

Using Requirements Engineering to support Non-Functional Requirements Elicitation for DAQ System

Kyung-Sik Kim*, Seok-Won Lee**

*Student, Dept. of Computer Engineering, Ajou University, Suwon, Korea

**Professor, Dept. of Software and Computer Engineering, Dept. of Artificial Intelligence,
Ajou University, Suwon, Korea

[Abstract]

In recent machine learning studies, in order to consider the quality and completeness of data, derivation of non-functional requirements for data has been proposed from the viewpoint of requirements engineering. In particular, requirements engineers have defined data requirements in machine learning. In this study, data requirements were derived at the data acquisition (DAQ) stage, where data is collected and stored before data preprocessing. Through this, it is possible to express the requirements of all data required in the existing DAQ system, the presence of tasks (functions) satisfying them, and the relationship between the requirements and functions. In addition, it is possible to elicit requirements and to define the relationship, so that a software design document can be produced, and a systematic approach and direction can be established in terms of software design and maintenance. This research using existing DAQ system cases, scenarios and use cases for requirements engineering approach are created, and data requirements for each case are extracted based on them, and the relationship between requirements, functions, and goals is illustrated through goal modeling. Through the research results, it was possible to extract the non-functional requirements of the system, especially the data requirements, from the DAQ system using requirements engineering.

▶ **Key words:** Requirements Engineering, Data Acquisition System, Goal Modelling,
Non-Functional Requirements, Data Quality Requirements

-
- First Author: Kyung-Sik Kim, Corresponding Author: Seok-Won Lee
 - *Kyung-Sik Kim (steve9242@ajou.ac.kr), Dept. of Computer Engineering, Ajou University
 - **Seok-Won Lee (leesw@ajou.ac.kr), Dept. of Software and Computer Engineering, Dept. of Artificial Intelligence, Ajou University
 - Received: 2020. 12. 23, Revised: 2021. 02. 16, Accepted: 2021. 02. 19.

[요 약]

최근의 기계 학습 연구에서는 데이터의 품질 및 완전성을 고려하기 위해 요구 공학 관점에서 데이터에 대한 비기능적 요구사항 도출이 제안되고 있으며, 특히 기계 학습에서의 데이터 요구사항을 요구 공학자들이 정의하였다. 본 연구에서는 데이터 전처리 이전에 데이터를 수집하고 저장하는 데이터 수집(DAQ) 단계에서 데이터 요구사항을 도출하였다. 이를 통해 기존 DAQ 시스템에서 필요로 하는 모든 데이터 요구사항과 이를 만족시키는 태스크(기능)의 유무, 그리고 요구사항과 기능 간의 관계를 표현할 수 있다. 또한 추출된 요구사항을 체계적으로 표현하고, 그 관계를 정의할 수 있어 소프트웨어 설계 문서의 작성이 가능하고 소프트웨어 설계 및 유지 보수 측면에서 체계적인 접근 및 방향성을 설정할 수 있다. 본 연구에서는 기존 DAQ 시스템 사례를 이용하여 요구 공학적 접근을 위한 시나리오와 유스 케이스(Use case)를 작성하고, 이를 기반으로 사례별 데이터 요구사항을 추출하고 목표 모델링을 통해 요구사항, 기능, 목표 간의 관계를 도식화한다. 연구 결과를 통해 요구 공학을 이용한 DAQ 시스템에서 시스템이 필요로 하는 비기능적 요구사항, 특히 데이터의 요구사항을 추출할 수 있었다.

▶ **주제어:** 요구 공학, 데이터 수집 시스템, 목표 모델링, 비기능적 요구사항, 데이터 품질 요구사항

I. Introduction

최근 딥 러닝을 필두로 기계 학습 시스템을 비롯한 인공지능 시스템에 대한 많은 수요가 있어 다양한 분야에서 개발이 이루어지고 있다. 기계 학습은 인공지능망을 비롯한 다양한 알고리즘을 이용하여 기존에 보유하고 있었던 데이터 집합 내에서 특징을 추출한 다음 새로운 데이터에서 얼마나 성공적으로 해당 특징을 추출하여 보여주는지, 그러한 특징의 결과가 무엇을 의미하는지를 성능의 평가 척도로 한다. 이러한 연유로 기계 학습에 있어서 데이터를 정제하는 과정인 데이터 전처리와 사용자가 원하는 결과를 얻을 수 있는 적절한 알고리즘의 선택이 중요하다. 그리고 지도 학습의 설계 과정에서의 문제점들을 지적하고 있는데, 여기서 데이터 세트를 생성할 때, 필요한 전문가가 있다면 그들로부터 어떠한 속성이나 특징들이 가장 정보를 많이 담고 있는 지 자문을 받을 수 있다고 한다. 만약 그렇지 못한 경우는 모든 특징을 추출하여 그 중에서 올바른 특징이 있기를 바라는 ‘무차별적’ 추출 방식을 가장 간단한 방식으로 소개하고 있다. 그러나 이러한 무차별 방식은 대부분 노이즈 값이나 잃어버리는 값을 포함할 수 있어서 전처리 과정의 중요함을 역설하고 있다.

최근 기계 학습 분야에서도 이러한 데이터 품질을 고려한 설계를 도입하기 위한 연구를 진행하고 있다. 데이터과학자로부터 인터뷰를 진행하여 요구공학자들이 기계 학습 분야에서 새로운 요구사항 항목을 인지하고 정의해야 할 필요가 있다는 점을 인터뷰 결과로 도출하였다. 우선 요구 공학적 프로세스를 향상하기 위해 기계 학습을 어떻

게 이용할 것인지 다양한 시도는 있었지만, 반대로 요구 공학적 프로세스가 기계 학습 분야를 향상하기 위한 접근은 많이 없다는 점을 지적하고 있다. 해당 연구에서는 요구 공학을 적용하기 위해 다음과 같은 3가지의 연구 질문을 정하여 인터뷰를 시행하였고, 그 결과로 다음의 도전 과제들을 정의하였다. [1]

1. 기계 학습 시스템에서 데이터과학자들이 어떻게 요구사항들을 추출, 문서화, 분석을 하였는가?
2. 데이터과학자들이 어떠한 프로세스로 설계를 하였으며, 이 중 요구사항에 관련된 부분이 있는가?
3. 현재 데이터과학자들이 당면한 도전 과제 중 요구사항에 관련된 과제가 있는가?

연구자들은 데이터의 질적 요구사항에서는 데이터의 수집, 데이터 형식, 데이터의 범위에 관한 요구사항들을 정의하고 식별하는 것이 중요하다고 언급하고 있다. 우선 첫 번째 인터뷰 대상자는 다양한 범위의 데이터를 포함해야 하는 완전성, 데이터 세트와 데이터의 표현과 형식이 서로 같아야 하는 일관성, 실제로 사실인 데이터에 의지할 수 있는 정확성, 이 3가지 요소를 가장 중요한 질적 요소라고 답변하였다. 또한 다른 인터뷰 결과에서는 각각 공개 데이터 세트에서의 낮은 신뢰성과 사람에 의한 직접적인 수집과 레이블링의 리스크, 그리고 틀린 데이터가 포함된 데이터 세트로 훈련을 하였을 때의 리스크 감소에 대해 논의하고 있다. 이러한 요구사항과 리스크를 종합하여 기계 학습

에서의 훈련 과정 이전에 수행되어야 하는 과정에 대한 중요성을 공통적으로 언급하고 있다.

