

## ◆특집◆ 웹기반 엔지니어링 협업 기술

## 온톨로지 매핑 기반 엔지니어링 정보 검색

정민\*, 서효원\*\*

## Engineering Information Search based on Ontology Mapping

Min Jung\* and Hyo Won Suh\*\*

## ABSTRACT

The participants in collaborative environment want to get the right information or documents which are intended to find. In general search systems, documents which contain only the keywords are retrieved. For searching different word-expressions for the same meaning, we perform mapping before searching. Our mapping-based search approach has two parts, ontology-based mapping logic and ontology libraries. The ontology-based mapping consists of three steps such as character matching (CM), definition comparing (DC) and similarity checking (SC). First, the character matching is the mapping of two terminologies that have identical character strings. Second, the definition comparing is the method that compares two terminologies' ontological definitions. Third, the similarity checking pairs two terminologies which were not mapped by two prior steps through evaluating the similarity of the ontological definitions. For the ontology libraries, document ontology library (DOL), keyword ontology library (KOL), and mapping result library (MRL) are defined. With these three libraries and three mapping steps, an ontology-based search engine (OntSE) is built, and a use case scenario is discussed to show the applicability.

**Key Words** : Ontology (온톨로지), Semantic (의미), Mapping (매핑), Search (검색)

## 1. 서론

Collaborative Product Commerce (CPC) 환경에서는 방대한 양의 정보 교환과 검색이 필요하다. 효과적인 협업을 위해서는 참여자들이 다른 참여자들이 생성해놓은 정보를 바탕으로 다음 작업을 진행해야 한다. 그러나 참여자들은 보통 같은 의미인데도 서로 다른 용어를 사용하기 때문에 정보

(문서)를 검색하는데 어려움이 존재한다.

정보시스템 내부에서 사용하는 용어들이 온톨로지 정의를 갖는다는 가정하에서, 본 연구는 정확성 있는 검색에 대한 해결책으로 온톨로지 매핑을 기반으로 한 검색을 제안한다. 이 검색 방법은 구축되어있는 온톨로지 라이브러리를 대상으로, 검색어에 대한 온톨로지 매핑을 통해 유사어를 찾은 뒤, 엔지니어링 정보(문서)를 검색하는 것이다. 이때, 온톨로지 매핑 로직은 문자열 비교(Character Matching), 정의 비교(Definition Comparing) 그리고 유사도 검사(Similarity Checking)의 3 단계로 구성된다. 온톨로지 검색은 문서 온톨로지 라이브러리(DOL: Document Ontology Library), 키워드 온톨로지 라이브러리(KOL: Keyword Ontology Library) 및 매

\* 한국과학기술원 산업공학과

\*\* 한국과학기술원 산업공학과

Tel. 042-869-3163

Email hw\_suh@kaist.ac.kr

온톨로지 매핑, PDM, 동시공학 분야에 관심을 두고 연구활동을 하고 있다.

핑 결과 온톨로지 라이브러리(MRL: Mapping Result Library)를 구축하여 위의 3 단계 매핑을 적용하여 이루어진다.

본 연구의 구성은 다음과 같다. 2 장에서는 관련 연구를 살펴보고, 3 장에서는 온톨로지 구조와 표현 언어를 소개한다. 4 장에서는 온톨로지 매핑 알고리즘을 소개하고, 5 장에서는 온톨로지 매핑 기반 엔지니어링 정보 검색 방법과 OntSE (Ontology Search Engine) 시스템에 대해서 설명한다. 6 장에서는 OntSE 시스템에서 온톨로지 매핑 알고리즘이 어떻게 적용되는지를 설명한다. 7 장에서는 연구의 결론, 제한점 등에 대해서 서술한다.

