

# 온톨로지 기반 전사적 아키텍처 상호운용성 방안 연구

황상규<sup>0</sup>, 김왕석, 변영태

한국국방연구원<sup>0</sup>, 홍익대학교

kid4@naver.com, kwsuck@cs.hongik.ac.kr, byun@cs.hongik.ac.kr

## The Study for Methodology of Ontology-based Enterprise Architecture Interoperability

Sang-Kyu Hwang<sup>0</sup>, Wang-Suk Kim, Young-Tae Byun

Korea Institute for Defense Analyses<sup>0</sup>, Hong-Ik Univ.

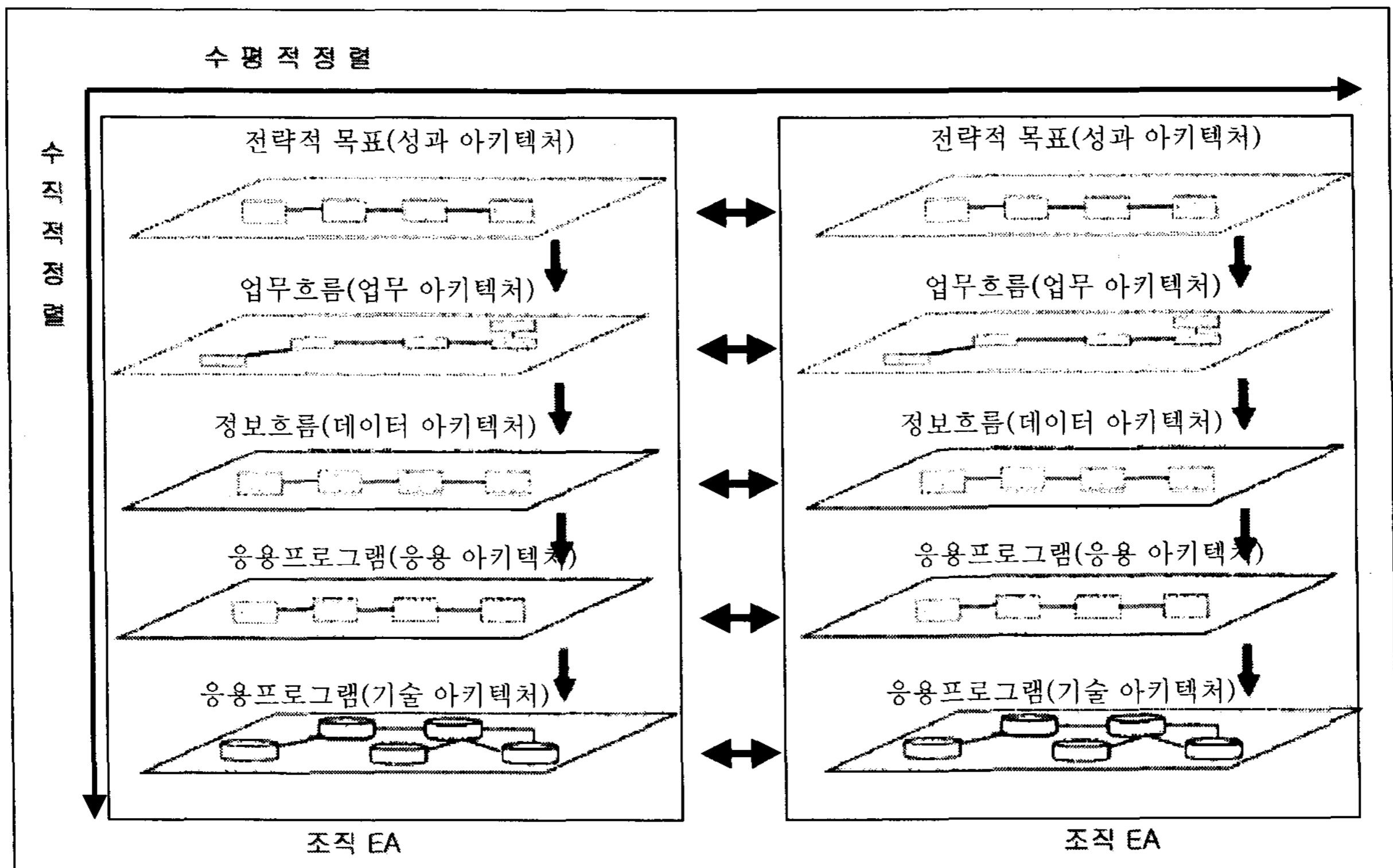
### 요 약

정부의 행정 서비스를 보다 효과적이고 효율적으로 제공하는 것은 정부혁신의 핵심키워드이다. 이를 위한 수단으로 정부 공공기관들은 전사적 아키텍처(Enterprise Architecture)를 적극적으로 도입, 활용해야 한다. 전사적 아키텍처란 아직까지도 진화중인 개념으로 해당 기관 정보화 업무의 전 영역을 업무 프로세스(Business Processes), 정보 흐름 및 관계(Information Flow and Relationships), 애플리케이션(Applications), 데이터 명세 및 관계(Data Descriptions and Relationships), 그리고 기술 하부구조(Technology Infrastructure)로 구분 짓고 각각을 구조화된 산출물로 정리함으로써 공공정보화의 청사진을 제시한다. 