

KT 용어 온토로지 구축

노덕근^o, 변동철, 박순철

KT^o, 전북대학교 전자정보공학부

ndk6357^o@kt.co.kr, bdrsea@chonbuk.ac.kr, scpark@chonbuk.ac.kr

Ontology construction for Korea Telecom(KT) Terms

Duck Keun Roh^o Dong Ryul Byun, Soon Cheol Park

Korea Telecom^o, Chonbuk National University Electronic Information Engineering

요약

본 논문에서는 한국통신(KT)에서 사용되는 주요 용어들을 추출하여 추출된 용어들 간의 고유성과 관계성을 기초로 한 용어 온토로지를 구축하였다. 또한 생성된 용어 온토로지를 이용한 검색질의 예를 통해서 기업의 다양한 분야를 관리하는데 도움을 줄 수 있는 방안을 모색했다. 온토로지 구축 툴로는 온토로지에더터, Protege를 사용하였으며, 온토로지는 최상위 클래스 Organization(기관), Employee(직원), Product(상품), Technique(기술) 등 4가지로 분류하여 구축하였다. 본 연구를 기초로 한국통신(KT)의 다양한 지식정보를 체계화하고 KT 데이터베이스를 효과적으로 관리할 수 있을 것이다. 또한 구축된 온토로지를 이용한 미래의 KT 시맨틱 검색시스템 구축에 기초가 되기를 기대한다.

1. 서론

기업의 목적은 이윤 추구에 있다. 영업을 해서 물건을 잘 팔고, 이윤을 많이 남기면 목적은 달성되는 것이다. 이를 위해서 기업은 외부적으로는 상품을 잘 만들어 고장이 없도록 하고, 고장이 생겼을 경우 신속한 A/S활동으로 고객을 상대해야 한다. 기업 내부적으로는 인력을 지위, 연봉, 지역, 취미 등의 중요 정보를 정확하고 신속히 파악하여 신 성장 사업이나 집중 투자 사업에 맞춤형으로 배치함으로써 최대한의 성과를 낼 수 있을 것이다. 이러한 목적으로 기업은 여러 분야의 많은 정보를 가지고 있으며 기업과 관련된 엄청난 지식 정보가 지속적으로 축적되고 있다.

기업은 대량의 정보를 이해하기 쉽고, 관리하기 쉽게 보여줄 것을 원한다. 정보를 잘 이용하면 기업의 기술, 상품, 조직, 인력을 파악할 수 있으며, 상품과 기술의 접목 방안과 인력의 적재적소 배치 및 최선을 다해야 되는 분야를 알 수 있다. 나아가 고객에 맞는 맞춤형 서비스를 해주어 궁극적으로 기업의 목적인 이윤을 얻을 수 있을 것이다

정보의 접근성과 유용성을 높이기 위하여 기업은 많은 양의 지식 정보를 체계화해야 한다. 이를 위해서는 각 분야별 지식 정보에 대한 명확한 학문적·이론적 지식을 비롯하여 기술적 지식, 응용적 지식 등 기업의 정보 관리에 활용하고자 하는 지식 정보에 따라 상당한 지식이 요구된다. 이러한 지식 정보 처리와 관련되어 최고의 핵심 기술로 떠오르고 있는 것은 시맨틱 웹(Semantic Web)이며, 온토로지(Ontology)는 시맨틱 웹을 구축하는 데 필수적으로 이용되고 있다[1,2,3].

본 논문에서는 한국통신(KT)에서 사용되는 수많은 용어들 간의 고유성과 관계성을 기초로 하여 한국통신 용어 온토로지를 구축한다. 또한 생성된 용어 온토로지가 기업의 다양한 분야를 관리하는데 어떻게 도움을 줄 수 있는가를 실제 예를 들어 설명할 것이다.

