

발 간 등 록 번 호

11-1371029-000234-01

BIBFRAME 2.0 해설서

2022. 09.



국립중앙도서관
National Library of Korea

BIBFRAME 2.0 해설서

2022. 09.

국립중앙도서관

제 출 문

국립중앙도서관장 귀하

본 보고서를 「차세대 서지 언어 BIBFRAME 명세서 번역 및 해설서 제작」 연구의 최종 보고서로 제출합니다.

2022년 9월

연구책임자: 박진호 (교수)

공동연구자: 박옥남 (교수)

송민선 (교수)

이 연구는 2022년도 국립중앙도서관 연구개발비로 수행되었으며, 본 연구에서 제시된 정책 제언이나 의견 등은 국립중앙도서관의 공식의견이 아니라 본 연구진들의 개인 견해를 밝혀 둡니다.

목차

제1장 들어가는 글	1
제2장 웹의 변화	5
제3장 링크드 오픈 데이터와 RDF	11
3.1 링크드 오픈 데이터	11
3.2 링크드 오픈 데이터의 규모	21
3.3 RDF	25
제4장 BIBFRAME	42
4.1 등장배경	42
4.2 BIBFRAME 2.0 모델	43
4.3 BIBFRAME 1.0과 2.0	45
4.4 BIBFRAME Editor	47
제5장 KORMARC의 BIBFRAME 변환	73
5.1 변환 명세서에 대한 참고사항	74
5.2 ORMARC 서지 변환 명세를 위한 프로세스 노트	75
5.3 숫자로 된 식별기호의 변환	89
5.4 KORMARC 서지변환-MARC21 필드와의 차이점들을 중심으로	98
5.5 국립중앙도서관 서지데이터 변환	111
5.6 BIBFRAME 국가도서관 사례	134
제6장 마치는 글	143

표 목차

〈표 1〉 별점 4점과 5점 차이 비교	17
〈표 2〉 BIBFRAME 편집기의 종류와 설명	48
〈표 3〉 리더부 변환 고려사항	98
〈표 4〉 001-005, 007(제어, 형태기술 필드) 변환 고려사항	99
〈표 5〉 006 변환 고려사항	99
〈표 6〉 008 변환 고려사항	100
〈표 7〉 010-049 변환 고려사항	104
〈표 8〉 050-088, 090 변환 고려사항	105
〈표 9〉 1XX, 7XX, 8XX 변환 고려사항	106
〈표 10〉 3XX 변환 고려사항	108
〈표 11〉 5XX 변환 고려사항	109
〈표 12〉 841-887 변환 고려사항	112
〈표 13〉 채식주의자 KORMARC와 BIBFRAME 대응요소 간략 정리본	113
〈표 14〉 미국의회도서관 링크드 데이터 서비스에서 제공하는 데이터세트 현황	126
〈표 15〉 KORMARC의 부호표와 어휘집 목록	128
〈표 16〉 MARC21에만 존재하는 필드	130
〈표 17〉 MARC21과 KORMARC의 동일 필드 내 다른 표시기호	132
〈표 18〉 MARC21 식별기호 중 KORMARC 미적용 식별기호 : 서지	132
〈표 19〉 전거통제 관점 KORMARC과 MARC21 필드 비교	134
〈표 20〉 MARC21 식별기호 중 KORMARC 미적용 전거용 식별기호	134

그림 목차

〈그림 1〉 문서 중심 웹(web of document) 개념도	6
〈그림 2〉 웹의 계층별 데이터 구조 예시	7
〈그림 3〉 문서정보를 활용한 보고서 작성 예시	7
〈그림 4〉 데이터를 활용한 보고서 작성 예시	8
〈그림 5〉 데이터 중심의 웹 정보 개념도	9
〈그림 6〉 Linked Open Data 개념도	11
〈그림 7〉 5 star open data 개념도	12
〈그림 8〉 별점 1점 예제	13
〈그림 9〉 별점 2점 예제	13
〈그림 10〉 별점 3점 예제	14
〈그림 11〉 별점 4점 예제	15
〈그림 12〉 별점 5점 예제	17
〈그림 13〉 위키피디아의 ‘Love’ 페이지 중 일부시	18
〈그림 14〉 위키피디아의 ‘Love’ 페이지 중 링크 정보 확인	18
〈그림 15〉 국립중앙도서관 선거 정보 중 ‘사랑’ 검색결과 화면	19
〈그림 16〉 미국의회도서관 주제명 표목 중 ‘love’ 페이지	20
〈그림 17〉 2007년 링크드 오픈 데이터 다이어그램	21
〈그림 18〉 2020년 링크드 오픈 데이터 다이어그램	22
〈그림 19〉 링크드 오픈 데이터 다이어그램의 VIAF 연결 모습	23
〈그림 20〉 DBpedia에서 Resource와 Page	24
〈그림 21〉 DBpedia의 SPARQL Endpoint 화면	25
〈그림 22〉 DBpedia 도서관의 유형(type)인 책(Book)으로 연결 화면	25
〈그림 23〉 국립중앙도서관 SPARQL Endpoint 검색결과 화면	27
〈그림 24〉 RDF 그래프 모델	29
〈그림 25〉 RDF SPO 구조의 두 가지 유형	29
〈그림 26〉 RDF SPO 구조의 확장 예시	30
〈그림 27〉 RDFS 표현 예제	33
〈그림 28〉 DBpedia에서 도서관 개체 페이지 중 일부	34
〈그림 29〉 abstract라는 속성의 정의와 기술규칙	35
〈그림 30〉 국립중앙도서관 LOD 서비스의 채식주의자 검색 결과 중 일부	35
〈그림 31〉 더블링크어의 title 요소에 대한 상세 기술정보	36

〈그림 32〉 국립중앙도서관에서 정의한 type의 정의와 기술규칙	36
〈그림 33〉 Schema.org 홈페이지 중 일부	37
〈그림 34〉 Schema.org의 Book 개체에 대한 속성 중 일부	38
〈그림 35〉 국립중앙도서관에서 검색한 채식주의자 결과 화면	39
〈그림 36〉 BIBFRAME Model	44
〈그림 37〉 BIBFRAME 1.0 모델	46
〈그림 38〉 BIBFRAME Model	47
〈그림 39〉 Bibframe Profile Editor 초기화면	49
〈그림 40〉 Bibframe Profile Editor 초기화면	50
〈그림 41〉 Bibframe Profile 편집기 화면	51
〈그림 42〉 bibframe.rdf 파일의 내용 중 일부	51
〈그림 43〉 Bibframe Profile 편집기의 Agent 요소 관리 화면〉 페이지	52
〈그림 44〉 Bibframe Profile 편집기의 Agent의 하위요소인 Person 관리 화면	52
〈그림 45〉 Person에 대한 RDF 문서	53
〈그림 46〉 Marva 편집기 초기화면	53
〈그림 47〉 MARVA 편집기의 신규레코드 생성 초기화면	54
〈그림 48〉 미국의회도서관 vegetarian 서지 검색결과 화면	55
〈그림 49〉 MARVA 편집기의 저작 입력화면 중 일부1	55
〈그림 50〉 MARVA 편집기의 저작 입력화면 중 일부2	56
〈그림 51〉 MARVA 편집기의 저작 입력화면 중 저자명 전거 검색 화면1	56
〈그림 52〉 MARVA 편집기의 저작 입력화면 중 저자명 전거 검색 화면2	56
〈그림 53〉 MARVA 편집기의 저작 입력화면 중 저자명 전거 검색 화면3	57
〈그림 54〉 MARVA 편집기의 저작 입력화면 중 저자명 입력 후 화면	57
〈그림 55〉 MARVA 편집기의 저작 입력화면 중 저자명을 문자로 입력했을 경우	58
〈그림 56〉 MARVA 편집기의 저작의 역할 검색 화면	59
〈그림 57〉 MARVA 편집기의 저작의 역할 입력 완료화면	59
〈그림 58〉 한강 작가의 RDF 그래프 표현1	61
〈그림 59〉 한강 작가의 RDF 그래프 표현2	61
〈그림 60〉 한강 작가의 RDF 그래프 표현3	62
〈그림 61〉 MARVA의 저작 중 제목 입력 화면	62
〈그림 62〉 MARVA의 저작의 장르, 형식 입력화면(LCGFT 검색 후 선택)	63
〈그림 63〉 MARVA의 저작의 장르, 형식 입력 완료 후 화면	63
〈그림 64〉 MARVA의 장르/형식 추가 버튼	64
〈그림 65〉 MARVA에서 저작의 출판년도, 원생산지 입력 후 화면	64
〈그림 66〉 MARVA에서 저작의 다른 기여자 입력 후 화면	64

〈그림 67〉 MARVA에서 저작의 다른 기여자 입력 후 화면	65
〈그림 68〉 MARVA에서 저작의 식별번호 입력 후 화면	66
〈그림 69〉 MARVA에서 저작의 콘텐츠 유형과 언어 정보 입력 후 화면	66
〈그림 70〉 MARVA에서 저작의 콘텐츠 유형과 언어 정보 입력 후 화면	67
〈그림 71〉 MARVA에서 인스턴스의 해당 저작 URI 입력화면	67
〈그림 72〉 MARVA에서 인스턴스의 제목, 책임사항, 판사항 입력화면	68
〈그림 73〉 MARVA에서 인스턴스의 제공자 활동 정보 입력화면	68
〈그림 74〉 MARVA에서 인스턴스의 저자 및 기여자 정보 입력화면	69
〈그림 75〉 MARVA에서 인스턴스의 간행방식 입력화면	69
〈그림 76〉 MARVA에서 인스턴스의 식별자 입력화면	70
〈그림 77〉 MARVA에서 인스턴스의 미디어유형, 형태사항, 수록매체 입력화면	70
〈그림 78〉 MARVA에서 인스턴스의 URI 정보 입력화면	71
〈그림 79〉 MARVA에서 인스턴스의 관리 메타데이터 정보 입력화면	71
〈그림 80〉 MARVA에서 입력 완료 후 코드보기 화면	72
〈그림 81〉 국립중앙도서관 홈페이지의 채식주의자 검색결과 화면	113
〈그림 82〉 미국의회도서관 링크드 데이터 서비스 페이지	126
〈그림 83〉 독일 국가도서관 목록검색화면	138
〈그림 84〉 Libris XL 검색화면	139
〈그림 85〉 미국의회도서관 BIBFRAME 저작에서 채식주의자 검색결과 화면	140
〈그림 86〉 BIBFRAME 채식주의자 검색결과 상세 화면	141
〈그림 87〉 BIBFRAME 채식주의자 인스턴스 상세 화면	141
〈그림 88〉 BIBFRAME 채식주의자 개별자료 상세 화면	142

01 들어가는 글

새로운 개념이 등장하거나, 새로운 사건이 발생했을 때 이를 이해하기 위해서는 맥락을 이해하는 것이 중요합니다. 우리가 '새롭다'라고 느끼는 많은 것들은 전에 없던 완전히 새로운 무엇인가보다는 이전에 문제가 되는 무엇을 해결할 수 있는 새로운 시도, 방법인 경우가 대부분일 것입니다. 즉, 새로운 개념을 이해하기 위해 이전의 문제가 무엇인지를 알고, 최근 일어나고 있는 새로운 상황이나 사건들을 함께 검토한다면 좀 더 완벽한 이해를 도울 수 있을 것입니다.

이전 문제가 무엇인지 아는 것과 비슷한 관점에서 이전에 유지했던 방식이 무엇인지를 아는 것도 중요합니다. 왜냐하면 새로운 무엇인가가 등장했다고 하더라도 모든 것이 완벽하게 새롭지는 않습니다. 최근 관심을 많이 받는 자율주행 자동차를 생각해봅시다. 내가 운전행위를 하지 않고, 기계가 알아서 운전을 해준다고 합니다. 분명히 이전에 없던 새로운 개념의 출현입니다. 하지만 여전히 자동차의 외관구조는 동일합니다. 운전대와 브레이크 페달, 가속페달, 운전석, 좌석, 계기판 등 모든 것이 이전의 운전환경을 동일하게 유지합니다. 물론 언젠가 가까운 미래에 기술이 좀 더 발전하면 자동차의 외관뿐만 아니라 내부 구성도 많은 부분 변할 것입니다. 또 여기서 중요한 것은 자율주행 자동차를 현실에서 구현하려면 법이나 관련 정책 등도 함께 변해야 하고, 사람들의 인식도 변해야 합니다. 어쩌면 도로 환경이나 신호등 체계 등 많은 교통 환경도 함께 변화가 일어날 것입니다. 그래도 여전히 많은 부분 기존의 질서와 체계를 유지합니다. 새롭고 낯선 것이 등장했지만 여전히 우리가 지키고 있는 규칙도 유효합니다. 쉽게 예전 방식을 버리기 어렵습니다. 왜냐하면 현재 우리가 지키고 있는 여러 가지 규칙들은 과거 시행착오를 거쳐 수정하면서 현재까지 이르렀고 많은 사람들이 타당하다고 받아들이고 있는 것들이기 때문입니다.

이제 BIBFRAME과 관련해서 우리(도서관, 사서)에게 익숙한 것과 새로운 것이 무엇인지 먼저 검토해보겠습니다. 정보자원을 관리(목록화)할 때 우리에게 익숙한 방식을 한번 생각해 보겠습니다. 우리는 표준의 중요성을 너무 잘 알고 있습니다. 임의 혹은 내가 속

한 기관, 내가 속한 지역에서만 통용되는 방식은 허용하지 않습니다. 우리가 목록작업을 할 때 사용하는 MARC는 엄격한 규칙으로 구성된 표준입니다. 우리는 엄격한 편목 규칙에 따릅니다. 우리가 따르는 편목 규칙에는 어떤 서지요소를 어떤 규칙에 따라 사용해야 하는지가 미리 정의되어 있습니다. 한마디로 우리는 표준 편목규칙에 따라 다양한 정보자원에 대한 목록작업을 수행합니다. 우리에게 표준 기반의 목록 작업은 너무나 당연하고 자연스러운 것입니다. 표준을 지킴으로써 데이터 교환 시 얻을 수 있는 혜택을 우리는 잘 알고 있습니다.

또 우리는 전거 통제의 중요성을 잘 알고 있습니다. 오랜 기간 우리는 표목명, 이름(사람, 기관 등), 주제 등을 통제하여 일관성을 유지하기 위해 노력해왔습니다. 이는 하나의 개념을 고유하고 유일하게 식별하도록 해줍니다. 또한 통제된 데이터를 기반으로 더 정확하거나 유사한 다른 정보로의 연결 등을 통해 이용자 서비스 품질을 높여 왔습니다. 자연스럽게 무엇인가를 유일하고 고유하게 식별하는 데 있어서 식별자의 중요성을 알고 있고, 사용해 왔습니다.

우리의 이런 경험은 굉장히 중요합니다. 우리는 이 경험을 자원을 관리하고 서비스하는 데 자연스럽게 사용합니다. 우리는 표준을 준수하지 않고 임의로 자원을 관리하거나, 통제 어휘없이 단순히 생긴 모양, 형태로 특정 개념을 구분하는 것을 경계합니다. 자원을 고유하게 관리하기 위해 기본적으로 식별자가 필요함을 알고 있습니다. 무엇보다 데이터 관리에 있어서 '나'가 아니라 '우리'라는 관점을 항상 유지하고자 노력합니다. 소위 상호운용성 확보가 정보자원과 데이터 관리에 있어서 얼마나 중요한 문제인지 잘 알고 있습니다.

본 해설서의 대상은 BIBFRAME입니다. 미국의회도서관은 2011년 5월에 공식적으로 서지 프레임워크 이니셔티브(Bibliographic Framework Initiative)를 시작했습니다. 미국의회도서관을 중심으로 개발되었고, 현재도 진행 중인 새로운 서지 프레임워크입니다. 실제 BIBFRAME의 구성은 간단하다고 볼 수 있습니다. BIBFRAME의 전체 개념(개체 간의 관계)과 이 개념을 표현하는 데 사용되는 표준 용어와 입력규칙이 전부라고 할 수 있습니다. 문제는 표현하는 방법이 RDF(Resource Description Framework)를 준수하고 있어 배경 지식이 없으면 이를 이해하는 데 어려움이 존재합니다. RDF를 설명할 때 빠지지 않는 개념이 링크드 데이터(Linked Data), 시맨틱 웹(Semantic Web)입니다. 링크드 데이터는 시맨틱 웹을 구현하기 위한 방법 중 하나이고, RDF는 링크드 데이터를 구현하기 위한(표현하기 위한) 방법 중 하나입니다. 이 RDF, 링크드 데이터가 우리에게 새로운 것들입니다.

이 새로운 것들을 기술 명세서와 같은 문건으로 접하게 되면 큰 어려움을 겪습니다. 이미 이 책을 읽는 많은 분들이 경험했으리라 생각합니다. 대부분의 기술 명세서는 직접 관련 기술내용만을 언급하기 때문입니다. 앞서 언급한 것처럼 우리에게 새로운 이 기술들을

이해하기 위해서는 맥락, 등장 배경을 알 필요가 있습니다. 그러면 왜 이런 기술과 개념들이 등장했는지 납득할 수 있습니다. 등장 배경을 이해하고 즉, 내가 납득하고 나서 기술명세를 접하면 이해하기 더 쉽습니다. RDF, 링크드 데이터의 경우 우리가 이해해야 하는 중요한 맥락, 등장배경은 다양한 인터넷 서비스 중 가장 성공한 웹(Web)의 변화입니다. 인터넷과 동의어처럼 사용할 정도로 우리에게 친숙한 기술입니다. 링크드 데이터는 한마디로 웹을 누구나 사용할 수 있는 데이터베이스로 만드는 일입니다. 지금까지의 웹이 정보 전달을 위한 플랫폼이었다면 이제 데이터 전달을 위한 플랫폼으로 변화하고 있다는 의미입니다. 데이터를 전달한다?, 생각해 보면 우리에게 새로운 것들을 실현하기 위해서는 다시 우리에게 친숙한 것들로 돌아가야 합니다. 누구나 사용할 수 있는 데이터 플랫폼을 만들기 위해서는 모두가 지켜야 할 규칙이 존재해야 합니다. 즉 표준이 있어야 한다는 말입니다. 표준 기반으로 데이터를 관리하기 위해서는 우리가 늘 해왔던 표준 어휘와 (목록)규칙이 필요합니다. 대상을 명확하게 그리고 무엇보다 대상을 유일하게 구분하기 위한 식별자의 도입도 필요합니다. 물론 이것보다 훨씬 많은 것들이 필요할 수도 있지만 이 정도의 필요성만 느껴도 링크드 데이터, RDF, BIBFRAME을 이해하는 데 큰 도움이 됩니다.

