

# A Study on Elicitation Procedures of the Entity for Data Model

Doyu Kim<sup>†</sup> · Jeongmo Yeo<sup>††</sup>

## ABSTRACT

The data model that can be said as skeleton of the information system constitutes important 2 axes in the information system together with the process model. There is entity, properties, relation as key factors of the data model, and entity is the most fundamental factor in the data model, and thus total data model becomes vague if not deriving entity definitely. This study dealt with entity deduction only. Deducing methods of existing entity depended on experiences, task knowledge of designers and clear procedures were not suggested, so there were many difficulties in approaching them from beginners or unskilled persons. For giving helps in solving the problem, this study proposes entity-deducing procedures based on tasks that can derive entity with a systematic process at previously derived target businesses through suggested methods from advancing researches. And the study enabled proposing procedures on imaginary tasks to be applied, objecting to undergraduates who had not experiences on the data modeling, and then verified suggesting process through a similarity checking between best answers with deduced entity by students after taking impossible points of comparing existing methods with suggesting process into consideration. By doing so, deducing entity closely to the best answer was confirmed accordingly. Therefore, a fact could be confirmed that beginners were able to deduce entity closely to the best answer even if letting beginners who had not experiences on the data modeling be applied to unfamiliar tasks. Regarding researches on properties and relation deduction besides entity, this study leaves them to next time.

**Keywords :** Entity, Data Modeling, Target Businesses, Elicitation

## 데이터 모델을 위한 엔터티 도출 절차에 관한 연구

김도유<sup>†</sup> · 여정모<sup>††</sup>

## 요약

정보 시스템의 골격이라고도 할 수 있는 데이터 모델은 프로세스 모델과 함께 정보 시스템에 있어서 중요한 두개의 축을 이룬다. 데이터 모델의 핵심요소로는 엔터티, 속성, 관계가 있으며, 이 중에서도 엔터티는 데이터 모델에서 가장 근본적인 요소로서, 엔터티를 명확하게 도출하지 못한다면 데이터 모델 전체가 모호해지게 된다. 본 연구에서는 엔터티 도출에 대해서만 다루었다. 기존의 엔터티를 도출하는 방법은 설계자의 경험과 업무지식에 많이 의존되고 명확한 절차가 제시되지 않아, 초보자나 미숙련자가 접근하기에는 많은 어려움이 있다. 이를 해결하는데 도움이 될 수 있도록 본 연구에서는 선행연구의 제안 절차를 통해 미리 도출된 대상업무에서 체계적인 절차로 엔터티를 도출할 수 있는 업무기반 엔터티 도출 절차를 제안한다. 그리고 데이터 모델링에 경험이 없는 학부생들을 대상으로 가상업무에 대하여 제안 절차를 적용하도록 하였고, 기존 방법과 제안 절차의 비교가 불가능하다는 점을 감안하여 학생들이 도출한 엔터티와 모범 답안 간의 유사도 검사로 제안 절차를 검증하였다. 그 결과, 모범 답안에 상당히 근접하게 엔터티를 도출한 것을 확인하였다. 따라서 본 연구에서 제안한 절차가 데이터 모델링에 경험이 없는 초보자가 익숙하지 않은 업무에 적용하더라도 모범 답안에 근접하게 엔터티를 도출할 수 있음이 확인되었다. 엔터티를 제외한 데이터 모델의 핵심요소인 속성과 관계의 도출에 대한 연구는 차후로 미룬다.

**키워드 :** 엔터티, 데이터 모델링, 대상업무, 도출

## 1. 서론

시스템의 골격이 되는 데이터모델에서 엔터티는 가장 중요한 요소이다. 엔터티를 도출하는 과정은 데이터 모델링에서 가장 먼저 수행되어야 하며, 모델링 작업에서 가장 핵심적인 부분이라고 할 수 있다[1]. 만약 이러한 엔터티가 명확하지 못하고 모호하다면 그 다음 과정이 아무리 완벽하다고

<sup>†</sup> 준회원: 부경대학교 컴퓨터공학과 석사과정

<sup>††</sup> 정회원: 부경대학교 컴퓨터공학과 교수

논문접수: 2012년 9월 11일

수정일: 1차 2012년 12월 26일

심사완료: 2013년 1월 16일

\* Corresponding Author: Doyu Kim(dykim729@naver.com)  
Jeongmo Yeo(yeo@pknu.ac.kr)

하더라도 데이터 모델 전체가 모호하게 설계될 수밖에 없다 [2]. 그만큼 엔터티가 중요하기 때문에 엔터티 도출은 매우 주의 깊게 진행되어야 한다.