이렇게 구축된 아키텍처 결과물은 기관 내 뿐 아니라 타 기관 EA와의 연계를 통해 상호 운용되어야 하며, 이 과정에서 어휘의미중의성(Word Sense Ambiguation)등 상호운용성 문제가 핵심 이슈로 대두되리라 예상된다. 공공분야로 한정짓더라도 200여 정부부처 EA간의 상호운용성 문제는 정부조직 간 경계를 넘어 통합된 EA정보를 취합, 활용하는데 있어 가장 큰 장애요인이 되리라 예상된다. 본 논문에서는 온톨로지를 사용하여 각기 서로 다른 EA간 상호운용성 문제를 해결하는 방안에 대해 논의하고자 한다.

### 1. 전사적 아키텍처

2005년 12월, 대한민국 국회 본회의에서 모든 공공기관의 전사적 아키텍처(Enterprise Architecture) 도입 및 활용을 의무화하는 '정보시스템 효율적 도입 및 운용 등에 관한 법률안'이 통과되었다[대한민국국회, 2005]. 이에 따라 금년부터 모든

공공기관은 의무적으로 EA를 구축, 활용해야 한다. EA를 구축하는 과정은 조직이 보유한 모든 정보자원을 조직의 전략적 목표(성과 아키텍처)와 목표달성을 위한 관련 업무(업무 아키텍처), 업무수행에 필요한 정보의 흐름(데이터 아키텍처), 정보획득을 위한 지원도구로써 응용프로그램(응용 아키텍처), 응용프로그램 구현에 필요한 IT기술(기술



[그림 1] 전사적 수준의 정렬(Enterprise Alignment)

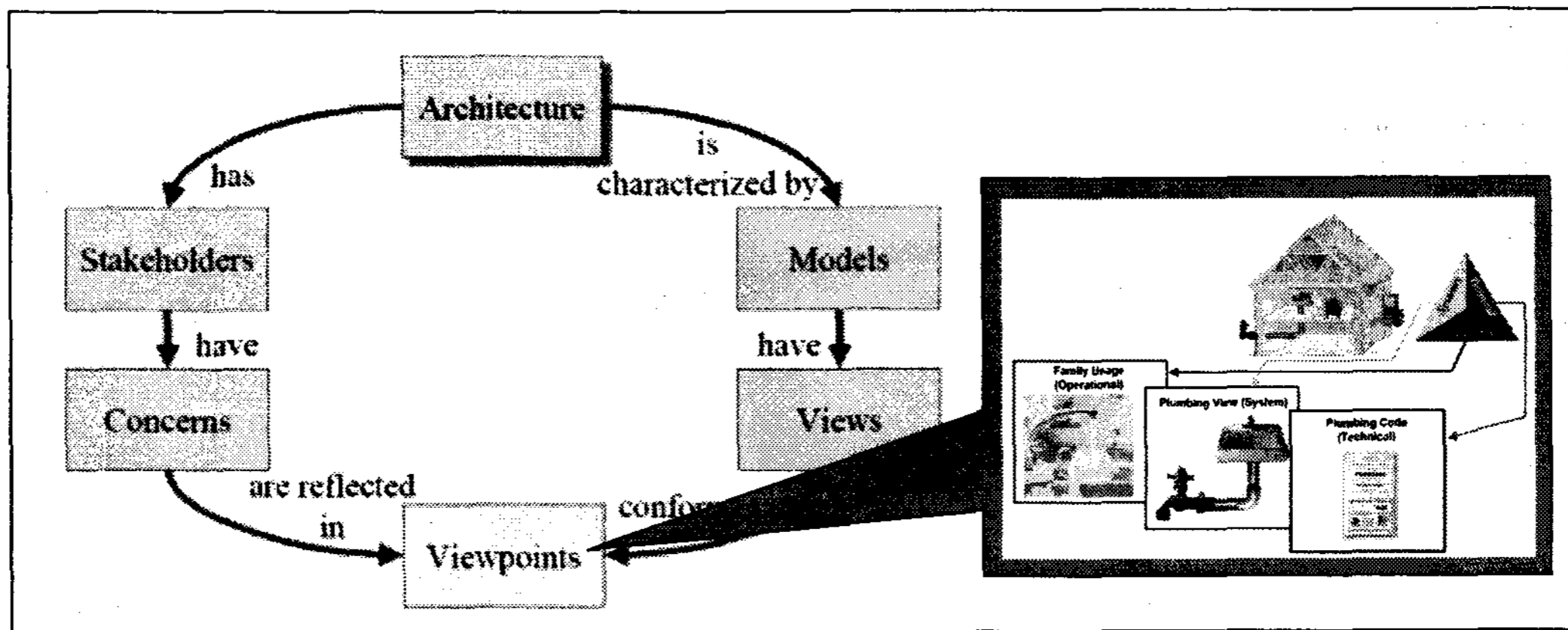
아키텍처)로 구분하는 단계를 거쳐, 각 요소별 엘리먼트 간 연관관계를 식별하는 전사적 정렬(Enterprise Alignment) 단계로 진행된다. 아키텍처 관점에서 정렬이란, '어떤 연관된 것들 상호간의 배열(arrangement)'이라고 정의할 수 있으며[UCS, 1999], 각기 서로 다른 EA간 연계 및 활용을 위해서는 수직적 정렬 뿐 아니라 수평적 정렬도 요구된다.