## 2. 관련 연구

전통적인 검색 엔진은 도메인 지식을 활용하지 않기 때문에 사용자의 검색 요청의 의미와 Web 문서에 들어있는 용어의 고유한 의미 관계를 활용하지 못했다. 이런 특징이 content-based 검색을 수행하는 데에 제한 점이 되었다. 일반적인 검색은 자연어에서 유사어, 동의어 등으로 인한 복잡도를 고려하지 못하기 때문에 검색의 결과를 보충하기 어렵다. 이 문제점은 검색 엔진에 LSA<sup>1</sup> 등의 여러 방법을 도입해서 정확성 있는 검색을 하려고 해도, 근본적으로 해결되지 않는다. 따라서 검색에 도움을 주기 위해서, 도메인 지식을 어떻게 이용할 것인지가 의미 검색의 핵심이다.<sup>2</sup>

또한, Web 상의 정보 검색 분야에서 질의어의 개선에 대한 연구들이 활발히 진행되어 왔다. 질의어 개선 연구는 일반적으로 질의어 확장(query expansion)과 추천 시스템(recommendation system)으로 크게 두 방향으로 나눌 수 있다.<sup>3</sup> 질의어 확장은 사용자에게 더 나은 질의어를 입력할 수 있도록 도와주는 데 목적이 있다. 이것은 여분의 용어를 처음의 질의어에 삭제 혹은 덧붙임으로서 검색의 효율성을 개선시킨다. 상호적인 질의어 확장(interactive query expansion)은 hyper-index<sup>4</sup> 또는 문서 저장소에서 자동적으로 구축된 컨셉 계층을 바탕으로, 사용자에게 후보 질의어를 제공한다. 한편 추천 시스템<sup>5</sup>은 과거에 사용자가 사용했던 비슷한 용어를 추천 (content-based recommendation)하거나, 현 사용자를 과거의 비슷한 성향이 있는 사용자로 분류해서 그들이 수행했던 용어를 추천 (collaborative recommendation) 한다. 그러나 이런

접근들은 질의어의 개선을 위해 도메인 모델을 이용하는 것이 아니기 때문에 한계점을 갖는다.

본 온톨로지 매핑 기반 검색 엔진은 용어의 온톨로지 정의를 이용하여 도메인 지식의 활용을 도모하고, 용어의 온톨로지를 추천하는 content-based recommendation 을 제공한다.

## 3. 온톨로지 구조 및 표현언어

온톨로지는 개념화 (conceptualization)에 대한 명시적인 명세사항 (specification)이다. 즉, 암묵적인 지식을 명시화시켜 정형화된 구조로 서술하고 있는 것이다.<sup>6</sup>

임의의 도메인을 온톨로지로 표현하기 위해서는 온톨로지의 정의, 구조 그리고 표현 언어의 결정이 선행되어야 한다. 온톨로지의 정의와 구조는 디자인에 따라 다르다. 본 연구에서는 KAON 프로젝트<sup>7</sup>에서 제시한 온톨로지의 구조를 선택하였다. KAON 프로젝트에서의 온톨로지 구조는 포괄적인 구조를 formal 하게 나타내고 있는데 그 기본 요소는 다음과 같다.

Ontology (O,A)

Core ontology  $O = \{ C, \leq_C, R, \sigma, \leq_R \}$

C: Concept

$\leq_C$ : Concept hierarchy

R: Relation

$\sigma$ : Signature (Relation function)

$\leq_R$ : Relation hierarchy

Axiom A

한편, 온톨로지는 다양한 언어로 표현될 수 있는데, 본 연구에서는 표현하기 쉽고, 표현 능력이 우수한 FOL (First-order Logic) 을 선택하였다.<sup>8</sup> Table 1 은 FOL 로 표현된 온톨로지의 예이다.

Table 1 Example of Ontology Representation in FOL

Element	Example
Concept (C)	Part(x), endProduct(x)
Relation (R)	isSubPartOf(x, y)
Concept Hierarchy (H <sup>c</sup> )	$\forall x \text{ endProduct}(x) \Rightarrow \text{part}(x)$
Relation Function (rel)	$\forall x, y \text{ isSubPartOf}(x, y) \Rightarrow \text{part}(x) \wedge \text{part}(y)$
Axiom (A <sup>o</sup> )	$\forall x, y \text{ isSubPartOf}(x, y) \Rightarrow \neg \text{isSubPartOf}(y, x)$

#### 4. 온톨로지 매핑 로직

본 연구에서는 의미는 같지만 다르게 표현된 용어의 짝을 찾는 의미 매핑 로직을 활용한다. 온톨로지 매핑 로직은 1) 문자열 비교(Character Matching, CM), 2) 정의 비교(Definition Comparing, DC), 3) 유사도 검사(Similarity Checking, SC)의 3 단계로 구성되어있다. 먼저 매핑을 하고자 하는 온톨로지 정의 A 와 B 의 데이터를 얻어서 문자열 비교를 통해 용어를 매칭시킨다. 만약 매칭되지 않으면 온톨로지 하위정의를 비교한다. 정의 비교 단계에서 매칭되지 않으면 유사도 검사를 수행한다.