2. 이론적 배경

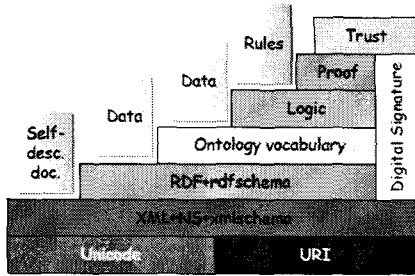
2.1 시맨틱 웹과 온토로지

시맨틱 웹(Semantic Web)[4,5,6,7]은 기존의 웹과 완전히 구별

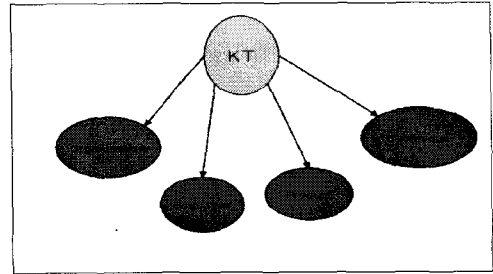
되는 새로운 개념이 아니라 현재의 웹을 확장하여 웹에 올라오는 정보에 잘 정의된 의미를 부여하고 이를 통해 컴퓨터와 사람이 협동적으로 작업을 수행할 수 있도록 하는 패러다임이다. 대표적인 웹 표준화 단체인, W3C(World Wide Web Consortium)에서는 시맨틱 웹을 RDF(Resource Description Framework)나 기타, 다른 표준을 기반으로 웹에 있는 데이터를 추상적으로 표현하는 것이라고 정의하였다[4,7,8]. 즉, 시맨틱 웹은 기존의 웹 데이터에 의미를 부여할 수 있는 표현 방법을 제시하는 것이다 [7]. 이러한 의미를 부여하기 위해서는 용어에 대한 온토로지 구축이 필수적이다.

시맨틱 웹은 [그림 1]과 같이 계층적인 구성요소로 이루어진다[8]. 가장 기본이 되는 층은 자원을 명시적으로 지칭하는 URI와 UNICODE로 구성되어 있다. 다음 층은 임의의 개념을 모듈 방식으로 정의할 수 있는 XML과 네임스페이스(Namespace)이며, 그 다음으로 자원을 기술하기 위한 RDF와 온토로지가 위치한다. 그 위의 계층인 Logic은 인공지능의 추론연구 분야이며, Proof와 Trust는 시맨틱 웹 정보의 신뢰성과 보안에 관한 분야로 지속적인 연구가 진행 중이다.

"온토로지" 단어의 어원은 'ontos(being)'와 'logos(word)'이다 [7]. 온토로지는 해당 분야에 따라서 다양한 정의가 내려지고 있다. 우선 철학적인 관점에서는 세상의 기본이 되는 구성 요소에 대한 명확한 이해와 정의를 나타낸다. 이론적으로는 어떤 영역에서 공통적으로 사용되는 어휘들의 집합을 표현하고 있다. 자연어 처리와 관련해서는 실세계 혹은 특정 도메인에 존재하는 모든 개념들과 개념들의 속성 그리고 개념들이 상호간에 의미적으로 어떻게 연결되어 있는가에 대한 정보를 가지고 있는 지식베이스를 의미한다. 인공지능 분야에서는 관심영역분야에 존재하는 객체, 개념, 개체를 어떻게 나타내는가에 대한 명시적이고 공식적인 사상을 의미한다. 시맨틱 웹에서의 온토로지는 두 개의 데이터베이스 간에 정보를 비교하고 연결하기를 원하는 프로그램으로서 응용 프로그램 간에 웹 기반 지식의 처리, 공유, 재사용을 위해서 아주 중요한 역할을 수행한다. 결론적으로 온토로지는 어떤 문제 영역에 대한 개념, 그 개념에 대한 특성과 속성, 제한조건, 그들 간의 관계에 대한 명시적인 서술이라고 정의해 볼 수 있다[3].



[그림 1] 시맨틱 웹의 계층구조



[그림 2] KT의 TOP CLASS 생성

2.2 OWL 기반 온토로지의 필수 요소

OWL[9,10]은 웹 온토로지와 연관된 지식창고(knowledge bases), 즉 추론시스템에 축적된 논리적 명제(assertions)들을 정의하는 언어로서, 이러한 명제는 하나의 온토로지에 근거하거나 또는 OWL이 명시하는 방식에 따라 수집한 다수의 분산형 온토로지에 근거할 수 있다. OWL 명제는 클래스의 구성원들에 관한 사실과 구성원들 간의 관계를 기술하며, 이러한 명제의 집합으로 이루어진 온토로지는 구문적으로 정의되지 않은 사실의 논리적 유추를 가능하게 한다. 클래스와 속성 및 이에 적용할 수 있는 제약사항들의 집합인 OWL 온토로지는 Class, Property, Individual 세 가지 요소를 포함한다[10,11].