## 2. 서로 다른 EA간 상호운용성

EA를 구축하는 과정에서 영역별로 아키텍처로 구분 짓고 영역별 엘리먼트 간 연관관계를 식별하는 정렬과정 뿐 아니라 EA를 활용하는 과정에서 각기 서로 다른 EA간 연계 시 발생하는 상호운용성 문제는 EA의 성공을 판가름 짓는 중요한 요소로 작용하게 된다. EA상호운용성 문제는 크게 어휘의미중의성(Word Sense Ambiguation)으로 대표되는 어휘적 수준의 상호운용성 문제와 EA구축 시 관점(View)의 차이에서 발생하는 개념적 수준의 상호운용성 문제로 구분해 볼 수 있다. 어휘의미중의성이란 정보검색에 있어 대상 어휘가 문서

상에서 여러 가지 다양한 의미로 사용됨에 따라 원래 의도와는 다른 형태의 문서가 검색됨으로써 검색정확율이 저하되는 현상을 의미한다[황상규, 1999]. 미 국방성의 경우, 어휘의미중의성에 의해 발생하는 일관성 문제를 해결하기 위하여 EA구축 시 필수산출물의 하나로 EA용어사전 구축을 의무화함으로써 참여자들의 어휘사용을 제한하고 이를 통해 데이터의 무결성을 보장하고 있다[BEA, 2006]. 그러나 EA용어사전 구축하더라도 서로 다른 EA간 연계(수평적 정렬) 시 각기 서로 다른 용어사전을 사용한다면 어휘적 수준의 상호운용성 문제는 그대로 남아있게 된다. 또한, 용어사전을 구축하더라도 개념적 수준의 상호운용성 문제는 해결되지 않는데, 개념적 수준의 상호운용성 문제는 어휘표현의 앞서 서로 다른 EA 개발 참여자들 간의 개념적(Concerns) 차이에서 비롯된 관점(View)의 차이에서 발생한 문제이기 때문이다.

집 배수구의 파이프가 막힌 상황을 예로 들면 집 주인의 입장에서는 '배관장애'라는 운영적 관점(Operational View)에서 배관공에게 수리를 요청할 것이며, 배관공은 가정 내 배관시스템(System



[그림 2] 개념적 수준의 상호운용성 문제

View)중 세면대와 연결된 '45mm PVC파이프'가 파손되었음을 확인 후 새로운 부품을 주문할 것이다. 부품 주문 요청을 받은 상점점원은 제품명세표를 참조(Technical View)하여 가정용 45mm 파이프의 제품번호 'AA0547'을 확인 후 해당 물품을 발송할 것이다. 결국 대상을 바라보는 이(Stakeholders)마다 각기 다른 개념(Concerns)을 가지고 각기 다른 형태로 표현함으로써 관점의 차이가 발생하게 된다. 아키텍트(Architect)는 각기 서로 다른 참여자별 관점(Views)에서 추상화된 모델링과정을 통해 관점별 아키텍처 산출물을 도출해 내고 아키텍처 산출물간의 관계를 전사적으로 정렬(Enterprise Alignment)함으로써 관점간의 합일점(Viewpoints)을 도출해낼 수 있다. [그림 2]의 사례에서처럼 운영 아키텍처 관점에서 '배관장애'는 시스템 아키텍처 관점에서 'PVC파이프'와 기술 아키텍처 관점에서 'AA0547'와 연관성을 가질 수 있다. 이 경우, EA산출물을 참조한 배관공은 집주인과 대화할 때 '45mm PVC파이프' 파손보다는 '배관장애'라고 보다 쉽게 설명할 수 있으며, 상점점원에게 전화 주문할 경우에는 '45mm PVC파이프' 2개라고 설명하기 보다는 'AA0547' 2개라고 요청함으로써 개념차이에서 발생할 수 있는 실수를 미연에 예방할 수 있다. 그러나 서로 다른 EA간 개념적 수준의 상호운용성 문제는 여전히 그대로 남아있게 된다. 물론 개념적 수준의 상호운용성 문제를 해결하기 위해 수작업을 통해 서로 다른 EA간 수평적 정렬화를 시도할 수 있으나, 앞으로 공공기관 EA만 200개 넘게 구축해야 하는 현실을 고려할 때 비용 대 효과측면에서 자동/반자동화된

