

디지털 인문 정보학 캡스톤 디자인 프로젝트

한국어 학습자료 분류 통합 기준 마련 프로젝트

팀명: 한교 도서관

팀원 명단:

이세희(O), 1911098, tpgml0502@naver.com, 010-3915-5150
박서정, 1911088, 0305sjpark@naver.com, 010-9838-8971
이수민, 1911068, sudl000925@naver.com, 010-6227-9154
이예지, 1911008, mary.lee07164@gmail.com, 010-5914-7164

목차

I. 프로젝트 수행 목적

II. 프로젝트 결과물의 개요

1. 프로젝트 결과물 설명
2. 프로젝트 결과물의 그림
3. 프로젝트 결과물의 구조
4. 현실적 제약조건
5. 관련 기술 소개
6. 개발 도구
7. 기존 소스나, 다른 과목 결과물 등등 자료 이용

III. 결과로서 제출할 실적물

IV. 프로젝트 수행 추진 체계 및 일정

표 목차

표 1. 각 조원의 조직도

표 2. 프로젝트 진행 일정표

그림 목차

그림 1. 홈페이지 예상 모양

I. 프로젝트 수행 목적

현재 한국어 학습자료를 출판하는 공식적 기관인 세종학당은 물론, 다양한 대학과 교육청 사이트에서는 각각 다른 체계와 방식을 사용하여 한국어 학습 자료의 정보를 전달하고 있다. 국립국어원의 경우 단계, 제공처, 등록일, 기능, 영역, 유형, 내용에 대한 정보를 제공하고 있고 세종학당에서는 난이도, 교육 과정, 저작권, 등록일, 학습 지원, 내용 소개, 목차, 구매 안내에 대한 정보를 제공하고 있다. 대학에서는 교재에 대한 정보를 제공하고 있지 않거나 교재 예시 사진, 혹은 mp3 파일 등의 정보 혹은 자료만을 제공하고 있었다.

이처럼 같은 한국어 학습자료임에도 소개하는 기관에 따라 제공되는 정보 및 제공 형식은 물론이고 정보 유형의 명칭 또한 다르다는 것을 알 수 있다. 한편, 대부분의 다문화 센터에서는 프로그램에 대한 소개만 할 뿐, 참고할 수 있는 한국어 학습자료에 대한 정보가 없어서 일정 및 사정상 한국어 교실이나 학원에 다닐 수 없는 학습자의 경우 직접 한국어 교재 및 학습자료를 찾아보고 공부해야 하는 불편한 상황으로 이어지고 있다.

따라서 이 프로젝트에서는 이러한 두 문제 상황을 해결하기 위해 한국어 학습자료 소개의 기준을 정형화하여 현재 통일되어 있지 않은 한국어 학습자료 소개 페이지를 표준화하고 그것을 배포하는 것을 목표로 한다. 이 기준이 될 온톨로지를 구축한 뒤, 한국어 학습자료를 소개하는 내용이 담긴 사이트인 국립국어원, 세종학당, 각 대학교의 한국어교육센터, 다문화센터 사이트에서 자유롭게 사용할 수 있도록 한다.