##### 4.1 문자열 비교 (Character Matching)

이 단계에서는 용어가 동일한 어휘를 갖는다면 동일한 의미를 갖는다고 가정한다. 따라서 두 용어가 어휘가 같으면 매칭시킨다. 예를 들면, 'Part'와 'Part' 는 같은 문자열을 가지고 있으므로 매칭이 되지만, 같은 의미를 나타내는 'CarModel' 과 'Marque' 는 매칭 되지 않는다.

##### 4.2 정의 비교 (Definition Comparing)

이 단계에서는 두 용어끼리 같은 의미를 갖는지 확인한다. 이를 위해 각 용어에 대한 하위 정의를 정의하고, 그 정의가 동일하면 매칭시킨다. 예를 들면 Table 2 와 같이 'CarModel' 과 'Marque' 의 하위 정의가 같으므로 매칭시킨다.

Table 2 Example for Definition Comparing

Term	Ontology sub-definition
Marque(x)	Model(x) ∧ typeOf(x,y) ∧ Car(y)
CarModel(x)	Model(x) ∧ typeOf(x,y) ∧ Car(y)

##### 4.3 유사도 검사 (Similarity Checking)

유사도 검사는 여러 방법을 사용하여 할 수 있는데, 크게 3 가지로 나누어 볼 수 있다. 첫째 앞서 정의한 어휘의 온톨로지 하위 정의를 바탕으로 유사도를 판단 할 수 있고,<sup>9</sup> 둘째 문서 내용에 대한 포괄적 온톨로지의 구조 비교를 통하여 유사도를 구할 수 있으며,<sup>10</sup> 셋째 도메인 내의 다른 관련 용어의 매칭 결과를 바탕으로 유사도를 판단 할 수 있다.<sup>11</sup> 본 연구에서는 가장 간단한 첫 번째 방법을 바탕으로 한 유사도 검사를 택하였다.

이 단계는 두 용어의 하위 정의가 같지 않은 경우에 적용하는 단계로서, 두 용어의 문자열이 비슷한 정도와 하위 정의가 비슷한 정도의 가중합을 이용하여 구한다. 문자열이 비슷한 정도는 Jaro Metric<sup>12</sup>을 이용하여 계산한다. Jaro Metric 은 두 문자열이 주어진 경우, 공통된 문자의 개수와 순서를 기반으로 유사도를 구한다.

하위 정의가 비슷한 정도는 두 용어의 하위 정의의 전체 엘리먼트 중 매핑된 엘리먼트의 비율로 계산한다. 두 용어 C<sub>a</sub> 와 C<sub>b</sub> 의 하위 정의가 다음과 같으면, 하위 정의 유사도(DefSim) 는 다음과 같다.

$$C_a = d_{a_1} \wedge d_{a_2} \wedge \dots \wedge d_{a_m}$$

$$C_b = d_{b_1} \wedge d_{b_2} \wedge \dots \wedge d_{b_n}$$

$$DefSim(C_a, C_b) = \frac{N_M \times 2}{N_{C_a} + N_{C_b}}$$

N<sub>C<sub>a</sub></sub> : C<sub>a</sub> 의 엘리먼트 개수

N<sub>C<sub>b</sub></sub> : C<sub>b</sub> 의 엘리먼트 개수

N<sub>M</sub> : 매핑된 엘리먼트 쌍의 수

유사도(Similarity)는 Jaro Metric(Jaro(C<sub>a</sub>,C<sub>b</sub>)) 과 DefSim 의 가중 합으로 다음과 같다.

$$Similarity = \alpha \times Jaro(C_a, C_b) + \beta \times DefSim(C_a, C_b)$$

$$\alpha + \beta = 1, \alpha, \beta \geq 0$$

이때 가중치는 일반적으로 DefSim 에 더 크게 준다. 이렇게 구한 유사도가 특정한 수치(본 연구에서는 0.5) 이상이라면 두 용어가 매칭이 된 것이라고 판단한다. 매핑 로직의 이해를 위해 Fig. 1 에는 간단한 예제를 나타내고 있다.