제니퍼 호커프(Jennifer Horkoff)는 특히 기계 학습 시스템을 향상하기 위한 비기능적 요구사항의 도전 과제와 연구 방향을 제시하였다.[2] 우선, 모든 기계 학습 연구자들이 해결하고자 하는 문제가 무엇인지는 정의하고 있지만, 문제를 해결하는 데 있어 정확성이나 높은 성능에 대해서 모델 설계 이전에 고려하지 않는다는 문제점을 지적하고 있다. 그리고 최근의 기계 학습 솔루션에서는 정확성, 공평성, 투명성, 보안성, 실험성, 신뢰성과 같은 품질 요소를 고려해야 한다고 역설하고 있다.

데이터 수집 시스템(Data Acquisition System: 이하 DAQ)은 사람이 주변에서 관측할 수 있는 물리적인 현상들에 대해 다양한 디바이스를 이용하여 데이터 형태로 저장하는 시스템이다.[3] DAQ를 이용하여 이와 같이 수집한 데이터는 기계 학습과 같은 다양한 데이터 처리 분야에서 활용될 수 있다. DAQ는 머신 러닝 시스템 구조에는 포함되지 않지만, 데이터의 수집과 데이터 세트의 생성 및 저장을 담당하고 있어, 기계 학습에서의 데이터 전처리 과정과 결과의 성능에 영향을 준다. 대부분의 DAQ 개선 연구에서는 기존의 DAQ에서 시스템 구성 요소의 추가 및 변경과 디바이스에서 수집하는 아날로그 신호 처리 알고리즘의 개선을 통해 기존보다 더 나은 데이터 수집 및 데이터 세트 생성을 통한 성능 개선을 이루었으며, 사용자가 필요로 하는 데이터의 품질 속성에 대한 분석 및 연구는 거의 수행되지 않고 있다. 본 연구는 요구 공학을 이용하여 사용자가 원하는 데이터 품질 속성을 추출하고, 목표 모델링(Goal Modelling) 방식을 통하여 DAQ에서의 비기능적 요구사항을 추출한다. 목표 모델링 기법은 소프트웨어 시스템의 목표를 설정하고, 이를 통해 요구사항을 체계적으로 분석하는 기법으로서, 사전 정의된 데이터 품질 요구사항을 반영하기 위해 사용되거나 기존 DAQ에서 도메인 전문가들과 같은 이해 관계자들의 요구사항을 구하는데 사용되고 있다. 그러나, 이 두가지를 합하여 DAQ 시스템 도메인에서의 데이터 품질 요구사항을 추출하는 연구는 진행되지 않는데, 이로 인해 DAQ 시스템이 수집하고자 하는 데이터의 요구사항과 같은 비기능적 요구사항을 체계적으로 추출하여 설계에 반영하고자 한다. 이러한 비기능적 요구사항은 기존 DAQ 품질과 비교를 통해 기존에는 고려되지 않았던 새로운 품질 속성으로 볼 수 있으며, 요구사항 중심의 DAQ 모델을 제안하여 이와 같은 새로운 품질 속성을 기존 DAQ에서 고려하는 품질 속성 리스트에 추가하여 시스템을 체계적으로 표현한다.

II. Preliminaries

1. Related works

1.1 Data Acquisition System(DAQ)

DAQ는 사람이 주변에서 관측할 수 있는 온도, 습도와 같은 물리적인 현상들을 다양한 센서 디바이스를 이용하여 아날로그 신호 형태로 추출하고, 해당 신호를 아날로그 디지털 변환기를 이용하여 디지털 데이터로 변환하여 데이터베이스에 저장하는 시스템이다[3]. 이와 같이 저장된 디지털 데이터는 사용자가 원하는 정보를 제공한다. 그림 1은 DAQ 시스템의 구조를 보여준다.

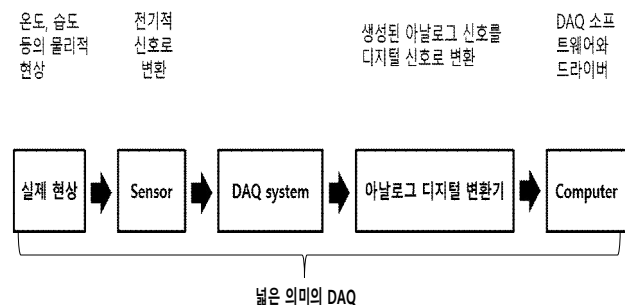


Fig. 1. Process of Data Acquisition System(DAQ)

DAQ는 데이터의 단순한 수집 및 저장을 담당하는 것 이외에도, 데이터를 사용자에게 어떻게 표현할 지 형식을 지정할 수 있다. DAQ를 통해 저장 및 표현되는 데이터는 데이터 전문가에 의해 전처리 과정을 거쳐서 데이터 분석이나 기계 학습과 같은 데이터 처리 분야에서 활용될 수 있다.

1.2 Non-Functional Requirements

소프트웨어 시스템에서의 요구사항은 시스템에 관계가 있는 이해 관계자들이 시스템에 필요하다고 보는 기능이나 조건들을 체계적으로 정리한 것으로, 설계 전에 요구사항 정의를 하여 시스템에 이를 반영할 수 있어야 한다.

요구 공학에서 요구사항은 크게 기능적 요구사항과 비기능적 요구사항 2가지로 분류된다. 기능적 요구사항이 시스템에 의해 제공되는 서비스나 기능에 대해 설명 및 기술하는 반면, 비기능적 요구사항은 시스템 속성이나 시스템에 의해 제공되는 서비스나 기능에 대한 제약사항에 대하여 기술하고 있다. 비기능적 요구사항의 예시로는 접근성, 신뢰성과 같은 요소들이 있으며, 시스템의 기능이나 서비스를 해당 요소들을 어떻게 만족할 지 결정한다. 최근 머신 러닝 연구에서는 요구 공학을 이용하여 이해 관계자들의 요구사항을 분석하여 더 많은 양질의 데이터를 추출하고 추출된 요구사항을 분석하여 성능을 평가하는 부분에

서 확장하여, 양질의 데이터를 얻기 위한 데이터 요구사항 명세와 데이터의 편향 및 비정상적인 데이터를 탐지하는 연구가 진행되고 있다. 요구 공학 연구자들은 데이터 과학자와 인터뷰를 통해 이와 같이 데이터의 요구사항을 파악해야 할 필요성을 거론하고 있다.[4] 본 연구에서의 DAQ의 품질 속성은 비기능적 요구사항을 추출하여 이를 기존 연구에서 고려하지 않은 품질 속성일 경우에는 속성 리스트에 새로 추가한다.