그러나 엔터티를 결정하는데 있어 정확한 판단 기준을 가지고 정확한 엔터티를 도출해 내기란 무척 어려운 일이다. 먼저 엔터티를 정확하게 도출하기 위해서는 해당 업무에 대한 지식이 있어야 한다[3]. 뿐만 아니라 엔터티를 정확하게 도출하기 위해서는 매우 전략적인 판단이 요구된다. 논리적으로만 생각한다면 어떠한 개체들을 묶어서 만든 집합이라도 엔터티가 될 자격이 있다. 그러나 엔터티를 어떤 집합으로 정의하느냐에 따라서 이후의 데이터 모델링에 미치는 영향은 매우 크다[4,5]. 얼마나 합리적이고 전략적으로 엔터티를 도출하느냐에 대한 문제는 보다 높은 차원의 문제이기 때문에 이해시키기도, 익혀서 활용하기에도 결코 쉽지 않다.

기존의 엔터티 도출 방법은 설계자의 많은 설계 경험과 업무지식을 필요로 한다[1,6]. 또한 체계적인 절차가 없을 뿐만 아니라 업무가 도출되지 않은 상태에서 시작되기 때문에 경험이 없는 초보자가 사용하기에는 많은 어려움이 있다. 초보자는 업무에 대한 지식이 없으므로 업무가 도출되지 않은 상황에서 데이터 모델링을 진행하기 어렵다. 그러므로 초보자가 데이터 모델링에 대한 개념과 경험을 습득하기 위해서는 업무가 완전히 도출된 상태에서 데이터 모델링을 시작하는 방법이 필요하다. 하지만 이러한 방법에 대해서 현재 나와 있는 방법이 없다. 이러한 이유로 인하여 유독 데이터 모델에 대한 전문성이 프로그램의 성숙도에 비해 잘 발전되지 않는 경향이 있다[7].

또한 현실 세계의 업무에 대한 데이터 모델링 방법이나 결과는 정답이 있을 수 없으며, 데이터 모델링 결과가 얼마나 업무를 잘 반영하고 있으며, 데이터 모델링 방법이 얼마나 효율적으로 진행될 수 있는가가 중요하다.

따라서 본 연구에서는 이러한 문제점을 해결하는데 도움이 될 수 있도록 대상업무가 명확하게 도출된 상태에서 시작하여 엔터티를 도출할 수 있는, 초보자들이 사용하기에 적합한 엔터티 도출 절차를 제안하고자 한다.

## 2. 관련 연구

### 2.1 업무기반 대상업무 도출 방법

먼저 업무기반 엔터티 도출을 위해서는 선행연구에서 제안한 업무기반 대상업무 도출 방법[9]을 이용하여 대상업무가 도출되어 있어야 한다. 업무기반 대상업무 도출 방법은 다음과 같이 5단계로 진행된다.

#### 단계1 : 업무후보 도출 및 1차 정제 단계

단계1에서는 수집된 요구사항을 문장 형태로 기술한 업무 기술서의 각 문장을 읽어가면서 문장 안에 있는 동사나 명사형 동사를 기준으로 업무후보들을 도출한다. 도출과정 중에 나타나는 시스템행위들은 업무행위에 관련된 후보를 제외하고 나머지는 제거한다. 그리고 복합적인 행위를 포함하

는 후보는 모두 단일 행위들로 분리하여 도출한다. ‘업무후보명’은 ‘행위대상+행위’의 형태로 기술하고, ‘업무후보 설명’은 행위를 명확하게 파악할 수 있도록 업무기술서의 내용을 바탕으로 육하원칙을 적용하여 기술한다. 그리고 데이터를 의미하는 부분도 같은 방법으로 데이터 업무후보를 도출하고, 도출된 모든 후보를 검증한다.

#### 단계2 : 업무 데이터 표준화 단계

단계2에서는 도출된 데이터에 대하여 업무 데이터 표준화하는 작업을 진행한다. 애플리케이션의 설계와 구축에 관련된 모든 이해당사자들 사이에 의사소통을 원활히 하도록 업무단어와 업무용어에 대하여 데이터 사전을 구성하고 적용한다.

#### 단계3 : 업무후보 2차 정제 단계

단계3에서는 동일한 의미의 행위 업무후보들을 모아서 그룹핑한다. 각 그룹에서 대표를 선정하고 대표가 아닌 후보는 제거한다. 그 다음에 데이터 업무후보에 대해 또 같은 방법으로 2차 정제를 진행한다.