어쩌면 낯설다, 어렵다고 느끼는 BIBFRAME의 문제들은 우리가 수 세기 동안 고민하고 쌓아왔던 우리의 데이터 구축 방식, 우리의 일하는 방식과 너무나 유사합니다. 다만 도서관을 벗어나 웹이라는 새롭지만 익숙한 환경에서 우리 데이터들이 공존할 수 있는 방법이 조금 다를 뿐입니다. 그럼에도 새로운 데이터 중심 웹을 만들기 위해 표준과 식별을 중요하게 생각해야 하는 이유는 우리가 데이터를 관리했던 이유와 같습니다.

이 책은 BIBFRAME 해설서라는 이름을 달고 있습니다. 해설서의 목적은 BIBFRAME 모델과 어휘를 다루는 표준문서의 이해를 돕기 위한 것입니다. 사실 모델과 어휘를 이해한다는 것 역시 우리에게 익숙한 방식입니다. 우리는 표준 편목규칙에 따라 데이터를 다룹니다. 모델과 어휘의 이해는 이 방식과 동일합니다. 우리는 이미 더블린 코어(dublin core), MODS와 같이 디지털 저작물과 웹 친화적이면서, 도서관 이외 도메인과 소통하기 위한 다양한 방식을 경험해왔고, 이들의 표준 어휘집을 사용하고 있습니다. 다만 자칫 단순해 보이는 이 작업이 RDF, 링크드 데이터를 만나면 어려운 일처럼 보이게 됩니다. 이견 오해입니다. 본 해설서는 이런 오해들을 줄여서 BIBFRAME을 이해하는 데 도움을 주기 위해 제작하였습니다. 그래서 해설서에는 BIBFRAME의 모델이나 어휘 명세보다는 도서관에서의 등장배경, 웹의 환경변화를 함께 설명합니다. 이는 우리가 BIBFRAME에 대해 납득하기 위해서입니다. 또한 왜 이런 개념이 필요한지, 그리고 RDF, 링크드 데이터가 무엇인지에 대해서도 설명하고 있습니다. 가장 어려울 수도 있는 부분이지만 사실 우리에게 가장 익숙한 방식이기도 합니다. 이 점을 잊지 않고 해설서를 접한다면 BIBFRAME에

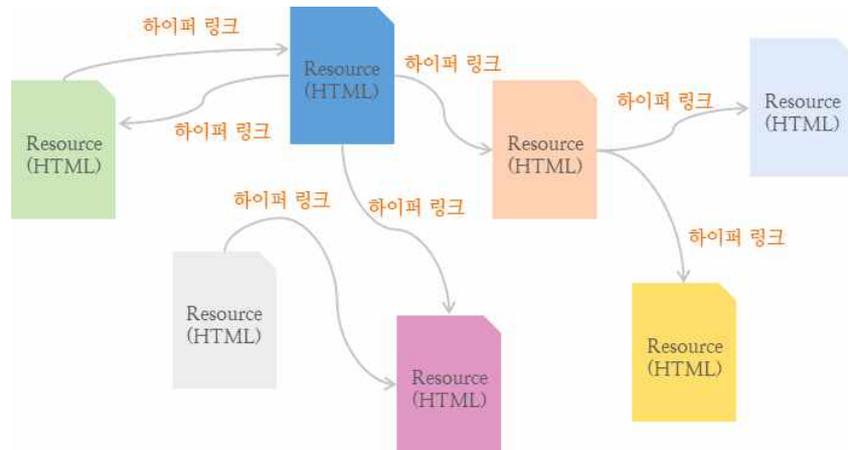
대해 더 잘 이해할 수 있을 뿐만 아니라 앞으로 어떤 새로운 방식의 데이터를 다루는 개념이나 기술이 등장하더라도 여러분이 느낄 장벽이 줄어들 것으로 기대합니다. 잊지 마십시오. BIBFRAME, 링크드 데이터는 낯설기도 하지만 우리에게 가장 익숙한 방식, 우리가 이미 알고 있는 원칙과 기초 하에 작동합니다.

02 웹의 변화

BIBFRAME의 이해를 위해서 반드시 알아야 하는 배경 중 하나는 웹의 변화입니다. 우리가 너무나 친숙하게 사용하고 있고, 이제는 웹이 없으면 일상생활뿐만 아니라 업무까지도 불가능할 정도입니다. 아마 웹이 변화하고 있다고 하면 크게 공감하는 사람이 적을 수도 있습니다. 왜냐하면 매일 사용하는 웹에서 우리가 직접 어떤 점이 변하고 있는지 알아차리기 어렵기 때문입니다. 또 이런 변화는 대부분 기술요소들의 변화로 단순히 웹을 이용하는 우리가 알 필요없는 것들이 대부분입니다.

웹의 변화에서 우리가 알아야 하는 것은 ‘문서 중심의 웹(web of documents)’과 ‘데이터 중심의 웹(web of data)’입니다. 자칫 어려워 보일 수 있지만 사실은 크게 어렵지 않습니다. 현재 문서 덩어리인 문서 중심의 웹을 데이터 덩어리인 데이터 중심 웹으로 만들고자 하는 것입니다. 여기서 자칫 오해가 발생하는 부분이 현재의 문서 중심 웹이 사라지고 데이터 중심 웹이 새롭게 등장하는 것이라는 생각입니다. 데이터 중심 웹이란 현재 문서 중심 웹이 유지되면서 데이터를 자유롭게 공유하고 활용할 수 있는 새로운 웹의 부분 또는 새로운 계층이 만들어지는 것이라고 생각하셔야 합니다.

그렇다면 문서 중심 웹은 무엇이고 데이터 중심 웹은 무엇일까요? 문서 중심 웹은 우리가 늘 사용하고 있는 현재의 웹을 표현한 말입니다. 즉, 우리가 흔히 웹 문서 혹은 웹 페이지라고 부르는 문서들이 서로 연결되어 구성된 현재의 웹을 말합니다. 사실 현재의 웹 생태계는 아주 단순한 3가지 기술요소에 기초합니다. 이 3가지는 ‘http라는 프로토콜’, ‘html이라는 언어’로 작성된 웹 페이지, 그리고 이 문서들을 연결하는 ‘하이퍼링크(hyperlink)’입니다. 이 세 가지 기술요소들로 현재의 웹, 문서 중심의 웹을 표현하면 <그림 1>과 같습니다.



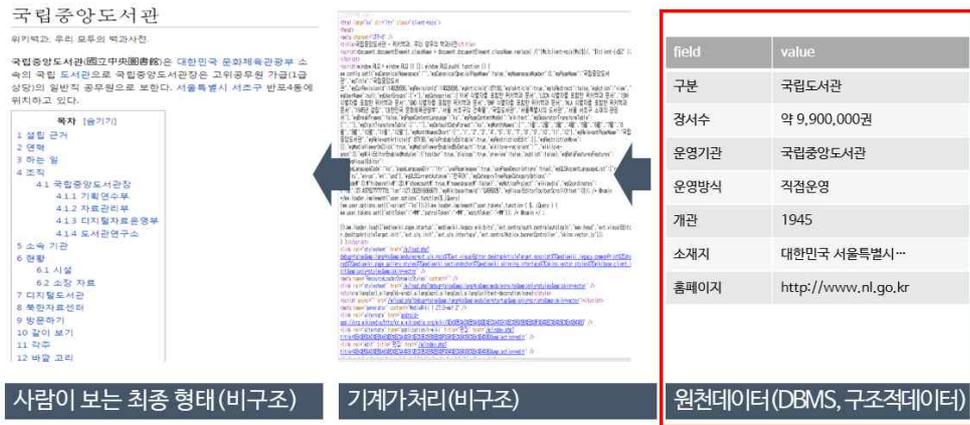
〈그림 1〉 문서 중심 웹(web of document) 개념도

〈그림 1〉에서 보듯이 현재의 웹 환경은 단순합니다. 그리고 문서, 즉 정보를 중심으로 문서와 문서를 연결하는 것이 핵심입니다. 데이터 중심 웹으로 넘어가기 전에 이런 문서 중심 웹 생태계 구조를 먼저 살펴보겠습니다. 흔히 3계층 구조(three-tier architecture)로 이를 설명합니다. 세 개의 계층과 각 계층이 하는 역할은 다음과 같습니다. 여기서는 일반적인 웹 사이트를 기준으로 간단하게 정리하였습니다.

- 프레젠테이션 계층(presentation tier): 최종 이용자가 응용 프로그램과 상호작용하는 계층. HTML 문서로 이용자에게 표현함.
- 애플리케이션 계층(application tier): 논리 계층(logic tier) 또는 중간 계층(middle tier) 이라고도 하는 계층. 데이터 계층에서 정보를 읽어 프레젠테이션 계층으로 전송(데이터 계층의 데이터를 추가, 수정, 삭제할 수 있음)
- 데이터 계층(data tier): 데이터베이스 계층, 백엔드(back-end) 계층으로도 칭하며, 애플리케이션에서 처리하는 정보가 저장되고 관리되는 곳. 일반적으로 데이터베이스를 의미함.

데이터 중심 웹을 설명하기 전에 웹의 구조를 언급한 이유는 간단합니다. 사실 데이터 중심 웹의 필요성 혹은 배경이라고 할 수 있는 문제 상황이 담겨 있기 때문입니다. 오늘날 우리가 익숙하게 사용하는 웹 사이트를 생각해 보면 우리는 늘 우리에게 이미 익숙한 문서(책) 형식의 정보 페이지를 접하고 있음을 알 수 있습니다. 문자와 이미지 혹은 영상이 포함된 이 문서는 우리가 정보를 접하는 아주 익숙한 형식을 갖추고 있습니다. 사람에게는 편리하고 친숙한 방식입니다. 생각해 보면 우리가 익숙하게 접하고 있는 이 문서는 비정형입니다. 그런데 이 비정형 정보의 원천은 데이터베이스에 있는 정형 데

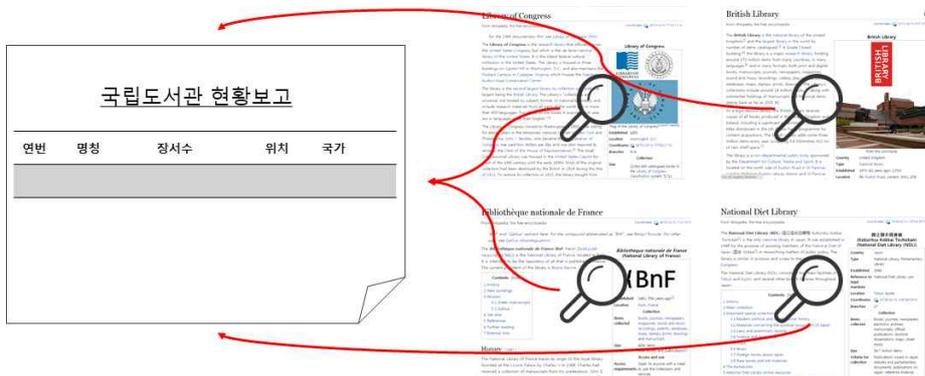
이터입니다. 잘 정리한 정형 데이터를 사람을 위해 다시 비정형 형태로 전달하고 우리는 이 정보를 활용합니다.



〈그림 2〉 웹의 계층별 데이터 구조 예시

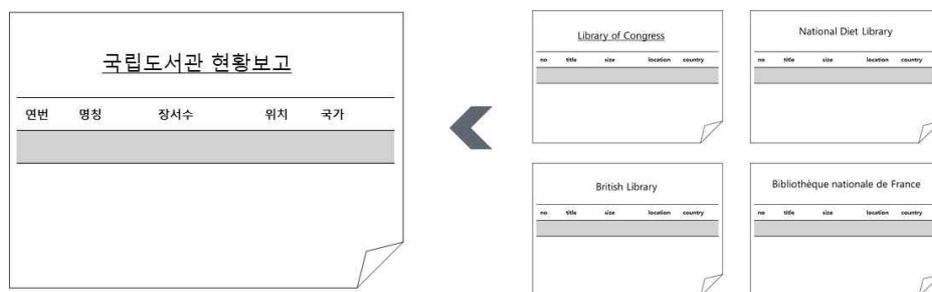
〈그림 2〉는 이해를 돕기 위한 그림이라는 점을 이해해 주십시오. 그림에서처럼 우리가 접하는 최종형태의 문서는 HTML 구성된 비정형 문서입니다. 그런데 이 비정형 문서의 원천데이터는 구조적데이터입니다. 여기서 이런 생각을 할 수 있습니다. 비정형 문서 정보도 중요하지만, 만약 데이터 계층, 즉 데이터베이스에 저장된 정형 데이터를 공유하고 사용할 수 있다면 어떨까요?, 그것도 이미 자유롭게 정보가 공유되고 있는 웹이라는 플랫폼을 그대로 사용하면서 정형 데이터를 공유할 수 있다면 어떤 일이 가능할까요?

한 가지 상황을 상상해보겠습니다. 전 세계 국립도서관의 현황을 정리하라는 업무지시가 있다고 생각해 보겠습니다. 현재 가장 일반적인 접근 방법은 각국의 도서관 현황을 담고 있는 문서를 찾는 것입니다. 최근 정보를 찾기 위해서는 각국 도서관의 홈페이지 혹은 위키피디아 같은 정보 서비스를 이용하는 것이 가장 좋은 방법일 것입니다.



〈그림 3〉 문서정보를 활용한 보고서 작성 예시

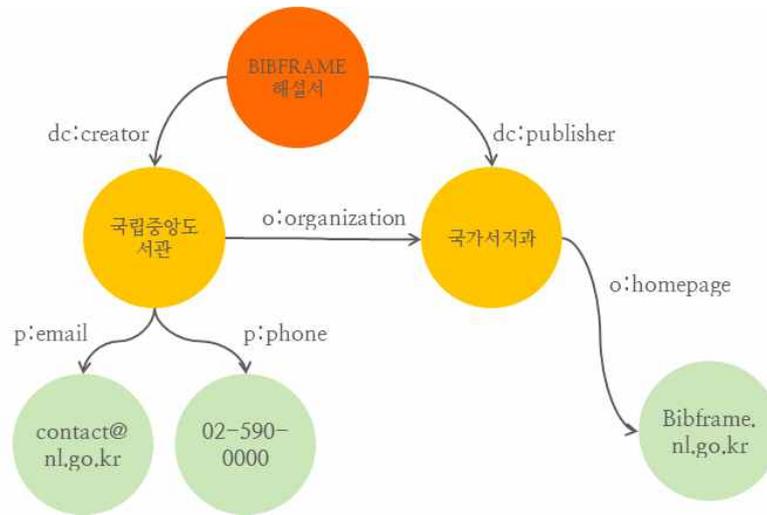
국립도서관 현황에서 조사할 항목들(예: 장서 수, 공식명칭, 위치, 해당 국가, 직원 수, 웹 사이트 주소, 전화번호 등)을 결정하고 각 국립도서관 홈페이지에서 해당 정보를 찾아야 합니다. 이런 정보들이 홈페이지 잘 모여있지 않은 이상은 원하는 데이터를 찾기 위해서 많은 시간과 노력을 기울여야 합니다. 만약 이 문서형태의 정보들의 원천인 데이터베이스가 존재하고 여기에 접근할 수 있다면 어떨까요?



〈그림 4〉 데이터를 활용한 보고서 작성 예시

아마도 보다 쉽게 원하는 데이터를 기반으로 보고서 작성이 가능할 것입니다. 문서가 아니라 데이터에 집중하고자 하는 이유를 쉽게 이해할 수 있을 겁니다. 데이터 중심 웹은 이런 웹을 이야기하는 것입니다. 웹에 정보가 아니라 데이터를 공유할 수 있는 환경을 만들자는 것입니다. 그리고 이미 웹이라는 플랫폼은 데이터베이스에 구조화된 데이터를 가지고 있습니다. 문제는 이 데이터를 어떻게 누구나 잘 사용할 수 있는 형식으로 발행하고 공유할 것인가입니다. 여기서 우리에게 익숙한 방식의 도입이 필요합니다. 앞에서 언급했던 내용을 다시 생각해 보면 표준기반 접근방식, 식별자, 통제어휘가 그 핵심이라 할 수 있습니다.

플랫폼이 핵심이기 때문에 웹을 생각해 보면 이미 웹에서 사용하고 있는 http 프로토콜, HTML, 하이퍼링크 기술을 그대로 활용하면 됩니다. 단, 우리는 사람을 위한 문서를 만드는 것이 아니기 때문에 사람 중심의 가독형 문서를 만들기 위한 HTML 대신 다른 표현방법이 필요합니다. http 프로토콜과 하이퍼링크 기술은 그대로 사용하고 HTML만 기계 처리가 가능한 형식의 RDF로 변경하면 이 문제를 해결할 수 있습니다. 이제 웹의 정보들을 구조화된 방식을 사용해 다시 개념적으로 표현하면 〈그림 5〉와 같은 형식으로 설명할 수 있습니다.



〈그림 5〉 데이터 중심의 웹 정보 개념도

〈그림 5〉는 그래프 형식으로 구조화된 정보의 형태를 표현한 것입니다. 원으로 표시된 노드들과 각 노드와 노드를 연결하는 속성은 방향성을 가지면서 연결된 노드가 무엇인지 설명해줍니다. 여러분에게 친숙한 메타데이터를 속성으로 사용하고 있다는 점을 눈여겨 봐 주십시오. 데이터 중심 웹을 만든다고 했을 때 기계 가독형식으로 데이터를 만들어서 발행한다면 이를 공유하기 위해 제일 중요한 것은 표준을 따르는 것일 겁니다. 앞선 국립도서관 현황 자료 보고서 작업을 다시 떠올려 보십시오. 만약 구조화된 형태로 각국별 국립도서관 현황 자료가 존재한다고 가정하면 작업시간이 많이 줄어들 것입니다. 그런데 국립도서관 현황을 나타내는 각 메타데이터, 즉 속성들의 이름이 다르면 어떨까요? 예를 들어 도서관명을 도서관명, name, organizationName, org_Name 등 서로 다르게 표현하고 있다면 이를 이해하기 위한 별도 노력이 필요할 것입니다. 또, 이 메타데이터가 같더라도 서로 표현방식이 다르다면 다시 문제가 됩니다. 예를 들어 개관일이라는 속성은 ‘YYYY-MM-DD’, ‘DD-MM-YYYY’, ‘YYYY/MM/DD’ 등 다양한 방식으로 표현 가능합니다. 표현방식을 서로 다르다면 이를 조사해서 정리해야 하는 사람은 많은 시간과 노력을 데이터 정제에 소모해야 합니다.