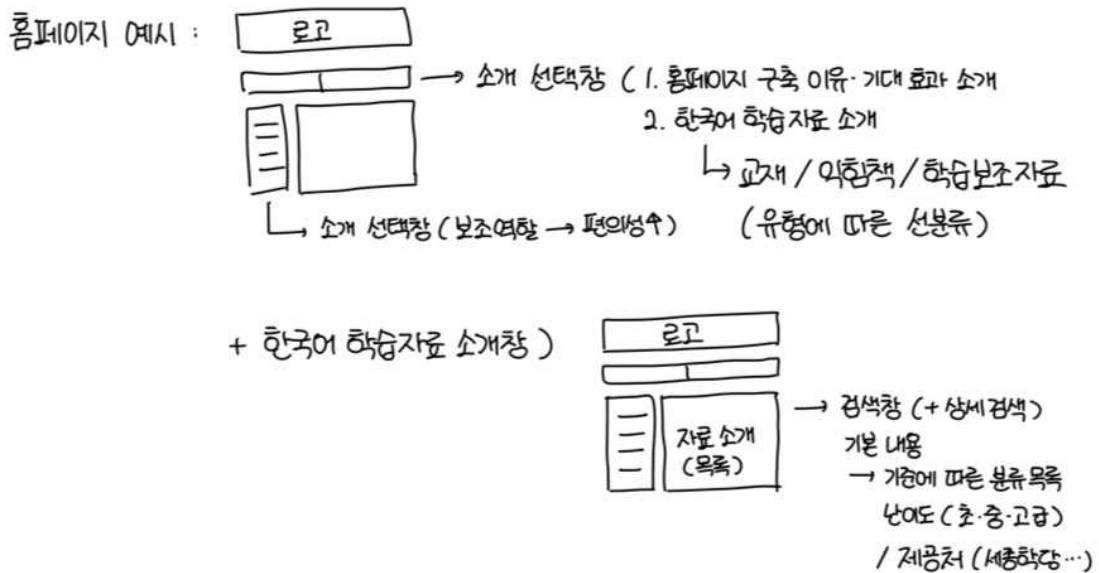
II. 프로젝트 결과물의 개요

1. 프로젝트 결과물 설명

한국어 학습자료에 대해 기본적으로 제공할 정보 유형을 통합시켜 하나의 기준점이 될 온톨로지를 구축한다. 해당 프로젝트에서 다루는 한국어 학습자료는 교재와 익힘책, 그리고 영상, 음원, 이미지 등의 학습 보조자료를 포함한다. 이 온톨로지는 한국어 학습자료를 소개하는 모든 공공기관과 대학의 사이트에 쓰일 수 있다. 특히, 한국에 거주하는 외국인 학습자에게 가장 적극적이고 가까운 기관인 지역 다문화 센터 및 중앙 다문화 포털 사이트 '다누리'에서도 이 온톨로지를 기반으로 하여 사이트에 한국어 학습자료에 대한 정보를 소개할 수 있도록 하는 예시 활용안으로써 홈페이지 예시안을 제작한다. 온톨로지는 이용자의 편의를 최대한 보장하기 위해 다양한 기준으로 검색에 용이한 형태로 구축하고 홈페이지는 각자의 상황에 맞추어 수정이 편리하게 이루어질 수 있도록 페이지를 구축한다.

이와 같은 하나의 온톨로지를 통해 한국어 학습자료에 대한 정보를 사이트마다 일정하게 전달할 수 있게 되면, 해당 페이지를 이용하여 한국어 학습자료에 대한 정보를 얻어야 하는 이용자 입장에서 사이트가 제공하는 정보를 받아들이는 것에 어려움이나 혼란 없이 일정한 정보를 제공받을 수 있다. 더불어, 지역 다문화 센터에서도 관련 정보를 제공하게 된다면 한국어 교실과 같은 지역 프로그램에 참여하는 한국어 교사 및 예비 교원, 봉사활동 참여자나 홀로 한국어를 독학해야 하는 상황에 놓인 외국인, 또는 한국어를 학습하는 과정에서 더 많은 학습자료를 얻고 싶은 외국인 학습자에게 여러 사이트에 흩어져 있는 정보를 직접 수집하고 비교해야 하는 어려움 없이 원하는 자료에 대한 정보를 찾아낼 수 있도록 도움을 줄 수 있다.

2. 프로젝트 결과물의 그림



3. 프로젝트 결과물의 구조

온톨로지는 아직 구축되지 않은 상황이기에 사진으로는 넣지 못했으나, 한국어 학습자료, 난이도, 저자, 영역, 기능의 클래스를 만들어 서로 링크시킬 것이며, 목차, 등록일, 시리즈 등의 속성 및 관계를 통해 링크하면서 내용을 작성하도록 한다.

