

u-지능정보

1. 개요

1.1. 추진경과 및 중점 추진방향

■ 추진경과

- 2004년(Ver.2005)~2005년(Ver.2006)에는 음성언어정보처리 분야의 표준화 대상항목을 도출하여 표준화로로드맵을 수립하였다.
- 2004년(Ver.2005)과 2005년(Ver.2006) 중점 표준화 대상항목의 차이점은 2005년 중점 표준화항목은 2004년과 비교하여, 국내표준화활동 능력과 실제 표준화 채택 가능성을 고려하여 보다 구체적으로 표준화항목을 선정하였다.

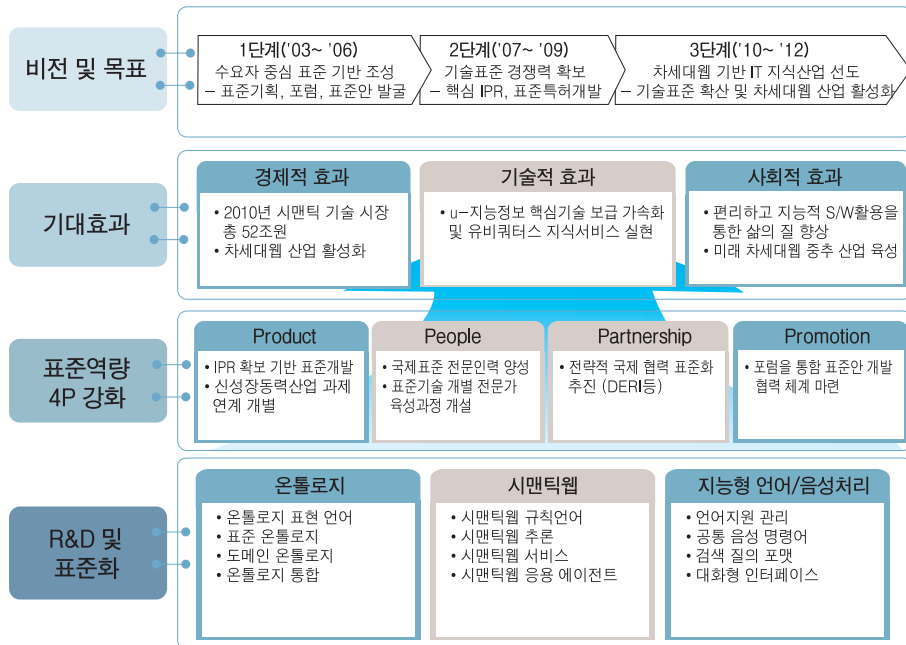
연도	2004년(Ver.2005)	2005년(Ver.2006)	차이점
중점 표준화항목	<ul style="list-style-type: none"> · 음성/언어 자원 · 음성언어처리 응용기술 · 음성언어처리 평가기술 	<ul style="list-style-type: none"> · 전문용어 표현 기술 · 개인미디어 규격 표현 기술 · 공통 음성 명령어 · 음성언어처리 평가 기술 	중점 표준화 대상항목을 구체적으로 선정함

*음성언어정보처리 분야의 표준화 로드맵은 2004년부터 수립하였음

■ 중점 추진방향

- 2006년(Ver.2007)은 소프트웨어 u-지능정보 기술 분야에서 표준화 대상항목을 도출하며, 중점 추진 방향은 해당 분야의 표준화 대상항목들 중에서 국내 전문가들의 표준화활동에서 표준화추진 성공 가능성 및 시급성을 고려하여 중점 표준화항목을 우선적으로 선정하도록 하였다.
- u-지능정보 기술 분야의 표준화로드맵 수립 및 표준 정립은 관련 국가 연구개발 과제를 수행하는 연구자를 중심으로 국내 산업체 및 대학, 연구소가 참여하는 표준화포럼을 구성하여 추진하도록 한다.

1.2. 표준화의 Vision 및 기대효과



(그림 1) U-지능정보 기술 표준화의 비전 및 기대효과

1.2.1. 표준화의 필요성

차세대 웹환경과 유비쿼터스 환경에서 컴퓨터가 이해할 수 있는 지능적 지식정보처리 서비스를 지원하기 위해 u-지능정보 기술에 대한 표준개발이 필요함

- 유비쿼터스 환경에서 다양한 지식정보의 의미처리를 통해 사용자는 보다 편리하고 지능적인 S/W를 활용하기 위해서는 이와 관련한 u-지능정보 기술들에 대한 표준 개발이 필요하다.
- u-지능정보 기술은 소프트웨어 타기술들의 응용에 필요한 컴포넌트 기술로 여러 응용에서 활용 가능토록 온톨로지 구축, 시맨틱 웹, 적응형 지능형 인터페이스의 시맨틱 기술별 표준화를 추진하여 신성장동력 산업 활성화에 기여한다.
 - 컴퓨터가 웹 정보자원의 의미를 이해하고 이를 바탕으로 추론이 가능하게 되어 웹과 사람의 상호 작용이 가능할 수 있는 시맨틱 웹이 차세대 웹으로 발전하기 위해 시맨틱 웹 지식표현, 추론 등에 대한 표준화 연구가 필요하다.
 - 시맨틱 웹이 차세대 웹 환경에서 컴퓨터가 이해하고 처리할 수 있는 지식을 표현하고 생성할 수 있는 근간이 되는 기술로 온톨로지 구축 기술이 핵심 역할을 수행하며, 표준화된 지식구조의 설정이 필수적이고 국가 차원의 표준 온톨로지 구축이 요구됨

- 유비쿼터스 환경에서 정보를 교환하는 상대는 현재 '사람과 사람' 중심에서 '사람과 기계'로 바뀌고 있으며, 언제 어디서나 편리하게 이용할 수 있는 적응형 사용자 인터페이스의 수요가 급증할 것으로 예상되며, 이와 관련된 표준화작업이 필요함
- 자연스런 사용자 인터페이스를 구현하기 위해서는 기술 특성상 대량의 음성/언어자원이 필요한 반면, 이의 구축에 많은 비용과 시간이 소요되며, 각 기관별로 규격이 달라 상호 활용이 되지 않고 있어 효율적 음성/언어자원 활용을 위한 표준화작업이 필요함
- 차세대 웹환경에서 컴퓨터가 이해할 수 있는 지능적 지식정보처리 서비스와 유비쿼터스 환경에서 다양한 디바이스간의 끊임없는 IT지식 서비스를 지원하기 위해 u-지능정보 기술에 대한 표준개발이 요구된다.

1.2.2. 표준화의 목표

유비쿼터스 환경의 다양한 디바이스간의 끊임없는 IT지식서비스를 지원하는 유비쿼터스 기반 시맨틱 웹 표준(안)을 개발

- 2007년까지 시맨틱 웹과 온톨로지 구축 기술의 표준(안)을 개발하여 W3C, ISO 국제표준화를 추진
- 2012년까지 유비쿼터스 환경의 다양한 디바이스간의 끊임없는 IT지식 서비스를 지원하는 유비쿼터스 기반 시맨틱 웹 표준(안) 개발을 적극적으로 주도하여 관련 표준 특허 20%-30% 획득
- u-지능정보 기술 분야 국내표준화 기반을 조성하고, 기고서 제출 등 국제기구의 지속적인 참여를 통하여 향후 국내 산업의 활성화를 도모하고 세계 기술 선도 입지를 강화

1.2.3. Vision 및 기대효과

u-지능정보 기술의 표준화를 통해 u-지능정보 기술 보급의 가속화와 유비쿼터스 IT지식 서비스 구현을 통한 신성장동력산업에 기여

- u-지능정보 기술의 표준화를 통하여 u-지능정보 기술의 보급을 가속화함으로써 유비쿼터스 IT지식서비스 산업의 시너지 효과와 함께 성장 동력화에 기여할 수 있다.
- u-지능정보 기술에서 시맨틱 웹, 온톨로지 구축 관련 핵심기술에 대한 기술개발 및 연구를 통해 향후 차세대 웹 기술 표준의 우위를 점전하고 그 기반을 구축할 수 있을 것이다. 또한 차세대 웹 기술에 대한 표준 기술 체계의 전문화, 국제화, 실용화를 목표로 한 중장기 기술개발 투자를 통해, 향후 미래 신성장 동력으로서 인터넷 산업을 국가의 중추 산업으로 체계적이며 효율적으로 육성할 수 있는 효과가 있다.
- 또한, 국내 산업계에서 필요로 하는 기술에 대한 국내표준 및 국제표준을 개발하여 관련 업체간의 기술 호환성을 확보함으로써 산업 효율성 증대 및 비용 감소가 가능하므로, 국내 기반 기술 보급, 핵심 표준 기술의 선행 연구 및 국제표준화추진을 통하여 관련 산업의 기술 경쟁력을 강화시키고 지적 재산권을 확보함으로써 국

내 업체의 사업 경쟁력 강화에 기여할 수 있다.

- 시맨틱 웹의 차세대 웹기술은 유비쿼터스 환경에서 다양한 지식 정보자원과 응용서비스를 지능적으로 가능케 하는 u-IT839 환경의 인프라 기술로 중요한 의의를 지닌다.
- u-IT839 환경에서의 공통 음성 명령어에 대한 표준화를 통하여 체계적이며 일관된 음성 인터페이스 개발을 유도함으로써 사용자가 서비스나 하드웨어 플랫폼 종류에 상관없이 보편적인 방법으로 편리하게 사용할 수 있도록 한다.
- 사용자의 자연스런 질의 처리와 음성언어 정보처리 엔진에 필요한 음성/언어자원관리의 표준화를 통해 음성/언어자원의 효율적 활용이 가능하게 된다.

2. 국내외 현황 분석

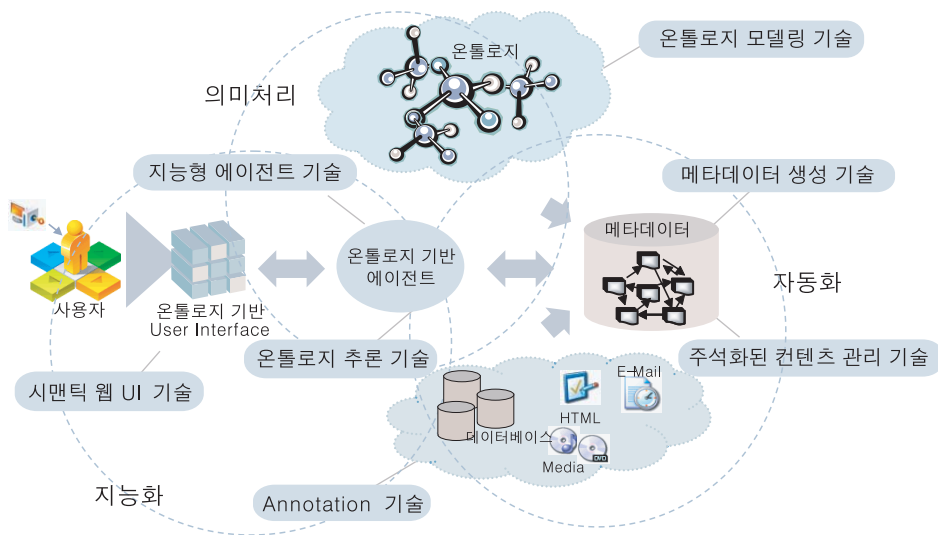
2.1. 중점기술 개요

2.1.1. 중점기술 및 표준화 대상항목의 정의

- 중점기술의 정의

컴퓨터가 이해할 수 있는 지능화된 정보를 구축하고 지식서비스의 품질을 향상시키는 기술

- 지능 정보 기술은 차세대 웹, 온톨로지, 메타데이터 생성 및 추론, 에이전트, 적응형 사용자 인터페이스 등의 시맨틱 기술과 관련된 지능형 지식처리 기술로서, 컴퓨터가 이해할 수 있는 지식의 지능적 처리를 통해 다양한 유비쿼터스 디바이스간의 끊김 없는 IT 지식 서비스를 지원한다.



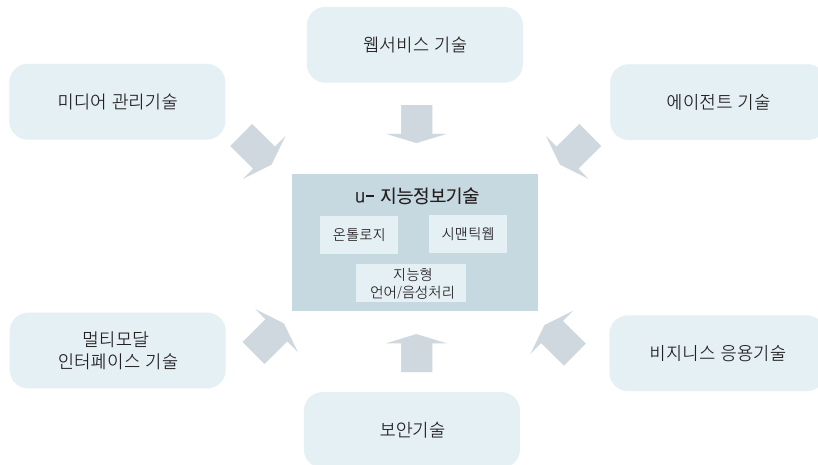
(그림 2) u-지능정보 기술 개념도

• 표준화 대상항목의 정의

구분	정의	표준화 대상항목	표준화 내용
u-지능정보 지식베이스	유비쿼터스 지식 서비스에 활용할 수 있는 온톨로지 기반 지식 베이스	온톨로지	- 온톨로지 표현 언어 및 검증 - 응용서비스에 적합한 표준 온톨로지 - OWL-S 등의 도메인 온톨로지 정의 - 온톨로지 매핑/통합 기술
u-지능정보 인프라	지능형 지식 처리에 필요한 온톨로지 상의 지식 획득 및 시맨틱 웹 인프라	시맨틱 웹	- RDF, RDF-S와 같은 시맨틱 웹 표현 기술 - 시맨틱 웹 규칙 표현 기술 - 대용량 온톨로지의 실시간 추론 - 시맨틱 웹 데이터를 검색하기 위한 검색 언어
지능형 사용자 인터페이스	유비쿼터스 환경에서 편리하고 지능적인 사용자 인터페이스	지능형 언어/음성 처리	- 언어자원 표현/생성/관리 방법론 - 다양한 정보기기 사용 편의를 위한 음성명령어 정의 - 미디어 검색 질의와 결과 제시를 위한 포맷 정의 - 사용자의 자연스런 질의를 위한 대화형 인터페이스