우리가 서지정보를 표준 편목규칙에 맞추어 생성하는 이유를 다시 생각해 보십시오. 도서관 간에 데이터 교환에 문제가 없는 이유는 무엇일까요? 이를 확장해서 도서관뿐만 아니라 모든 정보 생산 및 관리 기관 혹은 개인까지도 데이터를 공유할 수 있게 하려면 어떻게 해야 할까요? 단순히 공개만 하는 것이 아니라 누구나 사용할 수 있는 형태로 공개하려면 어떻게 해야 할까요? 우리가 지금까지 서지 데이터를 만들기 위해 했던 노력들이 똑같이 필요할 겁니다. 오늘날 데이터의 중요성이 점점 높아지고 있습니다. 웹은 정보뿐만 아니라 이런 데이터들을 공유할 수 있는 아주 중요한 플랫폼입니다. BIBFRAME도 이런 새로운 데

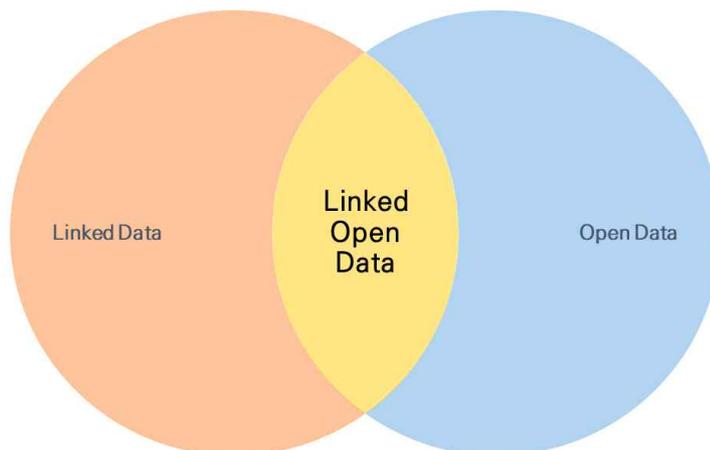
이더 기반 웹에 참여하고자 하는 노력입니다.

03 링크드 오픈 데이터와 RDF

BIBFRAME을 이해하기 위해서는 링크드 오픈 데이터, RDF에 대한 이해가 먼저 필요합니다. 이는 우리에게 낯선 기술에 대한 전문지식이 필요해서라기보다, 이 개념과 기술에서 사용하는 용어에 먼저 익숙해질 필요가 있기 때문입니다. 낯선 용어는 이해를 어렵게 합니다. 먼저 새로운 용어들에 익숙해지면 그 후에는 BIBFRAME에 대한 이해가 수월할 겁니다.

3.1 링크드 오픈 데이터

링크드 오픈 데이터(Linked Open Data)는 링크드 데이터(Linked Data)와 개방형 데이터(Open Data)의 합성어입니다. 즉, 링크드 오픈 데이터는 링크드 데이터이면서 오픈 데이터인 상태를 의미합니다(<그림 6>참조).



<그림 6> Linked Open Data 개념도

먼저 링크드 데이터에 대해 검토해보겠습니다. 링크드 데이터는 이미 거대한 정보생태계

를 갖추고 있는 웹(web)을 플랫폼으로 기계가 처리할 수 있는(machine processable) 형태로 데이터를 개방, 활용하는 것을 의미합니다. 구현되는 형태나 최종적인 목적은 웹을 데이터베이스로 활용하는 것입니다. 여기서 중요한 것은 ‘웹’과 ‘데이터베이스’입니다. 웹은 이미 누구나 친숙하게 사용하고 있는 대표적인 인터넷 서비스의 하나입니다. 웹은 인터넷과 동의어처럼 사용할 정도로 성공한 인터넷 서비스라고 할 수 있습니다. 링크드 데이터가 웹을 플랫폼으로 사용한다는 의미는 웹에서 현재 사용하고 있는 기술들을 그대로 사용한다는 의미입니다. 그래서 링크드 데이터는 HTTP, RDF, URIs와 같은 웹 표준 기술을 활용해서 기계 즉 컴퓨터가 자동으로 읽고, 처리할 수 있는 방식으로 정보를 공유합니다. 링크드 데이터는 서로 다른 출처의 데이터들을 서로 연결하고, 질의 가능하게 합니다.

오픈 데이터와 링크드 데이터가 합쳐진 링크드 오픈 데이터는 웹을 플랫폼으로 하는 가장 적극적인 형태의 개방형 데이터 기술, 활동이라고 볼 수 있습니다. 물론 이전에도 데이터, 정보를 개방하는 다양한 방식이 존재했습니다. 도서관의 경우 MARC 데이터 다운로드, OAI 프로토콜 등이 그 예라 할 수 있습니다. 단순하게 보면 CSV, Excel 형식의 데이터 서비스나 Open API도 그 예입니다. 이런 데이터 개방 방식과 링크드 오픈 데이터의 차이점은 웹의 창시자인 Tim Berners-Lee가 제시한 5 Star open data를 통해 손쉽게 이해할 수 있습니다(<그림 7> 참조).



<그림 7> 5 star open data 개념도(출처: <https://5stardata.info/ko/>)

<그림 7>에서 보는 것처럼 개방 수준이 가장 높은 단계는 별점 5점으로 링크드 데이터라는 속성을 포함하고 있습니다. 별점 1점부터 ‘5 star open data’ 홈페이지의 예제를 중심으로 검토해보면 별점 1점은 온라인 상태인 데이터를 의미합니다. 형식에 상관없이 온라인에서 이용할 수 있는 상태면 됩니다. 대표적인 형식은 PDF라는 문서형태로 <그림 5>와

같습니다.

서울 기온 예보	
날짜	최저 기온(°C)
2012년 2월 5일 일요일	-4
2012년 2월 6일 월요일	-3
2012년 2월 7일 화요일	-10

〈그림 8〉 별점 1점 예제

별점 1점의 경우 PDF 형식의 문서입니다. 데이터를 PDF 형태로 개방했다고 생각해 보면 왜 별점이 하나인지 이해할 수 있습니다. 서울 기온 예보 데이터가 필요한 누군가가 PDF 형태의 문서를 받았다면 이를 다시 다른 형태로 변환하고 정리하는 수고를 해야 합니다. 개방 데이터 입장에서 봤을 때 좋은 방법은 아니라고 할 수 있습니다.

별점 2점은 온라인으로 이용 가능하면서, 기계 가독형태인 데이터를 의미합니다. 〈그림 9〉처럼 마이크로소프트 엑셀 형식의 데이터가 예가 될 수 있습니다.

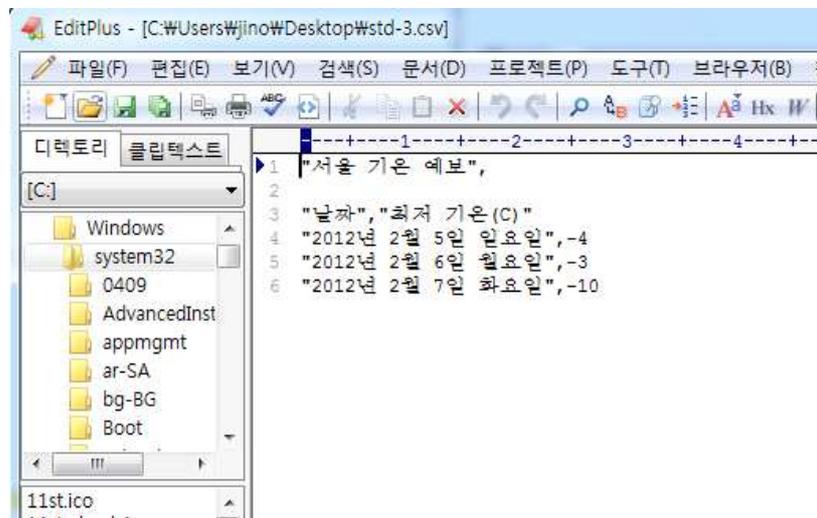
	A	B	C
1	서울 기온 예보		
2			
3	날짜	최저 기온(°C)	
4	2012년 2월 5일 일요일	-4	
5	2012년 2월 6일 월요일	-3	
6	2012년 2월 7일 화요일	-10	

〈그림 9〉 별점 2점 예제

엑셀 형식의 데이터는 활용하기에 무난합니다. 우리나라의 경우 대부분 마이크로소프트 오피스 제품이 사무환경에 설치되어 있습니다. 그런데 별점이 2개인 이유는 마이크로소프트

트라는 특정 회사에 종속되어 있는 소프트웨어를 사용해야 하기 때문입니다. 물론 요즘에는 오픈 오피스나 우리나라의 한컴소프트 제품을 사용하더라도 엑셀 형식의 데이터는 사용할 수 있습니다. 엑셀 형식의 데이터는 이해를 돕기 위한 예제일 뿐입니다. 별점 2개는 구조적인 형태로 데이터를 제공해주는 하지만 특정 소프트웨어가 필요한 상태를 의미합니다.

별점 3점은 온라인으로 이용 가능하고, 기계 가독형태이면서 개방 형식(open format)을 사용한 경우입니다. 대표적인 것이 CSV 형태로 데이터를 개방한 경우입니다.



〈그림 10〉 별점 3점 예제

〈그림 10〉처럼 CSV 형태의 데이터는 특정 운영체제나 소프트웨어에 종속되지 않고 자유롭게 열기와 편집이 가능한 상태입니다. 이용자 입장에서는 데이터를 활용하기에 이전보다 더 좋은 상태라고 할 수 있습니다.

별점 4점은 온라인으로 이용 가능, 기계가독형, 개방형 형식과 함께 URI로 개체 식별이 가능한 데이터 상태를 의미합니다. URI로 개체 식별이 가능하다는 의미는 식별자를 생각하면 쉽습니다. 도서관에서 생산하는 데이터들은 모두 식별자들이 부여되어 있습니다. 식별자를 부여하면 대상 개체는 고유하고 분명해집니다. 단순하게는 동명이인, 동음이의어를 식별하고 구분하기가 쉽습니다. 중요한 것은 오픈 데이터, 특히 링크드 데이터의 경우 플랫폼이 웹이라는 점입니다. 웹에서 개체를 고유하게 식별하는 방법은 URI를 사용하는 것입니다. 똑같은 URI가 서로 다른 정보나 개체에 부여될 수 없기 때문입니다.

서울 기온 예보

날짜	최저 기온(°C)
2012년 2월 5일 일요일	-4
2012년 2월 6일 월요일	-3
2012년 2월 7일 화요일	-10

최종 수정: 2012-01-22, 마이클 | 번역: 2012-02-03, 김보람 | 코드 공유: GitHub

<그림 11> 별점 4점 예제

<그림 11>은 별점 4점의 예제입니다. HTML로 작성된 예제입니다. 이 상태로는 식별자를 부여했는지 알기 어렵습니다. 실제 구성된 코드를 보면 아래와 같은 내용을 확인할 수 있습니다. 아래 코드는 전체 코드 중 RDFa로 인코딩된 부분만을 줄여서 표기한 것입니다. RDFa는 여기서 상세하게 다루지 않습니다. 이는 기존의 웹 표기 방식인 HTML을 그대로 유지하면서 사람, 장소, 사건, 리뷰, 레시피 등 다양한 항목들을 표현하여 검색 목록을 생성하기 위한 방법입니다. 링크드 데이터를 꼭 RDF 하나로만 구현하거나 표현할 수 있는 것은 아닙니다. RDFa처럼 HTML을 유지하면서 표현하는 방법들도 다양하게 존재합니다.

```

<html xmlns="http://www.w3.org/1999/xhtml"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:meteo="http://purl.org/ns/meteo#">
<head>
<h1 property="dcterms:title">서울 기온 예보</h1>
<div id="data" about="#Seoul" typeof="meteo:Place">
  <tr rel="meteo:forecast" resource="#forecast20120205">
    <td>
      <div about="#forecast20120205">
        <span property="meteo:predicted" content="2012-02-05T00:00:00Z"
          datatype="xsd:dateTime">2012년 2월 5일 일요일</span>
      </div>
    </td>
  </tr>

```

위 코드표에서 확인할 수 있는 사항은 PDF, 엑셀, CSV와 다르게 실제 값(인스턴스)이 들어가는 영역을 설명하는 메타데이터와 이 메타데이터가 어느 용어집 표준을 따르고 있는지 표기되어 있다는 점입니다. 코드에서 보면 상단에 다음과 같은 내용을 확인할 수 있습니다.

```
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:meteo="http://purl.org/ns/meteo#">
```

이는 서울의 기온을 표현하기 위해 표준용어집, 표준 메타데이터를 사용하겠다는 의미로 도서관에서는 친숙한 더블린코어 메타데이터와 날씨 정보를 표현하기 위해 기상관련 용어 집인 "http://purl.org/ns/meteo#"를 표준으로 사용하고 있음을 확인할 수 있습니다. 앞선 예제들과 비교해서 보다 명확해진 점은 서울의 날씨 정보를 표준 메타데이터(속성)를 활용해서 표현하고 있다는 점입니다. 쉽게 표라고 표현하면 서울의 날씨를 표현한 이전의 표에서 '날짜', '최저기온'은 어떤 규칙하에 어떤 값들을 넣어야 하는지 기준정보가 없습니다. 별점 4점은 이 부분이 개선되었다고 볼 수 있습니다. 표준에 맞추어 데이터를 입력했다는 점은 도서관, 사서 입장에서 보면 너무나 당연한 과정일 수 있습니다. 이런 당연한 일들이 웹에서 일어나고 있는 게 시맨틱 웹, 링크드 데이터 구현을 위한 노력이라고 보면 BIBFRAME을 이해하는 데 큰 도움이 될 수 있습니다.

어쩌면 별점 4점이면 표준을 잘 지키고 잘 구조화된 데이터를 제공해주고 있다고 볼 수도 있습니다. 하지만 링크드 데이터는 별점 5점으로 온라인 이용 가능, 기계가독형, 개방형 형식, URI로 개체 식별 외에 다른 데이터와의 연결이 필요합니다. 우리는 하나의 개체, 자원을 잘 이해하기 위해 같은 것들끼리 묶어주는 것, 유사한 것과 연결시키는 것이 얼마나 중요한지 잘 알고 있습니다. 별점 5점의 핵심은 여기에 있습니다.

좀 더 구체적으로 별점 5점의 의미를 알아보도록 하겠습니다. 사실 이 부분은 도서관에서는 너무나 당연하게 받아들이고 있는 부분입니다. 왜냐하면 도서관은 아주 오래전부터 하나의 표준을 모든 도서관들이 함께 사용해왔습니다. 우리는 표준을 준수하는 것이 얼마나 중요한 일인지 잘 알고 있습니다. 또한 전거 데이터(도서관의 이름전거와 주제명 전거)와 같이 다른 정보원과 연결함으로써 보다 풍부한 정보를 전달할 수 있고, 그 자체로도 훌륭한 정보원이 되는 데이터의 중요성을 잘 알고 있습니다. 또한 동일한 형태의 문자로 대상을 구분하는 것이 아니고 식별자를 부여하고 상위, 하위, 유사 등 다양한 관계로 하나의 개념을 관리하는 것의 중요성도 잘 알고 있습니다. 마지막으로 식별자를 활용해서 해당 개체를 독립적으로 관리하는 것의 중요성도 너무나 잘 알고 있습니다. 잘 알고 있다고 표현한 이유는 우리는 일상적으로 이런 활동을 수행하기 때문입니다.

임베디드 데이터 보기

서울 기온 예보

날짜	최저 기온(°C)
2012년 2월 5일 일요일	-4
2012년 2월 6일 월요일	-3
2012년 2월 7일 화요일	-10

최종 수정: 2012-01-22, 마이클 | 변경: 2012-02-03, 김보람 | 코드 공유: GitHub

〈그림 12〉 별점 5점 예제

별점 5점은 4점과 유사한 형태로 보입니다. 그런데 데이터 소스를 확인해 별점 4점과 비교해보면 〈표 1〉과 같은 차이를 발견할 수 있습니다.

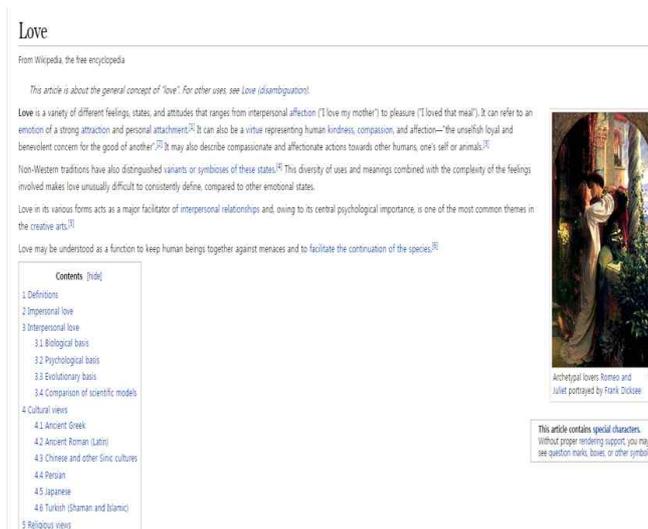
〈표 1〉 별점 4점과 5점 차이 비교

별점4	별점5
<pre><h1 property="dcterms:title">서울 기온 예보</h1> <div id="data" about="#Seoul" typeof="meteo:Place"></pre>	<pre><h1 property="dcterms:title">서울 기온 예보</h1> <div id="data" about="#Seoul" typeof="meteo:Place"> </pre>

〈표 1〉에서 보는 것처럼 별점 5점의 경우는 장소 “Seoul”에 대한 연관 정보로 dbpedia 의 Seoul이라는 개념을 링크하고 있습니다. 여기서 개념에 연결했다는 표현을 사용했습니다. 연결된 URL을 보면 “http://dbpedia.org/resource/Seoul” resource 아래에 Seoul이라는 페이지로 연결된 것처럼 보입니다. 그런데 실제로 이 URL로 접근하면 URL은 자동으로 “https://dbpedia.org/page/Seoul”로 이동합니다. resource에 표기된 Seoul은 개념적인 도시 서울을 의미합니다. 즉, 실체가 없습니다. 서울이라는 개념은 page 아래에 Seoul이라는 웹페이지에 표기된 다양한 정보들을 조합하면 완성됩니다. 즉 〈표 1〉에 연결된 것은 서울 기온 예보에서 설명하는 장소 ‘Seoul’의 개념이 무엇인지를 연결할 것입니다. 여기서 꼭 기억해야 하는 것은 하나입니다. 바로 ‘개념’을 연결했다는 것입니다.