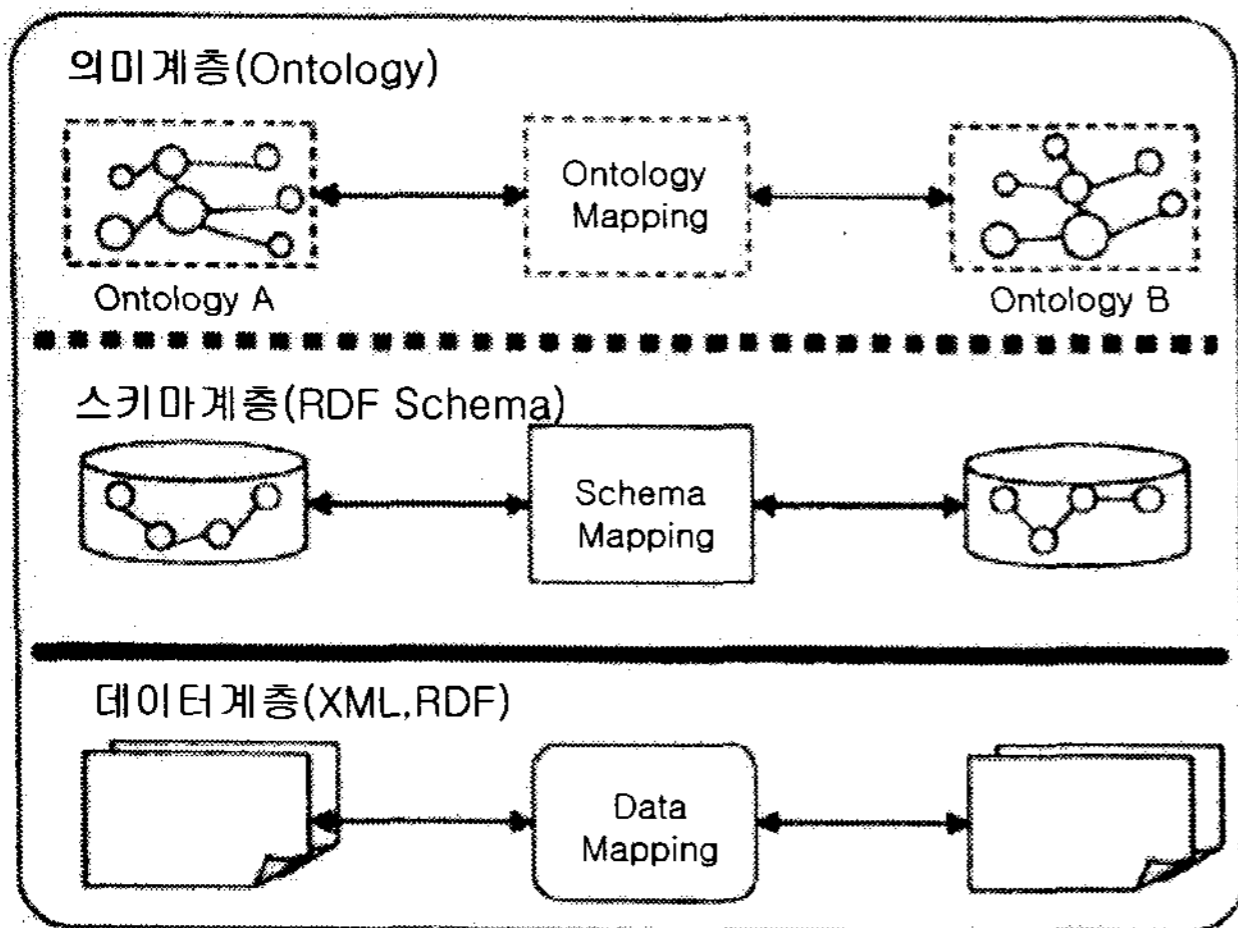
지원도구개발이 절실히 요구된다.

지금까지 서로 다른 EA간 연계(수평적 정렬)시 발생하는 어휘적 수준의 상호운용성 문제와 개념적 수준의 상호운용성 문제에 대해 살펴보았다. 두 가지 문제를 해결하기 위해서는 개념적 수준의 상호운용성 문제가 우선적으로 해결되어야 하는데, 개념적 수준의 상호운용성이 우선적으로 보장되어야 이를 기반으로 서로 다른 EA 어휘용어사전 간 정렬(Alignment)화가 이루어질 수 있기 때문이다. 결국, EA구축 시 관점(View)의 차이에서 발생하는 개념적 수준의 상호운용성 문제가 선결될 때 어휘적 수준의 상호운용성 문제도 해결될 수 있으며, 이를 위한 해결책으로써 온톨로지를 활용하여 EA 상호운용성 문제를 해결하는 방안을 제시하고자 한다.