2.1.2. 연관기술 분석

• 연관기술 관계도



(그림 3) U-지능정보 기술 표준화 연관기술 관계도

• 연관기술 분석표

연관기술	내용	표준화기구/단체		표준화수준		기술개발수준	
		국내	국외	국내	국외	국내	국외
웹서비스 기술	XML로 표현된 데이터들이 SOAP 등의 웹에서의 서비스 연동을 위한 기술	TTA	W3C, OASIS, WS-I	표준화 진행 중	표준화 진행 중	기술 개발 중	기술 개발 중
에이전트 기술	자율적으로 동작하면서 지능성과 적응성을 갖추도록 프로그래밍하는 기술	TTA	-	표준 미제정	표준화 진행 중	기술 개발 중	기술 개발 중
비즈니스 응용 기술	ebXML, 로제타넷 등과 같은 웹 기술을 이용하여 비즈니스 분야의 응용 관련 기술	TTA	OASIS, UN 등	표준 제/개정	표준 제/개정	상용화	상용화
보안 기술	암호화, 복호화, 키관리 등과 같은 안전하게 웹을 이용할 수 있도록 하는 기술	TTA	IETF	표준 제/개정	표준 제/개정	상용화	상용화
멀티모달 인터페이스 기술	음성, 펜, 제스처 등 여러 가지 방법으로 정보기기를 사용하는 사용자 인터페이스 기술	TTA	W3C	표준 미제정	표준화 진행 중	기술기획	기술 개발 중
미디어 관리 기술	대규모 웹/멀티미디어 콘텐츠에 대하여 효과적인 관리가 가능하도록 메타데이터를 생성/유통하는 기술	TTA	MPEG	표준 미제정	표준화 진행 중	기술기획	기술 개발 중

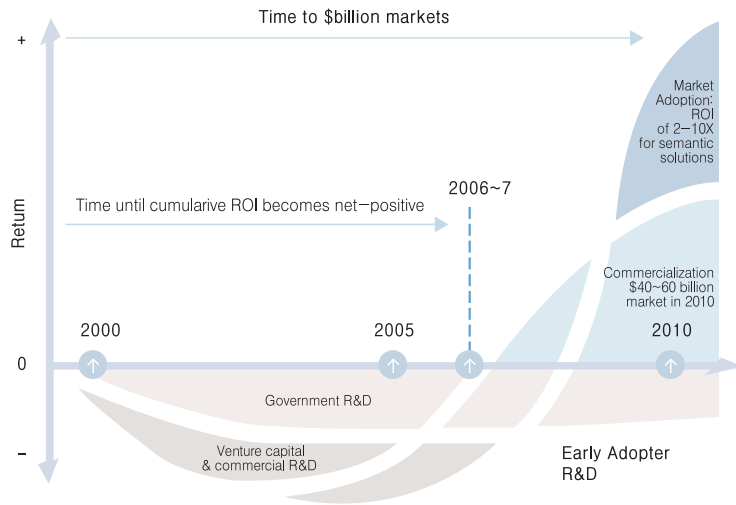
2.2. 시장 현황 및 전망

2.2.1. 국내 시장 현황 및 전망

- 온톨로지와 추론을 활용한 시맨틱 기술 관련 시장은 국내 IT 및 S/W 시장과 시맨틱 기술관련 세계 시장을 이용하여 추산한 결과 2010년 1조원에 달할 것으로 전망된다.(출처 : Saltlux 컨퍼런스, 시맨틱 기술 전문 업체인 Saltlux 추정, 2005년 10월)
- 시맨틱 웹과 관련된 국내 시장은 소비 시장 측면에서는 현재까지는 형성되어 있다고 보기 어렵다. 그러나 R&D 시장과 관련 그 규모는 꾸준히 증가하고 있으며, 시맨틱 웹 기술 자체 시장이라기보다는 관련 응용 시장의 부문 기술로서 시장을 형성해나갈 가능성이 매우 높다. 관련 응용 시장으로는 먼저 지식관리시스템(KMS) 시장, e-러닝 시장, 기업애플리케이션통합(EAI) 및 기업 간 통합 시장, 검색 시스템 시장 등에서 그 수요가 발생할 가능성이 크며, 이들 시장을 바탕으로 시맨틱 웹 기술은 점차 독자적 기술 및 응용 시장으로 인식되어 갈 것이다. 따라서 국내 시장 전망을 독자적으로 현 시점에 논하기는 어려우나, 이들 개별 응용 분야의 시장 규모에 따라 비례하여 증가할 것이며, 국외 시장 현황이나 전망에 비추어 거의 같은 비율로 성장하리라 예측된다.
- 언어 자원을 사용하는 분야로 텍스트 마이닝과 자동 번역 분야가 있다. 국내 자동 번역 시장은 연 평균 20% 이상으로 성장하고 있으며, 최근 중국 경제의 활성화를 예상하여 한중, 중한 번역 연구가 활성화되고 있다. 국내 텍스트 마이닝 시장과 관련하여 CRM시장은 2001년에는 2,000억~2,500억 원 정도가 될 것으로 추산되고 있으며, 2002년에는 3,500~4,000억 원으로 70~80% 가량 성장된 것으로 추산된다.
- 질의 처리 분야는 검색 시스템에서 활용된다. 국내 검색 포털 사이트 1위인 네이버의 NHN은 2006년 2분기 매출이 1329억으로 전분기보다 9.1% 늘고, 영업이익도 516억으로 10.8% 증가할 전망이다.(한국경제, 2006, 7월). 동영상 검색 관련 방문자 수는 100% 이상 증가하고 있으므로, 향후 동영상 관련 검색 시장이 확대될 것으로 전망된다.(ZDNet Korea, 2006, 4월)
- 음성 처리 분야의 경우 신성장 동력산업 추진으로 고부가가치, 경쟁력 확보를 위한 음성인터페이스 기술 수요가 증대됨에 따라 시장규모가 보다 빠른 속도로 확대될 전망이다.

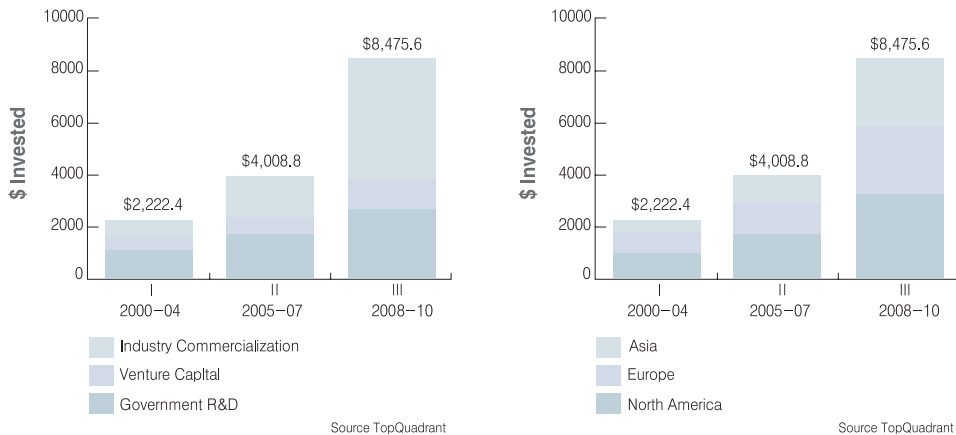
2.2.2. 국외 시장 현황 및 전망

- adobe, AT&T, Google, HP, IBM, Sony, Sun 등 65개 이상의 기업에서 시맨틱 기술 관련 제품을 개발하고 있으며, 2003년에는 12억 달러의 시장 규모이고, 55~75%의 연평균 성장(CAGR)을 통해서 2010년에는 52억 달러, 2015년에는 5,000억 달러에 달할 것으로 전망된다.("Semantic Wave 2006, Part-1", A Project10X Report, MILLS-DAVIS, 2006년 1월)



(출처: MILLS&DAVIS Semantic wave 2006)

(그림 4) 시맨틱 기술의 투자효과



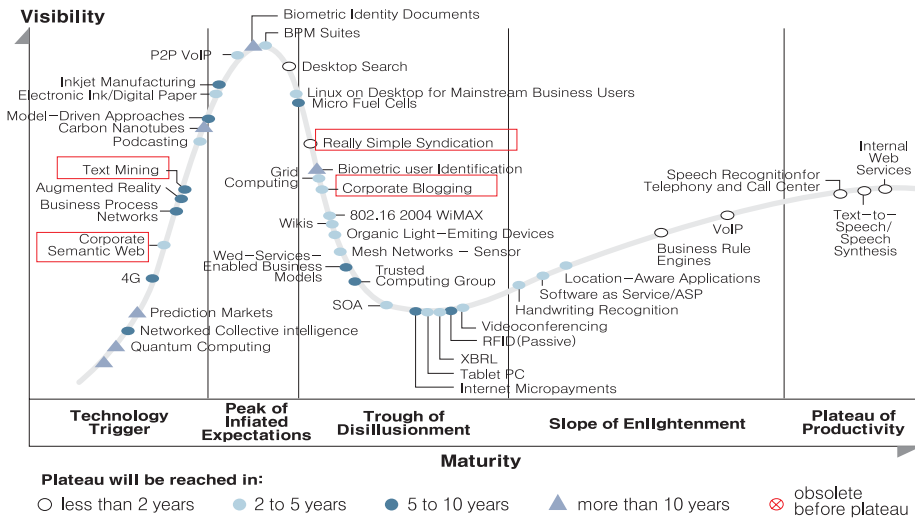
(그림 5) 2010년까지 시맨틱 기술 투자 : 투자 형태(왼쪽), 투자 지역(오른쪽) (TopQuadrant 2004)

- Mills-Davis의 Project10X 2006년 보고서에 의하면 온톨로지를 활용하는 시맨틱 시장의 성장속도는 2006년부터 2010년 사이에 10배의 성장을 기대하고, 2010년부터 2015년까지도 10배가 성장할 것으로 예측하고 있다.(<표 1>)

〈표 1〉 온톨로지를 활용하는 시맨틱 시장의 규모 (Source : Mills-Davis, Jan. 2006)

시장	2006년	2010년	2015년
Semantic Development	\$50M	\$0.4B	\$2.0B
Semantic Infrastructure	\$500M	\$17.0B	\$200.0B
Knowledge Work Automation	\$1,100M	\$30.0B	\$250.0B
- Information-intensive	\$350M	\$4.5B	\$40.0B
- Knowledge-intensive			
Systems That Know	\$100M	\$0.5B	\$10.0B
TOTAL	\$2,200M	\$52.4B	\$500.0B

- 2005년 가트너 그룹의 분석에 의하면 Corporate Semantic Web, RSS, Corporate Blogging, Text Mining 같은 기술이 조만간 성숙하여 경제적 기대 효과를 증대할 것으로 예상하고 있는 데, 이러한 기술은 기본적으로 온톨로지를 활용하고 있다.



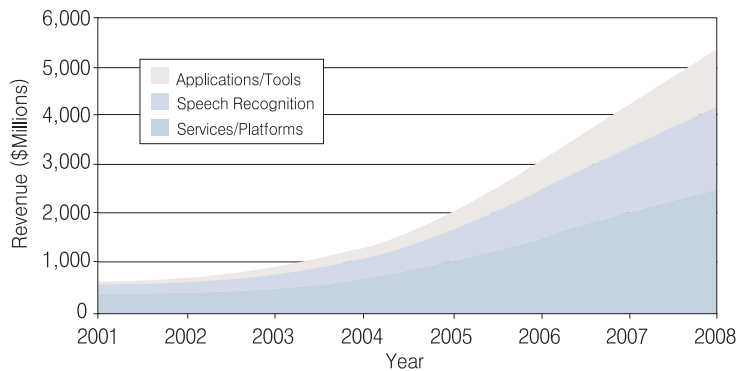
(그림 6) 가트너 그룹이 예측한 온톨로지 관련 주요 기술 (Source : Gartner Group, 2005)

- 온톨로지 기반 기술이 향후 핵심기술로 자리 잡을 것이며, 2007년경에는 간단한 온톨로지가 응용 프로젝트의 75%를 차지하고, 인공지능의 지식표현 기술을 활용하는 온톨로지가 응용 프로젝트의 80% 이상을 차지할 것으로 예상되고 있다.(Gartner 2002 보고서)
- 언어 자원을 사용하는 분야로는 텍스트 마이닝과 자동 번역 분야가 있다. 텍스트 마이닝 시장과 관련하여 e-mail marketing 시장은 4년간 10배 이상 증가가 이루어지고 있으며, eCRM 솔루션 시장에서는 연 평균 143%로 성장하고 있다. 자동 번역 시장은 2000년 6천 7백만 달러에서 2005년 2억 3천만불로 연평균 28% 이상의 성장이 이루어지고 있다.
- 질의 처리 기술과 밀접하게 관련되는 검색 시장의 경우, 미국에서 검색 사용자의 60%가 사용하는 구글

(google)의 경우, 2006년 2분기에 주식 기반의 보수 등의 특별 과목을 제외했을 경우 순이익은 7억 7200만 달러, 매출액은 트래픽 획득 비용을 제외하고 16억 8000만 달러이다. 특별과목을 포함했을 경우 순이익은 7억 2100만 달러로 전년 동기 대비 3억 4280만 달러의 2배에 달한다.(CNet news.com 2006, 7월)

- 음성인식 시장은 향후 10년 이내 성장 단계에 이르러 폭발적으로 시장규모가 커질 것으로 전망된다. 2003년 8억 9,780만 달러 규모에서 2008년에는 53억 달러까지 신장할 것이라고 예측되고 있다.(출처 : Allied Business Intelligence)

Speech Technology Market Revenue, World Market: 2001 to 2008
(Source: Allied Business Intelligence Inc)



2.3. 기술개발 현황 및 전망

2.3.1. 국내 기술개발 현황 및 전망

- 정부정책기조
 - 2004년 이후 한국전산원을 중심으로 시맨틱 웹에 대한 국가적 타당성과 표준 관련 주제들이 다루어지고 있으며, 2005년부터 시맨틱 웹 도입 전략에 대한 검토 과제와 시맨틱 웹서비스에 대한 연구를 지속적으로 추진하고 있다. 현재 유럽의 DERI 연구소와 시맨틱 웹서비스 표준안의 공동 연구를 추진 중이다.
 - 정보통신부에서 추진 중인 u-IT839 전략의 3대 인프라 중 하나인 소프트웨어 분야로 지능형 정보서비스 기반 구축을 정하고, 시맨틱 기술개발을 위한 기반을 조성하고 있다.(정보통신부, IT839 전략, 2006)
 - 시맨틱 웹과 관련하여 국내에서는 2010년까지 지식 기반 정보 유통 플랫폼 개발, 차세대 웹을 위한 시맨틱 서비스 에이전트 기술개발, 국가 IT 온톨로지 인프라 기술개발의 과제를 통해서 시맨틱 웹 기초 기술 및 활용 기술개발을 진행하고 있다.
 - 음성 처리 기술과 관련하여 2010년 2월까지 신성장 동력산업용 대용량 대화형 분산처리 음성 인터페이스 개발 과제를 통하여 지능형로봇, 텔레매틱스, 홈네트워크, 차세대PC 등 여러 분야에서 공통적으로 요구되

는 핵심 요소기능인 음성인터페이스 기술개발을 진행하고 있다.

- 2004년부터 국가 지식정보자원관리사업에서 한국어 언어/음성자원의 표준 구축을 통해 관련 신성장 동력 산업 기술개발에 공동 활용하고 있다.

- 온톨로지 관련 기초 지식으로, ETRI는 의미기반 정보검색 및 백과사전 질의응답에 활용된 45만 어휘개념망을 구축하였고, KAIST는 다국어 어휘 의미망인 CoreNet, 울산대학교는 40만 어휘에 기반한 보편적 온톨로지인 U-WIN을 구축하였다. 또한 KAIST와 ETRI, 포항공대, 부산대, 울산대 등 국내 산·학·연이 공동으로 2006년부터 국가IT 온톨로지 구축 사업을 통해 시맨틱 웹 상에서 실제적인 서비스가 가능한 코어 온톨로지 구축 및 추론 엔진 개발을 진행하고 있다.