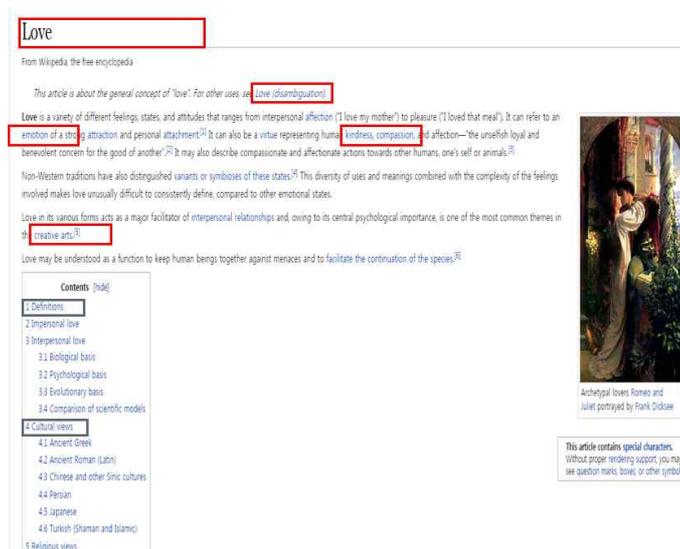
개념을 연결했다는 것은 어떤 의미일까요? 사실 링크드 데이터를 이해하는데 가장 핵심

적인 사항이 이 부분입니다. 앞서 문서 중심의 웹과 데이터 중심의 웹을 이야기했었습니다. 문서 중심의 웹과 데이터 중심 웹의 가장 큰 차이가 바로 ‘문자’냐 ‘개념’이냐에 있다고 볼 수 있습니다. <그림 13>은 위키피디아에 있는 ‘Love’를 설명하는 문서입니다.



<그림 13> 위키피디아의 ‘Love’ 페이지 중 일부

<그림 13>에서 보이는 것은 사실 문자의 나열입니다. 우리는 나열된 문자들로 구성된 문장을 읽고 이해할 수 있습니다. 사람은 가능합니다. 왜냐하면 사람은 각 각의 문자들이 의미하는 바가 무엇인지를 모두 이해하고 있기 때문입니다. 다음 그림을 다시 살펴보겠습니다.



<그림 14> 위키피디아의 ‘Love’ 페이지 중 링크 정보 확인

〈그림 14〉에서 테두리를 친 문자들은 링크가 걸려있습니다. 해당 링크를 클릭하면 해당 문자를 설명하는 또 다른 웹 페이지(문서)로 이동합니다. 해당 페이지로 넘어가면 해당 문자가 의미하는 바의 상세한 설명을 접하게 됩니다. 역시 우리가 사람이기 때문에 또 각각의 문자가 의미하는 바가 무엇인지를 알기 때문에 가능한 일입니다.

기계라면 어떨까요? 기계도 우리처럼 문자로 구성된 문서를 보면 이해하고 처리할 수 있을까요? 불가능합니다. 그렇다면 기계가 이해까지는 못하더라도 자동으로 의미를 찾아 연결하고 데이터 처리가 가능한 수준까지 가려면 어떤 것들이 필요할까요? 필요한 기술적 조치에 앞서 사람이 이해하는 방식과 비슷하게 생각하면 쉽습니다. 우리가 사용하는 단어, 용어들은 사전에 그 뜻과 용례들이 자세히 설명되어 있습니다. 문자가 아니라 해당 문자를 하나의 개념으로 기계가 인식하기 위해서는 우리가 사용하는 사전처럼 해당 문자에 대한 정의, 다른 문자 혹은 개념과의 관계, 이명 표기 등 다양한 정보를 담고 고유한 식별자를 부여하여 참조하도록 하면 가능합니다. 사실 이런 일들은 이미 도서관에서 오랫동안 수행해 왔습니다. 우리가 잘 알고 있는 이름 전거, 주제명 전거가 대표적 예입니다.

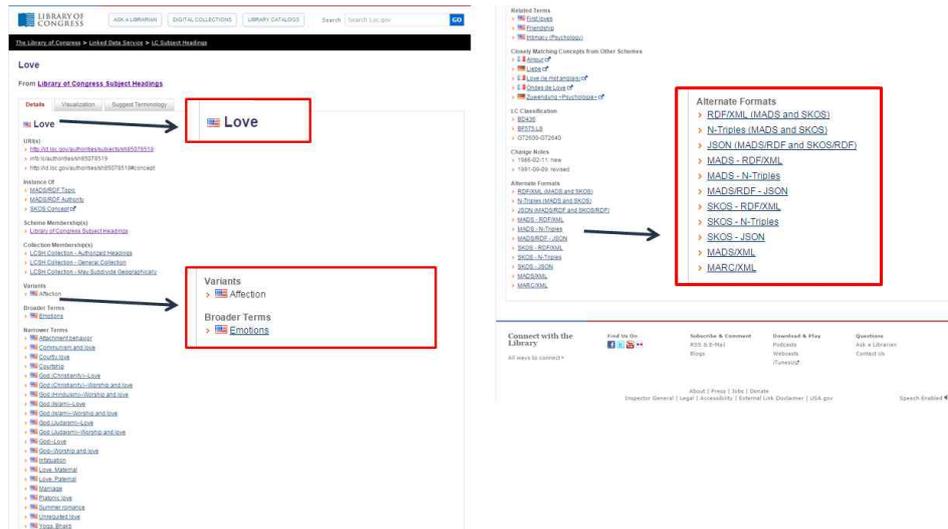


〈그림 15〉 국립중앙도서관 전거 정보 중 ‘사랑’ 검색결과 화면

〈그림 15〉는 국립중앙도서관 주제명 전거 홈페이지에서 ‘사랑’을 검색한 결과입니다. 그림에서 보듯이 ‘사랑’이라는 주제어의 식별자(URL과 주제어 식별번호)가 부여되어 있고, 용어설명, 다양한 언어표현, 상위어와 하위어 정보 확인이 가능합니다. 이렇게 관리하면 ‘사랑’은 문자열이 아니라 우리가 알고 있는 감정 중의 하나인 사랑이라는 개념으로 정립

가능합니다. 기계가 이를 참조하도록 하면 사람 이름, 가게 이름 등 다양한 ‘사랑’이라는 표현과 감정을 표현하는 ‘사랑’은 명확하게 구별됩니다.

미국의회도서관의 경우 이 서비스가 보다 잘 확장되어 있다고 볼 수 있습니다.



〈그림 16〉 미국의회도서관 주제명 표목 중 ‘love’ 페이지

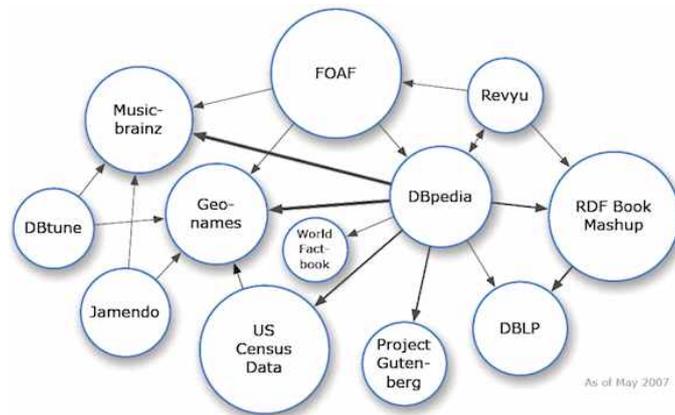
〈그림 16〉은 미국의회도서관 주제명 표목 서비스 중 ‘love’의 해당 페이지입니다. 국립중앙도서관과 마찬가지로 URL로 식별하고, 용어정의, 이형표기, 상/하 관계표현은 물론 다양한 용어집과 표현방법으로 해당 용어를 하나의 데이터로 활용하도록 서비스하고 있습니다. 국립중앙도서관의 경우도 LOD 링크정보를 제공하여 미국 의회도서관처럼 RDF 형식의 데이터를 서비스하고 있습니다. 우리가 그동안 구축한 이런 데이터는 링크드 데이터 생태계에서 아주 중요한 역할을 합니다.

링크드 오픈 데이터는 우리와는 전혀 상관없는 새로운 기술 트렌드라고 생각하는 도서관이나 사서들이 있을 수 있습니다. 지금까지 해설서에서 논의했던 내용을 다시 한번 생각해 보면 그렇지 않다는 점을 알 수 있습니다. 사실 도서관은 데이터 개방이나 링크드 데이터라는 개념이 등장하기 이전부터 기계 가독형 데이터에 대한 노력을 기울여왔고, 끊임없이 데이터를 개방해 왔습니다. 다만 지금은 보다 더 많은 데이터, 기관, 세상과 소통하기 위한 방법 중 하나를 기존의 노력에 더했을 뿐입니다. 다시 강조하지만 우리에게 낯선 기술요소도 있지만 역시 우리가 해왔던 방식이 필요하다는 점을 확인할 수 있습니다.

3.2 링크드 오픈 데이터의 규모

링크드 오픈 데이터 클라우드(Linked Open Data Cloud, <https://lod-cloud.net>)는 전 세계의 링크드 데이터 발행 현황을 도식화하여 꾸준히 발행하고 있습니다. 물론 여기에 포함되지 않은 수많은 데이터들이 존재합니다. 이 프로젝트에서는 그 중 데이터의 품질이 높고, 연결성이 강한 대표적인 링크드 데이터들을 선별하여 보여줍니다. 가장 최근의 그림은 2020년 5월에 작성된 것이며, 시작은 2007년 5월 1일이었습니다.

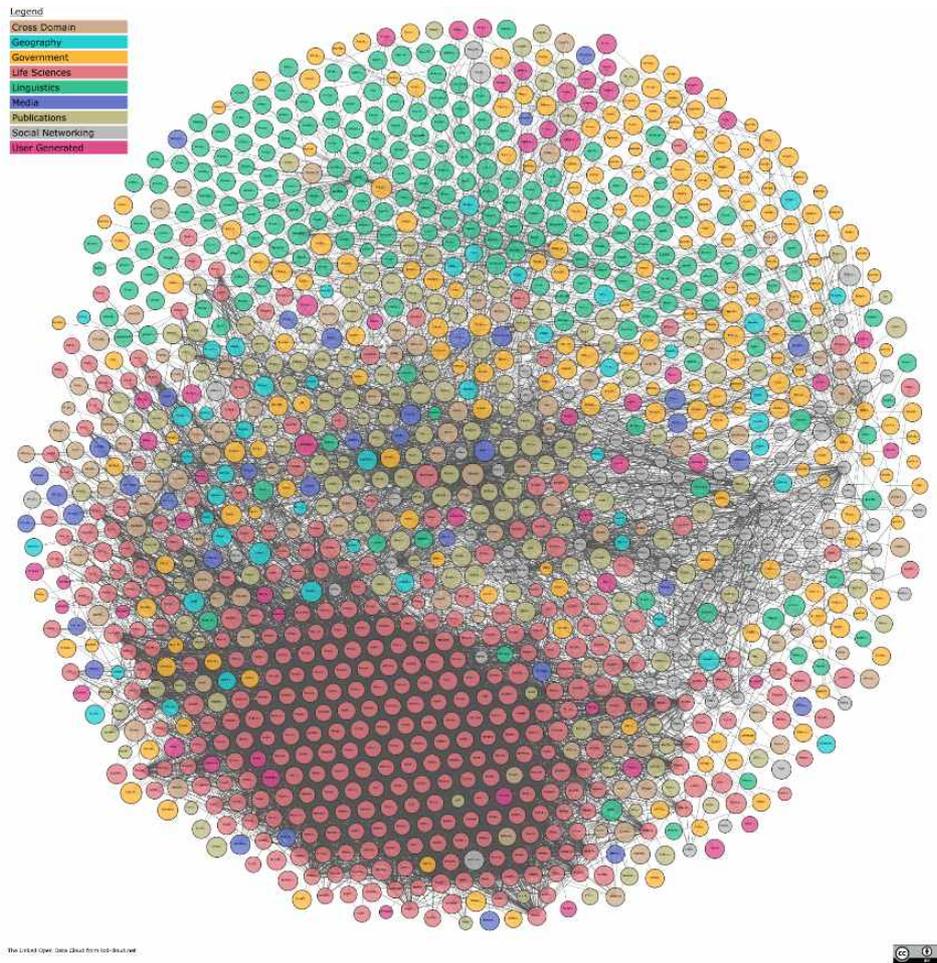
〈그림 17〉은 2007년 5월 1일에 발행된 첫 번째 링크드 오픈 데이터 다이어그램입니다.



〈그림 17〉 2007년 링크드 오픈 데이터 다이어그램

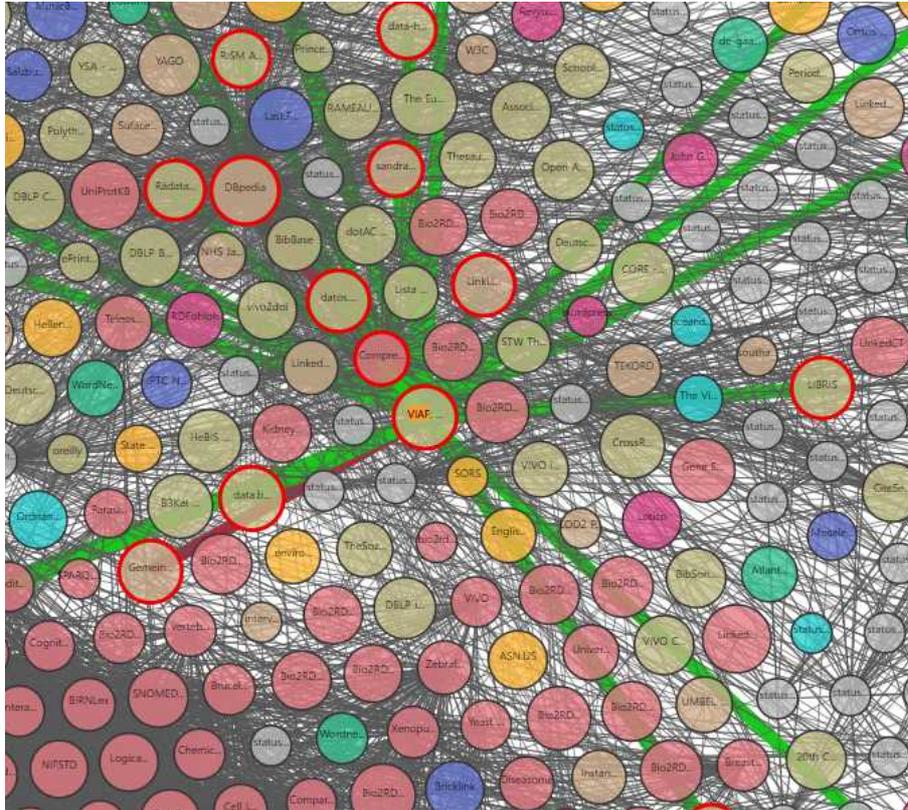
첫 다이어그램에는 12개의 데이터 세트가 서로 연결된 모습을 확인할 수 있습니다. 모든 링크드 데이터의 중심에 존재하는 것이 DBpedia로, 이는 Wikipedia의 정보를 링크드 데이터 형식으로 발행한 것입니다. 도서관의 전거 데이터와 같이 DBpedia는 웹 상의 자원들을 표현하고 연결하기 위한 기준, 중심 역할을 수행하고 있습니다.

가장 최근인 2020년 5월 20일에 발행된 다이어그램에는 1,255개의 데이터 세트가 표현되어 있으며 〈그림 18〉과 같습니다.



〈그림 18〉 2020년 링크드 오픈 데이터 다이어그램

〈그림 18〉에서 확인할 수 있듯이 현재는 범위가 너무 커서 전체 이미지로만 봐서는 상세 데이터세트 확인이 어려울 정도입니다. 이 중 대표적인 도서관 데이터 세트인 VIAF를 확대해보면 〈그림 19〉와 같습니다.



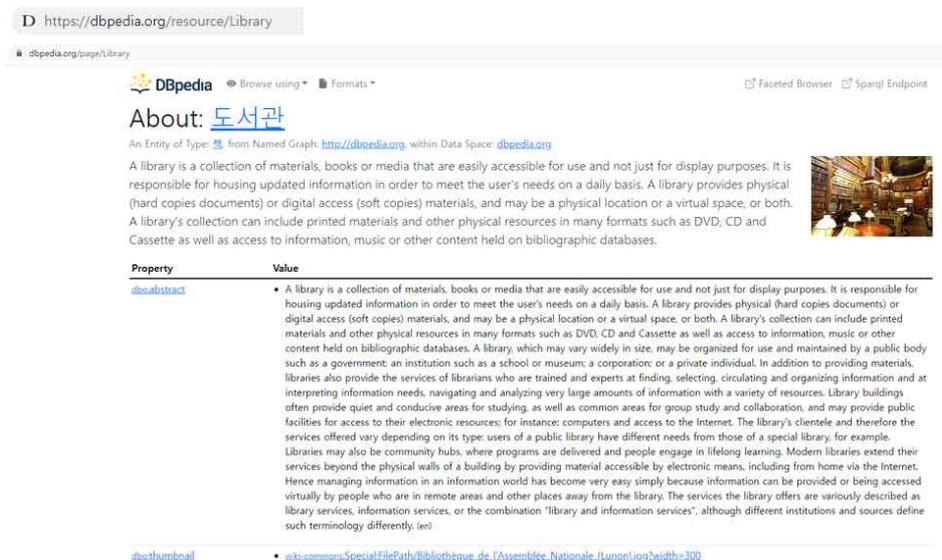
〈그림 19〉 링크드 오픈 데이터 다이어그램의 VIAF 연결 모습

VIAF 출판(publication) 분야에 속하는 데이터셋으로 출판은 물론 다양한 도메인과 강하게 연결되어 있는 모습을 확인할 수 있습니다. 특히 링크드 오픈 데이터의 핵심이라 할 수 있는 DBpedia와 강한 연결고리를 갖고 있음을 확인할 수 있습니다. 즉, 도서관의 전거 데이터가 링크드 오픈 데이터 세상에서 중요한 기준이자 토대 데이터임을 확인할 수 있습니다.

이런 링크드 데이터를 구축하는 방법은 사실 한 가지가 아닙니다. 여러 가지 방법이 존재하는데 어떤 방법을 사용하더라도 기본적으로 지켜야 하는 원칙은 존재합니다. Tim Berners-Lee는 4가지 원칙을 제시하였습니다. 이는 데이터 중심 웹(Web of Data)에서 누구나 지켜야 하는 아주 기본적인 약속입니다. 기술적인 원칙은 기존의 웹(Web of Documents)에서 활용되고 지금의 웹 생태계를 가능하게 한 핵심기술을 그대로 사용합니다.

첫 번째 원칙은 개체 식별을 위해 URI를 사용해야 한다는 것입니다. 플랫폼이 웹이니 무엇인가를 유일하고 의미있게 구분하기 위한 방법으로 URI를 사용하는 것은 당연한 일입니다. URI는 고유하기 때문에 식별자의 역할도 할 수 있고 해당 페이지에서 상세한 정보 제공 역시 가능합니다. 현재의 웹과 동일한 방식입니다. 앞서 설명했던 위키피디아의 링크

드 데이터 버전인 DBpedia를 예로 설명하겠습니다. ‘도서관’이라는 개념, 개체를 고유하게 구분하는 식별자, URI는 “https://dbpedia.org/resource/Library”입니다. 실제로 이 개념에 대한 구체적인 정보는 “https://dbpedia.org/page/Library”에서 확인할 수 있습니다. “https://dbpedia.org/resource/Library”로 이동하면 자동으로 “https://dbpedia.org/page/Library” 페이지로 URI가 이동합니다.