1.3 Goal Modelling

목표 모델링은 요구 공학 분야에서 목표(Goal) 지향 요구공학 기법은 추상적인 단계에서 구체적인 단계에 이르기까지 소프트웨어 시스템이 달성해야 하는 목표를 중심으로 시스템의 요구사항을 탐색 및 구조화할 수 있는 체계적인 기법이다.[5] 또한, 요구 공학 분야에서 높은 수준의 이해 관계자의 요구 사항을 포착하는 데 효과적인 방법론으로, 시스템이 이해 관계자 요구 사항을 충족시키기 위해 해야 할 작업을 파악하고, 그것들이 요구사항과 어떠한 관계를 가지는지 분석할 수 있다. 요구 공학에서의 목표는 요구사항 추출 및 명세화, 비즈니스 및 조직적인 문제 해결, 요구사항 갈등 해결, 시스템 디자인과 같이 다방면에서 유용하다.[6] 또한 목표는 고객의 요구사항(니즈)이나 우선순위, 상황이 변화함에 따라 수정되거나 진화할 수 있기 때문에 이러한 환경에 따라 목표를 어떻게 분석하고, 목표가 수정되면 요구사항이 어떻게 변화하는 지 연구되었다. 우선, 목표는 시스템에서 최종적인(High-level) 목표에 해당하며, 해당 시스템의 개발이 필요한 이유가 된다. 동시에 시스템 내의 다양한 단계에서 의사결정의 방향을 알려준다. 요구사항은 제안된 시스템에서 이러한 목표를 어떻게 달성해야 하는 지 명세하고 있다.[7] 목표 모델링에서는 달성해야 하는 최종 목표를 가장 위에 두고 그 아래에는 최종 목표를 달성하기 위해 만족해야 하는 세부 목표들을 분류한다. 또한 세부 목표의 아래에는 이를 달성하기 위해 필요한 소프트웨어 요구사항들을 정의하여 최종적으로 요구사항과 목표 간의 관계를 ++(생성), +(도움), -(손상), --(침해)와 같은 4가지의 형태로 표현 할 수 있다. 본 연구에서는 목표 모델링 기법 중에 목표를 달성하기 위한 조건이나 상황을 결정하기 위한 소프트웨어 골 모델링 방식을 사용하여 시나리오와 유스 케이스(Use case)들을 작성하여 이를 통해 요구사항들을 추출하였다.

III. The Proposed Scheme

3.1 List of Data Quality Attributes

제안 모델을 설명하기에 앞서 데이터 품질을 먼저 정의하였다. 우선 데이터 품질을 결정하는 요소에 대한 논의는 오래 전부터 존재하였으며, 2005년에 국제 데이터 품질 관련 표준 ISO/IEC 25012를 제정하였다. ISO/IEC 25012에서는 개발이나 운영에서 사용되는 모든 데이터의 품질에 대한 데이터품질 평가 모델을 제시하며, 아래 표 1와 같이 16개의 데이터 품질 평가항목을 통해 품질평가를 가이드하고 있다. [8]

Table 1. The List of Data Quality Attribute from ISO/IEC 25012 (Cited from [8])

No.	Data Quality Attribute	No.	Data Quality Attribute
1	Consistency	9	Understandability
2	Currentness	10	Efficiency
3	Completeness	11	Changeability
4	Precision	12	Portability
5	Accuracy	13	Traceability
6	Security	14	Creditability
7	Availability	15	Accessability
8	Recoverability	16	Compliance

그러나 [8]에서의 데이터 품질 평가항목은 개발이나 운영 시스템에서 나올 수 있는 모든 데이터품질 평가 항목들을 집합하여 평가 모델을 만들었기에 특정 도메인이나 시스템을 대상으로 한 것이 아니다. 이를 해결하기 위해 특정 도메인에서의 DAQ 시스템을 대상으로 데이터 요구사항을 추출할 수 있는 접근 방법이 필요하다.

또한 데이터 품질을 결정하는 요소들을 정확하게 추출하려면, 어느 도메인에서 어떠한 데이터를 얻고자 하는 지가 먼저 정의되어야 한다.

[9]에서는 데이터의 수집과 축적을 넘어서 데이터를 지식화하여 효과적으로 고객이 원하는 고부가가치를 도출하기 위하여 데이터품질 문제를 해결해야 한다고 지적하고 있다. 연구자들은 데이터품질의 확보를 위한 ISO/IEC 25012와 같은 국제 표준을 통해 데이터품질의 측정과 평가가 이루어져야 하지만, 소프트웨어 품질 표준에 비해 부족하다고 언급하고 있다. 이들은 데이터품질 평가속성에 가중치와 우선순위를 정량적 기법으로 도출하기 위하여 심리분석 방법인 RGT와 계층분석과정을 사용하였다. 그 결과 고객의 요구사항에 기반한 데이터 품질 속성 및 우선순위를 도출하였으나 모든 데이터 수집 과정에서 데이터 요구사항을 포괄적으로 도출하였기에 각각의 도메인에 개

Table 2. Improved Data Quality DAQ Modeling using Goal Modeling

단계	내용	산출물
목표 및 요구사항 도출	목표 모델링을 이용한 요구사항 식별	목표 모델링을 이용한 기존 DAQ의 기능적/ 비기능적 요구사항 식별
DAQ 시스템 속성 분석	기존 DAQ 시스템의 품질 분석	기존 DAQ 시스템이 가지고 있는 품질 속성을 분석 열거된 속성 중 데이터 품질 관련 속성을 분류
최종 모델	데이터 품질 속성을 포함하는 새로운 DAQ 모델 제안	전체 요구사항 리스트 중 기존 DAQ 시스템 품질 속성에 포함되지 않았던 요구사항을 추가 데이터 품질에 관련된 요구사항 명세

별적으로 적용할 수가 없다. 따라서 본 연구에서는 다음과 같이 제안 모델을 통해 특정 도메인에서의 데이터 품질 요구사항을 파악하고, 이를 요구 공학을 이용하여 데이터 품질 요구사항을 보완한 보다 완전하고 뛰어난 DAQ를 설계할 수 있다. 또한 제안 모델을 통해 요구 공학에서의 목표 모델링 기법을 활용한 데이터 품질 요구사항 도출과 목표 모델링을 사용하지 않았던 기존의 DAQ 시스템에서의 요구사항 도출을 비교한다.