분류	이름	연구기관	개요
어휘의미망	CoreNet	KAIST	- 하나의 개념체계를 공유하는 다국어 어휘의미망 (한국어, 중국어, 일본어) - 대용량 말뭉치에서의 상대 빈도에 따른 5등급의 기본 어휘를 단일어 사전을 이용하여 반자동 구축
	U-WIN	울산대	- 인간의 보편적인 인지체계와 개념관계를 기반으로 한 대규모 한국어 어휘 의미망 - 현재 40만 어휘 수준까지 구축되어 있으며 영어대역 어휘 네트워크 구축 중
	ETRI 어휘개념망	ETRI	- 어휘 개념들 간의 다양한 관계를 연결시켜 놓은 어휘 데이터베이스 - 백과사전 기반의 질의응답 시스템 및 의미 기반 정보검색 등에 활용
사전 분류체계	세종 의미분류체계	21세기 세종계획 국립국어원	- 다양한 문법범주정보를 포함한 국어정보학적 분류체계
분야별 시소러스	국가과학기술개발정보 시소러스	울산대, KISTI	- 연구인력, 연구사업, 사업비 규모, 학술성과물, 연구기관, 연구 토픽 등 다양한 과학기술 R&D 기반정보를 체계화할 수 있는 시소러스
	스마트메신저 영화정보 온톨로지	KT 연구소	- 영화와 관련된 서비스를 제공하기 위한 OWL 기반 온톨로지

- ETRI에서는 시맨틱 웹 관련 연구 결과로 OWL로 문서를 저작할 수 있게 해주는 비즈니스 온톨로지 모델러인 온투모(OntoMo), 존재하는 시맨틱 웹 문서들을 결합한 새로운 온톨로지를 짧은 시간 안에 구축할 수 있도록 해주는 비즈니스 온톨로지 병합 도구 모아(MoA), 온톨로지 기반의 추론엔진 아울러(OWLer), 규칙 기반 추론엔진 보쌈(BOSSAM) 등을 개발하였다. 또한 미국 스탠포드대 의료정보기관인 SMI(Stanford Medical Informatics)가 개발한 프로테제-2000을 기반으로 GUI 환경을 지원하는 온톨로지 저작도구인 이시아울(ezOWL)을 개발하였다.

- 정보통신부의 선도기술개발과제의 일환으로 “차세대 웹을 위한 시맨틱 서비스 에이전트 기술개발”을 위해 KT, K4M 등의 산업계와 연세대, 서울대, 원광대 등 학계가 힘을 모아 5개 년 계획을 바탕으로 시맨틱 웹의 상용 기술 확보와 시맨틱 웹 규칙 추론 기술의 국제적 선도 위치 확보를 목표로 개발을 진행 중이다.

- LG CNS는 온톨로지를 근간으로 시맨틱 웹 형태의 지식관리 시스템(KMS)을 구축하는 ‘시맨틱 지식관리 시스템(KMS) 구축’ 프로젝트를 진행하였으며, 한국교육학술정보원(KERIS)는 전국도서관 데이터베이스 통합프로젝트에 온톨로지를 이용해서 학술정보 지식데이터베이스를 구축하고 있으며, KTF와 문화콘텐츠진흥원에서도

각각 모바일 콘텐츠 서비스와 문화콘텐츠 유통 포털 등에 시맨틱 웹을 도입할 예정이다.(전자신문, 2006, 06)

- 언어자원은 한국전자통신연구원, KOTERM, 울산대, 부산대 등에서 만들어지고 있다. 일부 국내에서 만들어진 언어자원은 외국에서 만들어진 WordNet과 같은 언어 자원에 의존하고 있으며, 구축된 언어자원의 활용은 미진한 편이다.
- 국어정보화 중장기 사업으로 10년 간(1998년 ~ 2007년)의 계획으로 추진되고 있는 21세기 세종계획에서 1998년부터 2007년까지 1단계 개발환경 구축 및 기본자료 집성, 2단계 실용화, 3단계 운용으로 경제, 물리, 화학, 수학, 의학, 전산학, 공학 등의 다양한 분야의 전문용어를 구축하고 있다. 한국학술단체연합회에서는 2003년도부터 학술 전문용어 정비 및 표준화 사업을 통해서 생명과학, 물성과학, 인문사회 등의 학술 분야에 대한 전문용어를 정비하고 있다.
- 질의 처리는 모든 검색 시스템에서 사용되는 필수 항목으로, 국내 검색 포털 사이트인 네이버, 다음, 야후, 네이트 등에서는 자체 질의 처리 기술을 확보하고 있다. 네이버는 해외 검색 시장 진출과 검색 기술 역량 강화를 위해서 검색 기술 인원을 다수 확보하고 있는 첫눈을 인수했다.(K모바일, 2006, 6월)
- 음성 처리 분야에서는 ETRI에서 신성장 동력산업 분야에 공통 핵심적으로 요구되는 음성인터페이스 기술개발을 위하여 KT, 베스티안파트너스 등 산업체와 ICU, 서울대 등 대학과의 산·학·연 공동연구를 수행하고 있으며, 25만 개 이상의 POI를 처리하는 대용량 음성인식, 지식처리와 연계한 대화형 음성인식, BCN기반의 분산처리 음성인식 기술을 중점적으로 개발하고 있다.
- 국내 특허출원 현황 및 전망
 - 온톨로지와 관련된 국내 특허는 온톨로지를 이용한 응용 기술에 관한 특허가 15건 등록되어 있다.(전세계 특허 정보 검색 시스템인 wips(www2.wips.co.kr) 시스템에서 (온톨로지).TI. 로 검색한 결과)
 - 언어자원 관리와 관련된 국내 특허는 전문용어 사전 분석 방법 기술에 관한 특허가 1건 등록되어 있다.(전세계 특허 정보 검색 시스템인 wips(www2.wips.co.kr) 시스템에서 ((전문 and 용어) or terminology or jargon or (technical and term)).TI. 로 검색한 결과)
 - 질의 처리와 관련된 국내 특허로는 SQL 질의 분석, SQL 질의 관리, 웹상에서 SQL 처리, XML 질의 처리, XML 질의 기반 검색, 스트리밍 환경에서 XML 질의 처리 등이 있으며, 현재 20건 내외의 특허가 등록되어 있다.
 - 음성 처리와 관련된 국내 특허로는 음성인식, 음성합성, 음성응용시스템, 대화/질의 등 관련하여 LG전자가 96건, ETRI가 75건, 삼성전자가 64건, KT가 22건을 출원하여 주요 혁신 리더로 활동하고 있다.(2006.7. 음성인터페이스 특허 권리분석 중간결과 보고서, 한국특허정보원)

2.3.2. 국외 기술개발 현황 및 전망

- 주요 국가의 정책기조
 - 국제적으로 시맨틱 기술 분야는 아직 정부정책 차원에서 다루어지고 있지는 않지만, R&D 분야에는 정책적 지원을 아끼지 않고 있다.
 - 미국은 시맨틱 웹 분야에 대한 정책적 지원은 정부 차원에서 보안, 정보 및 국방 차원에서 이루어지고 있다. 공개적으로는 DARPA를 통해 주도되고 있으며, DARPA에서는 시맨틱 웹 기술개발을 위해서 7천만 달러를 지원하여, 시맨틱 웹 표현 언어인 DAML(DARPA Agent Markup Language)을 개발 하였다.([http : //wonderweb.semanticweb.org](http://wonderweb.semanticweb.org))
- 시맨틱 웹의 경우 현재 국제적 벤더들을 중심으로 시제품 제작을 하고 있긴 하지만, 완전한 상품화를 목표로 하기보다는 향후 제품을 위한 준비로 연구와 개발을 진행하는 수준이다.
- OWL 기반의 상황인지 온톨로지인 SOUPA (Standard Ontology for Ubiquitous and Pervasive Ontology), CoBra (Context Broker Agent Ontology) 등이 개발되어 프로토타입 수준의 지능화된 개인화 서비스로의 활용 연구가 활발히 진행되고 있다. 또한 코아 온톨로지로서 FOAF (Friend of A Friend), DAML-T (Darpa Agent Markup Language-Time)이 개발되어 활발히 이용되고 있다.

이름	연구기관	개요
SOUPA	UMBC	- 유비쿼터스 환경에서의 응용을 위해 개발된 OWL 온톨로지 - 코아 온톨로지를 기본으로 특정 응용을 위한 확장 온톨로지 구성됨
CoBrA	UMBC	- SOUPA의 온톨로지 상에서의 확장 - Privacy 표현도 포함
Cyc	Cyc Corp.	- 인간의 상식을 논리적 체계위에 표현

- 미국 DARPA를 중심으로 산·학 공동 연구가 활발히 진행되고 있으며, 미국의 우수 대학 및 기업이 참여하여 시맨틱 웹과 온톨로지 인스턴스 생성 기술을 개발하고 있다. 또한 EU에서는 Framework Programme 6의 지원을 받아 시맨틱 웹과 관련하여 IST에서 3개의 프로젝트에 3,550만 유로를 투자하여 연구를 진행하고 있다.

이름	연구기관	개요
SHOE	Univ. of Maryland	- 웹 페이지 저자들이 기계 가독형 의미를 웹 문서에 부착할 수 있도록 HTML을 확장하고자 했던 연구
SemTag	IBM	- 대용량의 코퍼스에 대해서 자동으로 의미 태깅을 하기 위해서 TAP 프로젝트에서 제공한 온톨로지를 활용
Onto Annotate	Ontoprise GmbH	- 온톨로지에 기반하여 반자동으로 의미를 부착하는 ontoprise 제품으로써, IE 기술을 이용하고 RDF와 RDF스키마를 사용
KIM	Sirma Group (Ontotext Lab.)	- 웹 문서를 포함한 다양한 문서들에서 의미정보를 추출하고 온톨로지 인스턴스를 자동 생성하는 시스템 - 영국의 GATE 시스템에 기반하고 있으며, 차세대 지식관리 시스템의 구축을 목적으로 하고 있음

- 학계에서는 온톨로지 분야에 스탠포드 대학에서 개발한 온톨로지 편집기가 2005년 8월 기준 Protege-2000 Version 3.1.1까지 릴리즈되었으며, 이는 온톨로지 저작 및 온톨로지 언어간 변환 작업을 기능을 지원하고 있다. 이외에도 CMU, MIT, UGA, UMBC 등을 중심으로 시맨틱 웹서비스 기술 연구가 활발히 진행되고

있으며, 시맨틱 웹 규칙 및 추론 기술에 대해서도 RuleML.org를 중심으로 SWRL이라는 시맨틱 웹 규칙 언어 개발에 많은 연구가 진행되고 있다.

- 미국과 유럽 중심으로 서술논리 기반의 실시간 추론 처리를 위한 연구와 DIG 중심으로 OWL 온톨로지의 실시간 추론 인터페이스를 정의, 활용하는 연구가 활발히 진행되고 있다. 또한 서술논리 추론을 실용화하기 위하여 RacerPro, Pellet, FACT++, KAON2와 같은 실용화 수준의 온톨로지 추론기에 대한 연구도 활발히 진행 중이다. F-logic 기반의 효율적인 온톨로지 추론이 가능한 OntoBroker는 독일의 Ontoprise 회사에 의해서 개발되어 실용적 응용 프로그램에 활발히 이용되고 있다.

이름	연구기관	개요
RacerPro	Franz, Racer	- 서술논리 기반 상용화 온톨로지 추론
FACT, FACT++	Univ. of Manchester, UK	- Tableaux 방식 기반 온톨로지 추론
Pellet	Univ. of Maryland	- Java 기반 온톨로지 추론
KAON2	Ontoprise	- 서술논리 기반 온톨로지 추론
Java Theorem Prover	Stanford Univ.	- FOPC기반 하이브리드 온톨로지 추론
OntoBroker	Ontoprise	- F-logic 기반의 효율적인 온톨로지 추론

- 언어 자원은 미국, 유럽, 일본 등에서 WordNet, EuroWordNet, CYC, EDR 등의 다수의 언어 자원이 만들어지고 있다. 이들 언어 자원은 범용으로 사용할 수 있도록 유/무상으로 공개되어 있으며, 이들을 이용한 언어 처리 기술개발이 활발히 진행되고 있다.
- 데이터베이스 검색을 위한 질의 언어로 SQL, SQL-MM (멀티미디어 데이터베이스 지원), XQuery (XML 데이터베이스 지원)가 정의되어 있으며, 시맨틱 웹 관련 질의 언어로 DQL, SparQL이 정의되어 있다. IBM에서는 '마블'이라는 코드명으로 MPEG-7 기술에 기반한 동영상/오디오 검색을 위한 차세대 검색엔진을 질의 처리 기술을 개발하고 있으며, 검색 사이트 구글(google)에서는 앞으로 모바일 검색, 개인화된 검색 등을 위한 질의 처리 기술을 개발할 것으로 전망된다.(아이뉴스24, 2006년 7월)
- 음성 처리 분야는 선진국에서 국가 주도로 기술 연구개발이 수행되어 왔다.
 - DARPA project(미국, 1992~1999년)에서 항공편 문의시스템, 전화망 대화 연속음성인식, 방송뉴스 인식, 지능형 대화 인터페이스 등의 기술을 개발
 - Verbmobil project(독일, 1993~2000년)에서 휴대형 자동 통역기를 개발
 - ATR(일본) : 음성번역통신연구소(ATR-SLT)를 설립, 음성 분야 연구 주도
 - European 6th Frame 프로젝트에서 최우선순위를 두어 휴먼인터페이스 음성인식 및 자동통역 기술 분야를 지원

- 주요 국가별 특허출원 동향

- 온톨로지와 관련된 특허는 온톨로지 생성, 자료변환, 온톨로지를 이용한 응용시스템 등의 특허가 등록되어 있으며, 미국에서는 60건, 일본에서는 5건, 유럽에서는 9건 정도의 특허가 등록되어 있다.(전세계 특허 정보 검색 시스템인 wips(www2.wips.co.kr) 시스템에서 (ontology or (ontology and inference)).TI. 로 검색한 결과)
- 전문용어와 관련된 특허는 전문용어 인식 기술, 전문용어 관리 기술, 전문용어 번역 기술 등의 특허가 등록되어 있으며, 미국에서는 20건, 일본에서는 5건, 유럽에서는 10건 정도의 특허가 등록되어 있다.(전세계 특허 정보 검색 시스템인 wips(www2.wips.co.kr) 시스템에서 (jargon or terminology or (technical and term)).TI. 로 검색한 결과)
- 질의 처리와 관련된 특허는 SQL 처리 분야, XML 질의 처리 분야, 검색 질의 처리로 구분할 수 있다. SQL 처리와 관련된 특허는 1995년부터 2006년까지 꾸준히 등록되고 있으며 현재 수백 개의 특허가 등록되어 있다. XML 질의 처리와 관련된 특허는 2003년부터 등록되어 현재는 30여건의 특허가 등록되어 있으며, 검색 질의 처리와 관련된 특허도 다수 등록되어 있다.
- 음성 처리와 관련된 특허로는 미국의 경우 IBM이 84건으로 가장 많은 등록특허를 보유하고 있으며, Matsushita가 42건, Microsoft가 33건, NEC가 32건을 특허 등록하여 주요 리더로서 활동하고 있으며, IBM, Microsoft, AT&T와 같은 미국 기업들 외에 일본 기업이 주요 출원인으로서 많은 수의 특허를 등록하여 활발한 특허 활동을 벌이고 있다.(2006.7. 음성인터페이스 특허 권리분석 중간결과 보고서, 한국특허정보원)

2.4. 표준화 현황 및 전망

2.4.1. 국내 표준화 현황 및 전망

- 웹코리아 포럼은 시맨틱 웹 분과를 설치하여 해당 분과를 중심으로 국내 관련 학계, 연구소, 산업계 등과 공동으로 웹온톨로지 언어, 시맨틱 웹 서비스, 시맨틱 웹 레지스트리 등의 시맨틱 웹 관련 최신 표준 동향을 연구하고, 장기적으로 시맨틱 웹 관련 신규 표준을 제안하기 위한 활동을 진행하고 있으며 '시맨틱 웹과 웹 서비스' 등을 주제로 하는 워크숍을 계획하고 있다.
- DERI Korea에서는 Ontology Working Group, Rule Language and Inference Working Group, Semantic Web Working Group을 통해서, 온톨로지 기반 추론, 온톨로지 도구 등의 개발과 관련된 온톨로지 공학과 시맨틱 웹 포털, 시맨틱 그리드 시스템 등의 개발과 관련된 시맨틱 웹 기술에 관한 연구를 진행하고 있다.
- 웹코리아 포럼, W3C 한국 사무국(ETRI) 등을 중심으로 TTA를 통해 시맨틱 웹 관련 기본규격들에 대한 국내표준(안) 개발을 수행하고 있고, RDF 관련 표준은 2002년 말 제정되었으며, 현재 10여건이 국내 단체표

준(안)으로 과제 채택 및 심의 중에 있다. 특기할 만한 사항은 XRM과 시맨틱 웹 추론엔진 관련 표준(안)이 국내에서 선도적으로 추진되고 있다.