홈페이지는 실제 완성된 모습을 하나의 예시이자 표준 모습으로서 구축한다. 기본 첫 페이지에는 상단에 페이지의 로고를 게시하고 바로 아래에 페이지 기본 선택창을 둔다. 선택창의 목록은 홈페이지의 기본 첫 화면이 될 '홈페이지 소개창'과 '한국어 학습자료 소개창'으로 구성한다. 선택창의 왼쪽 하단에는 선택창과 같은 내용을 담은 선택 목록을 두어 이용자의 편의성을 높이도록 한다. 이와 같은 방식은 다문화 센터 '다누리' 및 부산 다문화 교육 지원 센터 홈페이지 등 많은 홈페이지에서 사용하는 방식을 참고하였다. 이 목록 바로 오른쪽에는 실질적인 페이지의 내용을 담도록 한다. 첫 화면일 때는 이 내용창에 프로젝트 구축 목적과 기대효과 등을 간단히 소개하는 내용을 담는다. 한국어 학습자료에 대한 창일 때는 내용창의 맨 위에 검색창을 두고 바로 아래에는 기본적으로 다양한 기준에 따라 분류한 자료를 나열하여 제시한다. 그 기준에는 난이도, 제공처가 있다. 교재나 익힘책일 때는 도서 표지와 도서명을 함께 기본 제공하고 하나의 대상을 누르면 해당 도서에 대한 상세 정보를 제공하는 페이지로 넘어가게 된다. 보조 자료의 경우에는 게시글 나열 목록창과 같은 형식으로 제공한다.

4. 현실적 제약조건

1) 경제성: 온톨로지 구축 시 숫자로 표기할 수 있는 인스턴스는 되도록 숫자로 형태를 한정 짓고, 목적에 맞게 필요한 속성만 뽑아서 데이터베이스화하여 온톨로지 이용 시에 더 빠른 속도로 처리하여 경제성을 확보하고자 한다.

2) 윤리성: 홈페이지 제작 혹은 게시물 업로드 시 사용될 외부의 이미지 자료 및 기타 저작물은 저작권 침해를 방지하기 위하여 저작물의 이용허가 범위를 확인하고 그 범주 내에서 이용하거나 저작권자에게 사용허가를 받은 후 이용, 혹은 출처 표기를 하여 이용하고자 한다.

3) 안전성: 온톨로지 구축 시, 검색하거나 분류할 때 해당 유형만 모아서 보기 편하도록 클래스를 설정하고 서로 연결점이 있는 클래스끼리 링크시켜 관계도를 짜도록 한다. 이 과정 속에서 내용 작성 시, 관리자가 누구냐에 상관없이 일관성을 지닐 수 있도록 속성에 대한 설명을 미리 논의하여 확실히 한 후 작성하고자 한다.

4) 신뢰성: 구현된 결과물인 온톨로지는 실제 현장에서 사용되어도 될 정도의 기본적인 틀을 제공하는 것을 목표로 하며 홈페이지는 실제 사이트에서 이용할 시 참고할 수 있는 구성 틀을 제공하고자 한다. 또한, 홈페이지는 PC에

서 오류 없이 동작할 수 있도록 점검 과정을 거쳐 제작한다.

온톨로지의 내용적 측면으로는 속성의 구조와 속성명을 설정할 시 해당 프로젝트의 자체적 한국어 학습자료 분류에 대한 전문성, 신뢰성 확보를 위하여 국립국어원, 세종학당 등 한국어 교육에 있어 공신력을 갖춘 기관의 분류체계와 어휘를 참고한다.

5) 미학: 전문적인 기술력 부족과 저작권상 문제로 사이트 구성이 단순하고 꾸미는 데 있어 부족함이 생길 수 있다. 이는 직관적 화면 배치 및 체계적인 분류와 알아보기 쉬운 아이콘 설정으로 편의성에 초점을 맞추어 일부 시각적인 공백을 메운다. 한국어 학습자료를 이용하는 주요 이용자가 외국인이라는 특성 상 이용자의 편의성을 높이기 위해 직관적인 단어들로 구성하여 최대한 간편한 검색과 번역이 동원되었을 때 정확한 번역으로 이어질 수 있도록 지원한다.

6) 사회에 미치는 영향: 이 프로젝트의 한국어 학습자료 분류 온톨로지가 실제 여러 기관의 사이트에 적용될 경우 기존 홈페이지에 변화가 생겨 이용자가 적응하지 못하거나 이용에 어려움을 겪는 상황이 발생할 수 있다. 이를 해결하기 위해 홈페이지 이용 영상 매뉴얼을 제공하여 이용자들이 홈페이지를 직접 이용하기 이전에 영상을 보며 홈페이지에 적응하고, 그것을 실제 홈페이지 이용에 적용할 수 있도록 지원하고자 한다.