Case ⇔ UniPartForEndProduct(x) ∧ Material(y) ∧ Color(z) ∧ Weight(p) ∧ Function(q) ∧ madeOf(x,y) ∧ hasColor(x,z) ∧ weigh(x,p) ∧ provide(x,q)
Cover ⇔ Component(x) ∧ Matter(y) ∧ Hue(z) ∧ W(p) ∧ Function(q) ∧ madeOf(x,y) ∧ hasColor(x,z) ∧ weigh(x,p) ∧ provide(x,q)

Fig. 1 (a) Example for Similarity Checking

$$\begin{aligned}
 s &= \text{Case} \\
 t &= \text{Cover} \\
 H &= \frac{\min(4,5)}{2} = 2 \\
 i-2 &\leq j \leq i+2 \\
 s' &= ce \\
 t' &= ce \\
 \text{Transposition} &= 0 \\
 \text{Jaro}(Case, Cover) &= \frac{1}{3} \times \left( \frac{2}{4} + \frac{2}{5} + \frac{2-0}{2} \right) = 0.63 \\
 \text{DefSim}(Case, Cover) &= \frac{2 \times 5}{9+9} = 0.56 \\
 \text{Similarity}(Case, Cover) &= 0.4 \times 0.63 + 0.6 \times 0.56 = 0.59
 \end{aligned}$$

Fig. 1 (b) Similarity Calculation of the example

### 5. 온톨로지 매핑 기반 엔지니어링 정보 검색

#### 5.1 Ontology Search Engine 시스템의 구조

온톨로지 매핑 기반 엔지니어링 정보 검색 시스템(Ontology Search Engine, OntSE)의 구조는 Fig. 2에 나타나 있다. OntSE 시스템 크게 UI, 내부 로직, Ontology base로 구성되어 있다. 이를 기반으로 온톨로지 매핑 기반 검색은 6 가지 단계로 진행된다. 이에 대한 자세한 설명은 아래와 같다.

- 1) 통합시스템을 사용하면서 사용자가 엔지니어링 문서를 등록할 때, 문서에 대한 키워드의 온톨로지를 정의하는 과정
- 2) 사용자가 검색할 때, 검색 하고자 하는 키워드와 온톨로지를 정의하는 과정
- 3) 온톨로지 매핑을 통해 유사어를 찾아내는 과정 (이때 온톨로지 매핑 알고리즘은 4장에서 소개한 문자열 비교, 정의 비교, 유사도 검사의 3 단계로 진행된다.)
- 4) 찾아낸 유사어를 바탕으로 엔지니어링 정보가 존재하는 문서나 정보를 검색해 내는 과정
- 5) 찾아낸 유사어를 Ontology base에 업데이트 하는 과정
- 6) 검색된 엔지니어링 정보를 사용자에게 보여주는 과정

OntSE 시스템은 위의 1), 2)를 위해서 사용자

로부터 키워드와 온톨로지 정의를 받아오는 웹페이지를 제공한다. 또한 과정 3), 4), 5)를 위해서 프로그램 모듈을 제공한다. 그리고, 6)을 위해서 사용자에게 검색어에 대한 결과(문서)가 담겨있는 웹페이지를 보여준다.