- 한국정보과학회의 언어공학연구회에서는 2003년부터 국내의 온톨로지 관련 연구와 개발을 선도하기 위한 워크샵으로 KIPONTO 워크샵을 후원하고 있다.
- 한국어 품사 태그 세트 및 표기법 : 한국전자통신연구원을 중심으로 전산 언어학자, 국문학자들로 이루어진 관련 연구자들이 모여 1998년부터 1999년까지 2년 동안 품사 태그 세트 선정 표준화작업을 한 바 있다.
- 전자사전의 종류, 종류별 구조 및 내용 표기 방법(전문용어 포함) : 문화관광부와 국립국어연구원에서 추진 중인 '세종계획'에서 영역별 전문용어에 대한 표제어 선정 기준 및 수록할 정보의 종류와 표현 양식의 표준안이 추진되고 있고, 또한 KORTERM에서 전문용어와 기타 언어 자원에 대한 표준화 그룹인 ISO/TC37 분과에 한국어 전문용어에 대한 표준화가 추진되고 있다.
- 한국기술표준원에서 전문용어 및 기타 언어자원 국제표준화 회의(ISO/TC37)에서 전문용어 국제 규격에 한국어의 특성을 적극 반영하기 위해서 한국 대표단을 파견하고 있다.
- 한국전자통신연구원, 삼성종합기술원, LG, ICU 등에서 MPEG의 멀티미디어 내용 표현을 위한 메타데이터 표준인 MPEG-7에 대한 표준화작업에 활발히 참여하고 있다. 현재 표준화를 추진 중인 MPEG-7 Query Format 표준화작업에도 참여할 것으로 기대된다.
- 음성정보처리기술 포럼에서는 2003년 PDA 공통음성명령어 TTA 단체 표준화를 완료하였다.
- 텔레매틱스 분야에서 자동차 기기 제어와 네비게이션 명령어에 대한 기본적인 표준을 KATS가 2005년 3월 단체 표준으로 승인하였으며, AV 기기제어 등에 대한 추가 표준화작업이 추진 중이다.

2.4.2. 국외 표준화 현황 및 전망

- W3C에서의 Semantic Web 관련 표준화현황
 - W3C에서는 Semantic Web과 관련된 7개의 current working group이 표준과 관련된 작업을 진행하고 있으며, 표준제정과 관련된 업무가 종료된 3개의 past working group과 기타 타 분야(WebService 등)와 직접적으로 관련된 working group들이 존재하고 있다.

- Current Working Group

Working Group	주요 내용
Semantic Web Coordination Group	- Semantic web activity에서 이뤄지는 많은 working group 들의 일들 중에서 서로 중복되거나 하는 일이 없도록 전체적인 표준과 기술들을 중재하는 역할
RDF Data Access Working Group	- Web publishing, 개인 정보 관리 등 시맨틱 웹 응용에 유용한 RDF query language and data access protocol 정의 "SPARQL(SPARQL Protocol And RDF Query Language) Query Language for RDF" Candidate Recommendation, 6 Apr 2006 "SPARQL Protocol for RDF" Candidate Recommendation, 6 Apr 2006 "SPARQL Query Result XML Format" Candidate Recommendation, 6 Apr 2006
GRDDL(Gleaning Resource Descriptions from Dialects of Languages) Working Group	- XSLT를 이용하여 XML과 XHTML 문서에서 RDF 데이터를 추출하는 방식 GRDDL, W3C Team Submission 16 May 2005
Rule Interchange Format Working Group	- 다양한 시스템에서 사용될 수 있는 rule에 대한 format 정의, rule또는 language - OWL Web ontology language를 응용에 맞게 확장 또는 보충 - Major Milestones First Public Working Draft, Use Cases and Requirements (2006 February) First Public Working Draft, Technical Specification (2006 May) Last Call Working Draft, Technical Specification(2006 October) Recommendation (2007 May)
Semantic Web Deployment Working Group	- 시맨틱 웹 활용에 관한 연구를 진행하는 그룹으로, RDF 및 OWL 활용 기술 및 활용 방법, RDF와 HTML 간의 통합 방법에 관해서 연구를 진행함 - 기계가 이해할 수 있는 형태의 지식 구성 프레임워크로 RDF에 기반한 SKOS (Simple Knowledge Organization System) 정의 - 문서의 구조 정보를 표현하는 XHTML에 의미 정보를 표현할 수 있도록 RDF 구문을 사용하는 RDFa 정의 - RDFS vocabulary와 OWL vocabulary를 공표(publication)하는 방법 정의
Semantic Web Interest Group	- RDF, SparQL 활용에 관한 연구를 진행하는 그룹
Health Care and Life Sciences Interest Group	- 생물학과 의학 분야에서 발표된 논문, 소프트웨어, 데이터 등의 자료들을 통합 및 관리에 시맨틱 웹 기술 적용을 위한 연구를 진행하는 그룹 - 데이터 통합 및 협력 작업을 위한 온톨로지, Vocabulary 정의 - RDF와 OWL을 이용한 생명 과학 분야의 데모 시스템 구현을 위한 BioRDF 팀 운영 - 논문, 발표 자료, 강의 자료 등의 효과적인 관리, 공유, 검색 모델 개발

- Past Working Group

Working Group	주요 내용
Semantic Web Best Practices and Deployment(SWBPD) Working Group	- 시맨틱 웹 응용 개발자들에게 실제적인 도움이 되는 방법론들을 제공. Ontology/vocabulary repositories, 개발 방법론 등 교육적인 자료와 데모 시스템들을 제공 - "Best Practice Recipes for Publishing RDF Vocabularies", Working Draft 14 March 2006 - "Image Annotation on the Semantic Web", Working Draft 22 March 2006 - "RDF/OWL Representation of WordNet", Working Draft 19 June 2006 - "Quick Guide to Publishing a Thesaurus on the Semantic Web", Working Draft 17 May 2005
RDFCore Working Group (Closed)	- 제정 완료 : RDF, Recommendation 10 Feb 2004
Web-Ontology(WebOnt) Working Group (Closed)	- 제정 완료 : OWL Web Ontology Language, Recommendation 10 Feb 2004

- Nearby Working Group

Working Group	주요 내용
Semantic Annotations for WSDL Working Group	- WSDL 2.0 확장을 통해서 웹서비스에서 의미 정보를 기술 방법을 연구하는 그룹
Semantic Web Services Interest Group	- 시맨틱 웹과 웹서비스를 통합하는 방법에 관해서 연구하는 그룹 - OWL-S, WSMO, SWSF(Semantic Web Services Framework), WSDL-S (Web Service Semantics) 등이 제안되어 있음
Semantic Web Tools	- 그룹을 소개한 것은 아니고, Semantic Web과 관련된 툴들을 나열한다 - Altova의 SemanticWorks와 같은 편집도구, Aduna Metadata Server와 같은 triple store 시스템, Jena와 같은 RDF API 등이 제시되어 있음
Semantic Web Books	- 그룹을 소개한 것은 아니고, Semantic Web과 관련된 책들을 나열한다 - 2000년 이래로 현재까지 꾸준히 Semantic Web 관련 책들이 출판되고 있음
Semantic Web Wiki Front Page	- 그룹을 소개한 것은 아니고, Semantic Web Wiki를 제시함

- 최근 ISO TC37에서 온톨로지 분과를 만들어 W3C와 달리 온톨로지 콘텐츠 구축 및 관리 방법론에 대한 표준화추진을 계획하고 있다.

• 온톨로지 표현언어 및 추론 분야의 표준화

- 유럽과 미국에서는 특정 도메인을 위한 표준 온톨로지 정의 작업을 진행하고 있다. 유럽에서는 SEKT, DIP, Knowledge Web, ASG 등의 시맨틱 웹 관련 과제의 클러스터로 ESSI를 구성하고, ESSI를 중심으로 웹 서비스를 위한 온톨로지인 WSMO를 정의하고 있다. 미국에서는 DARPA를 중심으로 시맨틱 웹 온톨로지 표현 언어인 OWL에 기반해서 웹서비스 온톨로지인 OWL-S를 정의하고 있다.

표준	주요 내용
WSMO	- Semantic web service 언어 및 플랫폼 표준
WSMX	- 다양한 응용 및 웹서비스 환경에서의 실행 환경 모델링
WSML	- Semantic web을 위한 rule-based language
OWL-S	- OWL-based Web service ontology

- 미국과 유럽 중심의 W3C에서는 온톨로지 표현을 위한 표준 언어인 OWL(Web Ontology Language)가 제정되었고, 온톨로지의 규칙 표현을 위한 SWRL, 온톨로지 기반 Query 처리를 위한 SparQL과 같은 다양한 표준을 제안하고 있다.

표준	주요 내용
OWL/RDFS/RDF	- 서술논리를 지원하는 온톨로지 표현 언어 - 개념, 특성, Cardinality 기반의 모델 표현 방식
SWRL	- 온톨로지 기반 규칙 표현 언어 - 서술논리에 Horn 논리의 기능을 포함한 규칙 표현 언어
SparQL	- 온톨로지 기반 Query 처리 언어

- 시맨틱 웹에 기반한 표준 온톨로지에는 DC 온톨로지와 FOAF 온톨로지가 있다. DC는 dublin core initiative에서 데이터 기술을 위한 온톨로지로서 정의하였으며, FOAF는 FOAF 프로젝트에 의해서 개인 정보를 표현하는 온톨로지로서 정의되어 있다. DC와 FOAF는 현재 산업 표준처럼 다양한 분야에서 활용되고 있으며, 특히 DC는 RSS에서 활용되고 있다.
- 디바이스 온톨로지
 - 유비쿼터스 환경에서 디바이스 온톨로지에 대한 표준 활동을 시작하고 있으며, W3C와 FIPA에서는 모바일 환경에 적합한 온톨로지를 구축하는 단계이다.

표준	주요내용
W3C Device Independence	<ul style="list-style-type: none"> - 디바이스의 소프트웨어와 하드웨어의 프로파일을 기술하기 위한 RDF 기반의 프레임워크 - 웹 콘텐츠를 클라이언트 디바이스에 어떻게 적합하게 전달할 것인가에 관한 서버의 의사결정 과정을 용이하게 하기 위함
FIPA Device Ontology	<ul style="list-style-type: none"> - 디바이스를 기술하기 위한 프레임 기반의 구조 - 에이전트 커뮤니케이션을 용이하게 함

- 온톨로지 추론을 위한 표준화 및 현황
 - 온톨로지 추론 분야는 Tableaux Algorithm이 거의 표준으로 활용되고 있다. 현재 Tableaux Algorithm은 상용 프로그램인 RacerPro와 FaCT, Pellet과 같은 온톨로지 추론 엔진의 표준이다. 이들 추론 엔진 외에 F-logic(Frame Logic)를 기반으로 Onto Prize에서 구축한 온톨로지 추론 엔진 OntoBroker가 있다.
 - 최근 W3C에서는 기존의 시맨틱 웹 표준 구조에 규칙(Rule)과 질의 부분을 추가한 새로운 표준 구조를 발표하였다. 현재 온톨로지 추론을 위해 RuleML, SWRL 등이 시맨틱 규칙을 표현하는 표준으로서 제안되고 있다.
 - Rule Markup Initiative에 의해 정의된 RuleML는 XML을 이용하여 정방향 및 역방향 규칙들을 표현할 수 있다. RuleML과 관련하여 Object-oriented RuleML, Semantic Web Rule Language, RuleML Lite, Positional-Slotted Language 등과 같은 RuleML의 확장 또는 변형 언어들이 응용 분야별로 제시되었으며, RDF, 귀납적 추론(Induction), 번역기(Translator), 추론 엔진(Engine), 규칙 편집기 등 사용자 인터페이스 등에 관한 연구 및 개발이 이루어지고 있다.
- SWRL의 표준화 및 현황
 - SWRL는 OWL의 하부 언어인 OWL DL 및 OWL Lite와 RuleML의 하부 언어인 Unary/Binary Datalog RuleML을 통합한 언어로, SWRL은 DARPA DAML Program에서 발표된 2003년 10월 중순에 처음 소개된 이후 현재 버전 0.7까지 제안된 상태이다.
- 언어 자원 표준화
 - ISO(International Organization for Standardization) TC(Terminology and other language resources)37에서 전세계 25개 국의 참여로 전문용어와 언어 리소스의 원리, 방법, 응용에 대한 표준화를 추진하고 있다.

Working Group	주요 내용
ISO/TC 37/SC 4	<ul style="list-style-type: none"> - 언어자원에 사용되는 용어의 표준화 및 언어학적 온톨로지 인스턴스 생성과 표상을 위한 기본 메커니즘과 데이터 구조의 표준화 - 언어학적 정보 구조 및 질차를 문서화하기 위한 메타 데이터 표상 스키마의 표준화 - 수행 방식에 상관없이 언어학적 정보의 흐름과 연관된 형태 통사적 정보를 표현하기 위한 방법들을 제공 - 언어자원 타당성 검증 및 망 기반 분산 협동 작업을 위한 지침의 표준화

- 국외에서는 언어 자원 표준화와 관련해서 TEI(Text Encoding Initiative), EAGLES, LDC, GENELEX, GRAAL 등의 대규모 프로젝트의 결과물을 표준화하려고 한다.