〈그림 20〉 DBpedia에서 Resource와 Page

두 번째 원칙은 HTTP 프로토콜을 사용해서 데이터 정보에 접근을 허용해야 한다는 것입니다. 앞서 우리는 HTTP가 현재의 웹을 가능하게 한 핵심적인 기술요소라는 것을 확인했습니다. LOD에서도 마찬가지로 데이터에 대한 정보의 요청과 응답을 HTTP를 활용합니다.

세 번째 원칙은 RDF, SPARQL 사용과 제공에 관한 것입니다. 이는 URI로 정보를 찾았을 때 RDF(S), SPARQL과 같은 표준을 활용해서 유용한 정보를 제공해야 함을 의미합니다. RDF(S), SPARQL은 기존 웹에서 사용하지 않던 기술요소입니다. RDF에 대한 설명은 이후에 상세하게 다시 설명하겠습니다. SPARQL은 Simple Protocol and RDF Query Language의 약자입니다. LOD 원칙을 준수해서 웹에 공개된 데이터들은 거대한 데이터베이스와 같은 형태를 갖게 됩니다. 기존에 일반적인 정보화 환경에서는 데이터베이스에서 필요한 정보를 반출하기 위해서 SQL(Structured Query Language) 표준을 사용했습니다. 유사하게 데이터 웹에서는 필요한 정보들을 조회하고 반출하기 위해서 SPARQL 표준을 활용합니다.

RDF는 “Resource Description Framework”의 약자입니다. Resource는 URI를 부여할 수 있는 모든 개체, 개념을 의미합니다. 이미 자연에 존재하는 모든 것, 사람이 만들어 놓은 모든 것, 그리고 개념적으로 사람의 머릿속에 존재하는 모든 것들이 Resource입니다. Description은 이 Resource를 아주 상세하게 설명한다는 의미입니다. Framework은 ‘틀’, ‘표현방식’, ‘언어’ 등으로 이해할 수 있습니다. 즉 RDF는 URI를 부여해서 특정 개체나 개념을 아주 상세하게 설명하는 방식, 언어를 말합니다.

RDF의 구조는 아주 단순하고, 이미 우리에게 익숙한 방식입니다. 우리가 복잡하고 다양한 어떤 대상을 체계적으로 정리하고 설명할 때 사용하는 방식은 표를 만드는 것입니다. 우리에게 이미 익숙한 관계형 데이터베이스도 표형태를 갖고 있습니다. 다음과 같이 책 정보를 정리해 놓은 표를 한번 생각해 보겠습니다.

제목	저자	출판사	가격
제 꿈 꾸세요	김멜라	문학동네	14,500원
불편한 편의점 2	김호연	나무옆의자	14,000원
나는 불안할 때 논어를 읽는다	판명 저/이서연 역	미디어숲	17,800원

이 표는 책이라는 개체에 대한 정보를 담고 있고 다음과 같은 문장으로 표현할 수 있습니다.

“ 우리도서관에서 이번 달에 많이 대출된 책이야. 첫 번째는 제 꿈 꾸세요 라는 책이고, 작가는 김멜라야, 문학동네에서 출판된 책이지. 가격은 14,500 원인데 아마 온라인에서는 좀 더 할인된 가격에 살 수 있을거야”

우리가 데이터베이스를 만들고 저장하는 이유는 다수의 사람들에게 특정 목적에 맞는 정보를 전달하기 위해서입니다. 사실 우리가 늘 소통하는 많은 정보 중에 정보화, 데이터베이스화를 통해서 효율적인 소통이 가능하도록 하는 것이 데이터베이스입니다. 이런 표나 데이터베이스의 구조는 위의 예처럼 일반적인 언어생활의 정보 전달 메시지를 체계적으로 정리한 것이라고 볼 수 있습니다.

RDF도 마찬가지입니다. 그리고 RDF의 구조는 훨씬 단순합니다. RDF는 SPO라는 기본적인 구조로 구성되어 있습니다. SPO는 일반적으로 주어, 술어, 목적어로 번역합니다.

RDF의 구조 = 주어(Subject) - 술어(Predicate) - 목적어(Object)

그리고 결국 RDF로 표현되는 구문은 다음과 두 가지 형식만 존재합니다.

RDF의 구조 =	주어(Subject) -	술어(Predicate) -	목적어(Object)
RDF의 구조 =	URI -	URI -	URI
RDF의 구조 =	URI -	URI -	Literal

LOD의 플랫폼이 웹이다보니 무엇인가를 설명하기 위해서는 URI가 기본입니다. 그리고 설명하고자 하는 대상의 값은 다른 개체나 정보를 담고 있는 URI인 경우와 문자 형식으로 간단히 표현하고 종료되는 경우 두 가지밖에 없는 것입니다. 아래 그림은 국립중앙도서관 LOD 서비스에서 SPARQL Endpoint로 검색한 검색결과 화면입니다.

URI	schema:label	schema:CountryMark
http://lod.nl.go.kr/countries/AFG	http://www.w3.org/2002/07/owl#kameku	http://data.kata.kci/countries/AF
http://lod.nl.go.kr/countries/AFG	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://lod.nl.go.kr/ontology/Country
http://lod.nl.go.kr/countries/AGO	http://www.w3.org/2000/01/rdf-schema#label	Republic of Angola
http://lod.nl.go.kr/countries/AGO	http://purl.org/dc/terms/subject	http://lod.nl.go.kr/resource/CountryMark
http://lod.nl.go.kr/countries/AGO	http://www.w3.org/2002/07/owl#kameku	http://data.kata.kci/countries/AG
http://lod.nl.go.kr/countries/AGO	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://lod.nl.go.kr/ontology/Country
http://lod.nl.go.kr/countries/AGO	http://www.w3.org/2000/01/rdf-schema#label	Angola
http://lod.nl.go.kr/countries/AA	http://purl.org/dc/terms/subject	http://lod.nl.go.kr/resource/CountryMark
http://lod.nl.go.kr/countries/AA	http://www.w3.org/2002/07/owl#kameku	http://data.kata.kci/countries/AA
http://lod.nl.go.kr/countries/AA	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://lod.nl.go.kr/ontology/Country
http://lod.nl.go.kr/countries/ALB	http://www.w3.org/2000/01/rdf-schema#label	Republic of Albania

〈그림 23〉 국립중앙도서관 SPARQL Endpoint 검색결과 화면

〈그림 23〉에서 확인할 수 있는 것처럼 RDF로 표현된 구문은 URI - URI - URI, URI - URI - Literal 두 가지만 존재합니다. 그러면 이 문장을 어떻게 표처럼 읽을 수 있는지 알아보도록 하겠습니다. 다음과 같은 간단한 문장을 생각해 보겠습니다.

“제 꿈 꾸세요”의 “작가”는 “김멜라”이다.

앞선 표현식의 데이터 구조와 동일합니다. RDF라는 형식도 결국에는 특정 목적하에 다수와 소통하기 위한 방식 중 하나라고 생각하면 됩니다. 계속 강조하듯이 웹에서 소통하기 위해서는 웹 표준을 따라야 합니다. 즉 URI를 활용한 식별이 필요합니다. 이 문장을 URI를 활용해서 변경해 보겠습니다.

“<http://www.yes24.com/Product/Goods/111746531>”의 작가는

“https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802”이다.

위의 구문은 “제 꿈 꾸세요”에 해당하는 온라인 서점의 URI와 “김멜라”라는 작가의 국립중앙도서관 저자전거 URI를 사용해서 표현한 것입니다. 해당 URI를 사용함으로써 “제 꿈 꾸세요”와 “김멜라”는 세상에서 유일한 개체로 구분됩니다. 여기서 한 가지 작업을 더 수행해 보겠습니다. 작가라는 술어를 우리가 잘 알고 있는 더블링크어의 ‘creator’로 변경해 보겠습니다.

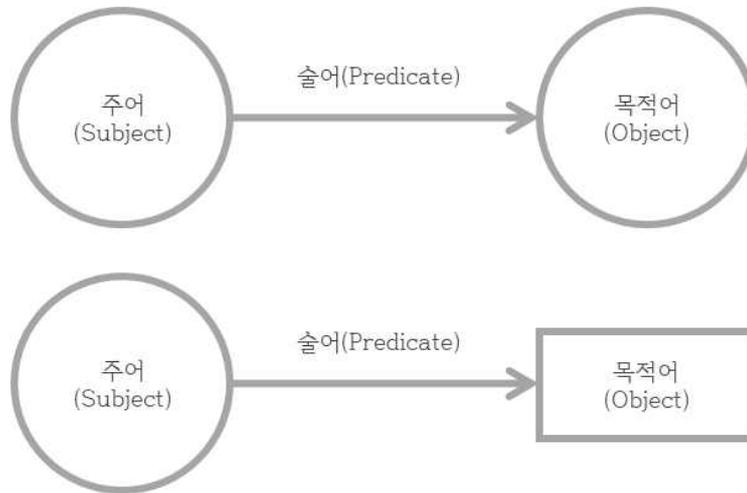
“http://www.yes24.com/Product/Goods/111746531”의

“http://purl.org/dc/terms/creator” 는

“https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802”이다.

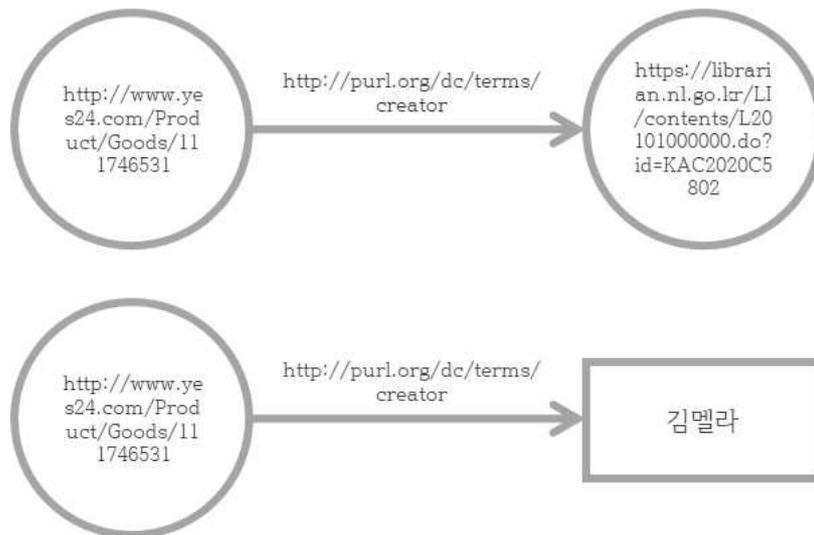
여기서 URI를 사용해서 사용했다는 의미를 다시 한번 생각해 보겠습니다. 웹 환경에서 URI는 다른 페이지로의 이동을 뜻합니다. 즉 RDF에서 URI로 무엇인가를 표현했다는 의미는 식별자의 의미와 함께 해당 링크를 선택하면 관련된 보다 풍부한 정보 확인이 가능하다는 것을 의미합니다. 즉 웹을 기반으로 풍부한 데이터 공유 환경이 만들어진다고 볼 수 있습니다. 예를 들어 우리의 예제에서 더블링크어 ‘creator’라는 URI를 사용함으로써 주어부의 저작물을 만들어낸 사람을 의미한다는 것을 알 수 있습니다. 왜냐하면 더블링크어의 creator는 특정 저작물을 만들어내는데 가장 큰 기여를 한 사람 즉, 작가를 표현할 때 사용할 수 있는 속성, 즉 메타데이터입니다. 해당 URI를 선택하면 ‘creator’가 갖는 의미가 무엇인지 어떻게 사용해야 하는지 정보 확인이 가능합니다. 또 ‘김멜라’라는 문자열로 작가 이름을 표현할 수도 있지만, URI를 사용함으로써 ‘김멜라’라는 동일 이름을 사용하는 다른 사람과 구분된 유일성을 갖게 됩니다. 또 작가에 대한 추가 정보, 상세한 정보는 물론 다른 책이나 관련 정보로 이동 가능하게 합니다. 이런 특징이 일반적인 데이터베이스와 다른 웹을 플랫폼으로 하는 글로벌 데이터베이스 즉 RDF의 장점이라 할 수 있습니다.

보통 RDF는 XML과 함께 RDF/XML 형식으로 표현하는데 이해를 위해서는 RDF 그래프 모델이 더 좋습니다. 주어, 술어, 목적어를 기준으로 표현하면 다음과 같습니다.



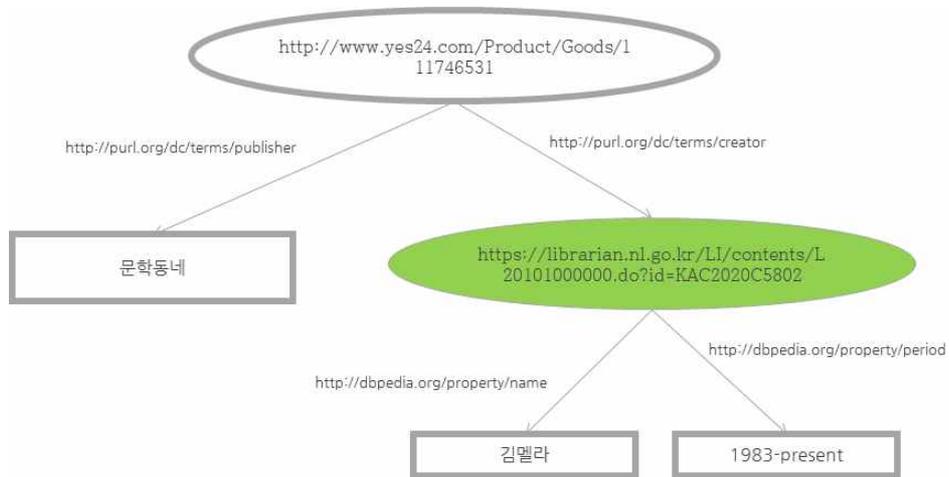
〈그림 24〉 RDF 그래프 모델

〈그림 24〉는 쉽게 이해 가능합니다. 위에서 이야기했던 두 가지 표현 형태를 그래프로 표현한 것입니다. 우리가 사용하는 예제를 이 그래프 형식에 맞추어서 변경하면 다음과 같습니다.



〈그림 25〉 RDF SPO 구조의 두 가지 유형

〈그림 25〉에서 보는 것처럼 목적어에 해당하는 작가 이름은 URI를 사용해서 표현할 수도 있고 문자열로 처리할 수도 있습니다. URI로 처리하면 보다 많은 개체와 연결이 가능합니다. RDF는 표와 다르게 이런 방식으로 “S-P-O”가 축적된다고 생각하면 됩니다. 예를 들면 다음과 같이 그래프는 확장됩니다.



〈그림 26〉 RDF SPO 구조의 확장 예시

위 확장된 형식을 SPO로 표현하면 다음과 같습니다. 보기에 편하도록 표 형식으로 정리 하였습니다.

주어(Subject)	술어(Predicate)	목적어(Object)
http://www.yes24.com/Product/Goods/111746531	http://purl.org/dc/terms/creator	https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802
https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802	http://dbpedia.org/property/name	김멜라
https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802	http://dbpedia.org/property/period	1983-present
http://www.yes24.com/Product/Goods/111746531	http://purl.org/dc/terms/publisher	문학동네

이렇게 SPO 세 개의 요소로 구성된 RDF 표현은 트리플(triple)이라고 부릅니다. 이 트리플을 XML 형식으로 표현하면 다음과 같습니다.

```
<RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:dc="http://purl.org/dc/terms/
```

```

xmlns:dbpprop="http://dbpedia.org/property/" >
  <Description about="http://www.yes24.com/Product/Goods/111746531">
    <dc:publisher>문학동네</dc:publisher>
    <dc:creator
      rdf:resource="https://librarian.nl.go.kr/LI/contents/L20101000000.do
?id=KAC2020C5802"/>
  </Description>
  <Description about="https://librarian.nl.go.kr/LI/contents/L20101000000.do
?id=KAC2020C5802">
    <dbpprop:name>김멜라</dbpprop:name>
    <dbpprop:period>1983-present</dbpprop:period >
  </Description>
</RDF>

```

자칫 난해해 보일 수 있지만 사실은 굉장히 단순한 형태입니다. 먼저 코드의 윗부분부터 차례대로 살펴보도록 하겠습니다.

```

<RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc=http://purl.org/dc/terms/
xmlns:dbpprop="http://dbpedia.org/property/" >

```

이 부분은 이 문서에서 사용하는 구문규칙, 표준 용어들의 출처를 밝히고 있습니다. w3c의 RDF 구문규칙을 따를 것이고, 더블린코어 메타데이터와 dbpedia의 속성을 사용해 서 개체를 설명하겠다는 의미입니다. 다음 구문은 실제 해당 개체의 설명을 담고 있습니다.

```

<Description about="http://www.yes24.com/Product/Goods/111746531">
  <dc:publisher>문학동네</dc:publisher>
  <dc:creator rdf:resource="https://librarian.nl.go.kr/LI/contents/L20101000000.do
?id=KAC2020C5802"/>
</Description>

```

<Description>은 RDF에서 정의하고 있는 요소입니다. 이제부터 RDF 문서로 기술할 대상이 무엇인지를 “about”을 활용해서 표현하고 있습니다. 주어부에 해당한다고 볼 수 있습

니다. 즉, 여기서 설명하는 개체는 “http://www.yes24.com/Product/Goods/111746531” 라는 URI로 식별되는 “제 꿈 꾸세요”를 의미합니다. 다음은 이 개체의 출판사 “문학동네”라는 것을 더블링크어의 publisher 요소를 활용해서 표현하고 있는 것입니다. 즉 술어와 목적어를 표현하는 부분입니다. 이렇게 하나의 트리플이 표현된 것입니다.

다음 부분도 해석하기 어렵지 않습니다. 책의 저자가 “https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802”라는 식별자로 구분되는 또 다른 개체임을 더블링크어 creator 요소를 활용해서 표현한 것입니다. 출판사인 “문학동네”처럼 문자열로 끝나지 않고, URI를 갖는 하나의 독립된 개체로 표현한 것입니다. 이 때문에 다음 구문에서는 이 저자에 대한 설명이 다시 이어집니다.

```
-----  
<Description about="https://librarian.nl.go.kr/LI/contents/L20101000000.do  
?id=KAC2020C5802">  
  <dbpprop:name>김멜라</dbpprop:name>  
  <dbpprop:period>1983-present</dbpprop:period >  
</Description>  
-----
```

이제 위 구문은 해석하기는 쉽습니다.