#### 단계4 : 대상업무후보 전환 단계

행위 업무후보들을 업무가 발생하는 순서대로 재배치한다. 이 중에 업무의 발생순서와 관련 없거나 대부분의 업무에 공통적인 업무후보들은 마지막부분에 위치하도록 한다. 업무후보가 빠짐없이 도출되었는지를 검증하기 위해서 데이터 업무후보들을 각각 자신을 사용하는 행위 업무후보 아래로 이동시키고 검증한다. 그리고 나서 다시 한 번 그룹핑으로 중복되거나 필요 없는 업무후보를 제거한다. 남아 있는 행위 ‘업무후보명’과 ‘업무후보 설명’을 ‘대상업무후보명’과 ‘대상업무후보 설명’으로 전환한다.

#### 단계5 : 대상업무 확정 단계

단계4의 결과에 대해서 유사한 대상업무후보들을 그룹핑하여 대표 대상업무후보를 선정한다. 대표가 되지 않은 것들을 대표 아래에 들여쓰기 한 후에 대표 대상업무후보를 그룹을 대표할 수 있도록 보완한다. 또한 그룹들도 좀 더 넓은 의미로 그룹핑될 수 있으면 그룹핑하여 계층구조로 구성할 수 있다. 개념과 의미가 명확해진 ‘대상업무후보명’과 ‘대상업무후보 설명’을 각각 ‘대상업무명’과 ‘대상업무 설명’으로 확정한다. 확정된 결과물을 근거로 사용자와 면담하여 검증하여 보완하도록 한다.

보완하여 확정된 결과들은 대상업무가 된다. 이렇게 도출된 대상업무를 통해서 업무분할도를 작성하고 데이터 모델링의 엔터티 도출, 속성 도출, 관계도출 등의 단계에서 업무의 범위를 규정해 주고 참조할 근거자료가 된다. 본 논문에서 제안하는 업무기반 엔터티 도출 방법은 대상업무가 확정된 상태에서 시작하기 때문에 업무기반 대상업무 도출 방법을 통해서 대상업무들을 확정하는 것이 본 연구의 선행연구이다.

## 2.2 데이터기반 엔터티 도출 방법

현장에서는 대상 업무가 전혀 도출되어 있지 않거나 일부 분만 도출된 상태에서 데이터 모델링을 시작한다. 데이터를 분석하여 업무를 도출하고 새로운 업무를 정의해 나가면서 데이터 모델을 설계해 간다. 현장에서 다루는 정보시스템의 규모가 크므로 프로젝트 팀원 간에 담당할 범위를 나눠서 일을 진행해야하고, 동시에 여러 작업이 진행되기 때문에 업무 분석을 선행하여 완전하게 업무를 도출한 후 데이터 모델링을 진행하기 어렵다. 그러므로 현장에서는 데이터를 분석하여 업무 분석과 데이터 모델링을 동시에 진행하는 데이터기반 엔터티 도출 방법을 사용한다. 데이터기반 엔터티 도출 방법에서는 체계적으로 정의된 절차 없이 요구사항을 수집하면서 엔터티 후보 도출을 함께 진행하는데, 일반적으로 각 요구사항에서 업무 수행 주체, 업무 수행 대상, 업무 수행 결과 등을 검토하여 관리대상인지 판단하고 즉시 엔터티 후보로 도출하기 때문에 설계자의 경험과 업무지식에 많이 의존하게 된다[1].

데이터기반 엔터티 도출 방법은 먼저 기존 시스템의 데이터 모델을 리버스(reverse)한 자료, 기업의 홈페이지, 기업의 장부나 보고서 등의 문서, 사용자 면담, 업무현장 방문, 관련 시스템 자료 참조, 관련 전문서적 참조 등을 통해서 업무를 분석하여 요구사항을 도출하면서 동시에 엔터티 후보를 도출한다. 엔터티 후보가 모두 도출되면 도출된 엔터티 후보들에 대하여 개념과 의미를 파악하여 엔터티의 자격이 있는지 식별하고[8], 엔터티 후보 정제 과정을 진행한다. 엔터티 후보들의 개념을 명확하게 정립하면서 엔터티 후보들 간의 동질성과 독립성을 파악하고 유사한 엔터티 후보들은 통합하고, 여러 개념을 복합적으로 포함하고 있는 엔터티 후보는 분리한다. 이후엔 결정된 데이터 집합들에 대하여 구체적인 서브타입들을 지정하고, 엔터티들이 확정해 나간다.