• 질의 포맷 표준화

- 멀티미디어 데이터베이스 검색 포맷인 SQL-MM은 ISO 표준이며, XML 데이터베이스 검색 포맷인 XQuery은 W3C에서 표준화 진행 중이다. 시맨틱 웹 질의 포맷인 SparQL은 W3C에서 표준으로 권고되고 있다.
- MPEG에서는 멀티미디어 내용 표현을 위한 표준으로 MPEG-7을 정의하고, MPEG-7에 기반한 멀티미디어 검색을 위해서 MPEG-7 Query Format을 정의하고 있다. JPEG에서는 이미지 검색을 위해서 JPSearch 질의 포맷을 정의하고 있다.

• MPEG-7에서는 멀티미디어의 메타데이터 표준화작업을 진행하고 있으며, 멀티미디어 내용정보 표현을 수용하는 표준을 진행하고 있다.

• Dublin Core Metadata Initiative는 지능형 검색 시스템을 위해 자원 표현 메타데이터와 메타데이터 언어(vocabulary)를 포함한 표준을 진행하고 있다.

• W3C에서는 음성 합성, 디지털 오디오, 음성인식, DTMF 입력 등을 사용하여 컴퓨터와 사람 간의 대화가 가능하도록 하는 마크업 언어(Markup Language)인 VXML(Voice Extensible Markup Language) 표준화를 추진하고 있다. 미국의 전화사업자를 중심으로 VoiceXML Forum을 결성하였고 W3C에서 이를 인정하여 2000년 초 표준 XML의 한 형태로 VXML 1.0 표준을 확정한 후 현재 상호 주도형 대화 처리가 가능한 VXML 2.0 표준을 발표하였다. VoiceXML 기반의 음성인터페이스는 현재 단어와 단문을 인식할 수 있는 단계에 있으나 향후 자연스러운 대화체 인식 및 합성을 지원하는 형상으로 표준화가 이루어질 것으로 전망된다.

• W3C는 VXML표준과 아울러 컴퓨터가 단어를 어떻게 말해야 할지 다루는 방법에 대한 SSML (Speech Synthesis Markup Language) 표준, 사람이 컴퓨터에 말하는 문장에 대한 문법을 기술하는 방법에 대한 SRGS(Speech Recognition Grammar Specification) 표준, 문서를 축약하고 인간 언어를 컴퓨터가 이해할 수 있도록 바꾸는 기술인 시맨틱 인터프리테이션 (Semantic Interpretation for Speech Recognition) 표준 등도 개발하고 있다.

- 마이크로소프트가 주도하는 SALT(Speech Application Language Tags) 포럼에서는 음성인터페이스를 포함하는 멀티모달 인터페이스 표준화를 추진하고 있다. 웹 애플리케이션에 음성인터페이스용 태그를 첨가할 수 있어 마우스나 키보드 사용 외에도 음성명령으로 소프트웨어를 제어할 수 있도록 한다.
- 분산 음성인식(Distributed Speech Recognition) 시스템의 핵심기술인 특징벡터 추출 알고리즘의 표준화를 위해 ETSI(European Telecommunications Standards Institute)에 의해 AURORA 프로젝트가 진행 중이다.

2.5. 표준화 대상항목별 현황 분석표

표준화 대상항목		온톨로지	시맨틱 웹	지능형 언어/음성처리
시장 현황 및 전망	국내	- R&D시장 관련 규모는 꾸준히 증가하고 있으며, 시맨틱 웹 기술 자체 시장보다는 관련 응용 시장의 부분 기술로서 시장 형성 전망 - 국내 시장도 국외의 시장 현황과 전망에 비추어, 거의 같은 비율로 성장하여 2010년 1조 시장 예측		- 언어자원은 자체 시장보다는 기술개발의 기반으로 이용되어, 한국 어정보처리 관련 자동번역, 텍스트 마이닝 분야의 시장 형성 - 국내 음성처리시장은 초기 단계로 기술력 확보와 함께 급성장할 것으로 전망
	국외	- adobe, AT&T, Google 등 65개 이상의 기업에서 시맨틱 기술 관련 제품을 개발하고 있으며, 온톨로지를 활용하는 시맨틱 시장의 성장속도는 2006년부터 2010년 사이 10배의 성장이 전망된다. - 연평균 55~75%의 연평균 성장을 통해 2010년 52억 달러, 2015년에는 5,000억 달러의 시맨틱 기술 시장 형성 전망		- 언어 자원을 활용하는 자동번역, 텍스트 마이닝 분야 시장 형성 뿐만 아니라, 구축된 언어 자원 판매 시장이 형성되어 있음 - 국외 음성처리시장은 CT기반 콜센터 응용을 중심으로 2008년 50억 달러 규모 예상 (출처 : ABI사)
기술 개발 현황 및 전망	국내	- KAIST, ETRI 등 산·학·연 중심으로 한 국가 IT 온톨로지 구축 사업을 통해 시맨틱 웹 상에서 실제 서비스가 가능한 IT 코어 온톨로지, 온톨로지 추론 엔진, 시맨틱 웹 응용서비스 구축 등의 인프라 기술개발에 착수함	- ETRI, KT, 솔트룩스, 송실대 등 연구소, 대학, 일부 기업을 중심으로 시맨틱 웹 기반 기술 개발 및 프로토타입 시스템 개발을 진행 중이다. - ETRI 중심으로 비즈니스 응용 분야에 적용할 수 있는 온톨로지 병합 도구 및 추론 엔진 개발이 진행되었음	- KAIST KORTERM, ETRI 등의 기관에서 세중계획 과제, 지식정보자원관리사업 등을 통해 전문용어 언어자원, 한국어 표준 언어/음성자원을 구축하여 관련 신성장 동력산업 기술개발에 공동 활용하고 있다. - 산업자원부 기술표준원 주관으로 한국전문용어 전문위원회가 구성되어, 한국어 전문용어 표준화 포럼을 구성 중에 있음 - 신성장 동력산업용 음성인터페이스기술을 정보통신부 선도기술개발과제로 ETRI주관으로 KT, ICU, 서울대 등 산·학·연 공동으로 개발하고 있음
	국외	- 스탠포드 대학에서 Protege 온톨로지 편집기 개발, 배포 - IBM, HP 등의 기업이 OntoWeb, Ontoprize 등 온톨로지 기반의 지식관리 및 전자상거래를 위한 기술개발	- 미국은 DARPA를 통해 시맨틱 웹 기술개발에 7천만 달러를 지원하여, 시맨틱 웹 표현언어 DAML 개발 - EU에서는 Framework Programme 6의 시맨틱 웹 관련 3개 프로젝트에 3,550만 유로 투자	- ISO/TC 37에 따른 전문용어 자원 및 컴퓨터 응용 표준화가 추진되고 있음 - 미국, 유럽, 일본 등에서 WordNet, EuroWordNet, CYC, EDR 등의 언어 자원을 구축함 - 미국의 경우, 정부 주도 음성처리 기술개발과제를 지속적으로 추진, 상용화 진입에 성공하였으며, 세계 기술개발을 주도하고 있음
기술 개발 수준	국내	설계	설계	설계
	국외	시제품	시제품	시제품
	기술격차	2년	2년	1년
	관련 제품	SHOE, SemTag, OntoAnnotage, KIM 등	Pellet, FaCT, Racer 등	WordNet, CYC, MARVEL 등
IPR 보유현황	국내	특허 확보 중	ETRI, 송실대, 솔트룩스, 한양대	KAIST, ETRI 등
	국외	HP, IBM 등	ontoprize, Racer 등	IBM
IPR확보 가능분야		없음	SVC	-
IPR확보 가능성		도메인 온톨로지	시맨틱 웹 추론 엔진	한국어 전문용어. 멀티미디어 질의처리. 공통 음성명령어
표준화현황 및 전망		W3C 온톨로지 표현언어 제정	RDF 및 OWL의 표준제정에 따른 후속조치로 질의 언어 및 추론 분야에 대한 지속적인 진행 전망	- ISO/TC37의 주도하에 언어별 국제표준 추진 - MPEG에서 질의 처리에 대한 표준화추진 중
표준화 기구/ 단체	국내	TTA	TTA	TTA, KATS, MPEG
	국외	W3C, ISO/TC37	W3C	ISO/TC37, MPEG
	국내 참여 업체 및 기관현황	KAIST, ETRI 등	시맨틱기술협회, 웹코리아 포럼, W3C 한국사무국 (ETRI)	KOTERM, ETRI, 삼성, 건국대
	국내 기여도	높음	높음	높음
표준화 수준	국내	표준안 항목 승인	표준안 항목 승인	표준안 항목 승인
	국외	표준안 개발/검토	표준안 개발/검토	표준안 개발/검토
국내표준화의 인프라수준 (시정요구정도 및 참여도)		높음	높음	높음

3. 중점 표준화 대상항목의 표준화 추진전략

3.1. 중점기술의 표준화 환경분석

3.1.1. 표준화 추진상의 문제점 및 현안사항

- 차세대 웹 인프라 구축을 위한 국가의 정책적 지원과 시맨틱 기술 분야의 급속한 시장 성장이 예상되나 아직 국내 인적, 기술적 역량은 미약한 실정이다.
- 지능정보처리를 위해 필수적인 언어 자원의 대규모 구축을 위해 언어 자원의 표현 및 관리에 관한 표준이 선결과제이나 국내외적으로 표준화 진행이 더디게 이루어지고 있다.
- u-지능정보 기술 표준화를 위해 정부, 산업체, 대학, 연구소 등의 관련자 및 전문가들이 끌고루 참여하여 각 분야의 입장을 반영하도록 노력하여야 하되, 최우선 목표는 순수한 학술적 연구들을 목표로 한 것이 아니라, 실용적인 산업화에 역점을 둘 필요가 있다.

3.1.2. SWOT 분석 및 표준화 추진방향

		국내 역량요인	강점요인 (S)		약점요인 (W)	
			시장	- 최고 통신 인프라 기반 차세대 웹 및 시맨틱 기술 시장 확산	시장	- 내수시장 규모 협소 및 시장 형성 한계
국외 환경요인			기술	- IT839 정책 및 신성장 동력산업 기술개발 국가적 추진	기술	- u-지능정보 핵심원천 기술 취약 및 해외 기술 의존
			표준	- ISO, MPEG 등 국제표준기구, 단체의 표준화활동에 조기 참여 및 대응	표준	- 산업계의 표준화 기반 기술 및 표준 전문인력 확보 미흡
기획요인 (O)	시장	- 차세대 웹 응용 확산 추세 및 시맨틱 기술 시장 급속 확대	<div><div>SO</div><div>WO</div><div>ST</div><div>WT</div><div>전략</div></div> <div>- 현황분석에 의한 우선순위 : 1</div> <div>- 차세대 웹 및 시맨틱 기술 적용을 위한 온톨로지 및 시맨틱 웹 분야 표준 개발 추진</div> <div>- IT839 국가 R&D과제 연계 Killer App. 시범사업을 통한 초기 시장 발굴 및 활성화</div> <div><SO전략 : 공격적 전략(강점사용-기회활용)></div> <div><ST전략 : 다각화 전략(강점사용-위협회피)></div>		<div>- 현황분석에 의한 우선순위 : 2</div> <div>- 유비쿼터스 환경의 지능형 사용자 인터페이스를 위한 지능적 언어/음성처리 원천 기술 확보 및 표준 개발 추진</div> <div><WO전략 : 만회 전략(약점극복-기회활용)></div> <div><WT전략 : 방어적 전략(약점최소화-위협회피)></div>	
	기술	- 시맨틱 기술 세계 선도 기술과 격차 미미				
	표준	- 세계적으로 표준화 초기 상태로 표준 추진 활발				
위협요인 (T)	시장	- 내수시장 규모 미비로 수출경쟁력 기반취약과 외국 선도 기술의 국내 초기 시장 공략	<div>- 현황분석에 의한 우선순위 : 3</div> <div>- 국내 통신 인프라를 활용하여 차세대 웹 유비쿼터스 분야 표준 개발 추진</div> <div>- 차세대 웹 분야 틈새 핵심IPR 확보를 통한 초기 시장 선점 강화</div>		<div>- 현황분석에 의한 우선순위 : 4</div> <div>- 해외 표준화단체, 기구의 적극적 참여를 통한 u-지능정보 표준 전문 인력 집중 양성</div> <div>- 표준기술개발 및 표준화 성공을 위해 전략적 국제 공동 추진</div>	
	기술	- 선도 국가, 기업의 핵심 원천 기술 IPR 대량 보유				
	표준	- IBM 등 다국적 세계 선도 기업들의 표준 기술 선점 우려				

- 현황 분석을 통한 우선순위 : SO ⇒ WO ⇒ ST ⇒ WT

- SO 전략 : 기술 표준화 분야별로 상이하지만, 뛰어난 정보통신 인프라의 장점을 발휘할 수 있는 유비쿼터스

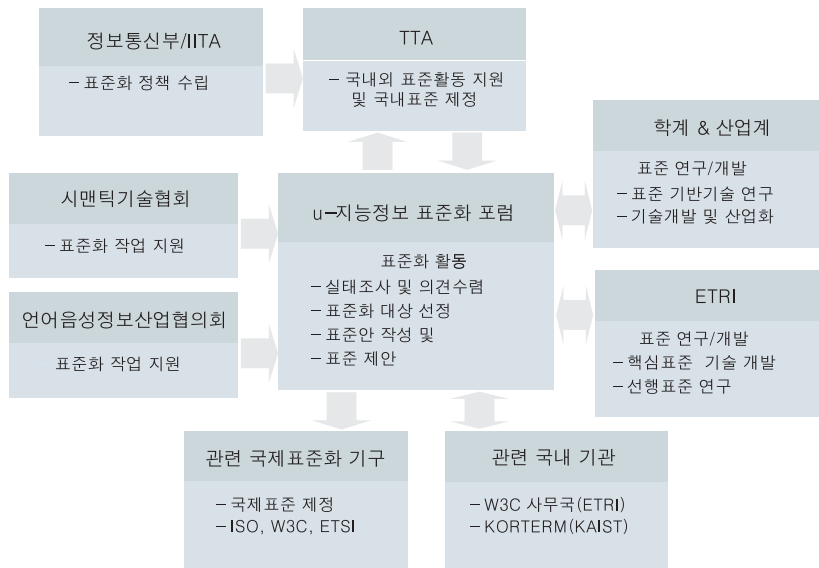
환경을 기반으로 하는 분야, 특히 그 중에서도 온톨로지와 시맨틱 웹 분야는 국제적으로 표준화 시작 단계이며, 따라서 선행 표준 기술을 수용하면서 신규 표준화활동을 추진하면 성공 가능성이 높다.

- WO 전략 : 세계적으로 표준화 초기상태이고 차세대 웹 응용의 기대와 시장이 확산 추세에 있으므로 지능형 인터페이스를 위한 질의 처리와 언어 자원 관리 분야의 표준 기술 확보를 통해 열악한 국내 산업의 돌파구를 찾도록 한다.
- ST 전략 : 시맨틱 웹 분야의 선도기술과 표준을 적용한 뚜렷한 제품이 아직까지 등장하지 않았기 때문에, 유비쿼터스 환경 인프라를 바탕으로 경쟁력 있는 시장 분석을 통해 틈새 시장의 핵심 원천 기술을 확보하고 표준화를 적극 추진하는 전략을 취해야 한다.
- WT 전략 : W3C, ISO 등 u-지능 정보 관련 국제표준화기구의 표준화활동에 적극 참여를 통한 표준 전문인력 양성과 표준화추진 역량을 배양하는 힘을 모아야 한다.