“https://librarian.nl.go.kr/LI/contents/L20101000000.do?id=KAC2020C5802”라는 개체가 주어부에 기술되고, dbpedia 속성의 name, period를 활용해서 해당 개체의 이름과 생몰년을 기술하고 있는 것입니다.

지금까지 검토내용을 이해했다면 RDF/XML 형식으로 구성된 구문을 해석하는 것이 어렵지는 않을 겁니다. 구문을 보기에 어렵다면 이를 그래프 형식으로 생각해 보면 훨씬 쉽게 해석할 수 있을 것입니다. 그런데 여기서 한 가지 더 생각해 볼 문제가 있습니다. 위의 문장을 다시 살펴보겠습니다.

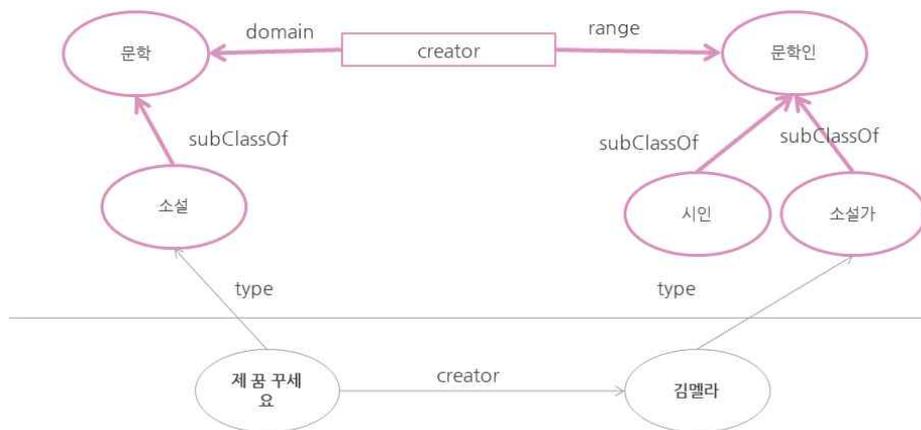
“제 꿈 꾸세요”의 “작가”는 “김멜라”이다.

사람은 이 문장을 읽으면 의미하는 바를 이해할 수 있습니다. 사람은 이 문장이 어떤 책의 작가를 설명하고 있다고 이해할 수 있습니다. 문장 전체를 보고 “제 꿈 꾸세요”는 책의 제목이고, “김멜라”는 사람 즉, 작가 이름이라고 판단합니다. 이 문장을 보고 “제 꿈 꾸세요”는 식당이름이고 “김멜라”는 식당 주인 이름이라고 판단하거나, 영화제목과 영화배우, 꽃집 이름과 꽃 이름이라고 생각하지 않습니다. 여기서 “작가”라는 술어가 그런 오해가 생

기지 않도록 중요한 역할을 담당합니다. 술어가 아니더라도 우리는 직관적으로 주어 부분은 책이라는 개체이고 목적어 부분은 사람 이름 개체라는 것을 인지합니다. 기계는 어떨까요? RDF는 기계 처리가 가능한 형식이라고 했는데, 위의 트리플 구조 하나만을 가지고 기계가 사람처럼 주어부는 책, 목적어부는 사람 이름이라고 판단하고 처리할 수 있을까요? 불가능합니다. 기계처리를 위해서는 좀 더 상세하고 구체적인 설계가 필요합니다. 즉, 주어부에 설명하고 있는 개체의 유형(type)이 무엇인지, 예를 들면 사람인지, 건물인지, 음식인지, 자동차인지 등을 설정해주어야 한다는 의미입니다. 마찬가지로 목적어부의 경우도 유형(type)이 무엇인지 설정해주는 것이 필요합니다. 좀 더 세밀하게 술어부에 해당하는 요소가 어떤 주어부, 목적어부에서 사용할 수 있는 것인지 설정해주어야 합니다. 사람이 직관적으로 판단할 수 있는 이런 요소들이 기계처리를 위해서는 별도의 세밀한 설계가 필요합니다.

일반적으로 링크드 데이터, RDF 형식의 표현을 위한 설계나 결과물을 온톨로지(ontology)라고 표현합니다. 다양한 정의가 가능하지만 우리 학문분야에서는 “개념화된 것을 형식적으로 명백하게 기술하는 명세(an explicit formal specification of a shared conceptualization)(Gruber, 1993)”라는 정의 사용합니다. 여기서 “shared conceptualization”은 인간과 기계(컴퓨터)가 공유하는 개념화를 의미하며, “Explicit Formal”이란 기계가 읽을 수 있는 형식으로 상세하게 개념, 개념 간의 관계 등을 표현한다는 의미입니다.

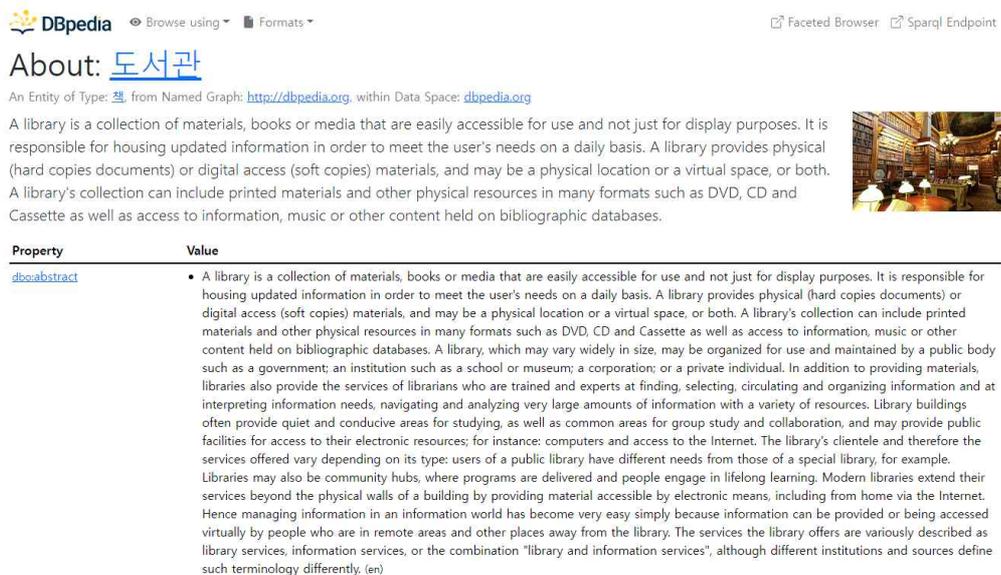
이제 우리의 예제를 기계가 처리할 수 있게 보다 명확하게 그래프로 표현해보겠습니다. 다음 <그림 27>과 같습니다.



<그림 27> RDFS 표현 예제

우리가 위의 예제에서 보았던 트리플 형태는 사실 <그림 27>과 같은 설계에 기반하여 최종적으로 만들어진 형태라고 볼 수 있습니다. 온톨로지에는 위 그림처럼 “제 꿈 꾸세요”

는 “소설”의 하나의 유형이고, “소설”은 “문학”이라는 클래스의 하위 클래스라고 정의된 부분이 존재할 수 있습니다. 마찬가지로 “김멜라”는 “소설가”의 하나의 유형이고 “소설가”는 “문학인”의 하위 클래스라는 정의가 존재할 수 있습니다. 사실 온톨로지에 정답은 없습니다. 특정 정보자원을 관리하고 서비스하는 기관의 목적이나 규모, 대상 개체에 따라 똑같은 문학작품을 온톨로지화 즉, 설계한다고 하더라도 서로 다르게 구성할 수 있습니다. 위 그림은 기계가 자동으로 처리한다는 의미가 단순히 RDF로 데이터를 표현하고 발행하기만 하면 가능한 것이 아니라는 점을 설명하기 위해 작성한 것입니다. 실제 RDF 온톨로지는 더 복잡하고 엄격한 규칙이 필요합니다. RDF Schema나 OWL(Web ontology language) 같은 용어집이 이런 규칙과 관련된 것들입니다. 흔히 RDF 보다 RDF Schema, OWL이 시맨틱스가 높다고 말하는데 이는 훨씬 정교하고 복잡한 규칙을 적용할 수 있기 때문입니다. 그러면 실제 DBpedia, 국립중앙도서관의 링크드 데이터 서비스에서 각 개체, 속성에 대한 의미를 어떻게 연결하여 보여주고 있는지 확인해보겠습니다. <그림 28>은 DBpedia에서 도서관(library)이라는 개체를 설명하고 있는 페이지입니다.



<그림 28> DBpedia에서 도서관 개체 페이지 중 일부

<그림 28>은 HTML 형식으로 변환하여 트리플 구조를 보여주고 있습니다. ‘About: 도서관’은 해당 개체의 주어부 Property는 술어부, Value는 목적어에 해당합니다. Property의 ‘dbo:abstract’를 눌러서 링크 정보를 확인하면 다음 <그림 29>와 같은 내용이 보여집니다.

About: [has abstract](#)

An Entity of Type: [Property](#), from Named Graph: <http://dbpedia.org/resource/classes#> within Data Space: [dbpedia.org](#)

Reserved for DBpedia.

Property	Value
rdf:type	<ul style="list-style-type: none"> rdf:Property owl:DatatypeProperty
rdfs:comment	<ul style="list-style-type: none"> Reserved for DBpedia. (en)
rdfs:isDefinedBy	<ul style="list-style-type: none"> http://dbpedia.org/ontology/
rdfs:label	<ul style="list-style-type: none"> has abstract (en)
rdfs:range	<ul style="list-style-type: none"> rdf:langString
wdrs:describedby	<ul style="list-style-type: none"> dbo:data/definitions.ttl
prev:wasDerivedFrom	<ul style="list-style-type: none"> http://mappings.dbpedia.org/index.php/OntologyProperty:abstract
is owl:defines of	<ul style="list-style-type: none"> http://dbpedia.org/ontology/
is owl:describes of	<ul style="list-style-type: none"> dbo:data/definitions.ttl

〈그림 29〉 abstract라는 속성의 정의와 기술규칙

〈그림 29〉는 ‘dbo:abstract’라는 속성의 정의와 입력규칙을 보여줍니다. dbo라는 네임 스페이스는 ‘http://dbpedia.org/ontology/’에서 정의하고 있는 요소입니다.

다음은 국립중앙도서관의 링크드 데이터 서비스를 살펴보겠습니다.

자료정보
목록보기

전자자료

채식주의자

<http://lod.nl.go.kr/resource/CNTS-00105771681>

국립중앙도서관 자료검색

f
t

자료정보	시각화	저자의 다른 저작물 (0)	외부정보 (0)
표제정보	제목	채식주의자	
	자료형식	전자자료 (Application)	
	자료유형	온라인자료 > 전자자료	
	장르	학술논문	
콘텐츠유형	온라인자료 유형	nk:text	
분류기호	한국십진분류기호 (KDC)	813	
출처정보	발행년	2004	
	발행처	장비	
형태기술정보	크기(수량) 및 위치	38 p	
	원시자료의 정보원	born digital	
주기사항	접근제한	nk:license2 [국립중앙도서관 공개] 국립중앙도서관내 (디지털열람실 예약 후 이용)에서 이용이 가능합니다.	
	이탈리아어	이탈리아어	

〈그림 30〉 국립중앙도서관 LOD 서비스의 채식주의자 검색 결과 중 일부

〈그림 30〉은 국립중앙도서관 링크드 데이터 서비스에서 검색한 서지정보입니다. 표제정보의 제목이라는 속성을 선택하면 속성으로 사용하고 있는 ‘더블린코어 title’ 정보로 이동합니다.

Term Name: title More details	
URI	http://purl.org/dc/terms/title
Label	Title
Definition	A name given to the resource.
Type of Term	Property
Has Range	http://www.w3.org/2000/01/rdf-schema#Literal
Subproperty of	· Title (http://purl.org/dc/elements/1.1/title)

Term Name: type More details	
URI	http://purl.org/dc/terms/type
Label	Type
Definition	The nature or genre of the resource.
Comment	Recommended practice is to use a controlled vocabulary such as the DCMI Type Vocabulary (DCMI-TYPE). To describe the file format, physical medium, or dimensions of the resource, use the property Format.
Type of Term	Property
Subproperty of	· Type (http://purl.org/dc/elements/1.1/type)

〈그림 31〉 더블린코어의 title 요소에 대한 상세 기술정보

또 온라인 자료 유형이라는 속성은 국립중앙도서관에서 정의한 속성으로 이를 선택하면 다음과 같은 속성 정보 확인이 가능합니다.

Type of Resource at lod.nl.go.kr http://lod.nl.go.kr/ontology/typeOfResource	
rdfs:domain	▪ nlon:OnlineMaterial
rdfs:isDefinedBy	▪ nlon:
rdfs:label	▪ type of resource (rdflangString) (en) ▪ 콘텐츠유형 (rdflangString) (ko)
rdfs:range	▪ nlon:OnlineMaterialType
vs:term_status	▪ stable (xsd:string)
rdf:type	▪ owl:ObjectProperty

This page shows information obtained from the SPARQL endpoint at <http://lod.nl.go.kr/sparql>.
[As Turtle](#) | [As RDF/XML](#) | [As JSON](#) | [As N3](#) | [As nTriple](#)

- Powered by National Library of Korea Linked Open Data using [QontoBase](#) (Triple Store)

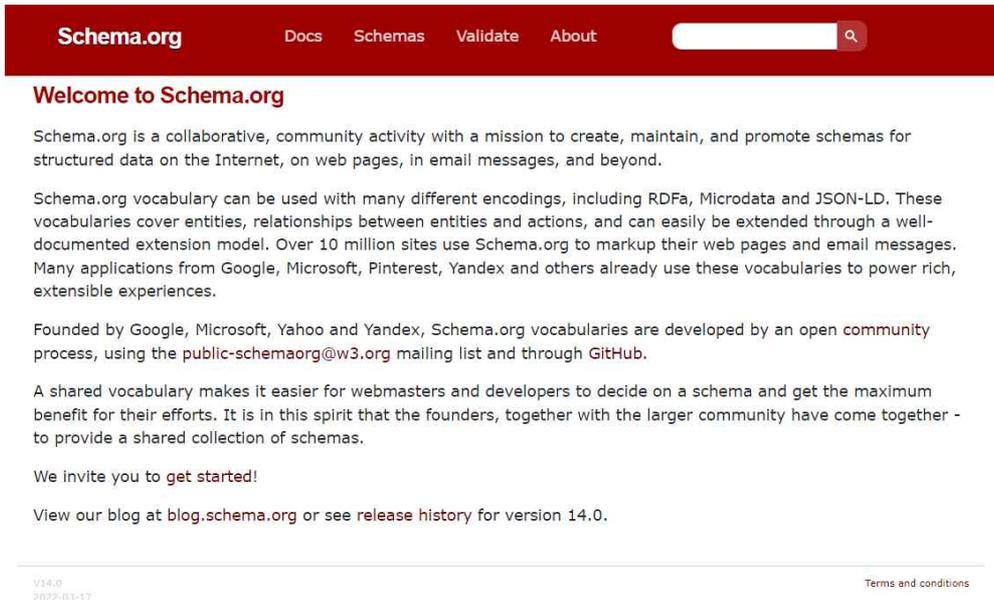
〈그림 32〉 국립중앙도서관에서 정의한 type의 정의와 기술규칙

자료유형은 더블린코어와 같은 표준용어집을 사용하지 않고 국립중앙도서관 자체의 목록규칙을 정의하여 사용하고 있습니다. 〈그림 32〉는 DBpedia처럼 해당 속성에 대한 정의와 입력규칙 등을 보여주고 있습니다.

이렇듯 RDF와 같이 웹을 플랫폼으로 하는 데이터 발행 환경에서는 소위 목록규칙, 즉

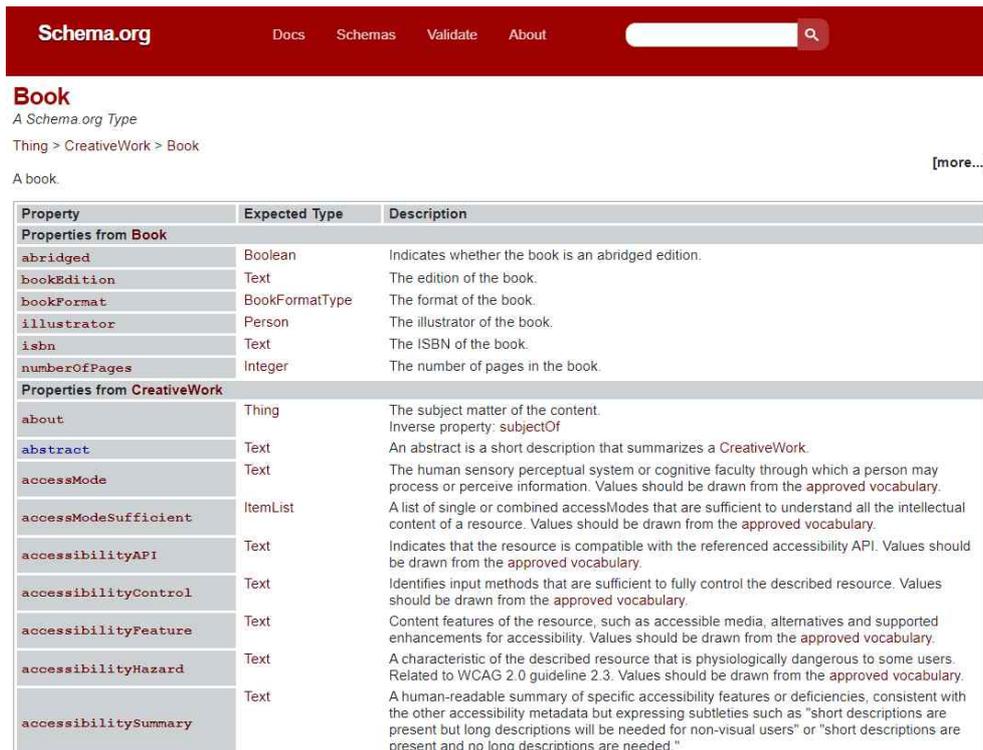
설계에 해당하는 부분 정보도 함께 조회하고 노출되어 있는 것을 알 수 있습니다. 우리에게 익숙한 목록작업 환경은 소위 목록규칙이라고 말할 수 있는 부분이 실제 데이터와 분리되어 있었습니다. 아주 예전의 카드형 목록 작업이나 MARC의 경우도 목록규칙은 문서 형태로 별도로 존재하고 사서는 이 목록 규칙을 숙지하여 목록작업을 수행했습니다. 물론 MARC의 경우 입력 규칙을 반영한 데이터베이스가 존재하고, 데이터베이스의 제약사항 혹은 프로그래밍 언어를 통한 목록 규칙 점검으로 오류를 제거하여 규칙에 맞는 목록작업 환경을 제공합니다. 하지만 이 경우에도 목록규칙은 일반이용자에게 즉시 노출되거나 공유되는 정보는 아닙니다. 반면 RDF의 경우는 내가 찾는 서지 정보의 값뿐만 아니라 이 값을 정의하는 메타데이터 즉, 속성 정보도 함께 제공합니다. 이는 데이터 정보를 필요로 하는 새로운 요구사항의 반영이라는 관점으로도 검토할 필요가 있습니다.