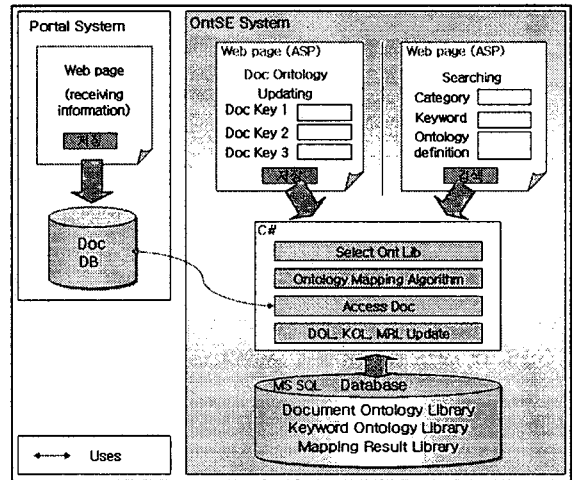


Fig. 2 Architecture of OntSE System

#### 5.2 OntSE 시스템의 Ontology Library

Fig. 2에서 볼 수 있듯이, OntSE 시스템의 Database는 Document Ontology Library(DOL), Keyword Ontology Library(KOL)와 Mapping Result Library(MRL)를 가지고 있다. 통합시스템에서 문서가 새로 등록될 때, 사용자는 문서에 해당하는 키워드와 온톨로지 정의를 입력한다. 이 키워드와 온톨로지 정의는 DOL에 저장이 되고, 사용자가 검색을 할 때 입력하는 키워드와 온톨로지 정의는 KOL에 저장된다. MRL에는 온톨로지 매핑을 통해 매칭이 된 용어들에 대한 정보가 존재한다.

KOL과 DOL은 내용적인 측면에서는 키워드와 온톨로지 정의가 저장되어 있으므로 같다고 생각할 수 있지만, 형식적인 측면에서 다른 양상을 보인다. 즉, DOL에는 엔지니어링 문서를 기반으로 한 키워드와 온톨로지 정의가 저장되어 있다. 따라서 DOL에는 키워드, 온톨로지 정의, 문서를 접근할 수 있는 정보가 함께 들어있다.

예를 들어, Fig. 3에서 볼 수 있듯이, 사용자가 통합시스템에서 업무를 수행하던 중에 문서("Conference for developing mobile phone cover.doc")를 업로드 한다면, 이 문서에 해당하는 키워드를 선정하고 그 키워드의 온톨로지 정의를 입력한다.