본 논문에서는 요구 공학 중심의 DAQ 모델을 제안한다. 제안 모델은 크게 3단계로 분류할 수 있는데, 처음에는 기존 DAQ 시스템에서 목표를 설정하여 이를 달성하기 위한 요구사항들을 추출한다. 그 결과, 기능적/비기능적 요구사항 구분없이 모든 요구사항 리스트를 추출해낼 수 있다. 두 번째로는 기존의 DAQ 시스템에서 품질에 관계된 속성을 분석한다. 기존 DAQ 시스템 사례를 가지고 와서 가지고 있었던 품질 속성을 추출하는데, 이 때 속성 중에 데이터 품질에 관련된 속성이 있으면 따로 분류하여 목표 모델링을 통한 요구사항과 비교 및 분석이 가능하게 한다. 마지막으로 목표 모델링 결과로 도출된 요구사항과 기존의 DAQ가 가지고 있었던 품질 속성을 통합하여 데이터 품질을 고려하는 개선된 형태의 DAQ 시스템을 제안한다. 여기서 새로 추가된 요구사항들은 주로 품질에 관련된 요구사항이며, 이들을 기존 DAQ 시스템 속성에서 고려하지 않았던 새로운 속성으로 추가할 수 있다.

목표 모델링을 이용하여 사용자의 목표와 이를 이루기 위해 사용자를 포함한 다양한 이해 관계자의 요구사항을 효과적으로 포착하고, 기존 DAQ 설계 시스템에서 데이터 품질 속성을 추가로 추출한다. 이러한 속성은 기존 DAQ의 목적을 달성하기 위해 어떠한 데이터를 추출해야 하는 지 조건이나 제약사항을 표현하고 있다. 제안한 방법은 목표 모델링 기법을 이용하여 기존의 DAQ가 달성하고자 하는 목표를 설정하고, 그 목표를 달성하기 위한 세부 목표와 최종적으로 이를 만족해야 하는 요구사항을 정의한다. 기존 DAQ에서는 설계 과정에서 목표를 달성하기 위한 기능

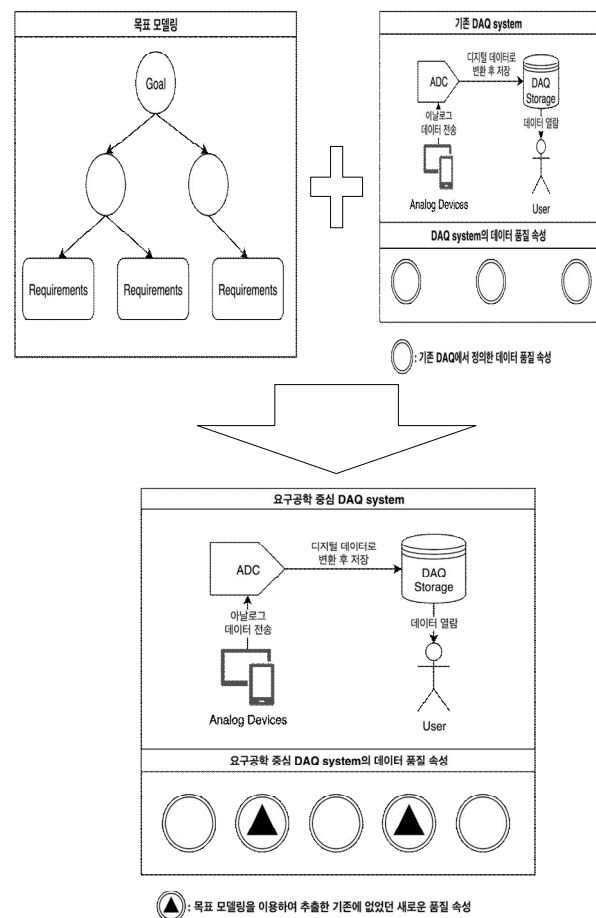


Fig. 2. DAQ Model Proposal using Requirements Engineering Methods

적 요구사항을 대부분 다루고 있으며, 비기능적 요구사항에 대해서는 거의 다루고 있지 않거나 만족하고 있다는 전제 하에 기능적 요구사항을 만족하기 위한 설계를 진행하고 있다. 따라서, 사례를 통해 기존 DAQ에서 비기능적 요구사항들을 정의한 후 기존 DAQ에서 만족하고 있는 데이터 품질 속성을 정의하고, 이를 목표 모델링의 결과로 추출된 요구사항들과 비교하여 기존 DAQ에서 정의하고 있지 않은 요구사항을 새로운 데이터 품질 속성으로 추가한다. 그 결과 해당 DAQ에서 이해 관계자의 필요를 만족하

는 DAQ를 만들기 위해 필요한 데이터 품질 속성을 모두 정의하여 DAQ의 개선 과정에서 수집해야 할 데이터가 품질을 향상할 수 있도록 한다. 제안 모델에 대한 이해를 돕기 위해, 다음 장에서는 예시 사례를 통해 요구공학 중심의 데이터 품질 속성 목표 모델을 제안한다.

3.2 Case Study

상야 빅데이터 의학프로젝트 (Xiangya Medical Big data Project)에서 수집한 다양한 형태의 의료 데이터를 하나의 플랫폼에서 저장하고 열람이 가능한 통합 데이터 DAQ 시스템을 개발하였다.[10] 또한 가상 시나리오에서는 간암을 측정에 사용되는 EMR, RIS 등 다양한 형태의 데이터를

통합하여 데이터 교환 및 수집을 구현할 수 있도록 다양한 의료 서비스 제공 업체의 임상 데이터를 통합 및 공유하는 것을 가정하였다. 시스템 품질 향상을 위해 설문과 인터뷰를 통해 의료 관계자들로부터 피드백을 받았으며, 두 단계 모델을 통해 메타 데이터 수집 시스템 내의 지식 수준을 의료 도메인 전문가들이 정의하는 단계와 개발자들은 도메인 전문가의 정의 형태에 맞게 자동적으로 양식을 개발하는 단계, 총 2단계로 나누어 개발하였다. 해당 시스템의 설명을 이용하여 구성한 가상 시나리오의 내용은 표 3과 같다.

여기서 사건의 흐름은 서술한 정상 사건과, 예외 흐름으로 사용된 오진 문제 이외에 다양한 예외 흐름들이 나올 수 있다. 이러한 흐름들을 소프트웨어 공학에서는 '유스 케이스

Table 3. Scenario of Liver Cancer Diagnosis System

Flow of Scenario	Content
Starting Situation	The liver cancer diagnosis system shows the meta data type and input format of biometric data to the user, and receives data from the user.
Normal Flow of Event	Using the data collected from the user, the diagnosis result is shown to the user based on the knowledge predefined by the expert group.
The Wrong Case within the Flow	When a patient who has been diagnosed with liver cancer and whose medical records have been entered his or her latest improved tumor treatment data, and suddenly shows the result that the cancer has been cured.
Concurrent Activities	Takes the existing hospital treatment data and checks the data recording cycle, and if the old data after 3~6 months has passed, the data update is required warning and a message recommending hospital treatment is output. If more than 6 months have passed, the input is restricted with a warning message about updating hospital medical records.
After the Scenario	The user can see the diagnosis result and reason of the system, and the user's result is transmitted to the system and stored as recent medical treatment data.