5. 관련 기술 소개

1) 온톨로지

온톨로지의 대표적인 정의로는 토마스 그루버(Thomas A. Gruber)의 &공유되는 개념화의 형식적이고 명시적인 명세&가 있다(1993). 즉 온톨로지란 사물 등 세상에 존재하는 것에 대하여 일반적으로 공유되고 합의된 개념, 관계를 컴퓨터가 읽고 처리할 수 있는 형태로 표현한 것을 말한다. 이는 문헌정보학을 비롯하여 시맨틱 웹, 자연어 처리 등 다양한 분야에서 활용된다.

온톨로지는 서로 다른 시스템 사이 기계적 처리가 가능한 공통 개념을 제공할 수 있도록 하여 데이터 공유나 재사용, 재활용이 가능하도록 한다. 정보검색의 측면에서 바라본다면 의미 기반의 검색이 이루어짐으로써 검색자의 의도를 반영하는 지능적 정보검색 서비스를 제공할 수 있도록 지원한다.

온톨로지는 클래스(class), 속성(property), 인스턴스(instance), 관계(relation)로 구성된다. 클래스는 사물 또는 개념에 해당하는 것으로 온톨로지의 핵심 요소이며 인스턴스는 클래스에 해당하는 실질적 사례를 말한다. 속성은 클래스나 인스턴스가 가진 특성, 성질 등을 의미하며 관계는 클래스, 인스턴스 간 관계를 나타낸다.

2) 데이터 베이스

여러 사람에 의해 공유되어 사용될 목적으로 통합하여 관리되는 데이터의 집합을 말한다. 자료항목의 중복을 없애고 자료를 구조화하여 저장함으로써 자료 검색과 갱신의 효율을 높인다. 논리적으로 연관된 하나 이상의 자료의 모음으로 그 내용을 고도로 구조함으로써 검색과 갱신의 효율화를 꾀한 것이다. 즉, 몇 개의 자료 파일을 조직적으로 통합하여 자료 항목의 중복을 없애고 자료를 구조화하여 기억시켜 놓은 자료의 집합체라고 할 수 있다.

데이터베이스가 갖는 몇 가지 특성을 살펴봄으로써 그 뜻을 보다 명확히 할 수 있다. 첫 번째로는 똑같은 자료를 중복하여 저장하지 않는 통합된 자료이고, 두 번째로는 컴퓨터가 액세스하여 처리할 수 있는 저장 장치에 수록된 자료이며, 세 번째로는 어떤 조직의 기능을 수행하는 데 없어서는 안 되며 존재 목적이 뚜렷하고 유용성 있는 운영 자료이기 때문에 임시로 필요해서 모아 놓은 데이터나 단순한 입출력 자료가 아니라는 점이며, 마지막으로 한 조직에서 가지는 데이터베이스는 그 조직 내의 모든 사람들이 소유하고 유지하며 이용하는 공동 자료로서 각 사용자는 같은 데이터라 할지라도 각자의 응용 목적에 따라 다르게 사용할 수 있다는 점이다.

6. 개발 도구

1) ATOM

웹사이트 구축 시 GitHub의 텍스트 에디터인 ATOM 64비트 운영 체제 지원 1.6 버전을 사용한다.

2) TopBraid Composer(TBC)

홈페이지에 사용될 온톨로지 구축 시 topquadrant의 온톨로지 모델링 도구인 TopBraid Composer(TBC) 6.0.1버전을 사용한다.