Class는 Individuals로 구성된 하나의 집합으로 생각할 수 있다. Class 생성에서 중요한 점은 특정 영역에 필요한 Class만을 생성해야 한다는 점이다. 이것은 Class 생성을 위한 특정 영역에 대한 분석이 이루어진 후 결정되어야 할 요소이며, 속성 및 Individuals에 대한 분석이 병행되어야 한다는 점을 의미한다.

Property는 Class(또는 Individual) 사이의 관계 및 Class(또는 Individual)와 데이터값 사이의 관계를 표현한다. 또한 Property는 Individuals에 대한 Binary relations로 생각할 수 있으며, Individual과 데이터값 사이의 Binary relations로 생각할 수도 있다. 즉 특정한 Class와 Class 사이의 관계를 표현한다는 것은 특정 Class들의 Individual 사이의 관계를 표현하는 것과 마찬가지로 때문이다.

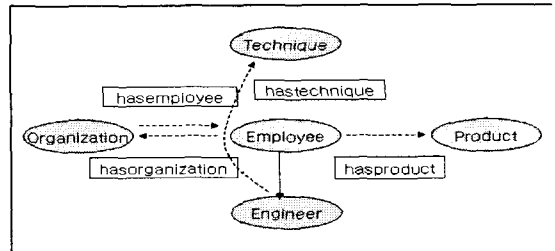
Individual은 Class의 인스턴스이자 특정 영역의 표현된 실제 객체들이며, Property를 통해 상호 관계를 맺는다. Individual은 특정 영역 온토로지의 Class 생성 및 Property 설정에 중요한 역할을 담당한다. 특정 Class의 Individual은 Property를 통해 다른 Class의 Individual 또는 Individual Datatype과 관계를 가진다. 이 Individual은 직접적 인스턴스(direct instance)와 Property를 통한 인스턴스(property instance)로 구분될 수도 있다.

3. KT 관련용어 온토로지 구축

3.1 KT관련 용어 온토로지 모델 개념

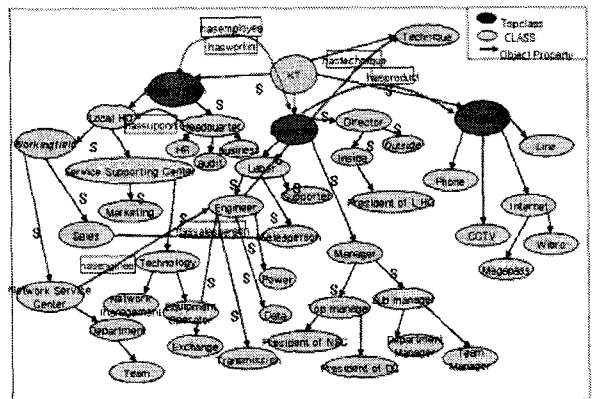
기업 KT는 [그림 2]처럼 4가지의 Top Class(Topclass)로 분류할 수 있다[12,13,14,15]. 분류된 Class는 Organization(기관), Employee(직원), Product(상품), Technique(기술)이며, 각 Class는 현재 KT조직에서 사용되고 있는 조직과 기술, 그리고 영업상품을 토대로 하여 만들었다.

Topclass는 상호간에 속성관계를 적용할 수 있다. Organization과 Employee 사이의 속성을 hasEmployee와 hasOrganization으로 정의하였다. 또한 Employee와 Product간의 속성은 hasproduct로 정의 했으며, Employee의 subclass인 Engineer와 Topclass의 technique간의 속성은 hastechnique로 하였다. [그림 3]은 Topclass간의 속성관계를 보여주고 있다.



[그림 3] Top Class간 속성관계 적용

KT관련 용어 클래스는 [그림 4]에서 보는 바와 같이 4개의 Topclass와 그 밑에 다수의 Subclass로 구성된다. 또한 각각의 Subclass는 역할과 지위, 용도 등으로 정의되었다. 각각의 독립성이 부여되고, 각 Topclass간의 속성관계를 [그림 4]와 같이 구성할 수 있다.