### 3. EA간 정보 매핑 방안

자동화된 정보 공유를 위한 정보 매핑 모델 연구는 크게 시맨틱 웹의 계층구조 상 데이터 계층, 스키마 계층, 의미 계층의 세 부류로 나누어 접근할 수 있다[Berners-Lee, 2003]. 데이터기반 매핑의 경우, 단순히 어휘 간 유사도 계산을 통해 대상 데이터 간의 매핑을 시도한다. 스키마기반 매핑의 경우, XML/RDF스키마 정보를 통해 대상 정보와 정보 간의 관계를 그래프적인 형태로 표현하고, 그래프로 표현된 노드간의 매칭을 통해 유사도가 높은 순으로 매핑 순위가 매겨진다. 마지막으로, 의미기반 매핑은 온톨로지를 활용하여 대상 개념 간의

매핑이 이루어진다. 온톨로지(ontology)란 '기술하고자 하는 대상을 개념화(conceptualization)를 통해 정형화(formal)되고, 명백하게(explicit)하게 설명해 놓은 명세서(specification)'이다. 의미기반 매핑의 경우, 대상 개념과 개념 간의 관계를 명시적으로 표현함으로써 개념사용에 있어 모호성을 제거할 수 있다는 점과 지식베이스의 일관성을 자동화된 방법에 의해 검사가 가능하다는 점, 명시된 데이터로부터 유추를 통해 특정하게 명시되지 않은 데이터를 도출해내는 자동화된 추론이 가능하다는 점 등을 장점을 가지고 있다.



[그림 3] 계층별 정보 매핑 방안

다음은 미연방정부 EA참조모델(FAE)를 중심으로 EA온톨로지 개발의 필요성을 살펴보기로 하겠다 [FEA RMO, 2005]. EA온톨로지 도입 시 가장 큰 장점은 자연어 형태로 기술되는 EA구축 시 발생하는 모호성(ambiguity)을 제거할 수 있다는 점이다. EA구축 시 발생하는 모호성은 비일관성(Inconsistency), 충돌(Conflict), 생략(Omission)이라는 세 가지 문제로 구분할 수 있으며, 각각의 경우 발생하게 되는 모호성 문제는 다음과 같다.

■ 비일관성(Inconsistency)

미연방정부 EA참조모델에서 각 영역별 참조모델은 계층구조를 가지고 있다. 기술참조모델(TRM)의 경우, 레벨 4의 계층구조로 되어 있으나, 필요에 따라 새로운 레벨을 더하거나 현재의 레벨의 정보들을 재배치하는 것이 요구되며, 이전에 정의

된 이름의 의미를 재정의하는 경우가 발생함에 따라 일관성이 깨질 수 있다.

■ 충돌(Conflict)

기술참조모델(TRM)에서 '오라클'이란 단어는 '업체이름'이라고 정의된 항목 외에도 'DBMS의 한 종류 기술'이라고 정의하는 등 여러 곳에서 각기 다른 용도로 정의되고 사용됨에 따라 충돌을 유발할 수 있다.

■ 생략(Omission)

