 만들어 온톨로지 패턴 매칭에 의해 문서를 분류하는 방법을 제시하고, 일반적인 통계적인 문서 분류에 온톨로지 패턴들을 이용한 분류 방법을 접목시킴으로써 분류의 정확도를 높이는 방법을 제안한다.

### 2.1 온톨로지 패턴분류

온톨로지 패턴에 의한 분류 방법의 특징은 패턴에 매치되지 않는 문서들은 분류되지 않으므로 재현률이 낮은 반면 매치된 문서들의 정확도는 매우 높다는 것이다. 따라서 온톨로지 패턴에 의하여 1차 분류를 시도한 후, 분류되지 않은 경우에 대해서 온톨로지 통계적인 분류를 함으로써 전체적인 분류 정확도를 높일 수 있다. 또한, 온톨로지 패턴 매칭의 경우는 빠른 처리 속도를 보장함으로써 실시간 회의 시스템에 가장 적합하다.

그리고, 온톨로지 분류를 위한 도메인 온톨로지의 구성은 기존 온톨로지를 재 사용하거나 기계 판독용 라이브러리를 주로 이용한다. 이러한 라이브러리는 온톨로지 생성에서 내용을 표현하기 위한 역할을 수행하게 된다.

### 2.2 온톨로지 통계 문서 요약

온톨로지 통계 문서 분류 방법은 수작업 혹은 경험에 의해 분류된 실험집단(training set) 문서에서의 단어 출현 빈도를 근거로 하여 새로운 문서가 분류될 가능성이 가장 높은 카테고리를 찾아내는 방법이다[3,5,6,7].

본 논문에서는 분류하려는 구문과 분류 대상 카테고리들을 온톨로지 벡터로 구성하고, 두 온톨로지 벡터 사이의 유사한 정도를 비교하여 유사도가 가장 높은 온톨로지 구문을 분류하는 벡터 유사도(similarity)를 이용한 방법[2,4]을 온톨로지 통계적인 분류방법으로 사용하였다.

### 2.3 온톨로지 구문 패턴 정의

구문 패턴을 정의하기 위하여 우선 형태소 분석기로 문장을 분석한다. 본 논문에서는 형태소 분석을 위하여 HAM(Hangul Analysis Module)[3,4]을 사용하였다. 형태소 분석된 결과에서 고유명사를 포함하여 분류의 단서가 된다고 볼 수 없는 요소들을 모두 제외시키고 핵심이 되는 명사와 동사들을 중심으로 '명사+조사', '동사+어미' 등이 연결된 형태로 패턴을 정의한다. 조사와 어미를 패턴에 포함시키는 이유는 우리말은 조사에 따라 문장의 의미가 변할 수 있기 때문이다.

## III. 결론

온톨로지 지식모델을 통하여 좀 더 효율적인 정보 통합 관점의 표현을 온톨로지의 표현방법인 개념과 관계로서 구성하였다. 기존 회의 관련 시스템들은 주로 멀티미디어 중심으로 시스템을 구성하여 멀티미디어 기반의 온라인 회의 시스템의 경우는 시공간의 제약이 없이 회의가 가능하다는 장점은 있지만 실제적인 회의에 대한 지식화가 불가능하다는 단점을 가지고 있어 회의 참석자들의 개별적인 역량에만 의존하여 회의가 진행되도록 하고 있다.

향후에는 다양한 의사 표현을 포함하며 일반 음성 정보의 검색 및 요약이 가능한 음성인식을 통해 활용될 수도 있을 것으로 사료된다.

## References

- [1] Gao, Wei, et al. "Distance learning techniques for ontology similarity measuring and ontology mapping." Cluster Computing, Vol. 20, No. 2, pp. 959-968, 2017.
- [2] Gao, Wei, et al. "Ontology learning algorithm for similarity measuring and ontology mapping using linear programming." Journal of Intelligent & Fuzzy Systems, Vol. 33, No. 5, pp. 3153-3163, 2017.
- [3] Ong, Edison, et al. "Ontobee: a linked ontology data server to support ontology term dereferencing, linkage, query and integration." Nucleic acids research Vol. 45, pp. 347-352, 2017.
- [4] Klopfenstein, D. V., et al. "GOATOOLS: A Python library for Gene Ontology analyses." Scientific reports, Vol. 8, No. 1, pp. 1-17, 2018.
- [5] Jo, Taeho. "Text mining." Studies in Big Data. Cham: Springer International Publishing, 2019.
- [6] Klusch, Matthias, et al. "Semantic web service

- search: a brief survey." *KI-Künstliche Intelligenz*, Vol. 30, No. 2, pp.139-147, 2018.
- [7] Fensel, Dieter, et al. *Enabling semantic web services: the web service modeling ontology*. Springer Science & Business Media, 2006.