이러한 데이터기반 엔터티 도출 방법의 장점은 빠른 시간 내 모범 답안에 근접한 데이터 모델을 설계할 수 있다는 것이다. 하지만 이러한 장점을 얻기 위해서는 오랜 시간의 경험과 업무 지식이 반드시 동반되어야 한다. 뿐만 아니라 데이터기반 엔터티 도출 방법은 각 과정에 대한 체계적인 절차가 명확하게 정의되어 있지 않아 데이터 모델링에 경험이 없는 초보자가 사용하기에는 적합하지 않다. 전문적인 데이터 모델 설계자의 경우에는 경험과 업무지식이 많기 때문에 체계적인 절차를 거치지 않아도 쉽게 엔터티를 도출할 수 있다. 하지만 대상 업무도 완전히 도출되지 않은 상황에서 시작하는 데이터기반 엔터티 도출 방법은 데이터 모델링 경험과 업무지식이 없는 초보자가 접근하기에는 많은 어려움이 있다. 데이터기반 엔터티 도출 방법은 절차가 체계적으로 정의되어 있지 않고, 설계자의 데이터 모델링 경험과 데이터 모델을 설계할 기업에 대한 업무지식에 많이 의존하기 때문이다. 데이터 모델링에 대한 개념 지식을 습득에도 어려움이 많은데다 거기에 기존의 방법은 경험과 업무지식을 필요로 하기 때문에 데이터 모델에 대한 전문성을 발전시키기 위해서는 많은 시간과 노력을 필요로 한다.

## 3. 업무기반 엔터티 도출 절차

본 연구에서는 데이터 모델링의 첫 단계인 엔터티 도출과정에서 경험과 숙련의 정도가 낮은 초보자도 어렵지 않게 진행할 수 있는 엔터티 도출 절차를 제안한다.

### 3.1 용어 정의

데이터영역(Data Area)은 데이터 친밀도가 높은 데이터들을 그룹핑한 것으로서, 단위 작업영역이 되고 데이터를 통합할 때 기준이 되기도 한다. 그리고 데이터영역은 데이터 표준과 데이터 활용 및 데이터 품질 단위가 되며, 한 데이터의 오너쉽(Ownership)일 가능성이 높다.

데이터주제영역(Data Subject Area)은 데이터영역을 주제별로 그룹화한 것으로서, 주로 서브시스템 단위가 데이터주제영역이 된다. 데이터주제영역은 규모에 따라 계층구조도 가능하다.

데이터 분할도(Data Hierarchical Diagram)는 시스템 노드, 데이터주제영역, 데이터영역, 엔터티 등을 구조적으로 계층화하여 도식화한 것으로서, 데이터 분할도를 이용하면 정보시스템에서 사용되는 데이터들이 어떠한 구조로 구성되어 있는지 그 개념을 용이하게 파악할 수 있다. 데이터 분할도를 구성하는 각 항목들은 계층 구조를 가질 수 있지만, 데이터들 사이의 관계는 표현하고 있지 않다.

### 3.2 업무기반 엔터티 도출 절차

본 연구에서는 선행연구[9]에서 제시한 업무기반 대상업무 도출 방법을 이용하여 대상업무가 완전히 도출되어 있다고 가정하고 미리 도출되어 있는 대상업무를 바탕으로 엔터티를 도출한다. 제안 절차는 모두 3단계로 구성되며, 선행과정에서 도출한 대상업무에서 업무를 고려하여 엔터티를 도출하기 때문에 업무기반 엔터티 도출 절차라고 할 수 있다.

제안 절차의 적용 사례는 ‘VOC(Voice of Customer) 시스템’이라는 가상적인 정보시스템을 구축한다고 했을 때 기반이 될 데이터 모델을 설계한다고 가정하고 제안 절차를 적용한 사례로, ‘VOC 시스템’은 사용자요구에 대한 대응과 체계적인 관리 및 조치를 위한 정보시스템이다. Fig. 1은 ‘VOC 시스템’의 요구사항명세서 일부이다.

Fig. 1과 같이 문장으로 기술된 업무기술서를 기준으로 선행연구의 업무기반 대상업무 도출 방법을 적용하여 대상



Fig. 1. A Business Specification Part of ‘VOC System’

Name of the Target Businesses	Description
SR (Service request)	Request help about CI using communication tool by users.
Consulting	Consulting with our representatives about how to use the information system.
Request for alterations	Request change a change in CI by users.
Trouble notice	Trouble notice about the information system by users.
Other requests	
Receiving SR	Receiving User's SR by SD representatives.

Fig. 2. Partial Previously Derived Target Businesses

업무를 도출하고, 이렇게 도출된 대상업무를 바탕으로 엔터티를 도출한다. Fig. 2는 미리 도출된 대상업무의 일부이다.

다음은 제안 절차의 각 단계에서의 작업내용과 적용사례이다.