[그림 4] KT 전체 Class에 속성 적용

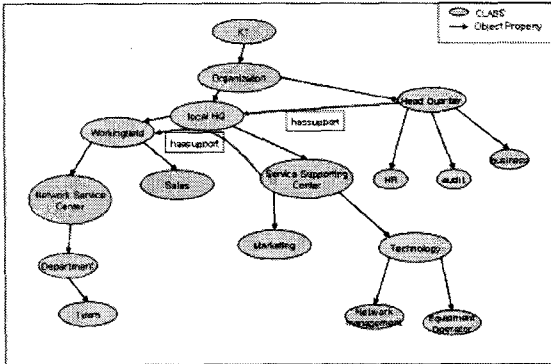
3.2 Organization(기관)의 Subclass

Organization(기관)은 localHQ(지역본부)와 headquarter(본사)로

2개의 Subclass로 나누어진다. localHQ(지역본부)는 workingfield(현장)와 service supporting center(서비스지원 센터)의 Subclass를 가지며, workingfield(현장)는 network service center(네트워크 서비스센터)와 sales(영업)를 subclass로 나눌 수 있고, department는 sales의 Subclass로 구성이 이루어진다.

headquarter(본사)는 hr(인력부서), audit(회계부서), business(사업부서)의 subclass를 가진다[15].

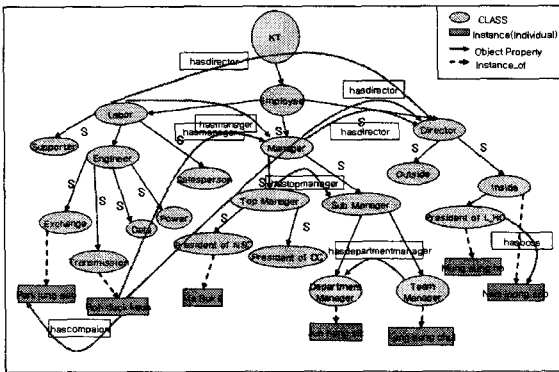
속성관계는 headquarter에서 localHQ를, service supporting center(서비스지원센터)가 workingfield를 hassupport의 속성을 가진다.(참조: [그림 5])



[그림 5] KT Organization의 Class에 속성 적용

3.3 Employee(직원)의 Subclass

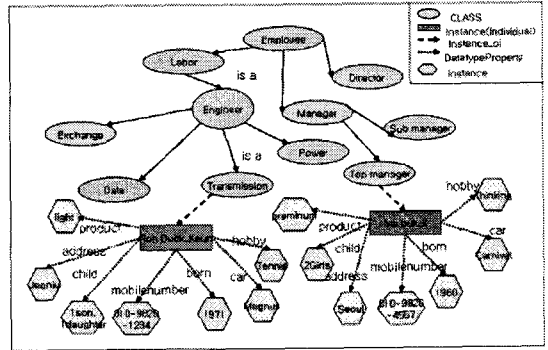
Employee는 [그림 6]에서 보듯이 labor(직원), manager(경영자), director(이사의) 3가지 Subclass를 가진다.[14,15] labor(직원)의 Subclass는 engineer(기술자)와 salesperson(판매사원) 그리고 supporter(지원)가 있으며, 특히 engineer(기술자)에서는 각 분야 별로 exchange(교환), power(전력), data(데이터), transmission(전송)으로 4개 부서로 구분을 두고 표시했다.



[그림 6] Employee의 Subclass와 Individual사이의 속성

manager(경영자)에서는 topmanager(고위경영자)와 submanager(실무경영자)로 구분되며, topmanager(고위경영자)는 president of NSC(NSC 센터장), president of DC(지역지사장)으로