• 표준화 추진방향

- u-지능정보 기술 중 국내표준화활동 능력에 따라 표준화추진 성공 가능성 및 시급성을 고려하여 중점 표준화항목을 우선적으로 선정하여 추진한다.
- u-지능정보 기술에 대한 국내표준 정립은 국내 산업체 및 대학, 연구소가 참여하는 표준화 포럼을 구성하여 추진한다. ETRI, KAIST, 숭실대 등 u-지능정보 기술 표준화작업에 참여했던 전문 인력들을 포함하여 먼저 표준화 포럼을 구성하여 국내표준을 정립하고, 이를 바탕으로 지속적으로 국제표준화활동이 계속될 수 있도록 적극 지원한다.
- u-지능정보에 공통적으로 활용할 수 있는 지식 표현 형태인 온톨로지는 다양한 IT지식 서비스의 상호 호환성과 운용성에 중요한 역할을 하므로, 이들 온톨로지의 구축과 표현을 중심으로 표준화를 추진한다.
- 최근 W3C에서는 웹과 사람의 상호작용 및 온톨로지 추론 기능을 강화한 새로운 시맨틱 웹 표준 구조를 발표하였으며, 국내 경쟁력이 어느 정도 있는 온톨로지 기반 추론 기술을 중심으로 표준화를 추진한다.
- 언어 자원 관리의 국외 표준화는 현재 전문용어 표현 양식을 중심으로 W3C, ETSI 등의 국제표준 기술에 참여하여 표준화를 추진 중이나 향후 확대하여 적극 참여한다.
- 지능적인 자연스런 사용자 인터페이스 구현을 위해 필수적인 사용자 질의 의도 파악과 질의 의미적 분석에 필요한 공통적인 질의 포맷의 표준화를 중심으로 추진한다.

3.1.3. 표준화 추진체계



(그림 7) u-지능정보 표준화 추진체계

- 정보통신부와 정보통신연구진흥원은 소프트웨어 u-지능정보 사업 추진과 연계한 표준화 정책방향을 설정하고, 관련 연구기관, 산업체, 대학은 u-지능정보 표준화 대상 기술을 연구/개발하고, 시맨틱기술협회, 언어음성산업협의회 등 유관협회는 u-지능정보 표준화작업을 지원한다.
- u-지능정보 기술 표준화 관련 기관들이 참여하는 u-지능정보 표준화포럼을 구성하여 구체적인 표준화 대상 항목 선정과 추진 일정 세부계획을 수립하고, 위 추진체계에 나타난 바와 같이 역할을 분담하여 표준화활동을 추진한다. 작성된 표준안은 ISO, W3C, TTA 등 국내 및 국외 표준 단체를 통하여 표준을 제정한다.

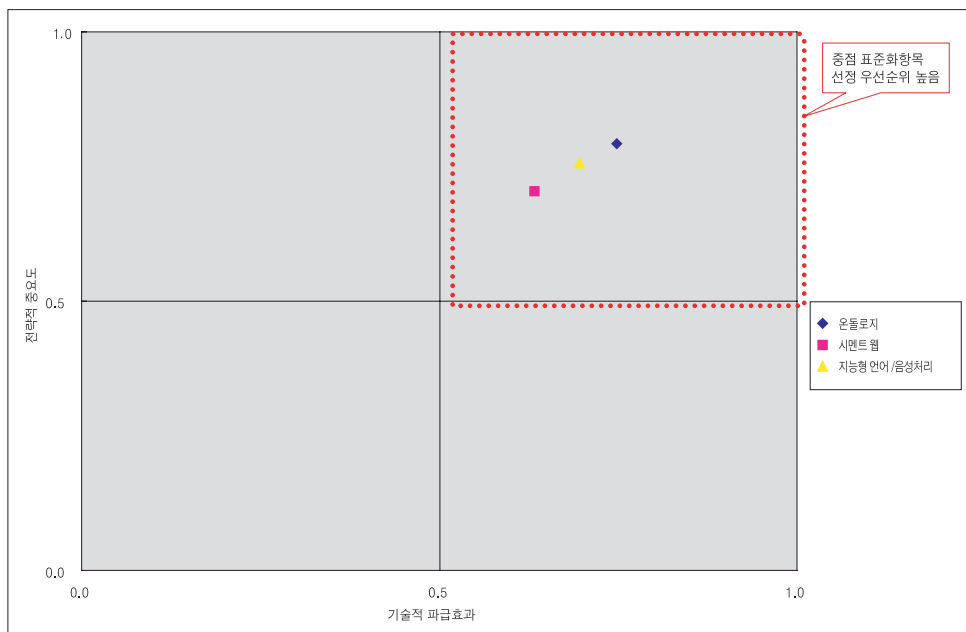
3.2. 중점 표준화 대상항목 선정

3.2.1. 중점 표준화 대상항목 선정방법

표준화 대상항목별 전략적 중요도 및 기술적 파급효과 분석								
고려요소	전략적 중요도					기술적 파급효과		
	P1 (정부의지)	P2 (산업체의지)	P3 (공공성)	P4 (적시성)	PI	E1 (기술 내 중요도)	E3 (타기술에 파급효과)	EI
고려요소별 가중치	0.21	0.17	0.09	0.11		0.28	0.26	
온톨로지	5	3	4	3	0.8	3	2	0.8
시맨틱 웹	4	3	4	2	0.6	4	4	0.7
지능형 언어/ 음성처리	5	3	4	3	0.7	4	4	0.8

* 표준화 대상항목의 각 고려요소별 평가점수는 해당 중점기술의 전문가들 의견을 종합하여 산출

* 각 고려요소별 평가점수는 1(매우 낮음), 2(낮음), 3(보통), 4(높음), 5(매우 높음)의 5점 척도



3.2.2. 중점 표준화 대상항목 선정사유

- 전략적 중요도 및 기술적 파급효과의 요소

- 지식화 사회로의 변화에 따라 지식 표현 기술 및 관련 메타 데이터 처리 방식에 대한 연구가 중심이 되는 시맨틱 웹의 표준 연구는 앞으로의 산업에서도 파급효과가 클 것으로 전망된다. 특히 시맨틱 웹 표현 언어 외에는 국제적으로도 선도 표준이 명확치 않으며, 그 수준도 초기 단계에 머물러 있다. 시맨틱 웹 규칙 표현/추론 세부 기술 분야의 표준화 및 기술 개발 필요성은 최근에 더욱더 중요시되고 있으며, 국내에서도 산·학·연을 중심으로 선도 기술개발 과제 등을 통해서 꾸준히 연구되고 있는 바 충분히 국제 경쟁력을 가질 수 있는 분야이다.
- 다양한 언어 표현을 담고 있는 언어 자원은 음성/언어 처리 분야에서 큰 걸림돌인 언어 표현의 다양성 문제를 극복할 수 있다. 언어 자원의 표준화는 향후 구축되는 언어 자원의 일관성을 제공하고, 다양한 언어 자원을 효과적으로 활용할 수 있는 토대를 제공할 것이다. 텍스트 및 멀티미디어 검색 분야는 꾸준히 각광을 받고 있는 분야로 향후에는 질의 처리 표준을 통한 발전이 기대되고 있다. MPEG 및 JPEG에서는 멀티미디어 검색을 위한 질의 처리 표준화를 2006년 현재 추진 중에 있으며, 이러한 시점에서 질의 처리에 대한 표준화 및 IPR 확보는 향후 검색 시장에서 유리한 위치를 선점할 수 있을 것이다.

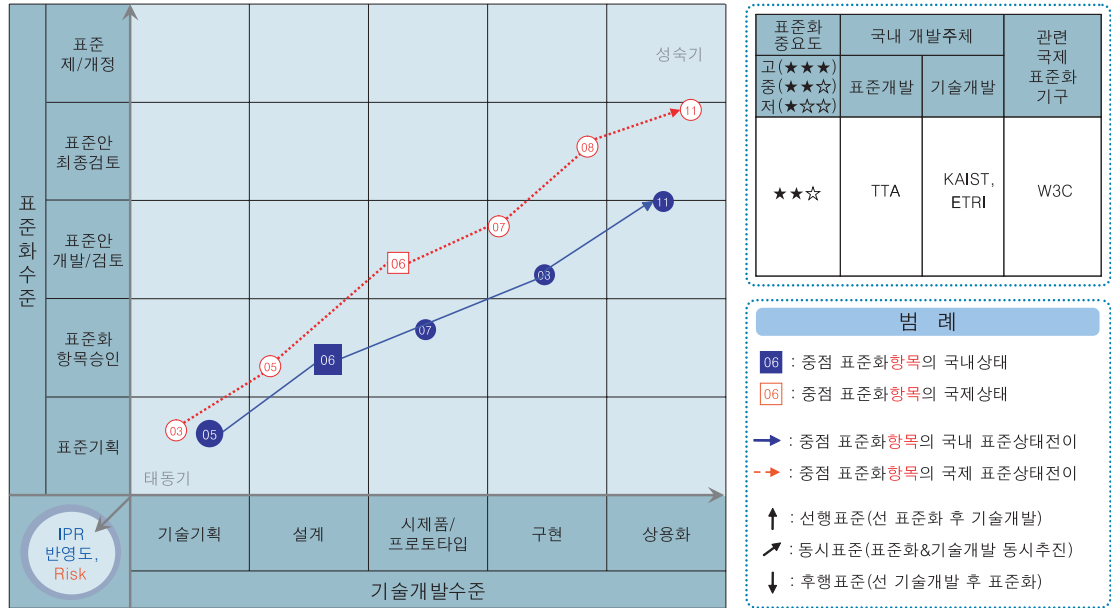
- 중점 표준화 대상항목별 선정사유

- u-지능정보 기술의 중점 표준화항목은 표준화추진 성공 가능성 및 산업적 파급성을 고려하여 온톨로지, 시맨틱 웹, 지능형 음성/언어 처리 기술을 도출하였다.
- 온톨로지 기술은 u-지능정보의 핵심기술인 대용량 온톨로지 구축과 관련된 기술로 보다 효과적이고 응용에 적합한 표준 및 도메인 온톨로지 구축, 매핑 통합에 관련된 분야이다. 온톨로지 구축 기술은 u-지능정보에서 공통적으로 활용될 수 있는 지식 형태로 표현되기 때문에 표준화된 형태로 제공함으로써 다양한 서비스들이 상호간 호환성 및 운용성에 중요한 역할을 한다. 따라서 표준 및 도메인 온톨로지 구축 기술을 중점 표준화항목으로 선정하였다.
- 시맨틱 웹 기술은 u-지능정보에서 다양한 서비스를 가능하게 하는 지능형 에이전트에서 필요한 규칙과 추론에 대한 요소기술로 최근 W3C에서 제안하는 시맨틱 웹 구조에 따르면 초기에 비해 추론과 관련된 부분이 강조되고 있다. 또한 산업계에서도 실용화 수준의 온톨로지 추론기에 대한 연구가 활발히 이뤄져 실용적 응용 프로그램에 적용되는 기술이므로 산업적 파급효과 측면에서 필요한 표준화항목이므로 중점 항목으로 도출하였다.
- 지능형 언어/음성 자원은 언어 표현에 사용되는 단어, 구 등의 의미 정보를 담고 있는 자원임. 지능형 언어/음성 자원은 u-지능 정보에서 인간과 기계 간의 의사 소통을 원활하게 할 수 있는 매개체가 된다. 질의 처리는 u-지능 정보 분야에서 사람이 기계로의 접근을 용이하게 할 뿐만 아니라, 정보 검색, 대화(dialog) 시스템, 지능형로봇 시스템 등에서도 사용될 것으로 전망된다. 그러므로 언어/음성 처리 기술 중에서 언어 자원과 질의 처리를 u-지능 정보 분야에서 표준화항목으로 선정하였다. 또한 u-IT839환경에서는 작고 가벼운 정보기가 도처에 존재하며 이를 음성으로 제어하게 되는데, 사용자가 다양한 기기를 혼란없이 사용하기 위해서는 체계적이고 일관성 있는 공통 음성 명령어 체계가 시급히 필요하기 때문에 표준화항목으로 선정하였다.

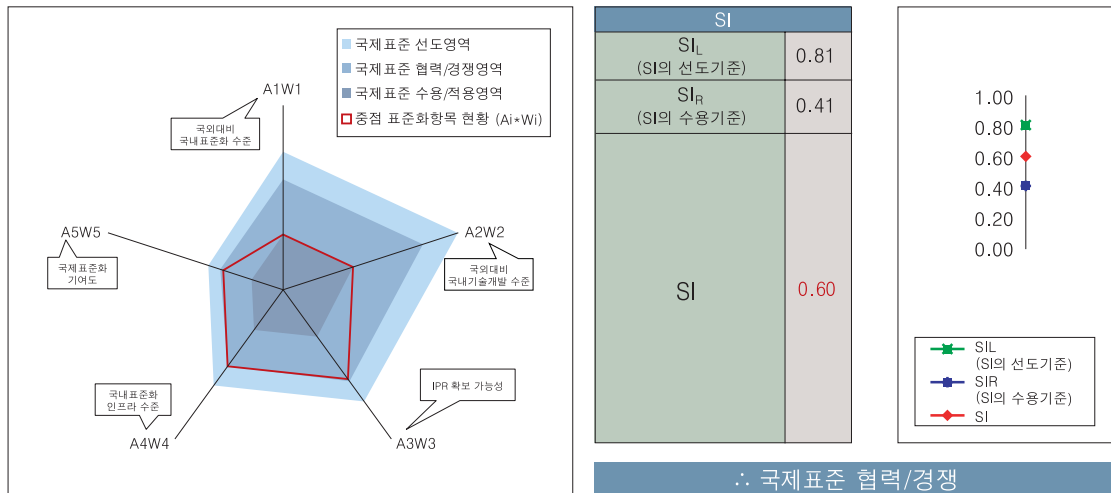
3.3. 중점 표준화 대상항목별 세부전략(안)

3.3.1. 온톨로지

- 표준상태전이도(표준화 & 기술 개발 연계 분석)