지금까지 RDF에 대해서 간략히 살펴보았습니다. 그런데 오직 RDF만이 웹에 이런 구조화된 데이터 정보를 공유할 수 있는 방식은 아닙니다. RDF가 지금과 같이 확산되기 전 혹은 비슷한 시기에 웹을 구조화시키고 데이터를 공유하기 위한 다양한 시도들이 있었습니다. 대표적인 것이 HTML5이 마이크로 데이터(MicroData)입니다. 마이크로 데이터는 RDF처럼 별도의 문법을 정의하지 않고 HTML 태그에 기계가 읽고 처리할 수 있는 정보를 포함시킵니다. 궁극적으로는 검색엔진이나 애플리케이션들이 보다 쉽게 웹페이지의 내용을 처리할 수 있도록 돕습니다. 그 형식이나 표현방법은 지금까지 우리가 검토했던 방식, 우리가 일해왔던 방식과 다르지 않습니다. 구체적인 형태를 살펴보기 전에 우리는 표준용어집이 존재할 것이고 이 용어집에서 정의한 속성(메타데이터)에 따라 해당 정보를 구조화시켜 표현한다고 추측할 수 있습니다. HTML5의 경우 다양한 표준용어집을 사용할 수 있지만 대표적으로 Schema.org를 표준으로 사용합니다. Schema.org는 Microsoft, Google, Yahoo! 등 검색엔진 기업들이 참여한 프로젝트입니다. 이는 궁극적으로 시맨틱 웹을 지향하는 프로젝트로 웹을 구조화된 형태로 표현하고 공유하기 위한 것입니다.



〈그림 33〉 Schema.org 홈페이지 중 일부

Schema.org에는 책을 포함한 다양한 개체와 메타데이터가 존재합니다. 책(Book)도 포함되어 있으며, 여기서 정의하고 있는 속성을 중심으로 도서관의 책 정보를 표현해보도록 하겠습니다. Schema.org의 Book에 대한 표준용어는 〈그림 34〉와 같습니다.



〈그림 34〉 Schema.org의 Book 개체에 대한 속성 중 일부

Book이라는 용어집 추상적인 개체는 URI ‘https://schema.org/Book’으로 식별됩니다. 이 용어집을 사용하겠다고 선언하면 기계는 이 URI 정보를 참조하여 해당 속성들을 구분하고 처리합니다.

[NL1]도서[도서]

채식주의자: 한강 연작소설



표제/저자사항	채식주의자: 한강 연작소설 / 한강 [1970-]
발행사항	파주 : 창비, 2007
형태사항	247 p.; 21 cm
주기사항	별표제: Vegetariana (La) vegetariana ISBN 9789871803316 이용가능한 다른 형태자료: 이용가능한 오프라인 자료
표준번호/부호	ISBN 9788936433598 UCI G701:B-00069251832
분류기호	한국십진분류법> 813.6 듀이십진분류법> 895.735
주제명	한국 현대 소설[韓國現代小說] 연작 소설[聯作小說]

〈그림 35〉 국립중앙도서관에서 검색한 채식주의자 결과 화면

국립중앙도서관에서 검색한 채식주의자 서지정보를 바탕으로 이를 표현하도록 하겠습니다(〈그림 35〉 참조). 이해하기 쉽도록 먼저 위와 같은 서지 정보 중 일부를 간단히 HTML로 구성하면 다음과 같습니다.

```

<div>
  <h1>제목: 채식주의자</h1>
  <span>저자: 한강</span>
  <span>출판사: 창비</span>
  <span>출판년: 2007</span>
  <span>ISBN: 9789871803316</span>
</div>

```

HTML로 구성된 정보 페이지의 가장 큰 특징은 문자로 표현된 서지정보의 의미를 알 수 없다는 것입니다. 위의 예제 역시 단순히 사람이 보기에 좋은 형태로 정보를 표현하기 위한 수단으로 문자의 크기, 레이아웃 구분, 영역 구분 정도만으로 구성되어 있습니다. 사람은 문자를 보고 책의 정보를 표현한 것이라고 알 수 있고 의미를 해석할 수 있지만 기계

처리는 어렵습니다. 이제 이 정보들을 HTML 태그는 유지하면서 책정보를 표현하고 있다는 것을 schema.org를 활용하여 표현해보도록 하겠습니다.

```
<div itemscope itemtype="http://schema.org/Book">
  <h1>제목: 채식주의자</h1>
  <span>저자: 한강</span>
  <span>출판사: 창비</span>
  <span>출판년: 2007</span>
  <span>ISBN: 9789871803316</span>
</div>
```

위 코드는 기존 코드에서 지금부터 설명할 개체가 책이고, 'schema.org/Book'이라는 표준용어집을 참조하여 기술하겠다고 선언한 것입니다. 지금부터 기술할 데이터의 표준용어집 위치를 기술한 것입니다. 이제 기계도 이 위치의 용어집을 참조할 것이고, 우리도 여기에 기술된 속성을 기준으로 정보를 표현할 것입니다.

```
<div itemscope itemtype="http://schema.org/Book">
  <h1 itemprop="name">제목: 채식주의자</h1>
  <span>저자: 한강</span>
  <span>출판사: 창비</span>
  <span>출판년: 2007</span>
  <span>ISBN: 9789871803316</span>
</div>
```

위 코드는 "제목: 채식주의자"라는 문자열이 해당 저작물의 이름이라는 것을 schema.org의 Book에서 정의한 속성 name을 활용하여 표현한 것입니다. 나머지 속성들도 포함하여 이를 다시 표기하면 다음과 같습니다.

```
<div itemscope itemtype="http://schema.org/Book">
  <h1 itemprop="name">제목: 채식주의자</h1>
  <span itemprop="author">저자: 한강</span>
  <span itemprop="publisher" itemtype="https://schema.org/Organization">출판사: 창비</span>
  <span itemprop="datePublished">출판년: 2007</span>
  <span itemprop="isbn">ISBN: 9789871803316</span>
</div>
```

위 코드를 보면 출판사의 경우 Book 용어집이 아니라 Organization 용어집을 따르도록 구성된 것을 알 수 있습니다. schema.org의 Book에는 publisher 속성이 존재하지 않습니다. publisher는 조직 개체를 표현하는 별도 용어집에 포함되어 있어 이를 표기한 것입니다. RDF와 완전히 같지는 않지만 표준용어집이 존재하고 이에 근거하여 기계가 자동으로 처리가능한 형태의 구조적 정보를 제공한다는 점에서는 다르지 않습니다.

이제 기계가 무엇인가를 자동으로 처리한다는 것의 의미와 함께 RDF에 대한 기본적인 이해가 가능할 것으로 봅니다. 서두에 언급했던 내용을 다시 한 번 언급하겠습니다. 우리가 익숙한 작업방식인 표준기반의 목록화, 식별자 부여와 통제어휘집 운영은 링크드 데이터의 세계에서든 똑같이 필요합니다. 그렇기 때문에 도서관 데이터의 품질은 높다고 할 수 있고 웹이라는 새로운 환경에서도 마찬가지로 중요하게 활용될 수 있습니다. 지금은 도서관뿐만 아니라 지식정보를 관리하는 기관, 개인연구자, 기업까지 개방 데이터 생태계에 참여하고, 해당 데이터를 활용하는 데 많은 노력을 기울이고 있습니다. BIBFRAME도 이런 노력의 일환으로 인지해야 합니다. 또 일부 기술이 낯설고 어려울 수 있지만 큰 관점에서 이런 기술을 활용하려는 노력은 우리가 그동안 일했던 방식과 이유 측면에서 동일하다는 점을 명심하면 BIBFRAME에 대한 접근이 훨씬 편안할 것입니다. 이제 본격적으로 BIBFRAME에 대한 내용을 다루어 보겠습니다. 지금까지 설명했던 내용들이 다시 한 번 반복적으로 등장할 것입니다. BIBFRAME에 대한 내용을 다루기 전에 지금까지 웹, 링크드 데이터, RDF에 대해 다룬 이유는 처음 등장하는 단어들에 익숙해지기 위함이며 배경과 맥락을 이해함으로써 BIBFRAME에 등장하는 동일 내용을 쉽게 이해할 수 있게 하기 위함입니다.

04 BIBFRAME

4.1 등장배경

미국의회도서관(Library of Congress)은 2011년 5월에 공식적으로 서지 프레임워크 이니셔티브(Bibliographic Framework Initiative)를 발족했습니다. 한마디로 BIBFRAME은 서지 기술을 위한 데이터 모델입니다. BIBFRAME은 전통적인 도서관의 데이터 표준인 MARC를 링크드 데이터 원칙을 활용한 새로운 형식으로 대체하기 위한 것입니다. MARC는 1960년대 미국의회도서관의 Henriette Avram이 개발한 것입니다. 1971년에 이르러 MARC 형식은 미국 서지 데이터 보급을 위한 국가표준이 되었고, 1973년에는 국제표준이 되었습니다. MARC라는 단일 형식의 표준은 도서관 간의 데이터 교환에 있어서 상호운용성을 보장해주는 역할을 담당해왔습니다. 사실 도서관 간에만 데이터 교환이 필요한 상황이라면 MARC가 문제 될 것은 없을지도 모릅니다. 하지만 웹이라는 새로운 정보생태계가 급속히 발전하고, 데이터가 새로운 정보자원으로 등장하면서 MARC의 한계가 드러나기 시작했습니다. Casey et al.(2008)은 MARC는 40년 된 데이터 관리 기술을 기반으로 하며 오늘날의 프로그래밍 스타일과 맞지 않는다고 지적하기도 했습니다.

링크드 데이터를 사용하는 이유는 도서관 간의 데이터 교환을 넘어 다양한 데이터, 도메인과의 연결이 가능하도록 하여 서지 데이터의 유용성을 높이기 위함입니다. 구체적으로 보면 BIBFRAME은 웹(web)을 플랫폼으로, 링크드 데이터(linked data) 기술을 사용하는 데이터 생태계에 서지 데이터가 함께할 수 있도록 하기 위한 것으로 볼 수 있습니다. 그렇다고 BIBFRAME이 완전히 새로운 업무 처리 절차를 필요로 하거나 완전히 새로운 기술로의 서지 기술환경 변화를 꾀한다고 보기는 어렵습니다. BIBFRAME의 핵심은 도서관이 그동안 노력해왔고 성과를 보았던 자원공유, 목록비용 절감이 가능한 데이터 교환 효과는 유지하면서 MARC 21 형식의 데이터를 새로운 형식으로 전환 가능하도록 하기 위한 것입니다. 여기서 새로운 형식으로서의 전환은 링크드 데이터 기술에 기반한 데이터 발행을 의미합

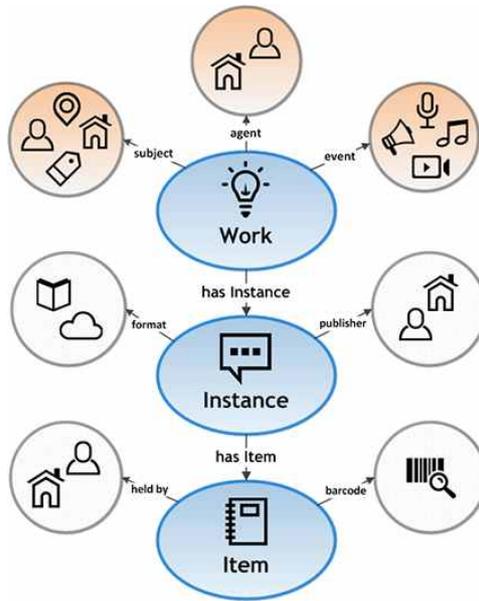
니다. 한마디로 BIBFRAME은 링크드 데이터 기술을 기반으로 하는 새로운 미래 서지기술 모형입니다. 이는 도서관 데이터의 품질을 높인다기 보다는 도서관 데이터가 다른 도메인의 데이터들과 공존하고 활용성을 높이기 위한 목적이라고 보면 좋습니다. 웹이라는 플랫폼이 점점 데이터화되고 있는 상황에서 웹에 도서관 데이터가 존재하지 않는다는 것은 도서관의 위기로 이어질 수도 있습니다.

4.2 BIBFRAME 2.0 모델

BIBFRAME의 핵심은 도서관은 물론 외부의 다양한 기관들이 도서관 데이터를 잘 사용할 수 있도록 하는 것입니다. 이를 위해 BIBFRAME은 서지 기술 표준을 링크드 데이터 모델로 표현했다고 볼 수 있습니다. 목록화된 자원(예, 책)은 저자, 내용 정보, 발행형식, 사본 정보 등 다양한 정보요소를 포함하고 있습니다. BIBFRAME 2.0은 이 정보를 저작(work), 인스턴스(instance), 개별자료(item) 세 가지의 핵심 추상 계층으로 구성되어 있습니다. 3가지 추상계층이 의미하는 바는 다음과 같이 정리할 수 있습니다.

- 저작(work): 저작은 BIBFRAME에서 가장 상위의 추상 계층으로 목록화된 자원의 개념적 본질(예, 저자, 언어, 주제)을 반영합니다.
- 인스턴스(instance): 특정 저작은 하나 이상의 개별적인 물리적 구현형(특정 출판 형식)을 가질 수 있습니다. 이것이 저작의 인스턴스입니다. 인스턴스는 발행자, 장소, 발행일, 형식(format)과 같은 정보를 반영합니다.
- 개별자료(item): 개별자료는 인스턴스의 실제 복본을 의미합니다. 여기서 사본은 물리적 형태 혹은 전자적 형태 모두를 의미합니다. 개별자료는 위치(물리적 혹은 가상), 서가기호, 바코드와 같은 정보를 반영합니다.

〈그림 36〉은 BIBFRAME의 핵심 추상 계층이 갖는 정보와 계층 간의 관계를 보여줍니다.



〈그림 36〉 BIBFRAME Model

〈그림 36〉을 보면 최상위 계층인 저작에 주제(subject), 행위자(agent), 사건(event) 세 개의 새로운 정보표현요소를 확인할 수 있습니다. BIBFRAME 2.0은 이 세 개 요소를 활용해서 저작과 관련된 핵심 개념을 정의합니다. 이 세 개 요소의 정의는 다음과 같습니다.

- 행위자(agent): 행위자는 저자, 편집자, 예술가, 사진작가, 작곡가, 삽화가 등과 같은 역할을 통해서 저작 혹은 인스턴스와 관련된 사람, 조직, 관할구역(jurisdictions) 등입니다.
- 주제(subject): 저작은 하나 이상의 개념(concept)에 대한 설명(about)입니다. 이런 개념을 저작의 주제(subject)라고 합니다. 주제가 될 수 있는 개념에는 주제(topic)¹⁾, 장소, 시간적 표현, 사건, 저작(work), 인스턴스(instance), 개별자료(item), 행위자(agent)가 모두 포함될 수 있습니다.
- 사건(event): 사건은 저작(work)의 내용이 될 수 있는 특정 사건들의 기록입니다.

사실 도서관에서 새로운 서지모델을 만드는 작업이 BIBFRAME이 처음은 아닙니다. 이미 사서들에게는 친숙한 개념인 FRBR가 존재합니다. BIBFRAME 모델을 보면 어느 정도 예상할 수 있듯이 BIBFRAME의 저작(work)은 FRBR에서 저작(work)과 표현형(expression)과 유사합니다. 또 BIBFRAME의 인스턴스(instance)는 FRBR의 구현형(manifestation)과

1) 'topic'과 'subject'는 모두 주제로 번역됩니다. 'topic'의 경우는 대화나 서면형식에서 다루어지는 특정 측면을 의미하는 반면 'subject'는 지식관점에서 보다 포괄적인 측면에서 사용합니다. 때문에 'subject'는 'topic'의 일부가 될 수 없고, 'subject'는 많은 'topic'을 포함할 수 있습니다.

유사합니다. 유사하다는 표현에 주목할 필요가 있습니다. 모델 관점, 즉 이해하기 쉬운 상위 계층에서 볼 때 유사한 관점이 있다는 것이지 BIBFRAME이 FRBR이나 RDA(Resource Description and Access)의 연장선에 있다고 볼 수는 없습니다. 그럼에도 BIBFRAME이나 FRBR 모두 RDF로 표현되기 때문에 두 모델 간의 상호운용성은 보장된다고 볼 수 있습니다(Schreur, 2017).

BIBFRAME은 전통적인 도서관의 표준인 MARC를 대체하기 위한 것입니다. 2012년 미국의회도서관은 데이터 관리 회사인 Zepheira와 MARC를 대체하기 위해 링크드 데이터를 개발하고 있음을 알렸습니다. 같은 해 말에 미국의회도서관은 MARC Resources(MARCR)라는 새로운 모델을 발표했습니다. 11월에 와서 미국의회도서관은 BIBFRAME이라는 이름으로 모델 초안을 발표하기에 이릅니다. 그리고 2016년 BIBFRAME 2.0 버전이 발행되었습니다(Schreur, 2017).

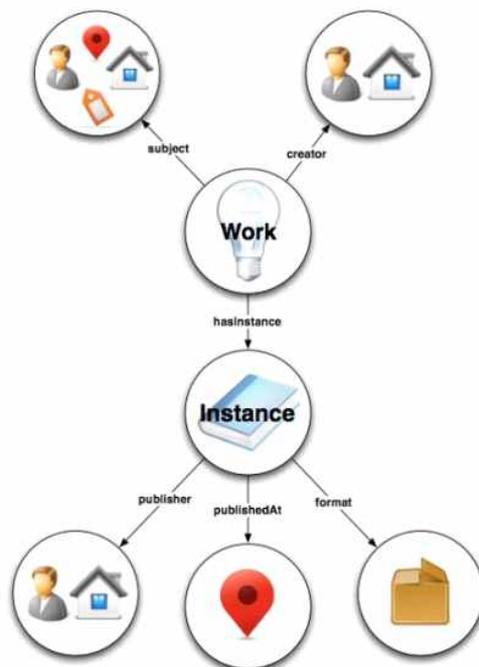
4.3 BIBFRAME 1.0과 2.0

BIBFRAME은 미국의회도서관과 Zepheira의 Eric Miller 간의 파트너십으로 개발되었으며 MARC 형식을 대체하도록 설계되었습니다. BIBFRAME의 목표는 도서관의 새로운 서지환경을 구성하는 것으로 상호연결성을 증진시키기 위한 것입니다. 어떤 면에서 보면 완전히 새로운 접근일 수 있는 이 시도는 전통적인 도서관의 MARC를 시맨틱 웹 기반의 언어로 변환하기 위한 것으로 새로운 메타데이터 표준을 준수함으로써 구현 가능합니다.