7. 기존 소스나, 다른 과목 결과물 등등 자료 이용

1) ATOM

ATOM은 온라인 코드 공유 및 협업 사이트 GitHub에서 2015년에 제작한 텍스트 에디터 프로그램이다. 다양한 프로그래밍 언어의 편집기로 사용할 수 있도록 고안된 도구로써, MIT License 하에 배포되는 자유 소프트웨어이자 오픈 소스 소프트웨어이다. Windows, Linux, Mac에서 동작 가능하며 HTML, JavaScript, CSS와 같은 웹 기술로 화면이 구현되어 있다. Node.js를 기반 코드로 하고 있고 Election이라는 플랫폼을 기반으로 만들어졌다. 자바스크립

트 변형 언어 ‘커피스크립트’로 개발되었고 기본 골격은 오픈소스 브라우저 ‘Chromium’에 기반한다. 2017년에는 Teletype이라는 동시 편집 기능이 추가되어 실시간으로 동시 수정이 가능하게 되었다. 이 프로그램은 2021년 1학기, 한상은 교수님의 ‘인문콘텐츠와 정보기술’에서 알게 된 프로그램으로써, 당시 최종 웹페이지 구축 과제로 이용했던 경험을 살려, 이 프로젝트에서는 목적에 맞게 편의성을 살리고, 더 발전시킨 웹페이지를 제작하고자 한다.

2) TopBraid Composer

TopBraid Composer는 미국의 시맨틱 웹 솔루션 업체 topquadrant에서 제작한 IDE로, 온톨로지 모델링 도구이자, SPARQL 쿼리 에디터, 데이터 통합 콘솔 및 시맨틱 어플리케이션 구축에 쓰이는 도구이다. W3C 표준에 준거하여 온톨로지 및 링크 데이터의 구성 개발, 관리 및 테스트를 지원한다. 이 IDE는 2021년 2학기, 윤소영 교수님의 ‘시맨틱데이터 구축’에서 알게 된 도구로써, 이 프로젝트만의 목적에 맞게 더 풍부한 내용을 담고 데이터 관리 및 검색에 용이한 온톨로지를 구축하고자 한다.

Ⅲ. 결과로서 제출할 실적물

온톨로지는 rdf 파일 형식으로 제작될 것이며 온톨로지는 원하는 정보의 유형에 따라 데이터를 모아서 한눈에 살필 수 있고 각 데이터가 어떻게 연결되어 있는지를 알 수 있는 시각화 정보도 제공 가능하다.

홈페이지 구축은 html 형식으로 제작될 것이다. 홈페이지는 링크를 통해 접속 가능하며 PC를 기본으로 하여 구성되어 있다. 이 두 가지의 실제 이용은 상황에 따라 이용 영상을 제작하여 이용에 관한 이해를 돕도록 한다.

IX. 프로젝트 수행 추진 체계 및 일정

각 조원의 조직도

조원	역할	담당
세희	팀장	회의 및 프로젝트 진행 주도 회의 알림 및 방 개설 면담 진행 및 내용 정리 온톨로지 제작 및 홈페이지 제작
수민	팀원	5주차 제안서 정리 및 발표 최종 ppt 작성
예지	팀원	매 회의 정리 및 제출 최종 보고서 작성 주도
서정	팀원	홈페이지 로고 제작 및 홈페이지에 들어갈 사진, 그림 자료 수집 최종 ppt 작성

주 단위의 프로젝트 수행 일정

주차	내용	주간 회의
3주 22.03.14 - 22.03.20	프로젝트 세부 내용 논의 및 정정 / 사전 자료 조사 후 상호 발표	

<p>4주</p> <p>22.03.21 - 22.03.27</p>	<p>설문지 작성 / 교수님 면담</p>	<p>03.22 (1차) - 설문조사 설계</p>
<p>5주</p> <p>22.03.28 - 22.04.03</p>	<p>피드백 수용하여 프로젝트 세부 내용 재논의 및 정정</p>	<p>03.30 (2차) - 주제 정정</p>
<p>6주</p> <p>22.04.04 - 22.04.10</p>	<p>추가 사전 조사/ 제안서 작성</p>	<p>04.05 (3차) - 프로젝트 세부 내용 아 이디어 회의</p>
<p>7주</p> <p>22.04.11 - 22.04.17</p>	<p>프로젝트 주제 최종 정리 / 온톨로지 개인 공부</p>	<p>04.12 (4차) - 프로젝트 주제 재조정 및 제안서 수정</p>
<p>8주</p> <p>22.04.18 - 22.04.24</p>	<p>온톨로지 구축 1 (관계도 완성)</p>	<p>04.19 (5차) - 온톨로지 틀 완성</p>
<p>9주</p> <p>22.04.25 - 22.05.01</p>	<p>온톨로지 구축 2 / 홈페이지 구축 1</p>	<p>04.26 (6차) - 인스턴스 수집 회의</p>
<p>10주</p> <p>22.05.02 - 22.05.08</p>	<p>온톨로지 구축 3 / 홈페이지 로고 제작</p>	<p>05.03 (7차) - 인스턴스 최종 선별 및 정리</p>