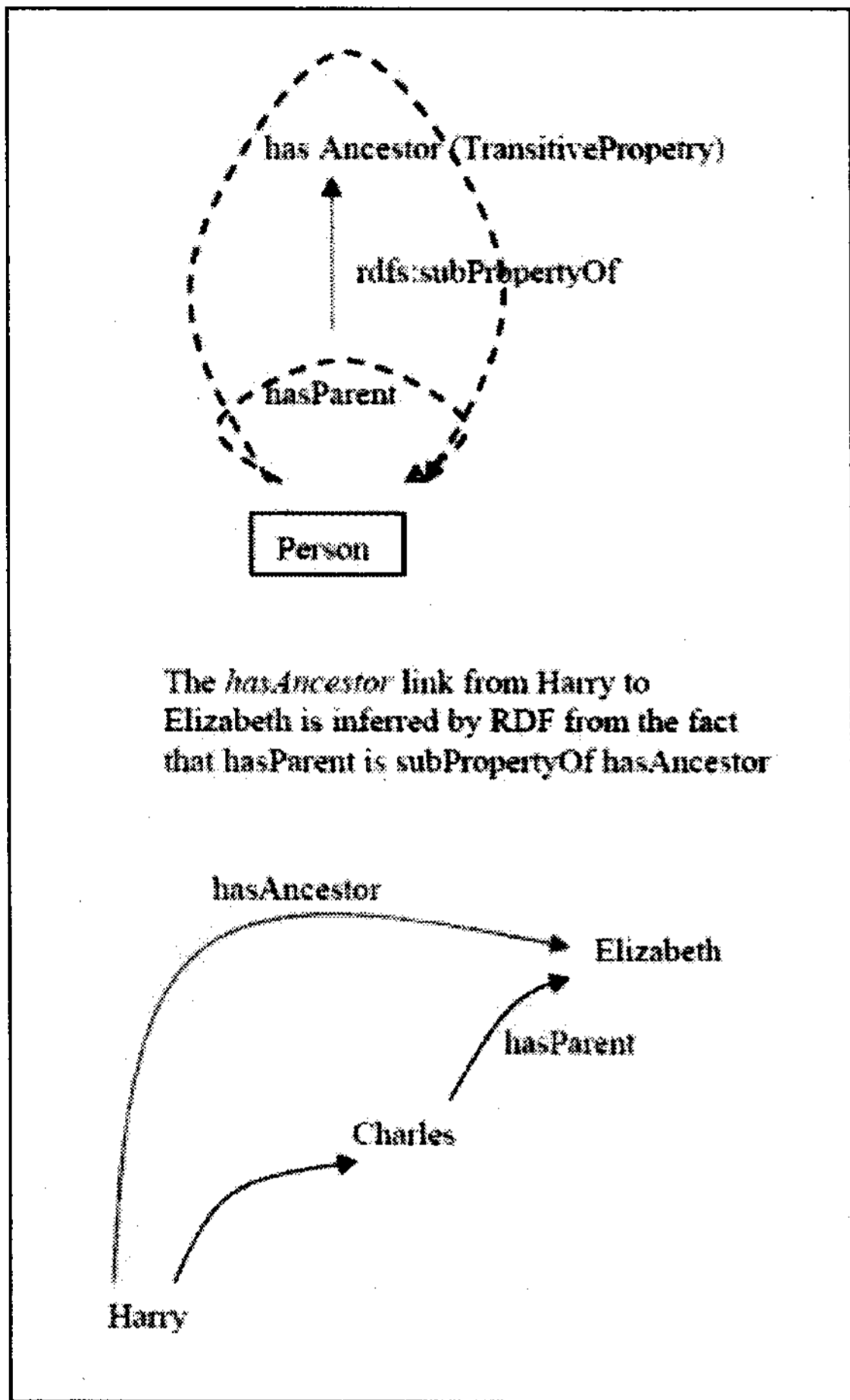
관계모델에서의 생략은 꽤 민감할 수 있다. 예를 들어 기술참조모델(TRM)의 '기술 표준 리스트'에서 DBMS란 항목을 명시하고, MySQL이 기술 표준 리스트의 DBMS란 항목의 하나임이 명시화되지 않고 '서비스 표준 리스트'에만 등재될 경우가 발생하기 쉽다. 이러한 경우, 사실상 MySQL이 DBMS기술 표준의 하나임을 식별하기란 사실상 불가능하며, EA활용 측면에서 SQL과 연관된 내역을 단순 키워드 매칭을 통해 검색하는 과정에서 MySQL을 SQL이라는 데이터베이스언어의 한 종류로 오인할 소지가 매우 높다.

미연방정부 EA온톨로지 모델인 FEA RMO는 온톨로지 설계하는 과정에서 유도된 5가지 온톨로지 디자인 패턴을 제시하고 있는데, 이러한 디자인 패턴을 FEA뿐 아니라 일반적인 EA구축 시 발생하는 모호성 문제를 해결하기 위한 도구로써 활용하는 방안을 새롭게 모색해 보았다.

■ axiom bridge and axiom bridge 패턴

일반적인 FRA 모델에서 하나의 참조모델은 다른 참조모델과의 관계가 성립될 수 있다. 대표적인 사례로 업무참조모델(BRM)과 성과참조모델(PRM)을 예로 들 수 있는데, BRM의 '교통안내서비스'란 업무 entity는 성과참조모델에서 성과지표의 대상 업무 중 하나인 '교통안내서비스 성과지표'의 대상업무entity로 재활용(재정의)될 수 있다. 이 경우 BRM모델에서 '교통안내서비스'란 업무entity는

PRM이라는 다른 참조모델에서 대상업무entity라는 다른 역할 수행하며, rdf:subClassOf를 사용하여 둘 간의 관계를 표현할 수 있다. 결국 axiom bridge 패턴은 비일관성의 문제를 해결하기 위한 방안으로 활용될 수 있다.



[그림 4] transitive parent패턴 예

■ common data property enforcer 패턴

미국회 도서관 분류체계(LOC)라는 도메인 내에서 과학관련분야 카테고리는 항상 'Q'임을 명시하기 위하여, OWL의 hasValue를 적용할 수 있다. OWL의 owl:hasValue는 클래스의 어떤 인스턴스가 특별한 특성을 위한 명확한 값을 가진다는 것을 나타낸다. common data property enforcer 패턴은 한정된 도메인 영역 내에서 특정 어휘의 사용용례를 제한적으로 명시화함으로써, 충돌(Conflict)로부터 유발되는 모호성 문제를 해결하기 위한 방안으로 활용될 수 있다.