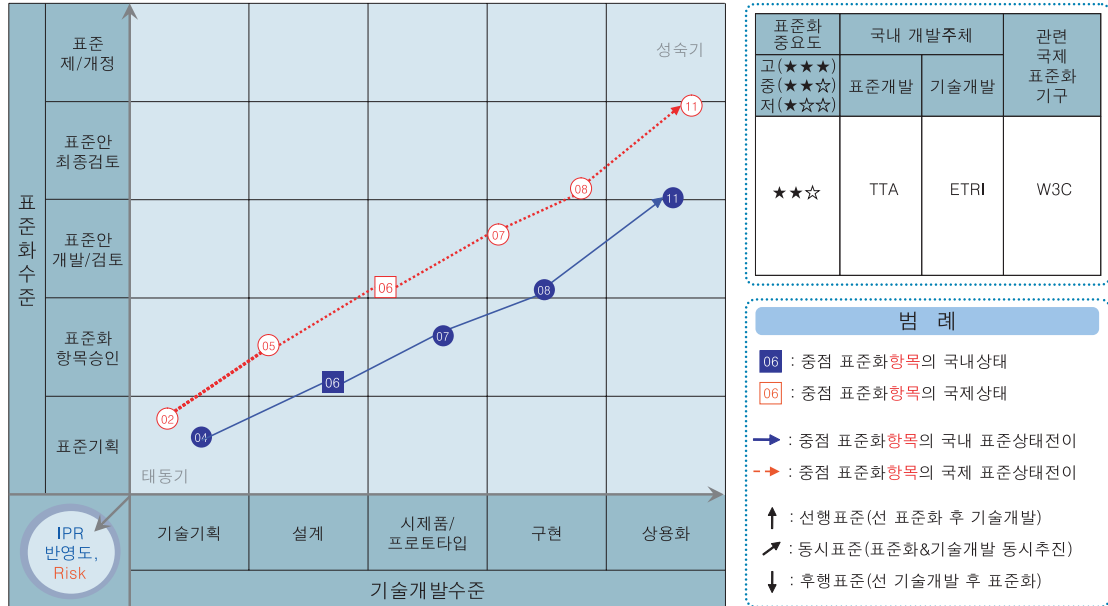
- 국제표준화 전략목표 도출



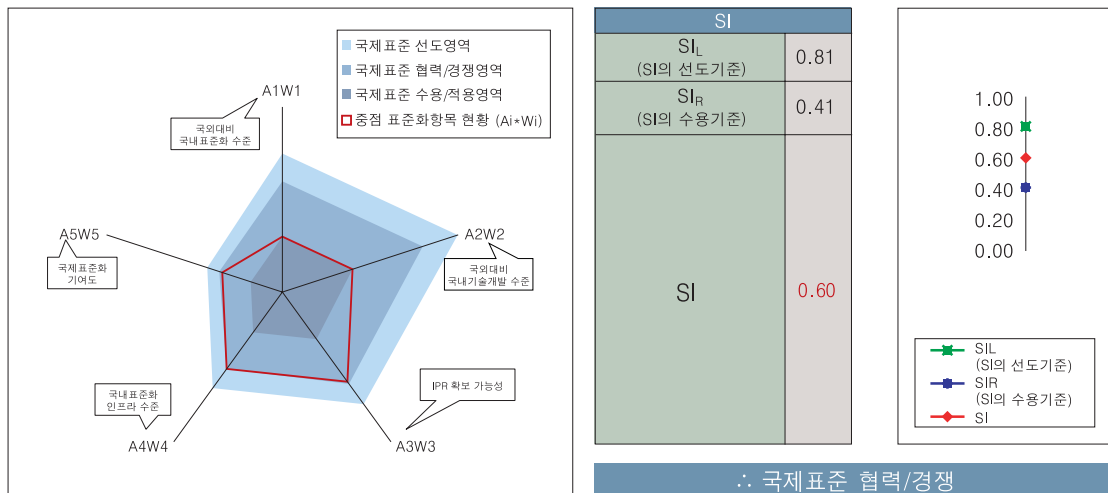
- 온톨로지 구축을 위한 표준 언어는 미국과 유럽 중심의 W3C에서 OWL (Web Ontology Language)가 이미 제정되었고, 표준 온톨로지 구축 분야에 있어서도 UMBC, CYC 등의 연구기관을 통해 활발히 개발되고 있다. 유럽의 ETSI 및 미국 DARPA 등에서 다양한 서비스들의 상호간 호환성 및 운용성에 중요한 역할을 하는 표준 온톨로지 및 도메인 온톨로지에 대한 연구도 진행 중이며, 국내에서도 KAIST와 ETRI를 중심으로 표준 온톨로지 구축 등에 대한 연구가 시작됨에 따라 충분히 경쟁력을 가질 수 있는 분야이다. 온톨로지 매핑/통합 등의 분야는 현재 국제적 표준화 노력은 미미하나, 온톨로지 응용서비스 및 기업 통합이나 지식 경영 분야에의 상용화가 빠르게 이루어지고 있는 분야에 활용되므로, 국제표준화항목 도출 및 발굴을 지원할 필요가 있다.
- IPR 확보 방안 : 표준 및 도메인 온톨로지 구축 분야는 국외에서는 연구소를 중심으로 표준이 진행되고 있고, 국내에서도 학계를 중심으로 연구가 시작되었으며, 실제 활용도를 높이고 산업적 파급력을 높이기 위해, 산업계에서의 실상용화를 통한 산업계 표준으로 자리 잡을 수 있도록 유도해야 한다. 또한 온톨로지 매핑/통합 세부 기술은 국제표준이 미미한 상황이고 개별 관심 기업들을 중심으로 상용화가 진행 중에 있으므로 이 분야의 벤처나 기술 투자를 유도하여 틈새 영역에 대한 국제적 IPR 확보가 용이하므로 이에 대한 기술 및 상품 개발 지원을 추진 하도록 한다.

3.3.2. 시맨틱 웹

- 표준상태전이(표준화 & 기술 개발 연계 분석)



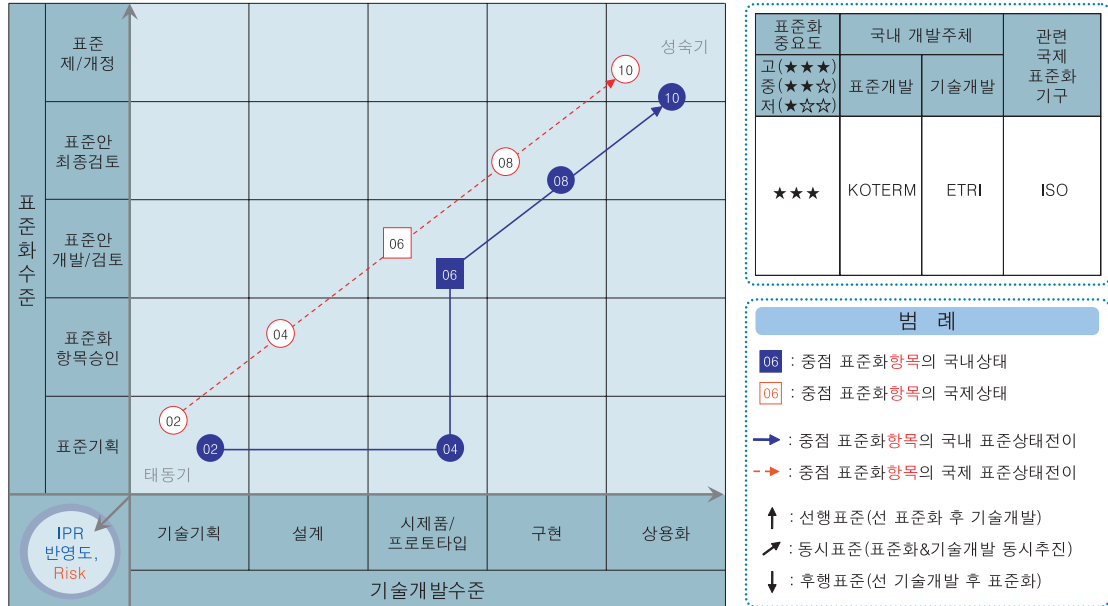
- 국제표준화 전략목표 도출



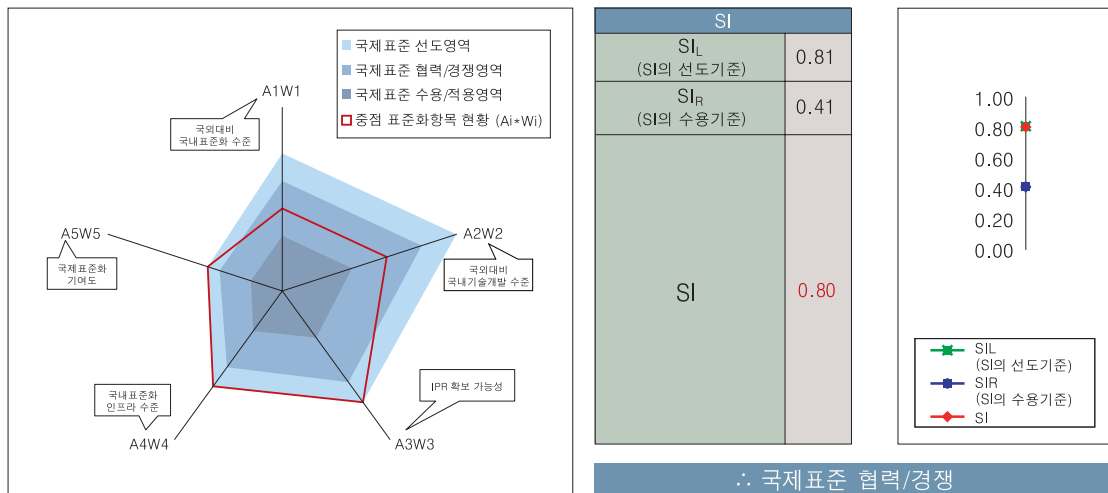
- 미국과 유럽을 중심으로 서술논리 기반의 실시간 추론 처리를 위한 연구와 DIG (Description logic Implementation Group) 중심으로 OWL 온톨로지의 실시간 추론 인터페이스를 정의, 활용하는 연구가 활발히 진행되고 있다. 시맨틱 웹 규칙 표현/추론 분야에 대한 국내 연구 및 개발 수준은 아직은 미미하나, 대학을 중심으로 연구 기반이 마련되고 있고, 정보통신부의 선도 기술개발 과제에서 산·학·연(KT, 연세대 등)을 통해 상용화 연구 및 개발이 2005년부터 시작되었다. 시맨틱 웹 규칙 표현/추론 기술 등은 웹 코리아 포럼의 시맨틱 웹 분과, NCA, ETRI를 중심으로 끊임없는 노력이 2000년 초반부터 이루어져 왔으며, 충분히 국제 경쟁력을 가질 수 있는 분야이다.
- IPR 확보 방안 : 이미 규칙 언어에 대해서는 SWRL이라는 주도적 표준 안에 존재하므로, SWRL 기반의 추론 기술 및 엔진 설계 등에 대한 IPR의 확보가 가장 유망하며 이에 대한 국내의 노력이 존재하므로 추가적 지원을 통해 신속한 IPR 확보를 독려하여야 할 것이다.

3.3.3. 지능형 언어/음성처리

- 표준상태전이도(표준화 & 기술 개발 연계 분석)



- 국제표준화 전략목표 도출



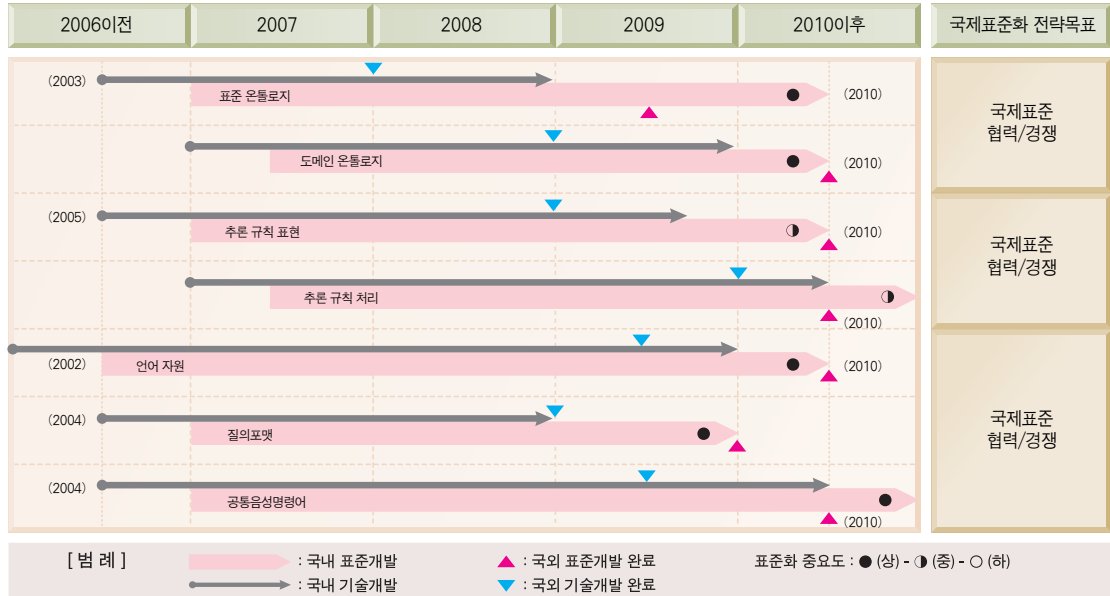
∴ 국제표준 협력/경쟁

- 세부전략(안)

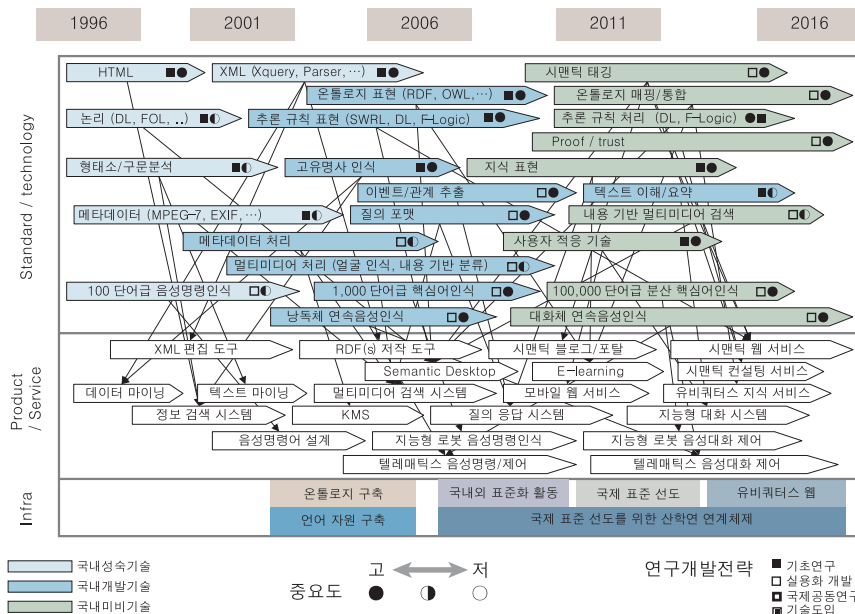
- 국내에서 구축되는 언어 자원은 표준화된 방법에 따르기 보다는 연구소별/업체별로 다양한 형식으로 구축된다. 그러므로 언어 자원을 구축하는 다양한 기술을 국내 연구소 및 업체에서 보유하고 있으나, 기관간의 정보 공유가 제대로 이루어지지 않고 있다.
- ISO/TC37에서 추진되는 언어 자원 데이터 범주와 레지스트리에 대한 국내 규격화를 시범 실시하고 이를 토대로 ISO/TC37에서 주도적으로 표준화활동을 실시한다.
- 질의 처리 분야에서 주요한 표준화 이슈는 질의를 표현하는 질의 포맷에 있다. 표준 질의 포맷으로 SQL, XQuery, SQLMM 등이 있으며, 현재 표준화가 진행 중인 질의 포맷으로 MPEG의 MPEG-7 질의 포맷과 JPEG의 JPSearch가 있다. 삼성, LG, 한국전자통신연구원 등의 국내 기관에서는 1990년대 초반부터 지금까지 MPEG 표준화활동에 참여하고 있다. 따라서, 각 기관에는 다수의 표준화 전문가가 있으며, 이들 중 다수는 표준 분과의 chair 활동을 통해서 MPEG 표준에 대한 충분한 노하우(knowhow)를 확보하고 있다. 이와 같은 국내표준 전문가들과의 협력을 통해서 질의 처리 관련 표준화를 추진하도록 한다.
- 현재 MPEG에서는 MPEG-7에 기반한 멀티미디어 검색을 위한 MPEG-7 질의 포맷에 대한 표준화를 진행하고 있다. 국내 다수의 MPEG-7 전문가들과의 협력을 통해서 MPEG-7 질의 포맷에 대한 표준화를 주도하도록 한다. MPEG-7은 XML Schema 형식으로 기술되어 있으므로, MPEG-7 질의 포맷을 작성하기 위해서는 XML 질의 언어인 XQuery에 관한 폭넓은 지식이 필요하며, MPEG-7 질의 포맷 표준화를 위해서는 XQuery 전문가와의 협력이 필요하다.
- 공통 음성 명령어는 현재 상용화 시작 단계에 있는 텔레매틱스와 디지털홈 분야에서 우선적으로 표준화한다. 최종 시제품에 적용한 후 일반 사용자를 대상으로 하는 usability test를 거쳐 최종 표준화 안을 작성한다.
- IPR 확보 방안
 - ISO/TC 37에서 대한민국이 주도하고 있는 SC4(언어자원 관리)의 표준화항목 중에서, MAF(Morphological Annotation Framework), LMF(Lexical Markup Framework), TDG(Thematic Domain Group)의 국내 대응 표준화를 국제표준으로 적극 추진하도록 한다.
 - 한국산업규격 및 각 산업계에서 사용되는 전문용어 표준화를 체계적으로 추진하여 강점을 가지는 표준화 부분을 국제표준으로 추진하도록 한다.
 - MPEG에서 MPEG-7으로 표현된 멀티미디어 검색하기 위한 질의 포맷에 대한 표준화를 2006년 4월부터 추진하고 있다. 국내 다수의 MPEG-7 및 XQuery 전문가들과의 공동 작업을 통해서 MPEG-7 질의 포맷에 대한 표준화를 주도한다.
 - MPEG-7 질의 포맷에 대한 표준화작업과 함께 MPEG-7 질의 포맷 처리 기술에 대한 IPR을 확보한다.
 - 사용자 usability가 최대화되는 공통 음성명령어 도출 방안을 조기에 확보하고, 인식을 개선을 위한 특정 음성처리 기술 확보를 통하여 실상용화에 기여한다.