단계1 : 엔터티후보 도출

엔터티를 도출하기 위하여 먼저 엔터티가 될 수 있는 엔터티후보를 도출이 필요하다. 엔터티후보 도출을 위하여 미리 도출된 대상업무를 정리한다. 엔터티 도출 이전 단계인 업무분석 단계에서 확정된 대상업무들을 데이터 업무후보들과 함께 복사하고 '엔터티후보' 란과 '육하원칙 적용 엔터티후보/데이터' 란을 추가한다. 또한 데이터 업무후보의 검토를 위해 해당 대상업무 아래로 이동시킨다.

준비 작업이 끝나면 행위나 행위결과를 엔터티후보로 도출한다. 이렇게 도출한 엔터티후보는 업무를 수행하는 행위를 표현한 행위 엔터티후보가 되며, 이후에 행위 엔터티후보에 육하원칙을 적용하여 또 다른 엔터티후보를 도출하므로 가장 먼저 진행한다. '대상업무명' 란과 '대상업무 설명' 란에서 행위를 찾아 엔터티후보로 도출하고 '엔터티후보' 란에 기재한다. 행위결과가 불명확한 경우에는 행위만을 기재

하고 행위와 행위결과를 함께 도출한 경우에는 '행위/행위결과' 형태로 기재한다.

이렇게 행위/행위결과로 도출된 엔터티후보에 육하원칙을 적용하면 해당 행위의 주체, 대상, 시간, 장소, 방법, 근거 등을 나타내는 엔터티후보를 도출할 수 있다. 행위/행위결과에 대하여 육하원칙을 적용하여 도출되는 엔터티후보는 '육하원칙 적용 엔터티후보/데이터' 란에 기재한다. 가능한 한 육하원칙을 직접 적용한 내용, 명확한 내용, 확인된 내용 등을 바탕으로 엔터티후보를 도출하여야 한다.

그리고 엔터티후보를 도출하면서 엔터티후보 간에 관계가 발견되면 표시하여 엔터티후보 간의 관계 파악에 도움이 되도록 한다. 엔터티후보에 종속 관계가 있는 경우는 '부모>자식'의 형태로 '엔터티후보/데이터집합'란에 표기한다. 만약 상호간에 종속 관계가 있는 경우는 '<>'기호를 사용하여 표기한다.

다음은 업무분석 단계에서 도출된 데이터업무후보에서 엔터티후보를 도출한다. 먼저 데이터업무후보에서 엔터티, 속성, 서브타입이 함께 모여 있기 때문에 데이터업무후보가 엔터티, 속성, 서브타입 중 어디에 속하는지 파악한 후, 종류별로 정리하는 작업이 필요하다. 정리된 데이터에서 엔터티로 판단한 엔터티후보를 도출하여 '엔터티후보'란에 기재한다. 슈퍼타입과 서브타입의 표현에서 슈퍼타입명, 서브타입 세트 구분 속성명, 서브타입명 등을 추출하여 '슈퍼타입명.서브타입세트 구분 속성명(서브타입명1, 서브타입명2, ...)'형태로 기재한다.

부가적 처리 작업이 있는 경우에는 추가적으로 엔터티후보가 도출될 수 있으므로 '<부가처리명>'형태로 '데이터'란에 표기하고 부가적인 처리에 사용된 엔터티후보를 '엔터티후보'란에 기재한다. 명시되지 않은 내용이지만 처리해야 하는 내용이 필수적으로 관리되어야 한다면 도출하여 표현한다.

Fig. 3은 적용 사례에서 단계1 결과의 일부이다.

단계2 : 엔터티후보 정제

Name of the Target Businesses	Description	Entity Candidate	Entity Candidate / Data
SR(Service request)	Users request help on CI using communication tools	Requests/SR	Users communication tools, CI
A formation of communication tools	Communication tools are Call, e-mail, screen for process requests		A(Communication tools):V(Call/e-mail screen for process requests)
A formation of requesting alteration	The alteration request include a requested content and a demanded handling date	request for alteration	E(request for alteration) A(Contents, expected process date)
A formation of Hindrance reports	Hindrance reports include a date, content and classification of the hindrance.	Hindrance reports	E(Hindrance reports):A(date, content and classification)
Other requests		Other requests	

Fig. 3. A Partial Result in Step 1 of the Example

단계1에서 엔터티가 될 가능성이 있다면 모두 후보로 도출하기 때문에 불필요한 후보를 정제하는 작업이 필요하다. 단계2에서는 도출된 엔터티후보에 대해 자격 검증을 수행하고 역할별로 분류하여 동일하거나 유사한 후보들을 통합한다.