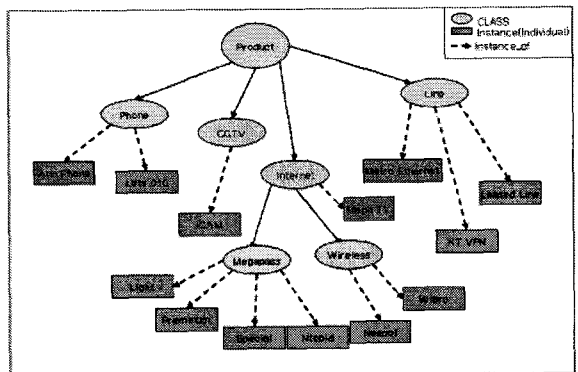
두 개의 Subclass를 가진다. submanager(실무경영자)는 departmentmanager(부서장)와 team manager(팀장)로 Subclass를 구분했다. director(이사)의 class에서는 inside(사내이사)와 outside(사외이사)로 Subclass를 정하였다. Employee(직원)의 individual에서 객체 roh duck keun은 manager(경영자) class와의 관계는 hasmanager, director(이사)의 class에서는 hasdirector, 그리고 exchange(교환)에 있는 객체 park jung soo와의 관계는 hascompanion의 속성관계를 맺었다.



[그림 7] KT Employee Class와 Individuals

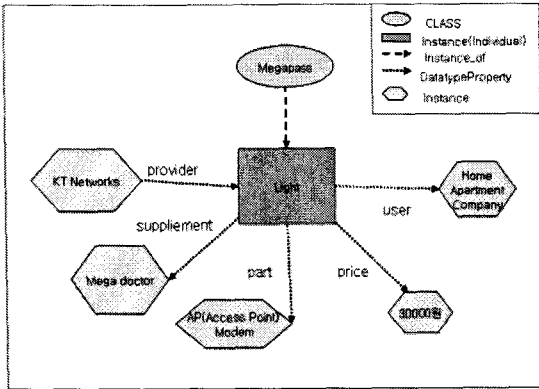
3.4 Product(상품)의 Subclass

[그림 8]의 KT Product Class와 IndividualsProduct(상품)은 KT 고유의 상품으로 여기서는 직원이 보유한 상품으로만 한정을 짓기로 했다.[12,13,16] 상품 상호간에는 서로 다르다는 전제 조건을 둔다. 상품은 Phone(전화), CCTV(영상감시), Internet(인터넷 상품), Line(회선) 4가지의 Subclass를 가진다.



[그림 8] KT Product Class와 Individuals

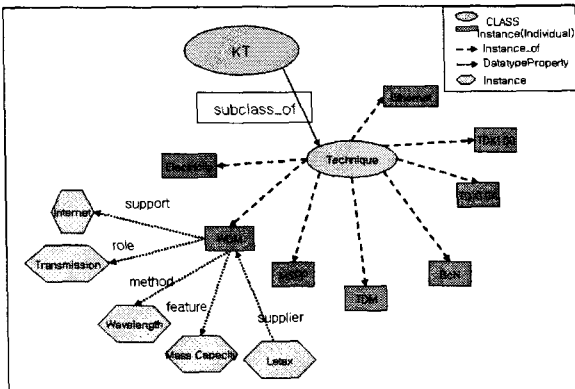
KT Product에 대한 Individuals에 대한 세부적인 내용을 할 살펴보기로 하자. 우선 KT Megapass Light와 ANN Phone에 대한 data 값은 provider, supplement, part, price, user 등 5가지로 나누어 입력하였다.



[그림 9] KT Product Megapass Light의 Instance

3.5 Technique(기술)의 Subclass

Technique(기술)은 별도의 Subclass를 보유하고 있지 않으며, Individual에서 Instance(객체)를 부여하는 방식으로 전개하고 있다. (참조: [그림 10])



[그림 10] KT Technique Class와 Individuals

4. 구축된 KT 관련용어 온토로지를 이용한 검색

구축된 온토로지로서 KT 기업에서는 인력관리, 상품판매 전략 수립, 조직 관리 등에 유용하게 사용될 수 있다. 본 논문에서는 인력과 기술 그리고 상품이 어떻게 연관성 있게 Query 될 수 있는지를 보여줄 것이다. 예를 들어 KT 조직의 한 구성원의 상관은 누구이며, 어느 기관에 소속 되어있는지 또한 그 직원은 어느 상품을 보유하고 있는지에 대해 살펴보겠다.