Table 4. The List of Use Case from Liver Cancer Diagnosis System

	Eliciting Requirements using Use Case
Use Case #1	Liver cancer should be diagnosed using patient data input through a meta-data format defined by a medical professional.
Use Case #2	It is necessary to obtain a sufficient amount of clinical data from existing medical data and use it to diagnose liver cancer diagnosis results and current status.
Use Case #3	As new measures of liver cancer diagnosis are discovered, medical professionals should be able to create or modify meta-data.
Use Case #4	Only licensed medical professionals and patients should be able to view the data in the diagnostic system.
Use Case #5	It is necessary to express the current progression of liver cancer in accurate by establishing an expression format showing stored data.
Use Case #6	Liver cancer treatment data in the existing hospital database must be imported in accordance with the predefined meta-data expression format.
Use Case #7	There should be no omission of existing data in the process of importing hospital treatment data.
Use Case #8	Check the latest update date of hospital treatment data and print out a message box asking for update when discovered data older than 3 months
Use Case #9	Check the latest treatment date of hospital treatment data and restrict input when checking data older than 6 months.
Use Case #10	When updating to the latest hospital treatment data, the input restriction is released to enable input.
Use Case #11	Patients should be able to see and understand the results and reasons for liver cancer diagnosis using the system.
Use Case #12	The system must have an expression format to show the diagnosis result and reason.
Use Case #13	If an improved liver cancer status (e.g., reduced tumor size) is inputted, and there is an exception that the patient is cured, the patient should be able to request reasons and feedback on the outcome from the system.
Use Case #14	For exceptions sent by the patient, the system developer should be able to modify these results.

Table 5. List of Data Requirements elicited from Use cases of Liver Cancer Diagnosis System

Data Requirements	Eliciting Data Requirements using Use Cases
Completeness	Hospital treatment data and user-entered data must have a sufficient amount to produce diagnostic results.
Accuracy	It is necessary to accurately express the current progression of liver cancer by establishing an expression format showing stored data. (Existed in original DAQ)
Precision	In the process of importing hospital treatment data, there should be no omissions or errors in existing data. (Existed in original DAQ)
Consistency	Hospital treatment data should be correctly delivered and there should be no contradictions in the data.
Currentness	Hospital care data must be kept up to date. (Existed in original DAQ)
Easy Understanding	Patients should be able to easily understand the metadata display format and input their liver cancer data.
Managability	Liver cancer data entered by patients and previously entered treatment data should be stored in an appropriate format.
Reliability	Patients should be able to trust the results of liver cancer diagnosis.
Compliance	The input data must comply with the existing data standards.

스(Use case)' 라고 부르며, 이들은 고객의 요구사항을 정리하고 고객의 요청사항을 전달하는데 유용하게 사용된다. [11]에서 정의된 의료 DAQ 시스템의 데이터 요구사항 각각의 품질 요소와 생성된 유스 케이스(Use case)들을 서로 연결하여 도출한 요구사항을 매칭하고 비교 분석하였다.

이 가상 시나리오에서 주목해야 할 부분은 지식 정의 및 개발 후에 임상 데이터를 실제로 이용하는 부분이다. 지식 정의 단계에서 지식의 관계 정의 및 표현을 위한 메타데이터 모델링과 개발 단계에서 다양한 사용자의 접근 및 편의를 위한 레이아웃, GUI 디자인에 주목하여 수집한 데이터의 정확성 및 품질 평가를 고려하고 있지 않다. 암 진단 시스템에서는 환자의 상태를 진단하기 때문에 품질 요구사항 중 보안성, 정밀성, 완전성 등이 중요한 요구사항으로 볼 수 있으나 기존 연구에서는 언급되지 않았다. 이에 대해 [11]의 데이터 요구사항을 반영하여 간암 진단 시스템에서의 데이터 요구사항 목록을 다음과 같이 표현하였다.

여기서 정밀성, 일관성, 쉬운 이해, 신뢰성과 같은 데이터 요구사항은 기존의 DAQ 예시 아래의 그림 3에서 이미 언급하고 있다. 그러나, 다른 요구사항들은 시나리오를 통한 유스 케이스(Use case)에서 도출하여 목표 모델링으로 이를 표현하였고, 결론적으로 기존 DAQ에서는 데이터의 요구사항을 자세히 고려하지 않은 채 설계를 진행하였으며 목표 모델링과 같은 요구 공학적 접근을 통해 설계 이전에 데이터의 요구사항을 파악하고 체계화할 수 있다.

요구사항 리스트 중에 정밀성과 정확성은 기존 DAQ 시스템 설계 단계에서 연구자들이 중요하게 고려한 요구사항이며, 처음 단일 단계 모델링으로는 원하는 결과를 얻을 수 없었기에, 두 단계 모델링을 이용하여 의료 전문가에 의해 메타 데이터 속성을 정의한다. 또한 개발자는 사용자 친화적인 유저 인터페이스 개발과 기존 의료 기록 데이터를 불러올 수 있는 기능을 추가하였다.

이 가상 시나리오의 DAQ 시스템 필요 태스크와 데이터 품질 요구사항과의 관계를 목표 모델로 그림 3과 같이 표현하였다. 이는 기존의 DAQ에서 고려하고 있는 목표와 세부 목표를 이용하여 태스크와 요구사항을 간단히 추출한 결과이다. 연구에서 제시하고 있는 기능을 토대로 요구사항을 분석하였을 때는 정확성, 접근성, 고품질의 데이터 3가지로 표현할 수 있다. 그러나, 좋은 데이터의 품질이란 구체적으로 무엇인지 알 수가 없으며, 기존 DAQ에서 데이터의 품질에 대한 논의는 없었기에 위의 시나리오와 유스 케이스(Use case) 분석을 통해 목표 달성을 위해 필요한 태스크와 요구사항들을 구체적으로 정의하고, 요구사항이 태스크에 영향을 미치는 관계 정도를 표현할 수 있다. 따라서 [11]에서 제시한 의료 데이터 품질 요구사항 리스트를 이용한 목표 모델링 결과는 다음의 그림 4와 같다.