11주 22.05.09 - 22.05.15	홈페이지 구축 2 (기본 틀 완성)	05.10 (8차) - 홈페이지 내용 정리 및 공유
12주 22.05.16 - 22.05.22	홈페이지 구축 3 / 보고서 제작 시작	05.17 (9차) - 보고서 내용 작성
13주 22.05.23 - 22.05.29	온톨로지 구축 완료 / 최종 발표 ppt 제작 시작	05.24 (10차) - ppt 구성 의견 공유
14주 22.05.30 - 22.06.05	홈페이지 구축 완료 / 최종 발표 ppt 제작 완료 / 최종 보고서 제작 완료	05.31 (11차) - 온톨로지, 홈페이지 관 련 의견 공유
15주 22.06.06 - 22.06.12	온톨로지, 홈페이지 최종 이용 체 크 / 최종 보고서 체크 / 최종 발표 연습	06.07 (12차) - 발표 연습

IV. 참고 자료

ATOM 사이트, <https://atom.io/>, 2022.04.04

data base, IT용어사전, <https://terms.naver.com/entry.naver?docId=1082446&cid=40942&categoryId=32840>

‘Ontology’, IT용어사전, <https://terms.naver.com/entry.naver?docId=863084&cid=42346&categoryId=42346> (2022.03.31)

‘Ontology’, 지형 공간정보체계 용어사전, <https://terms.naver.com/entry.naver?docId=3480314&cid=58439&categoryId=58439> (2022.03.31)

Topbraid Composer 사이트, <https://www.topquadrant.com/#>, 2022.04.04

‘다문화학생 한글교육·진학 지원… 동등한 출발선 보장’, 뉴시스, 2022.02.04, https://newsis.com/view/?id=NISX20220204_0001746659&clD=10201&pID=10200#

백운용 기자, ‘대구시, 대구거주 외국인주민 대상 한국어 무료교육 실시’, 국제뉴스, 2022.01.28.,

<https://www.gukjenews.com/news/articleView.html?idxno=2398858>

윤소영, 2021, ‘4-1 지식표현-온톨로지’, 디지털인문정보학트랙, 한성대학교

이소희 기자, ‘수원시, 다문화가족 자녀 한국어 학습 방문 교육 지원…대상자 모집’, 민중의 소리, 2022.02.03, <https://vop.co.kr/A00001608411.html>

이희용 기자, ‘외국인 유학생 유치 위해 해외 한국어 공교육 강화해야’, 연합뉴스, 2017.07.07, <https://www.yna.co.kr/view/AKR20170707083200371?input=1195m>

이희용 기자, “해외 한국어교육 문제없나…유학생 75% ‘사교육 독학 의존’”, 연합뉴스, 2017.07.07, <https://www.yna.co.kr/view/AKR20170707101600371?input=1195m>

임민철 기자, ‘기트허브, 아톰 에디터 첫 정식판 공개’, ZDNet Korea, 2015.06.26, <https://zdnet.co.kr/view/?no=20150626095402>

임성준 기자, “세종시교육청, 다문화 학생 ‘한국어 맞춤형 캠프’ 운영”, 매일경제TV, 2022.01.17, http://mbnmoney.mbn.co.kr/news/view?news_no=M1004533486

임성준 기자, ‘제주교육청, 다문화 가정 등에 맞춤형 통합 지원 강화’, 세계일보, 2022.01.27, <http://www.segye.com/newsView/20220126506528?OutUrl=naver>

임영은 기자, ‘소외됐던 외국인 주민 체계적 지원’, 충북일보, 2022.03.16, <https://www.inews365.com/news/article.html?no=706617>

임유경 기자, “아톰 에디터에 코드 ‘동시 편집’ 기능 추가”, ZDNet Korea, 2017.11.17, <https://zdnet.co.kr/view/?no=20171117082315>