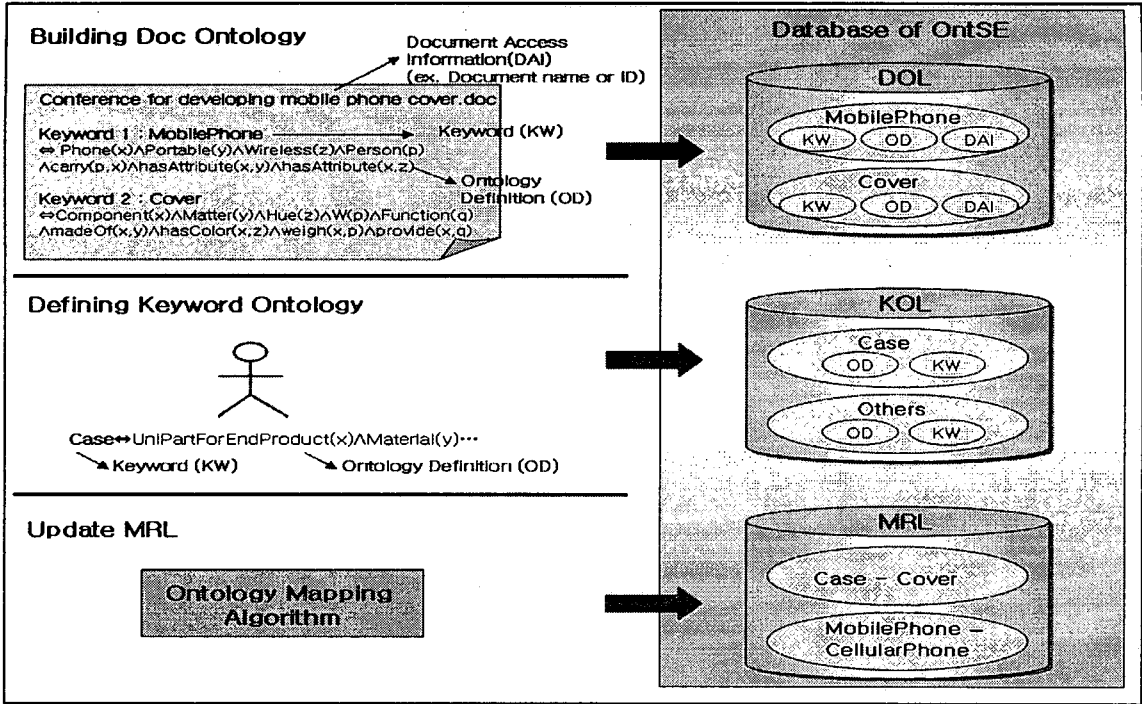


Fig.3 Difference between DOL and KOL and MRL

이 때에는 DOL 에 저장이 되는 반면, 추 후에 사용자가 “Case” 라는 키워드로 검색을 하면서 온톨로지 정의를 입력한다면 이것은 KOL 에 저장 되는 것이다.

5.3 OntSE 시스템의 사용자 시나리오

Fig. 4 는 사용자의 행동 시나리오이다. Fig 4 를 보면, 사용자가 검색할 때 시나리오 1 과 2 로 나뉘어 있음을 알 수 있다. 만약 사용자가 키워드의 온톨로지 정의를 하는 데에 있어서 미숙하다면, 일반 검색 시스템처럼 키워드를 입력해서 자료를 검색할 수 있게 하기 위해서이다.

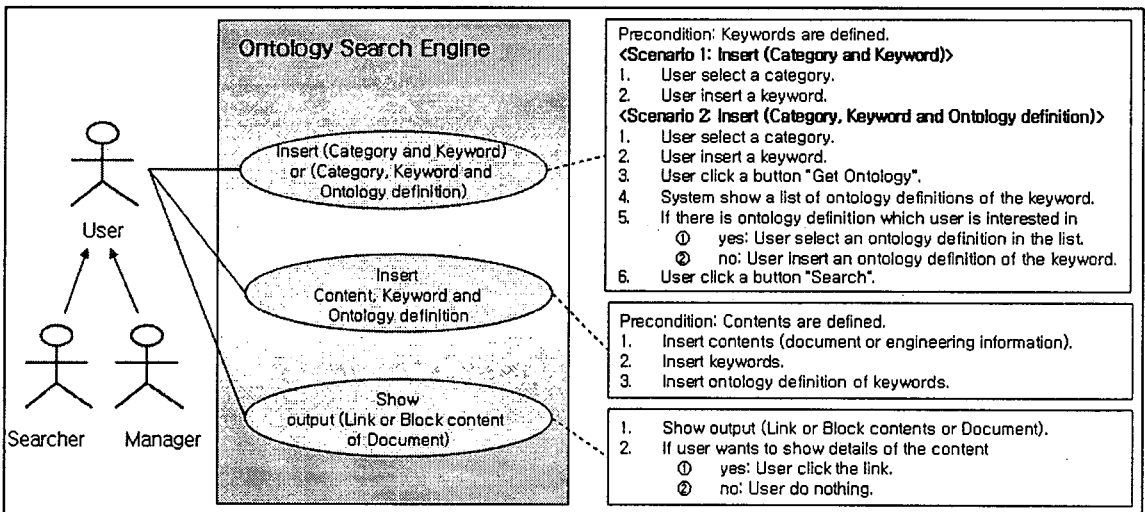


Fig. 4 Use case diagram of OntSE System

사용자가 키워드만 입력하고 검색했을 때에는, OntSE 시스템은 온톨로지 매핑 알고리즘을 수행하지 않는다. OntSE 시스템은 Database 의 Mapping Result Library 를 참고해서 해당 검색어에 대해서 매칭이 완료된 용어가 있는지 검색한다. 만약 용어가 있다면 검색어와 그 용어들에 대해서 엔지니어링 문서를 검색하지만, 매칭된 용어가 없다면 일반 검색과 마찬가지로 해당 검색어로만 문서를 검색한다.

시나리오 2 에서, 사용자가 키워드를 입력하고 “정의 찾기” 버튼을 클릭하면, 만약 그 키워드가 이미 Database 에 저장되어 있다면 저장되어 있는 온톨로지 정의들을 사용자에게 보여준다. 이 기능을 통해 사용자는 온톨로지를 정의 하는 방법의 가이드를 받을 수 있다.