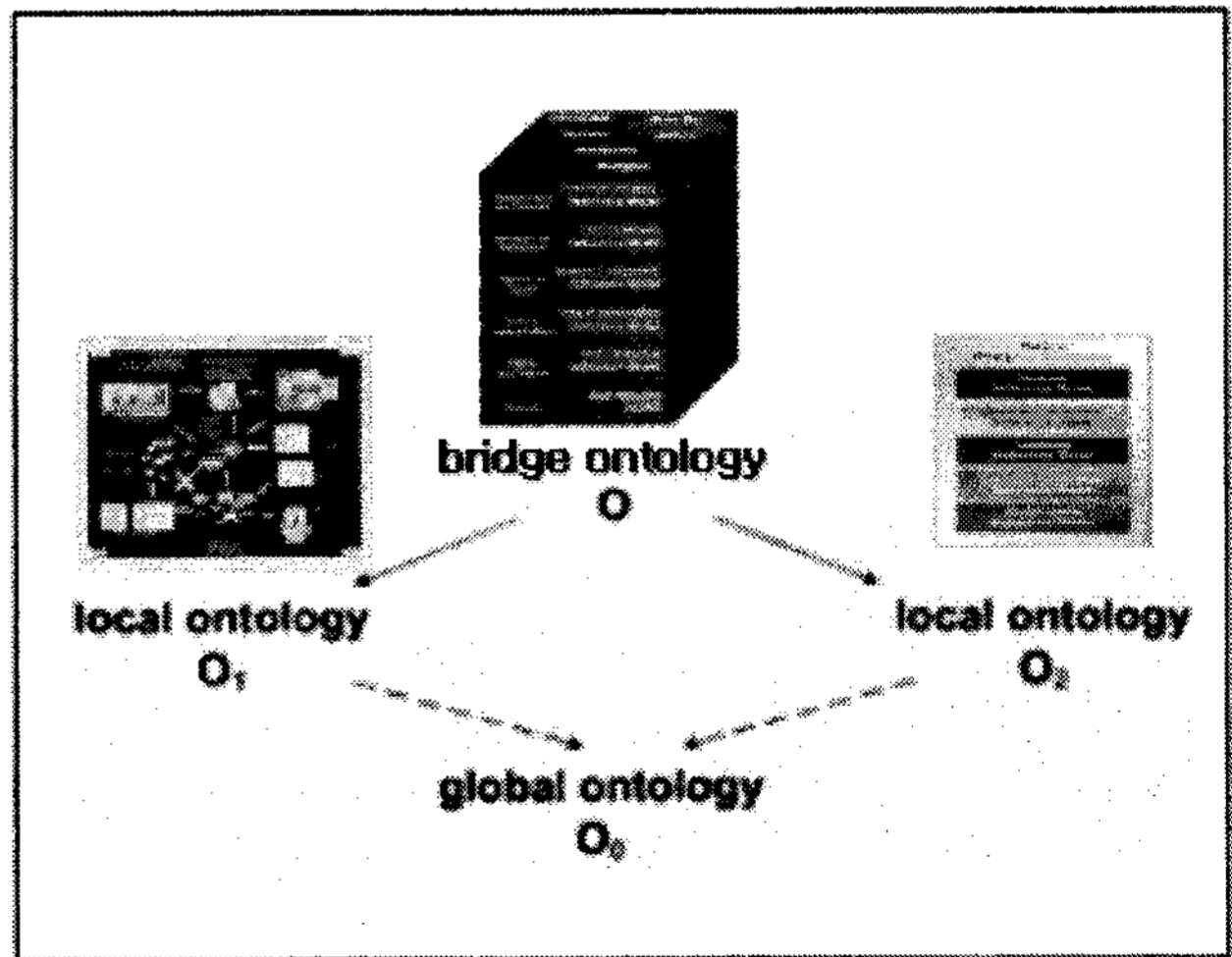
■ transitive parent 패턴

묵시적인 정보를 도출해내기 위해 RDF의 rdfs:subPropertyOf와 OWL의 owl:TransitiveProperty 정의를 함께 결합하여, hasAncestor라는 내포되어진 특성을 hasParent로부터 유도해낼 수 있다.

[그림 4]의 예제에서 Harry가 Charles이라는 부모를 가지고 있고, Charles가 Elizabeth이라는 부모를 가지면 Harry는 Elizabeth라는 조상을 가지고 있음을 유추해 낼 수 있게 된다. 이러한 transitive parent패턴은 생략(omission)에 의해 유발되는 모호성을 해결하기 위한 방안으로 활용될 수 있다.