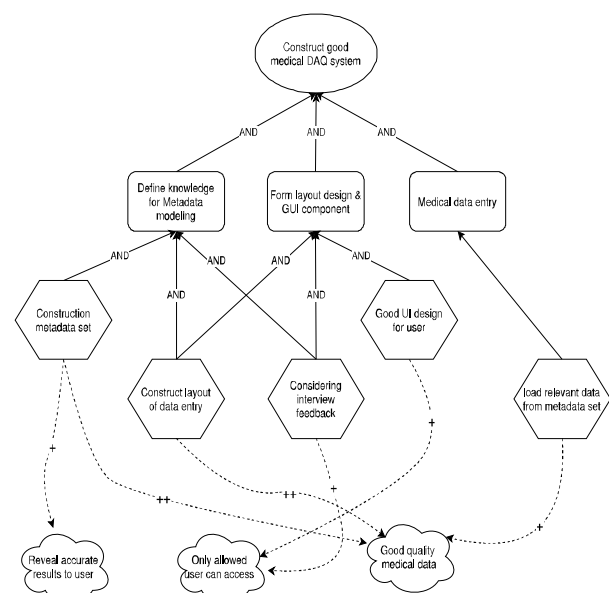


Fig. 3. Initial Goal Modeling of Liver Cancer Diagnosis System

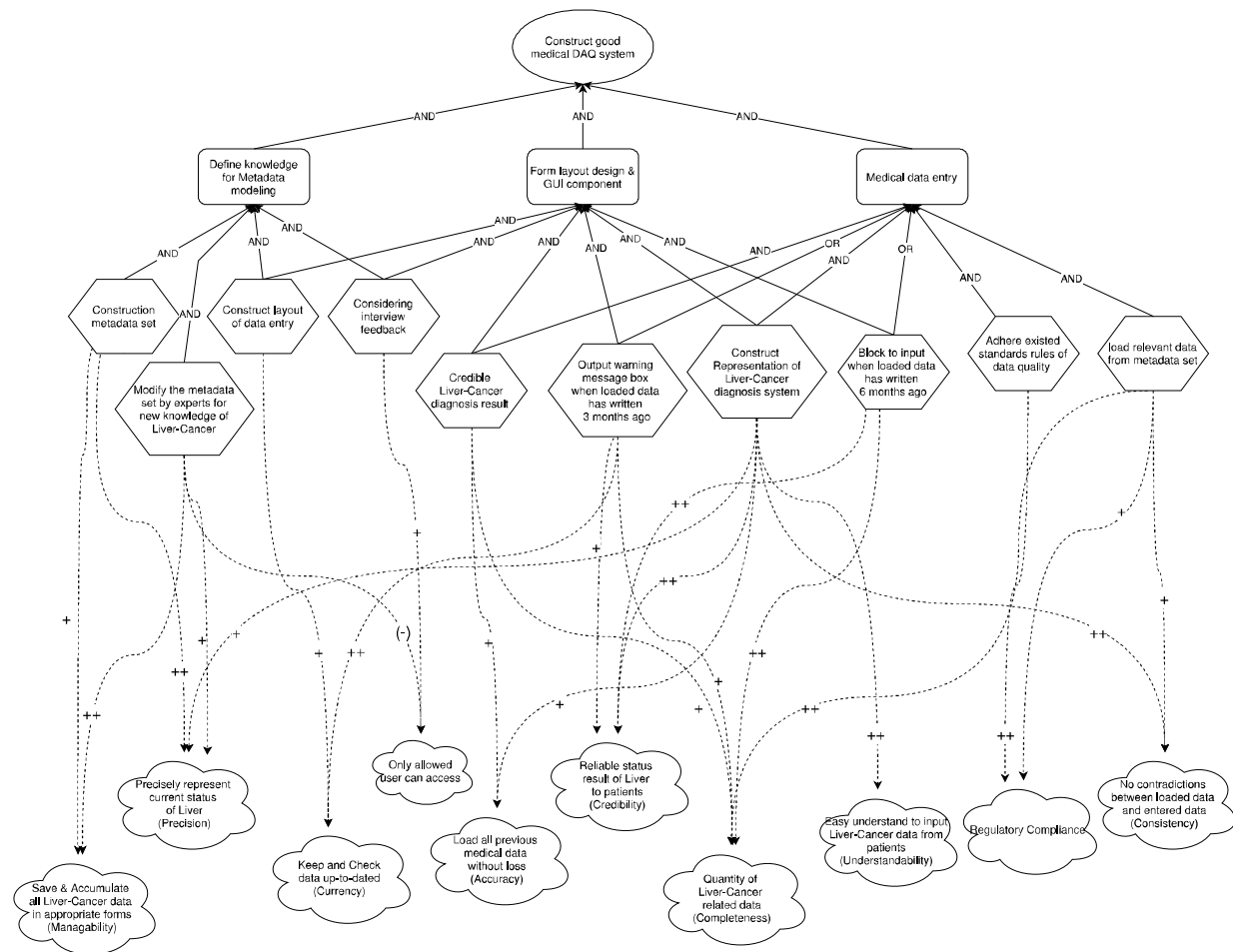


Fig. 4. Goal Modeling of Liver Cancer Diagnosis System using Requirements Engineering

위의 모델에서는 기존 연구자들이 간략히 언급한 good quality of medical data에서 구체적으로 어떤 조건을 갖 추어야 하고, good UI design for user에서 데이터의 최신화를 확인하여 입력을 제한하는 기능을 예시로 들었으며, 이러한 조건 및 기능을 새로운 태스크로 정의하고 그 아래에 태스크와 새로운 요구사항들 간의 관계를 정의하였다. 그림 3과 4를 서로 비교해보면, 데이터의 정확성은 공통적으로 확인할 수 있었다. 그리고 데이터의 표현 체계, 최신화가 필요한 기존 의료 데이터에 대한 경고 메시지, 데이터 품질 규칙 등을 태스크로 정의하여 이들과 요구사항들 간의 관계를 정의하였다. 기존의 요구 공학 프레임워크를 이용하여 ISO 데이터 품질 요구사항 표준에 존재하는 항목으로 요구사항들을 표현하였으며, 기존 DAQ 시스템 사례에서 필요한 데이터 품질 관련 요구사항을 추출하였다.

실제로 [10]에서는 메타데이터 모델링과 레이아웃 디자인 예시를 통해 저자들이 이 2개의 세부 목표인 소프트웨어를 달성하기 위해 무엇을 고려하였는지 기술하고 있다. 그

리고 제안 모델을 통해 실제 데이터 수집 및 적용에 대해 어떤 소프트웨어 골이 있으며, 이를 달성하기 위한 요구사항으로 좋은 품질의 의료 데이터가 필요하다는 것을 보여준다. 그리고, 좋은 품질의 의료 데이터를 얻기 위한 요구사항이 무엇인지 설계 이전에 정의하였으며, 특히 세부 목표인 레이아웃 디자인과 의료 데이터 입력을 위해 데이터의 표현 형식과 표준 형식의 필요와 같은 요구사항도 추출하였다.

목표 모델링 기법을 통해 완전성과 일관성, 정밀성, 현재성, 적시성 등은 DAQ 시스템에서 목표를 달성하는 데 중요한 요구사항들이라고 볼 수 있다. 첫 번째 사례에서 요구사항 분석을 하지 않고 시스템 설계를 하였거나, 비기능적 요구사항은 고려하지 않았다. 연구자들이 기존 DAQ 설계에서 이를 반영하였는지는 언급하고 있지 않기 때문에, 다른 사용자나 개발자가 신뢰성이나 이해성과 같은 데이터의 요구사항을 만족하고 있는 지 또는 이러한 요구사항을 만족하기 위해 어떤 작업이 수행되어야 하는 지 목표 모델링 도식을 활용해 시스템의 보완 방향을 파악할 수 있고, 이를 통한 시스템 설계 명세화를 통해 설계 문서를 만들 수 있다.