### 6. OntSE 시스템에서의 온톨로지 매핑

온톨로지 매핑 알고리즘은 문자열 비교(Character Matching, CM), 정의 비교(Definition Comparing, DC) 과 유사도 검사(Similarity Checking, SC) 의 세 단계로 진행된다.

Fig. 5 에서와 같이, CM 은 MRL 에 저장되어 있는 용어에 대해서 진행한다. 이는 매핑 로직의 수행 시간을 줄일 수 있게 한다. 한편, DC 와

SC 단계는 검색어의 온톨로지 정의와 KOL, DOL 에 저장되어 있으면서 MRL 에 존재하지 않은 용어들에 대해 매칭을 수행한다.

DC 에서는 용어의 온톨로지 정의를 컨셉과 릴레이션으로 분류하여 각각을 CM 으로 매칭 여부를 판단한다. CM 과 DC 단계에서 매칭이 되지 않은 용어에 대해서 SC 를 수행한다. 이 매핑 로직을 통해 매칭된 용어들을 MRL 에 입력하여, 다음 검색에서 활용할 수 있도록 한다.

4 장에서 예로 들은 ‘Case 와 ‘Cover’로 설명해 보자. ‘Case’ 의 온톨로지가 DOL 에 저장되어 있고, 사용자가 “Mobile phone cover development” 라는 키워드와 ‘Cover’ 의 온톨로지 정의를 입력하고 검색했다고 가정하자. Fig. 1 과 같이 ‘Case’와 ‘Cover’ 는 similarity 가 0.59 이므로 매칭이 되고, MRL 에 매칭 결과가 업데이트 된다. 따라서 사용자는 입력한 검색어인 “Mobile phone cover development” 뿐만 아니라 “Mobile phone case development” 라는 용어에 대해서도 같은 정보를 얻을 수 있다.

### 7. 결론

본 연구에서는 CPC 환경 하에서 참여자들 간에 엔지니어링 문서를 검색할 때, 더 효과적인 검색

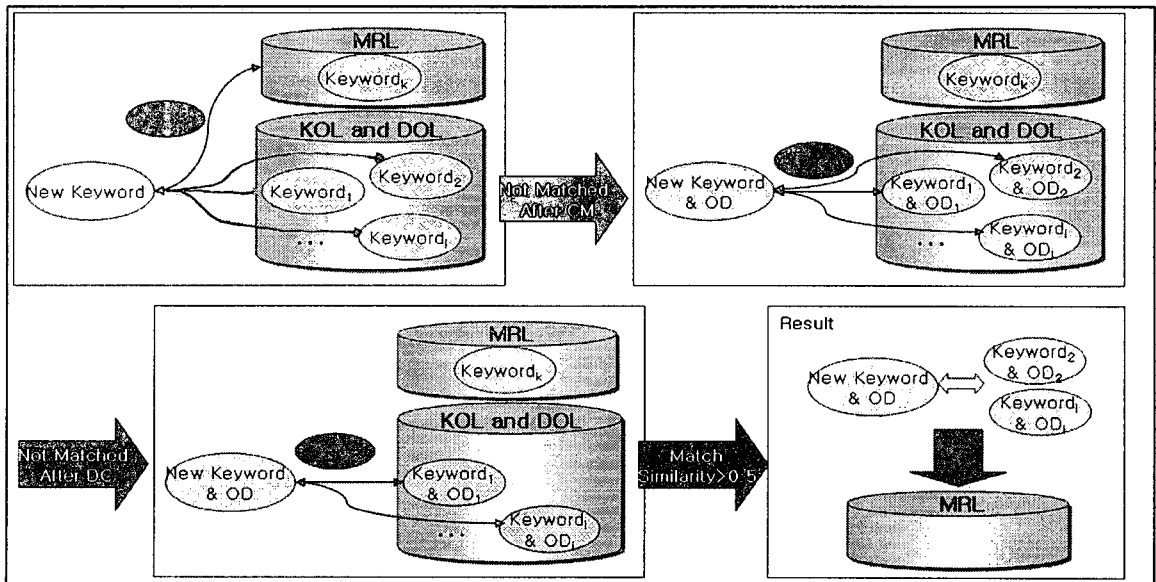


Fig. 5 Ontology Mapping Algorithm on OntSE System

색을 하기 위한 방안으로 온톨로지 매핑을 기반으로 한 검색을 연구하였다.