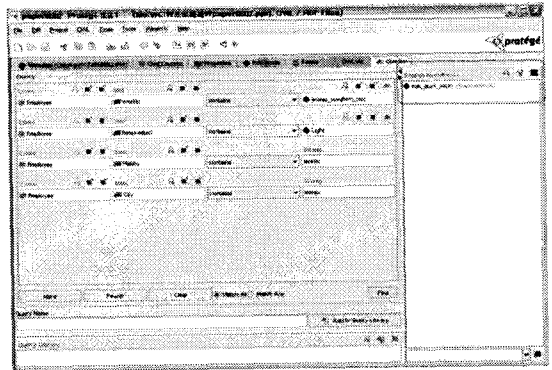
현재 온토로지를 구현하기 위해 적용된 DB는 조직 120 기관 중 20개 정도를 입력하였으며, 인력 38,000명 중 200명, 상품 40가지 중 18개 정도, 그리고 기술에 대한 자료는 KT내의 100 가지 정도의 기술 중에 20가지 정도를 입력하여 구현하고 있다.

4.1 Employee 정보 검색과 가족 검색

질의어: “KT 전북서부NSC에 근무하고, 직원이 보유한 상품

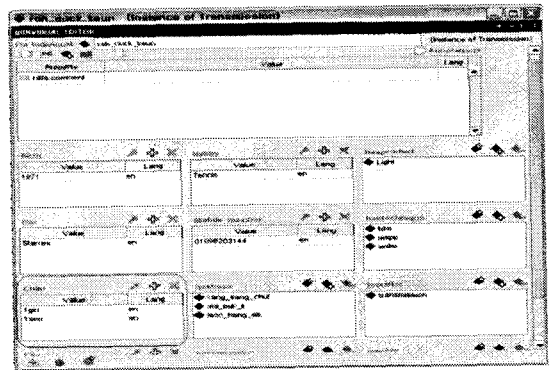
은 인터넷 Megapass Light를 사용하고 있으며, 취미는 tennis, 거주지는 전주에 사는 직원을 찾아 가족이 몇 명인지 검색하세요”.

KT 직원 중에서 전북서부NSC에서 근무하는 직원을 검색해보자. 검색 결과에 동시에 Product 중 인터넷 상품 Megapass Light를 보유한 직원을 찾게 되며, 또 다른 검색으로 hobby에 tennis를 입력하고, livein에 jeonju를 입력하여, 동시에 만족하는 결과 값을 찾게 된다. Individual 값으로 roh duck keun이라는 직원 1명이 검색되어 검색결과 창은 [그림 11]과 같이 보인다.



[그림 11] KT 직원 검색

[그림 11]에서 검색된 직원을 보고 그 직원의 가족을 찾으면 [그림 12]의 Child 라는 항목에서 원하는 결과를 찾을 수 있을 것이다. 검색결과는 1 girl, 1 son 을 [그림 12]에서 볼 수 있다.



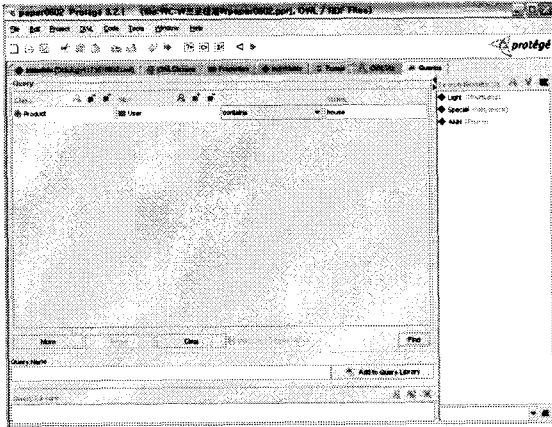
[그림 12] KT직원의 가족 검색

4.2 복합 검색

질의어: “House에 거주하는 일반인에게 추천해주고 싶은 인터넷 상품 중 저렴한 비용이어야 한다. 이 상품을 담당하는 KT 직원과 그 팀장의 차 이름과 핸드폰 번호를 검색하세요.”