4. Bridge 온톨로지 모델

지금까지 EA상호운용성 문제를 해결하기 위한 방안으로 EA 온톨로지 개발 및 활용방안에 대해 살펴보았다. 다음은 EA 온톨로지를 토대로 추론 때 카니즘을 통해 EA상호운용성 문제를 해결하는 EA-bridge 온톨로지 모델을 개발, 제안하고자 한다.



[그림 5] Bridge 온톨로지 모델

서로 다른 유형의 EA 온톨로지 모델을 하나의 구조로 통합하는 방법은 크게 두 가지 유형으로 구분해 볼 수 있다. 첫째, 표준 EA온톨로지 모델의 클래스를 정의하고, 표준 EA온톨로지의 서브클래스로서 각 기관별 EA를 구축하는 방법이다. 대부분의 공공기관의 경우, 미연방정부 EA모델

(FEA)을 표준모델로 하여 각 기관별 EA모델을 구축하는 형태로 얼핏 보기에는 바람직해 보이나, 전혀 상이한 형태의 EA모델이 요구되는 미 국방부 경우에는 첫 번째 방안을 적용하기가 어렵다는 문제점을 가지고 있다. 무엇보다도 EA 온톨로지 모델이 서로 매우 이질적인 경우, 변환된 결과가 전혀 다르게 왜곡된 형태로 나타날 수 있다. 또 다른 방안으로는, [그림 5]에서처럼 각각의 EA 온톨로지 모델들을 별도로 유지하고, 중간 매개체 역할을 하는 bridge 온톨로지를 매개체로 양쪽의 매타모델을 매핑시키는 방안이 있다. 이미 미 연방정부 EA와 미 국방 EA간 상호운용성 문제를 해결하기 위하여 DoD EA RM이라는 EA상호운용성 모델이 개발, 보완 중에 있으며[DoD EA RM, 2005], DoD EA RM을 토대로 Bridge 온톨로지를 개발함으로써 서로 다른 EA간 상호운용성 문제를 해결할 수 있으리라 예상한다.

각 EA온톨로지 별로 주요 핵심개념이 클래스형태로 정리된다면 OWL의 sameClassAs를 통해 여러 온톨로지 간의 Bridge Ontology정보를 생성해 낼 수 있다. 이는 A, B(Bridge), C라는 세 가지 온톨로지가 존재할 때 sameClassAs(A[X], B[Y])와 sameClassAs(B[Y], C[Z])에서 sameClassAs(A[X], C[Z])를 추론해 낼 수 있기 때문이다. 또한, 다수의 온톨로지 간의 복잡한 연관정보를 일일이 저장, 관리할 필요도 없는데 sameClassAs(A[X], B[Y])와 sameClassAs(B[Y], C[Z])가 존재할 때 sameClassAs(A[X], C[?])와 같은 형태의 질의를 통해 sameClassAs(A[X], C[Z])와 같은 Bridge Ontology정보를 자동으로 도출해 낼 수 있기 때문이다.

## 5. 결론 및 향후계획

금년부터 모든 공공기관의 EA 구축 및 활용을 강제화하는 법률안이 통과됨에 따라 공공분야 내에서만 약 200여개가 넘는 EA가 구축, 활용되어질 예정이며, 기관별 EA구축이 완료되는 2008년 시점부터는 EA활용고도화 측면에서 각기 서로 다른 EA상호운용성 문제가 핵심 화두로 떠오르리라 예상된다. 이러한 EA상호운용성 문제를 해결하기 위

한 방안으로 EA 온톨로지 개발의 필요성 및 활용 방안에 대해 설명하였으며, 온톨로지를 토대로 추론 매카니즘을 통해 실제 EA상호운용성 문제를 해결하는 구체적인 방안으로 Bridge 온톨로지 모델을 새롭게 제안하였다. 앞으로 이번 개념연구를 기반으로 공공기관 EA 표준산출물관리시스템인 ITAMS와 실제 연계, 운용이 가능한 온톨로지 상호운용성 지원시스템을 설계, 개발해나갈 예정이다.

## 참고문헌

- [대한민국국회, 2005] 대한민국국회, "정보시스템의 효율적 도입 및 운영 등에 관한 법률안", 법률 제7816호, 2005.12.
- [UCS, 1999] US Customs Service, "Enterprise Architecture Blueprint", Technical Architecture Group, 1999.
- [황상규, 1999] 황상규, 윤세진, 오경목, 변영태, 어휘의미 중의성이 인터넷기반 정보검색에 미치는 영향, 제6회 한국정보관리학회 학술대회 논문집.1999.
- [BEA, 2006] Business Enterprise Architecture, DoD BTA, 2006.
- [Berners-Lee, 2003] Berners-Lee, Tim. "WWW Past, present and future".
- [FEA RMO, 2005] FEA Reference Model Ontologies, TopQuadrant White Paper, 2005.
- [DoD EA RM, 2005] DoD Enterprise Architecture Reference Models, Department of Defense, 2005.