제안된 목표 모델을 통해 정의한 요구사항 집합에는 제안된 목표 모델을 통해서 데이터에 대한 품질 속성을 새로 추가할 수 있으며, 기존 DAQ가 가지고 있는 품질 속성 또한 모두 표현할 수 있기에 해당 도메인에서 요구되는 비기능적 요구사항을 최종적으로 추출할 수 있다. 제안 모델을 통해 추출한 데이터 품질 속성을 기존 DAQ와 비교하고 평가하였다. 기존의 DAQ 시스템 설계에서는 요구 공학을 사용하지 않은 채 설계를 진행하였기에 체계적으로 이해관계자들의 요구 사항을 확인할 수 없다. 그러나 제안 모델에서는 목표 모델링을 이용하여 시스템에서 달성하고자 하는 목표를 설정하고, 목표를 달성하기 위한 세부 목표를 단계적으로 정의하고, 최종적으로 시나리오를 이용하여 요구 사항들을 찾아낼 수 있었다. 정의된 요구사항 목록을 이용하여 실제 개발 단계 이전에 사용자나 개발자와 같은 다양한 이해 관계자들은 시스템의 기능(기능적 요구사항)과 제약 조건(비기능적 요구사항)을 알 수 있다. 기존의 DAQ 설계와 비교했을 때 설계 이전 단계에서 이해관계자들의 요구를 파악하고 이를 반영할 수 있어 그들의 입장에서 만족도가 높은 시스템을 개발할 수 있다.

IV. Conclusions

4. Results & Future Works

4.1 Results

위의 의료 DAQ 사례 연구에서 의료 도메인의 데이터 요구사항 리스트를 반영하여 기존의 DAQ 연구에 요구 공학적 시나리오와 유스 케이스(Use case) 기법을 이용하여, 첫 번째 사례의 시스템 설계 시에 필요한 요구사항들을 추출하였다. 두 번째 사례에서는 기존 DAQ 설계 이전에 연구자들이 생태학자들의 요구사항을 파악하여 이를 만족하는 기능들을 정의한 상태에서, 시나리오와 유스 케이스(Use case) 기법을 이용하여 데이터의 요구사항을 추출하였다. [8]에서 분류한 데이터 요구사항 목록을 이용하여 각 사례별로 요구사항을 정의하고, 이를 목표 모델링 기법을 이용하여 요구사항과 태스크 간의 관계를 표현하였다. 3가지 사례 연구 결과를 통해 본문에서 제안한 요구 공학적 기법을 적용한 DAQ 시스템의 비기능적 요구사항, 특히 데이터의 요구사항을 도출할 수 있었으며, 요구사항을 도출하고 목표 모델링 기법을 이용하여 사용자의 '목표-세부 목표-태스크-요구사항'을 도식화하는 과정에서, 기존 DAQ 시스템에서 언급하고 있지는 않지만 요구사항을 정의하는 과정에서 필요한 세부 목표나 태스크를 식별하

고 표현할 수 있다. 이를 통해 기존 DAQ 시스템은 비기능적 요구사항, 특히 데이터의 품질에 관련된 요구사항을 크게 고려하고 있지 않으며, 이해관계자들의 요구사항을 고려하여 설계를 한 연구에서도 특정 기능을 추가하여 요구사항을 만족하게 하는 기능적 요구사항을 고려하고 있었다. 그러나 본 연구를 통해 기존 DAQ 시스템의 목표를 설정하여 이를 달성하기 위한 대부분의 비기능적 요구사항들을 추출하여 기존의 요구사항 리스트에 추가할 수 있었으며, 특히 데이터 품질에 관련된 요구사항들을 기존의 ISO 25012 요구사항 리스트 항목을 참조하여 추출하였다.

현재의 DAQ에서는 수집하려는 데이터를 미리 선정하고 그에 맞추어서 시스템 설계를 하거나, 기존 DAQ의 수집 능력 향상을 위한 수집 알고리즘 개선에 집중하고 있다. 제안한 요구 공학 중심의 DAQ 모델을 이용하면 다양한 이해관계자의 요구사항을 먼저 파악하고 추출하여 어떤 데이터가 필요한 지 체계적으로 정리할 수 있다. 수집된 데이터는 기계 학습과 같은 다양한 데이터 처리 분야에서 사용되는데 사용되는 도메인에 따라 요구되는 데이터 품질 속성이 다르다. 제안 모델을 이용하면 시스템의 목표를 달성하기 위한 요구사항 집합을 정의할 수 있고, 설계 과정에서 고려하지 않는 데이터 품질을 시스템 목표를 달성하기 위한 소프트웨어 골에 어떻게 관련되어 있는 지 그 관계를 보여줄 수 있다. 또한 이를 이용하여 DAQ를 설계할 때 데이터 품질을 고려한 요구사항 명세가 가능하여 데이터 수집 과정에서 고품질의 데이터를 확보할 수 있으며, 의료 도메인에서의 보안성, 신뢰성과 같이 도메인에 따라 DAQ 설계 단계에서 중요하게 요구되는 품질 속성들을 집합화하여 설계 명세에 반영할 수 있다. 마지막으로 데이터의 요구사항을 만족하기 위해 어떠한 기능이나 조건이 필요한지 제시하고 이를 설계 문서에 체계적으로 반영할 수 있다. 해당 설계 문서를 통해 시스템 운영이나 유지 보수 측면에서 시스템 운영이나 유지 보수 측면에서 DAQ 시스템의 운영자나 처음 사용하는 다른 사용자가 자신이 얻고자 하는 데이터나 목표에 부합하는지 사용 전에 확인 가능하며, 또한 시스템의 기능 수정 및 보안을 하고자 할 때 문서를 통해 다루고 있는 기능의 목록을 확인하고 그 기능이 만족하는 요구사항과 그 결과물을 한 번에 확인할 수 있다.

4.2 Future Works

우선 제안 모델에서 이용한 목표 모델링 기법에서는 세부 목표를 만족하는 태스크와 이를 만족하는 요구사항을 추출하였는데, 이는 소프트웨어 골을 추출하는 유스 케이스(Use case) 방식을 통해 태스크를 주관적으로 정의하여

요구사항을 도출하였다. 그러나 [12]에서는 목표 모델링을 목표 달성 기준이 명확하게 정의된 하드 골 모델링, 목표의 달성 시기를 결정하기 위해 방법을 정의하는 소프트 골 모델링, 그리고 시스템의 조직적인 구조를 모델링하는 것, 이 3가지로 구분하였다. 이를 바탕으로 이해 관계자들과 요구 공학자들 간의 의사소통을 통해 요구사항을 추출, 검증, 그리고 개발하는 것을 요구 공학 프레임워크로 본다.