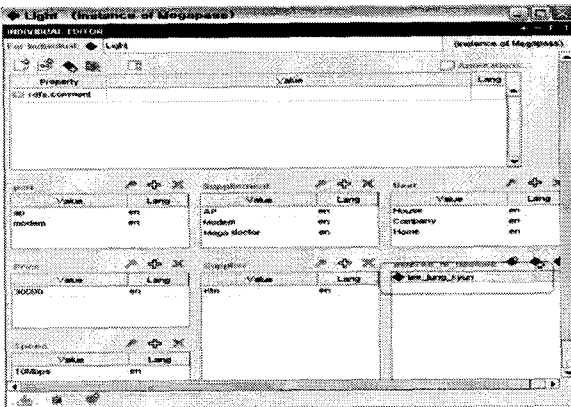
KT Product 에서 User가 House 라는 곳이였을 때, 해당 되는 상품을 검색하여 보자. 그러면 3가지 상품이 검색이 되었다.

그 중에서 인터넷 상품을 선택을 하게 되면 2가지 상품이 되며, 가격을 비교 하게 되면, Megapass Light라는 상품이 최종적으로 선택되어지게 된다. 결과는 [그림 13]에서 보인다.



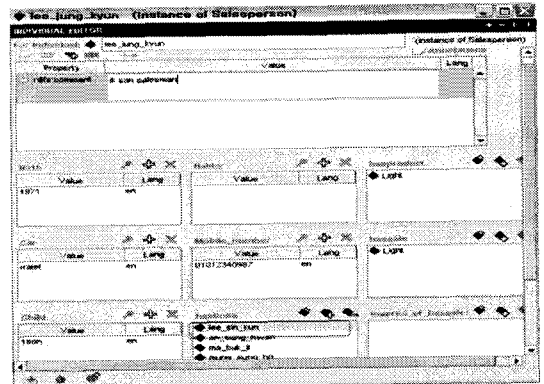
[그림 13] KT 인터넷 상품 검색

[그림 14]에서는 해당 상품을 Sale하는 담당 직원을 볼 수 있다. Inverse_of_hassle에서 lee_jung_kyun을 찾을 수 있다.



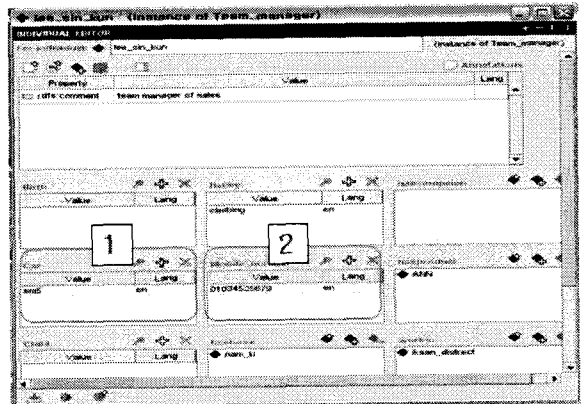
[그림 14] 담당 직원을 검색

Employee 중에서 담당 직원을 검색하였으며, 그 직원의 상세화면에서 hasboss라는 항목에서 담당 직원의 팀장인 lee_shin_kun 을 찾을 수 있었다. (참조: [그림 15])



[그림 15] 영업 담당의 상세화면

검색된 결과(참조:[그림 16])에서 해당 Individual 의 상세화면을 보면 담당 팀장의 상세내역을 볼 수 있다. 첫 번째로는 Car 에 대한 정보는 sm5 라는 정보를 알 수 있었으며, Mobile_number에 대한 두 번째 정보는 01034525679 라는 걸 알 수 있다.



[그림 16] 담당 팀장의 상세내역

5. 결론 및 향후과제

5.1 결론

지금까지 KT관련 용어를 분석하여 온토로지를 구축하였으며, Class에서 속성을 적용한 후 질의를 하여 검색됨을 살펴보았다.

온토로지를 이용하여 Employee(직원)를 기준으로 Organization과 Product 그리고 Technique과의 연관성을 볼 수 있었다. 즉 직원 한 사람이 속해 있는 Organization(조직)과 소유하고 있는 Product(상품) 그리고 Employee 중 Engineer가 소유하고 있는 Technique(기술) 과의 연관성도 함께 볼 수 있었다.