먼저 ‘엔터티후보’ 란과 ‘유하원칙 적용 엔터티후보/데이터’ 란을 복사하여 엔터티후보들을 정제할 준비를 한다. 동일한 대상업무에 대하여 행위 엔터티후보와 행위결과 엔터티후보가 모두 도출된 경우에는 둘 중 하나를 선택한다. 결론적으로 행위가 무엇을 처리하기 위한 것이고 어떤 것이 업무를 수행한 결과로 나타나는지를 파악하여 선택한다. 선택되지 못한 엔터티후보는 제거 표시로 붉은 색 처리한다.

그리고 도출된 엔터티후보를 대상으로 엔터티의 자격여부를 판단하여 엔터티 자격이 없는 경우에는 붉은 색으로 처리한다.

엔터티후보의 엔터티 자격검증이 완료되면, 엔터티후보를 분류하여 엔터티를 확정할 준비를 한다. 엔터티들을 통합하거나 분리하기 위해서는 각 엔터티의 개념과 의미를 먼저 파악하여야 하는데, 이를 위하여 엔터티후보들의 분류가 필요하다. 1차적으로 각 엔터티후보가 어떤 역할을 하는지 역할별로 분류한다. 그러면 각 엔터티후보에 대하여 핵심 엔터티(key entity)인지, 주요 엔터티(main entity)인지, 활동 엔터티(action entity)인지 또는 코드 엔터티인지 파악이 가능하다.

역할별 분류가 완료되면, 2차적으로 동일하거나 유사한 개념의 엔터티후보들을 분류한다. 이렇게 분류된 엔터티후보를 확인해보면 동일한 엔터티후보가 중복 도출되거나 서브타입 엔터티가 도출된 것을 확인할 수 있다. 명칭이 동일하지만 서로 다르게 표현된 엔터티후보들은 대표를 세워 하나로 통합할 수 있다. 서브타입 엔터티가 엔터티후보로 도출된 경우에는 슈퍼타입 엔터티가 도출되어 있을 것이므로, 서브타입들을 해당 슈퍼타입 엔터티에 소속시켜 통합할 수 있다.

Fig. 4는 적용 사례에서 단계2 결과의 일부이다.

Entity Candidate	Cassification
CI alteration.CI alteration basis (SR occurrence, PM, Upgrade, etc)	Main
CI alteration	Main
CI alteration.CI alteration basis(SR occurrence)	Main
Department	Key

Fig. 4. A Partial Result in Step 2 of the Example

단계3 : 엔터티 확정

엔터티는 엔터티의 개념을 어떻게 정립하느냐에 따라서 여러 엔터티로 분리할 수도 있고, 여러 엔터티를 하나의 엔터티로 통합할 수도 있으므로, 개념정립을 명확하게 해야 한다. 단계3에서는 가장 먼저 엔터티후보에 대하여 개념을

Entity Candidate	Subtype
Department	
Employee.Employee code	Full time employee
	Contract worker
	Temporary worker

Fig. 5. A Partial Result in Step 3 of the Example

Entity	Cassification	Definition	Type of the Data Management	Notice
Department	Key	the organization of the company		
Employee	Key	staff working in the company		
Full time employee	Sub	One of the category of contracts for employee	Contract code=1	
Contract worker	Sub	One of the category of contracts for employee	Contract code=2	
Temporary worker	Sub	One of the category of contracts for employee	Contract code=3	

Fig. 6. A Partial Entity Definition Specification of the Example

명확하게 정립하여 엔터티후보의 순수성과 동질성 및 독립성을 파악하고, 엔터티후보들을 통합 또는 분리한다. 그리고 나서 이력 엔터티와 서브타입을 확정함으로써 엔터티로 확정해 나간다. 엔터티후보를 엔터티로 확정하기 위해 이전 과정의 결과에서 제거되지 않은 엔터티후보와 서브타입을 새로운 시트로 복사하여 엔터티를 확정할 준비를 한다.

먼저 유사 엔터티후보의 개념 조정과 함께 엔터티후보의 개념을 정립한다. 엔터티후보의 개념을 정립할 때는 핵심 엔터티, 주요 엔터티, 활동 엔터티, 코드 엔터티 순서로 중요한 엔터티후보의 개념을 먼저 정립하는 것이 좋다.

엔터티후보의 개념을 정립하면서 서브타입을 가지는 엔터티후보를 정리한다. 분류된 엔터티후보를 확인하여, 서브타입 엔터티로 판단되면 서브타입 구분 속성을 찾아 슈퍼타입 엔터티명과 함께 ‘슈퍼타입엔터티명.서브타입 구분 속성명’과 같은 형태로 표시하여 ‘엔터티후보’ 란으로 이동한다.