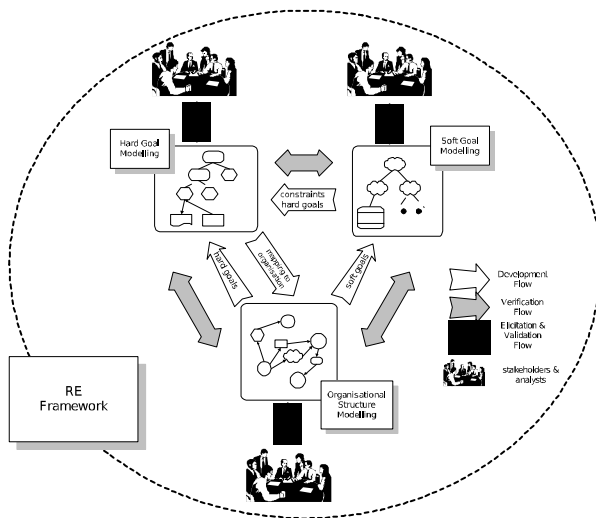


Fig. 5. Requirements Engineering Framework
(Cited from [12])

그림 5에서는 우선 요구 공학자가 초기 조직 모델을 구축하여 초기의 이해 관계자 대상들과 목표를 식별한다. 목표를 소프트/하드 골 모델링을 통해 구체화하면 더욱 많은 정보를 얻고, 새로운 이해 관계자 및 목표를 추가할 수 있다. 해당 프레임워크는 특히 소프트 골이 요구사항 해결 방안의 정성적인 추론을 지원하는 결정적인 역할을 하며, 비기능적 요구사항을 만족하는 체계적인 방법을 제공하는 부분을 프레임워크의 장점으로 기술하였다. 그러나 제안 연구는 기존에 정의된 DAQ 시스템에서 비기능적 요구사항(소프트 골)을 중점적으로 추출하였으며, 조직적인 구조나 하드 골에 대한 구축을 하지 않았고, 이해 관계자들과의 의사소통 과정을 생략하였다. 이에 그림 5의 프레임워크에서는 3가지 모델링을 반복적으로 상호작용하여 소프트 골에서 발견된 정보가 다른 2가지 모델링에서 동일하게 적용되어 추가적인 목표와 요구사항을 추출한다. 결과적으로 요구 공학자들에게 세부적인 모델링을 가능하게 하는 방법을 해당 프레임워크는 제공하고 있다. [12]

프레임워크에서의 소프트 골에 국한하여 요구사항을 추출하였지만 기존 DAQ 시스템에서 간과하고 있는 요구사항들, 특히 데이터의 요구사항에 대한 중요성을 언급하며

이를 제안 연구 결과로 증명하고 있다. 데이터 수집 효율을 향상하기 위해 요구 공학에서의 목표 모델링 기법을 적용하고, 사례 연구를 통해 몇 가지 도메인에서 DAQ의 데이터 품질 요구사항을 추출한 것에 연구 의의를 가지고 있다. 차후에는 해당 DAQ 시스템의 이해 관계자들을 정의하고 이들에게서 보다 정확한 요구사항을 추출하여, [12]의 요구 공학 프레임워크의 하드 골 모델링과 조직적 구조 모델링을 이용하여 추출한 요구사항들에 대한 검증, 마지막으로 이 결과를 통해 DAQ 시스템에 대한 설계 내용을 문서화를 할 예정이다.

또한 사례 연구를 진행하면서 3가지 사례 도메인에 대한 전문가의 자문이나 실제 데이터를 없이 진행하여 검증 과정이 시행되지 않았다. 차후에는 실제로 운용되고 있는 해당 사례의 DAQ 시스템의 도메인 전문가와의 인터뷰를 통해 요구사항을 정확하게 추출하여 요구사항 내용에 대한 검증과, 해당 데이터 요구사항이 도메인에서 어떠한 유스 케이스(Use case)를 통해 도출되는 지 요구사항 도출 과정에 대한 검증을 진행할 예정이다.

ACKNOWLEDGEMENT

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2020R1F1A1075605).

REFERENCES

- [1] Vogelsang, Andreas, and Markus Borg. "Requirements Engineering for Machine Learning: Perspectives from Data Scientists." 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW). IEEE, 2019.
- [2] Horkoff, Jennifer. "Non-functional requirements for machine learning: Challenges and new directions." 2019 IEEE 27th International Requirements Engineering Conference (RE). pp.386-391, 2019.
- [3] Abdallah, Mohammed, and Omar Elkeelany. "A survey on data acquisition systems DAQ." 2009 International Conference on Computing, Engineering and Information. IEEE, 2009.
- [4] Vogelsang, Andreas, and Markus Borg. "Requirements Engineering for Machine Learning: Perspectives from Data Scientists." 2019 IEEE 27th International Requirements Engineering Conference

Workshops (REW). IEEE, 2019.

- [5] Van Lamsweerde, Axel. "Goal-oriented requirements engineering: A guided tour." Proceedings fifth IEEE international symposium on requirements engineering. IEEE, 2001.
- [6] Yu, Eric, and John Mylopoulos. "Why goal-oriented requirements engineering." Proceedings of the 4th International Workshop on Requirements Engineering: Foundations of Software Quality. Vol. 15. 1998.
- [7] Anton, Annie I. "Goal-based requirements analysis." Proceedings of the second international conference on requirements engineering. IEEE, 1996.
- [8] ISO, ISO/IEC 25012:2008-Software engineering. Software product quality requirements and evaluation (SQuaRE). Data quality model, Report, International Organization for Standardization, 2009
- [9] Jang, K., Kim, J. H., & Kim, W. J. "Derivation of data quality attributes and their priorities based on customer requirements." KIPS Transactions on Software and Data Engineering, Vol.4, No.12, pp.549-560, 2015.
- [10] Li, Bei, et al. "Experiences of building a medical data acquisition system based on two-level modeling." International journal of medical informatics, vol.112. pp.114-122, 2018
- [11] Gresse von Wangenheim, C., and A. von Wangenheim. A Software Quality Model for Asynchronous Store-and-Forward Telemedicine Systems. Technical Report INCoD/UFSC 005/2011-E-GQS, GQS/INCoD/UFSC, Florianópolis/Brazil. Available at: <http://www.incod.ufsc.br/a-software-quality-model-for-asynchronous-store-andforward-telemedicine-systems>, 2011.
- [12] Donzelli, Paolo, and Roberto Setola. "Handling the knowledge acquired during the requirements engineering process: a case study." Proceedings of the 14th international conference on Software engineering and knowledge engineering. 2002.

Authors



Kyung-Sik Kim received the BS in Computer and Information Science from Korea University in 2018 and MS in Computer Engineering from Ajou University in 2020. His area of interests are Software

Engineering, Data Quality Improvement, Goal Modelling, and Requirements Engineering.



Seok-Won Lee received the BS in Computer Science from Dongkuk University at Seoul, Republic of Korea, MS in Computer Science from University of Pittsburgh at Pittsburgh, Pennsylvania, and PhD in Information

Technology from George Mason University at Fairfax, Virginia. He has been a faculty member at the University of North Carolina at Charlotte and the University of Texas at San Antonio. He was a research scientist at IBM T.J.Watson Research and Science Applications International Corporation in US. He is currently a Full Professor and Chair of the Departments of Software and Compute Engineering, and Applied Artificial Intelligence at Ajou University, Republic of Korea. His areas of specialization include software engineering with specific expertise in ontological requirement engineering and domain modeling, and knowledge engineering with specific expertise in knowledge acquisition, machine learning and knowledge-based systems.