온톨로지 매핑 로직은 문자열 비교(Character Matching), 정의 비교(Definition Comparing), 유사도 검사(Similarity Checking)의 3 단계 알고리즘 및 DOL, KOL 과 MRL 등의 온톨로지 라이브러리를 기반으로 이루어진다. 앞의 로직을 기반으로 한 온톨로지 기반 검색 엔진인 OntSE 시스템에서는 매핑 로직을 통해 유사어를 찾은 뒤 엔지니어링 정보가 저장되어있는 문서와 Database 를 검색한다. 이 시스템은 엔지니어링 정보를 열람할 수 있는 링크가 담긴 웹페이지를 결과로 보여준다. 따라서 사용자는 원하는 정보가 담긴 문서를 참고하여 업무에 임할 수 있다.

본 연구에서 제안된 온톨로지 매핑 기반 검색 방법은 기본적인 접근법을 제안 하였다. 따라서 일반적인 경우에 대한 매핑 로직의 개발이 유연한 정보 공유를 위해서 필요하다.

온톨로지 매핑 기반 검색의 적용을 위해서는 먼저 온톨로지를 구축해야 한다. 그러나 온톨로지 구축은 방대한 양의 지식공학 작업, 많은 시간과 노력을 필요로 한다. 따라서 이에 대한 간편하고 효율적인 접근법이 개발되어야 할 것이다.

## 후 기

본 연구는 산업자원부에서 추진하는 차세대신 기술개발사업의 하나로 수행되고 있는 '지식 프로세스 기반의 제품 개발 협업 프레임워크 개발' 과제 의 지원을 받아 수행 되었습니다

## 참고문헌

- Dumais, S. T., Furnas, G. W., Landauer, T. K. and Deerwester, S., "Using latent semantic analysis to improve information retrieval," Proceedings of CHI'88: Conference on Human Factors in Computing, New York, pp. 281-285, 1998.
- Fang, W. D., Zhang, L., Wang, Y. X. and Dong, S. B., "Toward a semantic search engine based on ontologies," Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, Vol. 3, pp. 1913 - 1918, 2005.
- Stojanovic, N., "On the query refinement in the ontology-based searching for information," Information Systems, Vol. 30, No. 7, pp. 543 - 563, 2005.
- Bruza, P. D. and Dennis, S., "Query Reformulation on the Internet: Empirical Data and the Hyperindex Search Engine," Computer-Assisted Information Searching on Internet RIAO97, Montreal, 1997.
- Balabanovic, M. and Shoham, Y., "Content based collaborative recommendation," Communications Of the ACM, Vol. 40, No. 3, pp. 66-72, 1997.
- Gruber, T. R., "What is an ontology?," <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>, Stanford University, 1993.
- Bozsak, E., Ehrig, M., Handschuh, S., Hotho, A., Maedche, A., Motik, B., Oberle, D., Schmitz, C., Staab, S. and Stojanovic, L., "KAON: Towards a large scale Semantic Web," E-commerce and Web technology, pp. 304-313, 2002.
- Stuart, R. and Peter, N., "Artificial Intelligence," Prentice Hall, pp. 240-266, 1995.
- Lee, M. J., "Ontology-based Semantic Mapping Approach using a Bayesian Network for Communication in CPC Environment," Master thesis, KAIST, pp. 28-31, 2006.
- Lee, M. J., Jung, W. C., Lee, J. H. and Suh, H. W., "Ontology Semantic Mapping based Data Integration of CAD and PDM System," Proceeding of 2005 Spring Conference of KSPE, 2005.
- Mitra, P., Noy, N. F. and Jaiswal, A. R., "OMEN: A Probabilistic Ontology Mapping Tool," Workshop on Meaning coordination and negotiation at the Third International Conference on the Semantic Web (ISWC-2004), Hisroshima, Japan, 2004.
- Cohen, W. W., Ravikumar, P. and Fienberg, S. E., "A Comparison of String Distance Metrics for Name-Matching Tasks," American Association for Artificial Intelligence, 2003.