3.4. 중장기 표준화로드맵

3.4.1. 중기(2007~2009) 표준화로드맵



3.4.2. 장기 표준화로드맵(10년 기술 예측)



[국내외 관련 표준 대응리스트]

구분	표준화 항목	표준명	기구 (업체)	제정 연도	재개정 현황	국내 관련 표준	국내 추진기구
온톨로지	온톨로지	OWL Web Ontology Language XML Presentation Syntax	W3C	2003	제정		
		LBase : Semantics for Languages of the Semantic Web	W3C	2003	초안		
		RDF Primer	W3C	2004	제정	TTAE.OT-10.0013	
		RDF Semantics W3C 2004 제정 TTAE.OT-10.0011 RDF Test Cases	W3C	2004	제정	TTAE.OT-10.0010	
		RDF Vocabulary Description Language 1.0 : RDF Schema	W3C	2004	제정	TTAE.OT-10.0009	
		RDF/XML Syntax Specification	W3C	2004	제정	TTAE.OT-10.0008	
		Resource Description Framework (RDF) : Concepts and Abstract Syntax	W3C	2004	제정	TTAE.OT-10.0007	
		자원 기술 체계(RDF) 기본서	TTA	2004	제정		
		자원 기술 체계(RDF) 의미 체계	TTA	2004	제정		
		자원 기술 체계(RDF) 테스트 케이스	TTA	2004	제정		
		자원 기술 체계(RDF) 어휘 기술 언어 1.0 : RDF 스키마	TTA	2004	제정		
		자원 기술 체계/확장성생성언어(RDF/XML) 구문 규격	TTA	2004	제정		
		자원 기술 체계(RDF) : 개념 및 추상 구문	TTA	2004	제정		
		OWL Web Ontology Language Guide	W3C	2004	제정		
		OWL Web Ontology Language Overview	W3C	2004	제정		
		OWL Web Ontology Language Reference	W3C	2004	제정		
		OWL Web Ontology Language Semantics and Abstract Syntax	W3C	2004	제정		
		OWL Web Ontology Language Test Cases	W3C	2004	제정		
		OWL Web Ontology Language Use Cases and Requirements	W3C	2004	제정		
시맨틱 웹	시맨틱 웹	SWRL : A Semantic Web Rule Language Combining OWL and RuleML	W3C	2004	제안		
		FOL RuleML : The First-Order Logic Web Language	W3C	2005	제안		
		Web Rule Language (WRL)	W3C	2005	제안		
		WRL XML Schemas	W3C	2005	제안		
		Relationship of WRL to relevant other technologies	W3C	2005	제안		
		Semantic Web Rule Language First-Order Logic (SWRL FOL)	W3C	2005	제안		
지능형 언어/ 음성 처리	언어 자원 관리	Basic descriptors and mechanisms for language resource	ISO/TC37	개발 중		표준없음	
		Feature Structures- Part 1 : Feature Structure Representation	ISO/TC37	개발 중		표준없음	
		Lexical database	ISO/TC37	개발 중		표준없음	
		Lexicographical symbols and typographical conventions	ISO				
		for use in terminograp	1951 : 1997	1997		개발 중	TTA
		Morphosyntactic annotation framework	ISO/TC37	개발 중		개발 중	TTA
		Linguistic annotation framework	ISO/TC37	개발 중		개발 중	TTA
		Representation schema	ISO/TC37	개발 중		표준없음	
		공통 음성DB 메타데이터 표현 방안	TTA	개발 중	초안	국내개발	음성정보처리 기술 포럼
		산업용 음성 DB 메타데이터 규격	KATS	개발 중	초안	국내개발	SIT 표준화 포럼
	질의 처리	질의 처리 XML Schema Part 0 : Primer	W3C	2004		표준없음	
		XML Schema Part 1 : Structure	W3C	2004		표준없음	
		XML Schema Part 2 : Datatypes	W3C	2004		표준없음	

구분	표준화 항목	표준명	기구 (업체)	제정 연도	재개정 현황	국내 관련 표준	국내 추진기구
지능형 언어/ 음성 처리	질의 처리	XQuery 1.0 : An XML Query Language	W3C	개발 중	후보 권고안	표준 없음	
		XML Path Language (XPath) Version 1.0	W3C	개발 중	권고안	표준 없음	
		Information technology — multimedia content description interface	ISO	2002		표준 없음	MPEG 포럼
		MPEG-7 Query Format	ISO	개발 중		표준 없음	MPEG 포럼
	음성 처리	PDA 공통음성명령어 표준	TTA	2003	제정	국내 개발	음성정보 처리기술 포럼
		차량 네비게이션을 위한 음성명령어 규격	KATS	2005	제정	국내 개발	SIT 표준화 포럼
		자동차기기제어 음성명령어 규격	KATS	2005	제정	국내 개발	SIT 표준화 포럼
		VoiceXML 2.0	W3C	개발 중	개정	표준 없음	없음
		Speech synthesis markup language	W3C	개발 중	초안	표준 없음	없음

[참고문헌]

- [1] 김희린, “음성정보처리 기술개발 현황 및 전망,” 정보처리학회지, 제11권, 제 2호, pp.25~32, 2004.
- [2] 구명완, 김재인, “음성정보처리기술 응용서비스,” 정보처리학회지, 제11권, 제2호, pp.17~24, 2004.
- [3] 국가 IT 온톨로지 인프라 기술개발 수행계획서, 정보통신부, 2006. 3.
- [4] IT839 표준화로드맵 Ver.2006 종합보고서, TTA, 2005.12.
- [5] 음성/언어정보산업현황 실태조사, 한국전자통신연구원, 2004.10.
- [6] 이용주, 김상훈, “대규모 공통음성DB 구축 현황,” 전자공학회지, 제30권, 제 7호, pp.69~77, 2003.
- [7] 정보통신부, IT839 전략, 2006.
- [8] 한국전문용어전문위원회, “Language resource management”, ISO/TC 37/SC 4.
- [9] 홍기형, “음성정보기술 응용 개발 표준화,” 전자공학회지, 제30권, 제 7호, pp.45~52, 2003.
- [10] Christian Ohlms, “The Future of the Semantic Web”, 9. AIK-Symposium, 19 April 2002.
- [11] Gartner, Semantic Web Technologies Take Middleware to the Next Level, 2002.
- [12] Gartner, Gartner’s Hype Cycle Special Report for 2005, 2005.
- [13] Michael Uschold, “Where is the Semantics in the Semantic Web?”, 5th International. Conference on Autonomous Agents, 2001.
- [14] Mills-Davis, “The business value of semantic technology”, 2004.
- [15] Mills-Davis, “Semantic Wave 2006, Part-1”, A Project10X Report, 2006.
- [16] Peer J. “Bringing together semantic Web and Web services.”, The Semantic Web – ISWC 2002. First International Web Conference. Proceedings LNCS Vol.2342, Springer-Verlag. 2002, pp.279~91.
- [17] Saltlux Annual Conference Proceedings, 2005.
- [18] Sheila A. McIlraith, Tran Cao Son, and Hoglei Zeon, “Semantic Web Services”, IEEE Intelligent Systems, March/April 2001, pp.46~53.
- [19] Sheila A. McIlraith. David L. Martin, “Bringing Semantics to Web Services”, IEEE Intelligent systems, 2003.
- [20] Staab S, Studer R, Schnurr H-P, Sure Y., “Knowledge processes and ontologies”, IEEE Intelligent Systems, vol.16, no.1, Jan.-Feb. 2001, pp.26~34.
- [21] Tim Berners-Lee, “Standards, Semantics and Survival”, <http://www.w3.org/2003/Talks/01-siia-tbl/>
- [22] Tim Berners-Lee, “The Semantic Web”, Scientific America, vol. 501., 2001.
- [23] Tim Berners-Lee, “Weaving the Web”, San Francisco., 1999.
- [24] Val Tsourikov, “Semantic Causal Reasoning : Applications in Patent Intelligence and Concept Extraction”, International Chemical Information Conference & Exhibition, October 2001.
- [25] Vipul Kashyap, “The (Semantic) Web in the Real World”, THE 11th Int. WORLD WIDE WEB CONFERENCE, 2002.
- [26] Voice Extensible Markup Language (VoiceXML) Version 2.0, <http://www.w3.org/TR /2004/REC-voicexml20-20040316>, March 2004.
- [27] Ying Ding, Dieter Fensel, Michel Klein, Borys Omelayenko, “The semantic web : yet another hip?”, Data & Knowledge

Engineering 41 (2002), pp. 205~227.

[28] York Sure, Victor Iosif, "First Results of a Semantic Web Technologies Evaluation", DOA' 02, October 2002.

[29] Yu Li, Biqing Huang, Wenhuan Liu, Hongmei Gou, "Ontology for modeling and analyzing of enterprise competence."

2001 IEEE International Conference on Systems, Man and Cybernetics. IEEE. Part vol.5, pp.2941-6, 2001.

[30] <http://cobra.umbc.edu/ontologies-2003-11.html>

[31] <http://deri-korea.org/main/main.asp>

[32] <http://kaon2.semanticweb.org>

[33] http://korterm.kaist.ac.kr/kor_iso_tc_37

[34] <http://mp7.watson.ibm.com>

[35] <http://mpeg.nist.gov>

[36] <http://owl.man.ac.uk/factplusplus>

[37] <http://pervasive.semanticweb.org/soupa-2004-06.html>

[38] <http://tap.stanford.edu/semtag/index.html>

[39] <http://wonderweb.semanticweb.org>

[40] <http://wordnet.princeton.edu>

[41] <http://www.cnet.com>

[42] <http://www.cs.umd.edu/projects/plus/SHOE>

[43] <http://www.cyc.com>

[44] <http://www.daml.org>

[45] <http://www.deri.ie>

[46] <http://www.etnews.co.kr>

[47] <http://www.foaf-project.org>

[48] <http://www.franz.com/products/racer>

[49] <http://www.illc.uva.nl/EuroWordNet>

[50] <http://www.inews24.com>

[51] <http://www.kmobile.co.kr>

[52] <http://www.korterm.or.kr>

[53] <http://www.essi-cluster.org>

[54] <http://www ldc.upenn.edu>

[55] <http://www.mindswap.org/2003/pellet>

[56] <http://www.mpeg.or.kr>

[57] <http://www.ontoprise.de>

[58] <http://www.ontotext.com/kim>

[59] <http://www.ruleml.org>

-
- [60] <http://www.sejong.or.kr>
 - [61] <http://www.tc37sc4.org>
 - [62] <http://www.w3.org/2004/OWL>
 - [63] <http://www.w3.org/RDF>
 - [64] <http://www.w3.org/Submission/SWRL>
 - [65] <http://www.w3.org/TR/rdf-schema>
 - [66] <http://www.w3.org/TR/rdf-sparql-query>
 - [67] <http://www.webkorea.or.kr/index.jsp>
 - [68] <http://www.zdnet.co.kr>
 - [69] <http://www2.wips.co.kr>

[약어]

ASG	Adaptive Services Grid
CAGR	Compound Annual Growth Rate
CoBra	Context Broker Agent Ontology
CRM	Customer Relationship Management
DAML	DARPA Agent Markup Language
DAML-T	Darpa Agent Markup Language-Time
DARPA	Defense Advanced Research Project Agency
DC	Dublin Core
DIG	Description logic Implementation Group
DIP	Data Information and Process Integration with Semantic Web Services
DQL	DAML Query Language
EAI	Enterprise Application Integration
ESSI	European Semantic Systems Initiative
ETSI	European Telecommunications Standards Institute
EU	European Union
FOAF	Friend of A Friend
FP5	the Fifth Framework Program
FP6	the Sixth Framework Program
FP7	the Seventh Framework Program
HTML	HyperText Markup Language
IETF	Internet Engineering Task Force
ISO	International Standard Organization
IST	Information Society Technologies
JPEG	Joint Photographic Experts Group
KMS	Knowledge Management System
MPEG	Moving Picture Experts Group
OASIS	Organization for the Advancement of Structured Information Standards
OWL	Ontology Web Language
RDF	Resource Description Framework
RDFS	RDF Schema
RSS	RDF Site Summary/Rich Site Summary
RuleML	Rule Markup Language
SALT	Speech Application Language Tags

SEKT	SEmantic Knowledge Technologies
SMI	Standard Medical Informatics
SOUPA	Standard Ontology for Ubiquitous and Pervasive Ontology
SSML	Speech Synthesis Markup Language
SWRL	Semantic Web Rule Language
TTA	Telecommunications Technology Association (한국정보통신기술협회)
UMBC	University of Maryland, Baltimore County
VXML	Voice eXtensible Markup Language
W3C	World Wide Web Consortium
WSML	Web Service Modeling Language
WSMO	Web Service Modeling Ontology
WSMX	Web Searvice eXecution environment
XML	Extensible Markup Language
XRML	eXtensible rights Markup Language
XQuery	XML Query