엔터티후보를 확인하다 보면 하나의 물리적 성격의 엔터티후보와 관계되는 논리적 성격의 엔터티후보가 나타날 수 있다. 예를 들어 ‘사원’이라는 물리적 성격의 엔터티후보와 관계되어 ‘담당’이라는 논리적인 엔터티후보가 나타날 수 있다. 이러한 논리적 성격의 엔터티후보는 동일한 개념의 역할을 하는 것들을 모아 표현한다. 이러한 엔터티후보를 표현하는 방법은 논리적 성격의 엔터티후보에 대응되는

물리적 성격의 엔터티후보를 찾고, 논리적 성격의 엔터티후보를 물리적 성격의 엔터티후보의 서브타입으로 구성하는 방법으로 나타내거나, 논리적 성격의 엔터티후보를 독립적으로 구성하고 물리적 성격의 엔터티후보와 관계를 맺는 방법이 있다.

정립된 엔터티후보의 개념을 확인하여 각 엔터티후보들의 순수성과 동질성 및 독립성을 검증하고, 엔터티후보를 통합하거나 분리한다.

엔터티후보의 통합·분리가 완료되면, M:N 관계를 해소할 것인지를 결정하고, 각 엔터티후보에 대하여 이력을 고려하며, 서브타입을 구체적으로 확정한다. 서브타입을 모두 확정하였으면 엔터티후보가 모두 엔터티로 확정한다.

Fig. 5는 적용 사례에서 단계3 결과의 일부이다.

확정한 엔터티에 대하여 엔터티 정의서를 작성한다. Fig. 6은 적용 사례에서 엔터티 정의서의 일부이다.

각 단계를 진행하면서 사용자 면담을 통하여 중간 과정을 확인할 수 있으며, 최종적으로 도출된 엔터티에 대하여 사용자의 검증이 필요하다.

#### 4. 제안 절차의 검증

본 연구에서 제안한 절차는 초보자를 대상으로 한다. 초보자가 각 절차를 진행하면서 자연스럽게 데이터 모델링에 대한 개념을 잡을 수 있고, 그러면서 모범 답안에 근접하게 엔터티를 도출하도록 하는 방법으로서, 기존 방법으로 소개

Team Number	Key Entity Similarity	Main Entity Similarity	Action Entity Similarity	Whole Entity Similarity
1	86%	80%	67%	75%
2	100%	80%	69%	79%
3	100%	67%	63%	73%
4	100%	67%	71%	80%
5	80%	75%	71%	75%
6	80%	80%	75%	78%
7	100%	100%	75%	87%
8	100%	75%	71%	80%
9	100%	100%	83%	93%
10	75%	67%	63%	67%
11	100%	100%	67%	80%
12	75%	67%	75%	73%
13	80%	67%	67%	71%
14	75%	67%	57%	64%
15	80%	75%	71%	75%
16	80%	80%	75%	78%
17	100%	80%	63%	75%
18	100%	100%	75%	86%
19	100%	100%	71%	83%
20	75%	67%	63%	67%
21	75%	67%	71%	71%
22	80%	80%	63%	72%
23	100%	100%	86%	94%
24	100%	80%	71%	82%
25	100%	100%	83%	91%
Average	90%	81%	70%	78%

Fig. 7. Result of Similarity Checking

했던 데이터 기반 엔터티 도출 방법과는 다른 목적을 가지고 있다. 제안 절차는 기존에는 없는 새로운 방법을 만들어 낸 것으로서, 비교 분석을 할 만한 대상을 찾을 수 없으므로 비교 분석이 불가능하였다. 그러므로 멘토가 도출한 엔터티를 모범 답안으로 하여 학생들이 도출한 엔터티와의 유사도검사로 제안 절차의 검증에 대체하였다.

본 연구에서 제안한 엔터티 도출 절차의 검증에 위하여 데이터 모델링에 경험이 없는 컴퓨터공학과 4학년 학부생들을 대상으로 3~4명으로 구성된 25개의 프로젝트 팀을 구성하고, 각 팀별로 가상의 업무기술서를 바탕으로 본 연구에서 제안한 절차를 적용하도록 하였다. 프로젝트 팀에는 현업에서 전문 IT 컨설턴트로 활동 중인 데이터 모델링 전문가들이 멘토의 역할로 각 팀에 참여하여, 제안 절차로 데이터 모델링을 함께 진행하였다. 프로젝트가 종료된 후 각 프로젝트 팀에서 학생들이 도출한 엔터티와 멘토가 모범 답안으로 도출한 엔터티 간의 유사도를 검사하였다. 엔터티 간의 유사도는 멘토가 도출한 엔터티를 기준으로 학생들이 도출한 엔터티의 확인한 것으로, 기준이 되는 모범 답안의 각 엔터티의 집합과 일치하는 집합을 가진 엔터티가 학생들이 도출한 답안에서 얼마나 존재하는지를 검사한 것이다.