현재 구축 되어있는 DB 시스템에서는 단편적인 한 분야에 관한 정보만을 볼 수 있었지만, 온토로지를 통하여 검색을 진

행하면 한 분야의 정보뿐만 아니라 관련된 모든 자료를 같이 볼 수 있게 되었다. 본 논문에는 인력관련 분야에서 적용한 예를 들어 설명하였지만, 예를 보면 구축된 온토로지를 통하여 다른 많은 부분에서 사용이 가능하다는 것을 알 수 있다. 이것은 KT가 가지고 있는 방대한 지식 정보를 온토로지를 통하여 체계화 하면서 효율적으로 KT 전체 DB를 관리 할 수 있다는 것을 보여준다. 또한 본 논문에서 사용한 KT용어 온토로지 DB는 3~5%입력으로 구축되어 추후 많은 정보가 입력이 되면, DB의 종합적 관리방안이 기대된다. 동시에 상호 연계하는 방안으로 유용환경 구축을 위한 기초적인 문제해결 방안을 제시하였다.

5.2 향후 연구 과제

향후 연구 과제로 현재 KT내부에서 관리하고 있는 “KT insa” 인력관리 시스템에 구축된 온토로지를 적용하여 지능형 검색 서비스를 구현할 수 있다. 또한 KT는 고객의 정보를 관리하는 CRM(고객관계관리)과의 연계를 고려할 수 있다. CRM(고객관계관리)은 고객이 가지고 있는 모든 정보를 기록하고 관리하고 있는 시스템이다. 즉 CRM에 온토로지를 구축하여 고객이 필요로 하는 상품은 무엇이며, 고객이 원하는 상품에 어느 정도 가격을 지불할 것이며, 어느 상품이 어떤 분야에 필요로 하는지를 고객에게 물어 보지 않고도 맞춤형 서비스가 가능하다. 영업의 비효율적인 자원의 중복 소요를 막고, 기다리는 서비스가 아닌 찾아가는 서비스를 지향하는 것이다. 그리고 고객이 필요한 것을 찾아주고 제시해주는, 한 층 업그레이드 된 영업이 가능할 것이다.

본 논문은 KT 내부에서 인력과 기술, 상품을 접목 시키는 초석을 마련하였으며, 외부적으로는 고객과 상품의 관계를 규정하는 기초를 다지고 있다. 향후 시멘텍 웹의 온토로지를 이용하여 KT CRM(고객관계관리)과 연계성을 확보한 방안을 모색하는 연구가 필요하다.

6. 참고문헌

- [1] A. Gomez-Perez and O. Corcho, "Ontology Language for the Semantic Web", IEEE Intelligent Systems, vol.17, no.1, pp.54-60, 2002.
- [2] 최훈섭, 임지희, 배영준, 최수일, 옥철영, "온토로지 구축 방법과 사례", 정보과학회지, vol.24, no.4, pp.31-44, 2006.
- [3] 이경중, "온토로지를 이용한 제품가족 분석", 경성대학교 석사학위논문, 2006.
- [4] 최중민, "시멘텍 웹의 개요와 연구동향", 정보과학회지, vol.21, no.3, pp.4-10, 2003.
- [5] Tim Berners-Lee, James Hendler and Ora Lassila, "The Semantic Web", Scientific American, 2001.
- [6] M-R Koivunen and E. Miller, "W3C Semantic Web activity", Proceedings of the Semantic Web. Kick-off Seminar in Finland, 2001.
- [7] 오현목, "시멘텍 웹 발전 방향 및 표준화 개발전략 연구", 한국전산원, 2005
- [8] D. Fensel and F. van Harmelen and I. Horrocks and D. McGuinness and P. Patel-Schneider, "An Ontology Infrastructure for the Semantic Web". 2001.
- [9] Deborah L. McGuinness and Frank van Harmelen, "OWL Web Ontology Language Overview", 2004, <http://www.w3.org/TR/owl-features/>.
- [10] 김기훈, "OWL 기반의 사용자 요구사항 Annotation 시스템", 한국 대학교 석사학위논문, 2005
- [11] Protege, <http://protege.stanford.edu/>
- [12] KT Homepage (<http://www.kt.co.kr>)
- [13] KT 사내 사이트 (<http://kate.kt.co.kr>)
- [14] KT 사내 메신저 (ktiman)
- [15] KT 인사 시스템 (<http://insa.kt.co.kr>)
- [16] KT Megapass 사이트 (<http://www.megapass.net>)