검사 결과, 핵심 엔터티 간의 평균유사도가 90%, 주요 엔터티 간의 평균유사도가 87%, 활동 엔터티 간의 유사도가 70%로 나타났고, 전체 엔터티 간의 평균유사도는 78%의 유사도를 보였다.

Fig. 7은 학생들이 도출한 엔터티와 모범 답안 간의 유사도검사 결과를 보여준다.

### 5. 결 론

본 연구에서는 미리 도출된 대상업무에서 엔터티를 도출하는, 초보자에게 적합한 절차를 제안하였다.

‘VOC 시스템’이라는 가상의 업무용 정보시스템의 대상업무로부터 엔터티를 도출하는 과정을 보였으며, 학생들을 대상으로 프로젝트 팀을 구성하여 본 연구에서 제안한 절차로 가상 업무에서 엔터티를 도출해 보도록 하였다. 그리고 나서 학생들이 도출한 엔터티와 멘토가 도출한 모범 답안 간의 유사도 검사를 실시하였다. 그 결과, 모범 답안에 전체 엔터티를 대상으로 78%의 유사도를 나타냈으며, 이를 통해 제안 절차를 이용하여 모범 답안에 상당히 근접하게 엔터티를 도출할 수 있다는 것을 확인하였다.

따라서 본 연구에서 제안한 절차를 활용한다면 데이터 모델링에 경험이 없는 초보자가 익숙하지 않은 업무에 적용하여도 엔터티를 도출할 수 있음을 보였다. 그러나 제안 절차는 여러 단계를 수행하여야 하기 때문에 학생들이 체감하는 소요시간은 다소 긴 것으로 나타났다. 하지만 업무 경험이 전혀 없는 학생들이 처음 접하는 데이터 모델링의 과정을 익히면서 엔터티를 도출할 수 있었다는 것을 생각한다면 그

렇게 많은 시간이 소요된 것은 아니라고 판단된다. 뿐만 아니라 제안 절차를 경험하여 익숙해진다면 소요시간을 많이 단축할 수 있을 것이다.

본 연구에서 제안한 절차는 초보자의 데이터 모델링 교육을 위한 것으로 경험과 업무지식이 많은 전문가들에게는 적합하지 않다. 하지만 현업에서도 업무를 파악하는 용도로 활용이 가능하다.

또한 본 연구의 엔터티 도출 절차는 데이터 모델 설계의 한 과정으로서, 제안 절차에서 도출된 결과를 이용하여 데이터 모델 설계의 이후 과정인 속성 도출 과정이나 관계 도출 과정을 진행할 수 있을 것으로 기대한다. 이러한 속성 도출, 관계 도출에 관한 연구는 차후로 미룬다.

### 참 고 문 헌

- [1] Hwasik Lee, “Data Architecture Solution I”, En-core, 2003.
- [2] Hwasik Lee, Hyunsan Kim, Haesul Yang, “Expression Method and Application Plan of the Entity for a Data Modeling”, Korea Information Processing Society, Vol.12, No2, 2005.
- [3] Korea Database Agency, “Data Architecture Professional Guide”, Korea Database Agency, 2010.
- [4] Jonghyun Kim, “Information Engineering and Data Modeling”, Korean Institute of Information Scientists and Engineers, Vol.12, No.11, 1994.
- [5] Steve Hoberman, “Data Modeling Master Class”, Technics Pubn Llc, 2011.
- [6] Michael C. Reingruber, “The Data Modeling Handbook”, Intervision, 2002.
- [7] Graeme Simsion, “Data Modeling Essentials”, Morgan KaufmannPub, 2005.
- [8] Jan Speelpenning, “Data Modeling and Relational Database Design”, Oracle, 2001.
- [9] Jeongmo Yeo, Sunyi Park, “A Study on Elicitation Method of Target Businesses for Application Development”, Korea Institute of Information and Communication Engineering, Vol.15, No.12, pp.2599-2608, 2012.



김도유

e-mail : dykim729@naver.com

2011년 부경대학교 컴퓨터멀티미디어 공학과(학사)

2011년~현 재 부경대학교 컴퓨터공학과 석사과정

관심분야: 데이터 아키텍처, ITA/EA



## 여 정 모

e-mail : yeo@pknu.ac.kr

1980년 동아대학교 전자공학과(학사)

1982년 부산대학교 전자공학과(공학석사)

1993년 울산대학교 대학원 전자 및 전산  
기공학과(공학박사)

1986년~현 재 부경대학교 컴퓨터공학과  
교수, (주)엔코아 사외이사

관심분야: 데이터 아키텍처, ITA/EA