1.0 모델은 저작(Works), 인스턴스(Instances), 전거(Authorities), 주석(Annotations)의 네 가지 클래스로 구분되는 구조를 갖고 있습니다. 1.0의 핵심 클래스는 저작(Works)과 인스턴스(Instances)입니다(〈그림 37〉 참조). 저작은 목록화하기 위한 자원의 개념입니다. 저작(work)은 FRBR의 저작(work)과 표현형(expression)개념으로 구성된 추상개체입니다. 인스턴스(instance)는 BIBFRAME 저작물의 물리적 구현형입니다. 각각의 BIBFRAME 인스턴스는 하나의 BIBFRAME 저작(work)을 나타내며, FRBR의 구현형(manifestion)과 유사합니다. 전거(authority)는 사람, 장소, 주제(topic), 조직과 같은 개념으로 저작(work) 혹은 인스턴스(instance)와 연결됩니다.

주석(annotation)은 저작 혹은 인스턴스와 관련될 수 있는 전거 이외의 개념들을 나타냅니다. 예를 들면 저자의 리뷰, 특정 기관의 특정 자원에 대한 소장정보는 주석에 해당합니다. BIBFRAME 1.0 모델의 가장 큰 특징 중 하나는 음악, 시청각 자료에도 적용할 수 있다는 점입니다. 사실 모델만을 본다면 이미 FRBR라는 서지모형에 대해 알고 있는 사서

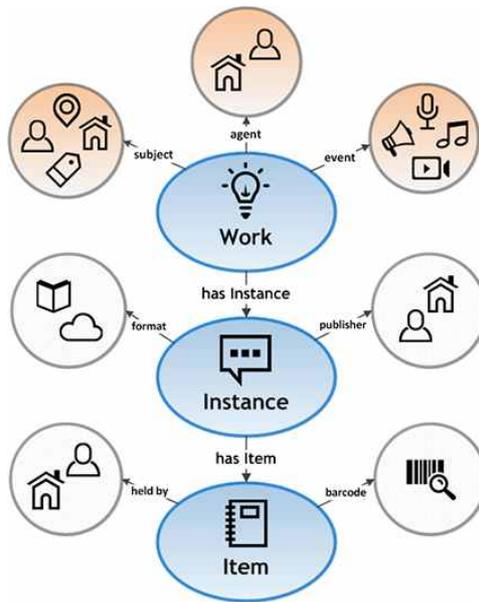
입장에서 이해하기 쉽습니다.



〈그림 37〉 BIBFRAME 1.0 모델

Schreur(2017)은 초기 BIBFRAME 모델의 문제점을 두 가지로 제시하였습니다. 첫 번째는 모델 자체의 구조와 관련된 것으로 BIBFRAME이 도서관뿐만 아니라 다양한 정보 전송을 위한 것이지만 여전히 MARC 형식은 도서관에서만 이해할 수 있고, 기존의 RDA가 BIBFRAME에서 잘 표현될 수 있는 방법이 불명확한 하다는 점입니다. 두 번째는 BIBFRAME이 MARC 형식의 방대한 자료를 활용함으로써 단순했던 모델이 혼란스러워지고 데이터 모델링 어휘를 제공하는 W3C의 RDF 모범사례를 벗어나기 시작했다는 점입니다.

미국의회도서관의 네트워크 표준 사무국(Office of Network Standards)은 대규모 MARC 데이터를 링크드 데이터로 발행하면서 BIBFRAME 1.0의 문제점 파악하기 위해 Robert Sanderson에게 연구를 의뢰하였습니다. Sanderson 연구팀은 기존 MARC로 표현된 도서관 데이터의 가치와 의미를 유지하면서 링크드 데이터로 변환 시 완전성을 유지하기 위한 방안을 제시하였습니다. 간략하게 Sanderson(2015)은 술어(predicate) 병합, 전거의 사용, 주석(annotation) 구조, 소장여부 표현, 사건(event)의 역할을 포함한 다양한 문제를 제기하였으며, 이를 참조하여 BIBFRAME2.0을 발행하였습니다. 새로운 BIBFRAME 모형은 다음 그림과 같습니다.



〈그림 38〉 BIBFRAME Model

BIBFRAME 1.0모델과 2.0 사이의 가장 큰 변화는 FRBR 모델을 더 잘 반영하기 위해 개별자료(item)을 포함하고, 전거(authority)와 주석(annotation) 클래스를 제거한 것입니다.

BIBFRAME 2.0이 이전의 문제들을 해결하고 보다 간결한 모델로 탄생한 것으로 사실입니다. 하지만 1.0에서 지니고 있던 도서관 데이터 중심적인 구조에서 완벽하게 탈피했다고 보기는 어려울 것 같습니다. 또 모든 도서관이 BIBFRAME 2.0 명세에 따라 기존 데이터를 쉽게 LOD로 발행할 수 있는 것도 아닙니다. LOD로 데이터를 발행하기 위해서는 단순한 변환 외에 준비하고 설계해야 하는 영역이 많습니다. 이런 부분들은 실제 BIBFRAME 도구 예시를 보면 자연스럽게 이해할 수 있습니다.

4.4 BIBFRAME Editor

미국의회도서관은 BIBFRAME 2.0 도구를 온라인 상에서 직접 활용해 볼 수 있는 별도의 서비스 '<https://bibframe.org>'를 구축하였습니다. 여기서 제공하는 편집도구는 총 3가지입니다. 2가지는 실제 BIBFRAME 용어집에 맞추어 서지 데이터를 입력할 수 있는 도구이고, 다른 하나는 BIBFRAME 프로파일 편집기(profile editor)로 실제 입력기에서 사용할 어휘집 등을 미리 지정해 놓는 환경설정 기능을 포함하고 있습니다.

〈표 2〉 BIBFRAME 편집기의 종류와 설명

구분	설명
Marva Editor	<ul style="list-style-type: none"> 가장 최근의 편집기 편집기의 기능 구현이 어떻게 되어야 하는지에 대한 이해와 직접 실습할 수 있도록 구성 지속적으로 업데이트 예정으로 향후에 기능 개선이 이루어질 수 있음
Bibframe Editor (BFE)	<ul style="list-style-type: none"> 미국 의회도서관에서 입력한 샘플과 설명을 포함하고 있는 데모 사이트 누구나 자유롭게 데이터 입력이 가능하지만 일정 시간 후에 삭제됨
Bibframe Profile Editor (BFPE)	<ul style="list-style-type: none"> 편집기의 프로파일 설정을 위한 일종의 관리 기능 이용자가 임의로 변경할 수 있으나 저장하지 않음

이제 실제 BIBFRAME 용어집 표준을 준수하는 편집기를 실제 서지정보를 입력해보도록 하겠습니다. 데이터 입력 전 한 가지 기억해야 하는 사항은 BIBFRAME은 링크드 데이터 원칙을 준수한다는 점입니다. 즉, URI를 기반으로 모든 개체를 구분하고 표현합니다. 그러기 위해서는 사전에 URI를 기반으로 서지 정보를 표현하기 위한 다양한 준비가 필요합니다. 물론 MARC 데이터를 링크드 데이터 형식으로 변환하기 위한 모듈의 경우 URI 정책을 수립하고 기존 데이터 중 개체 개념을 표현하기 위한 기준을 설정하면 해결할 수 있습니다. 대용량의 데이터를 한 번에 링크드 데이터 형식으로 변환하는 것과 데이터가 존재한다는 전제하에 편집기를 활용해서 결과물을 확인하는 것은 전혀 다른 문제입니다. 대용량 데이터를 한 번에 변환할 경우는 매핑 규칙에 맞추어 일괄적으로 변환작업이 이루어지게 됩니다. 즉 초기에 BIBFRAME 기반 데이터를 구축할 경우 가능한 방식으로 사서가 직접 결과물을 확인하고 입력하거나 편집할 여지가 없습니다.

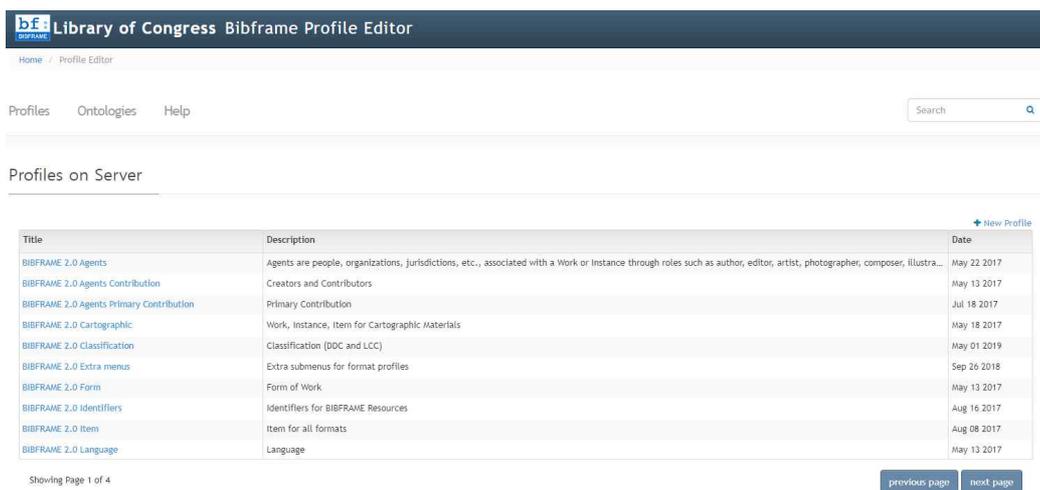
지금부터 검토할 도구는 Bibframe Profile Editor(BFPE)와 Marva Editor입니다. 전자는 일종의 환경설정 도구입니다. 그런데 환경설정 도구의 역할은 BIBFRAME 온톨로지를 시각적으로 편집가능하도록 구성한 것입니다. 즉, RDF 형식의 설계 파일을 GUI 환경의 편집환경으로 개선한 것에 불과합니다. 후자인 Marva Editor는 실제 서지데이터를 입력할 수 있는 편집도구입니다. BIBFRAME 기반이라고 해서 서지 데이터를 입력하는 환경이 현재와 크게 달라지지 않습니다. MARC 혹은 MODS 등 다양한 형식의 표준에 따라 서지 데이터를 입력한다고 하더라도 입력자인 사서가 마주하는 편집화면은 크게 다르지 않습니다. 어떤 환경에서도 우리는 표제, 책임사항, 저자명, 출판사명, 발행일 등의 정보를 입력합니다. 다만 차이점이 있다면 저장형식이 다릅니다. 이렇게 다른 저장형식을 우리가 직접 열어보거나 편집할 일은 거의 없습니다. 이 형식대로 데이터를 입력하는 일도 거의 없습니다. 작업환경은 기존과 크게 달라지지 않을 것입니다.

즉, 새로운 데이터 형식 혹은 모델이 등장한다고 하더라도 우리가 하는 일이 갑자기 변하거나 작업환경이 완전히 변화하는 일은 발생하지 않습니다. 그렇다면 새로운 것들이 등

장했을 때 우리가 정말 중요하게 알아야 하는 일들은 앞에서 설명했던 등장배경과 맥락에 있다고 볼 수 있습니다. 우리가 다루어야 할 개체(예, 책, 전자책, 지도 등)를 이제는 어떻게 바라보고 표현해야 하는 것인가가 더 중요합니다. 그리고 그 결과물인 데이터를 누가 어떻게 사용할 것인지 이해하는 것이 중요합니다.

앞으로 편집도구 화면을 중심으로 설명이 이어집니다. 늘 마주하던 화면과 형태만 조금 다를 뿐 큰 이질감을 느끼기 어려울 겁니다. 눈여겨봐야 하는 항목들이 있다면 기존 MARC 입력 시 중요하게 생각하지 않았거나 통제하지 않았던 부분들은 무엇인지와 RDF/XML 형태의 결과물에 대한 이해입니다. RDF/XML을 읽고 이해하는 방법은 앞선 설명을 참조하면 크게 어렵지 않습니다.

먼저 Bibframe Profile Editor를 살펴보도록 하겠습니다. 이 도구는 실제 서지 데이터를 입력하기 위한 기능을 포함하고 있지 않습니다. BIBFRAME 규격에 맞추어 데이터를 입력하기 위한 관리자 기능, 데이터 입력에 필요한 환경설정을 하기 위한 기능입니다. 사실 Profile Editor는 Bibframe 온톨로지 정보를 GUI 환경에서 검토하고 관리할 수 있도록 구현했다고 보면 좋습니다. BIBFRAME 온톨로지는 RDF 형식으로 구성되어 있으며 표준 용어의 정의와 관계, 입력규칙 등이 상세하게 기술되어 있습니다. Bibframe Profile Editor는 이 규칙들을 GUI 환경으로 옮겨놓은 것입니다. 웹 기반 도구이기 때문에 별도의 로그인 절차 없이 해당 사이트에 접속하여 바로 활용 가능합니다. <그림 39>는 프로파일 편집기의 초기화면입니다.



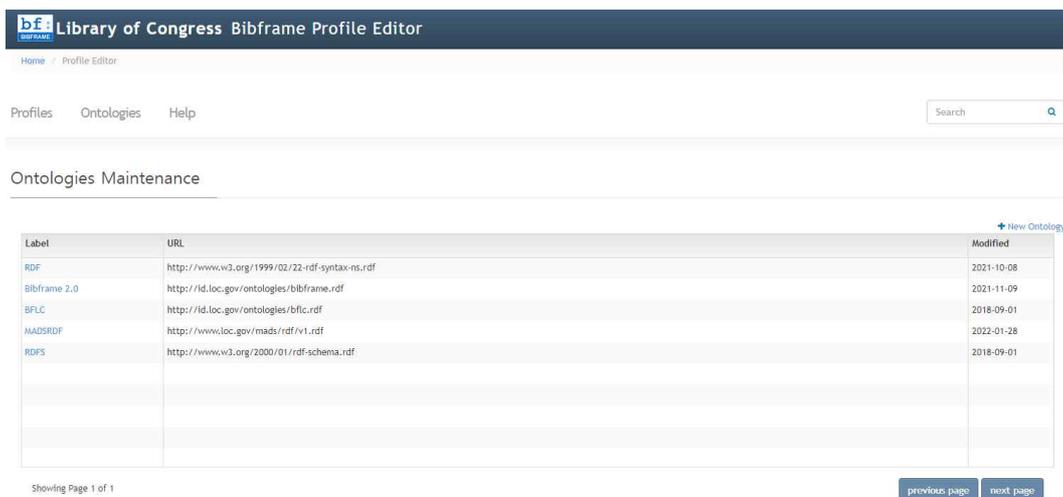
<그림 39> Bibframe Profile Editor 초기화면

현재 Profile Editor에는 39개의 입력 프로파일이 존재하며, 이와 별도로 온톨로지 정보도 함께 제공하고 있습니다. 온톨로지 정보에서는 모든 온톨로지 정보 전체를 확인할 수는

없으며, 단지 참조하고 있는 용어집의 종류와 URL 정보를 제공해줍니다. 해당 URL로 참조하여 RDF 형식의 용어집 전체의 설계 내용을 확인할 수 있습니다.

해당 용어집의 Label을 선택하면 간단한 등록화면을 확인할 수 있습니다. 등록화면에서 필요한 정보는 온톨로지에서 활용할 표준 용어집 명세를 확인할 수 있는 URL 정보입니다. ID 값은 시스템에서 자동으로 부여하도록 되어 있으며, 삭제 기능이 있는 것을 확인할 수 있습니다. 만약 여러분 도서관에서 BIBFRAME 관련 모델을 설계하는데 BIBFRAME 용어 외에 다른 용어의 사용이 필요하다면 여기에 등록해야 합니다. 실제로 이런 용어들이 있다면 미국의회도서관이 BIBFRAME 온톨로지와 용어를 개방한 것처럼, 여러분 도서관에서 정의한 용어들도 웹에 공개해야 합니다.

이는 단순한 등록 기능을 수행하는 것뿐이며, 실제 온톨로지를 구성하는 작업은 별도로 진행하고 온라인 상에 해당 용어 명세를 확인할 수 있는 문서가 활성화되어야 합니다. 과거에는 온톨로지 파일의 경우 owl 파일 형식으로 저장하는 추세가 있었으나 최근에는 설계파일이나 실제 인스턴스가 포함된 문서 모두 형식은 RDF이므로 RDF 형식으로 만들고 발행하는 것이 일반적입니다.



〈그림 40〉 Bibframe Profile Editor 초기화면



〈그림 41〉 Bibframe Profile 편집기 화면

```

1 <rdf:RDF xmlns:base="http://id.loc.gov/ontologies/bibframe/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/01-rdf-syntax-ns#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:dcterms="http://purl.org/dc/terms/">
2 <owl:Ontology rdf:about="">
3 <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#string">2.1.0</owl:versionInfo>
4 <owl:versionIRI rdf:resource="http://id.loc.gov/ontologies/bibframe-2-1-0/">
5 <owl:priorVersion rdf:resource="http://id.loc.gov/ontologies/bibframe-2-0-1/">
6 <rdfs:label>BIBFRAME vocabulary</rdfs:label>
7 <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2021-06-10T12:00:00.000-05:00</dcterms:issued>
8 <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2021-06-24T13:27:14.395-04:00</dcterms:modified>
9 <dcterms:description xml:lang="en">The Bibframe vocabulary consists of RDF classes and properties used for the description of
10 items cataloged principally by libraries, but may also be used to describe items cataloged by museums and archives.
11 Classes include the three core classes - Work, Instance, and Item - in addition to many more
12 classes to support description. Properties describe characteristics of the resource being
13 described as well as relationships among resources. For example: one Work
14 might be a "translation of" another Work; an Instance may be an
15 "instance of" a particular Bibframe Work. Other properties describe attributes of Works and Instances. For
16 example: the Bibframe property "subject" expresses an important attribute of a Work
17 (what the Work is about), and the property "extent" (e.g. number of pages) expresses an
18 attribute of an Instance.</dcterms:description>
19 <dcterms:creator rdfs:resource="http://id.loc.gov/vocabulary/organizations/dlc"/>
20 <dcterms:publisher rdfs:resource="http://id.loc.gov/vocabulary/organizations/dlcnc"/>
21 <dcterms:rights rdf:resource="https://creativecommons.org/publicdomain/zero/1.0/">
22 <cc:license rdf:resource="https://creativecommons.org/publicdomain/zero/1.0/">
23 </owl:Ontology>
24 <owl:Class rdf:about="http://id.loc.gov/ontologies/bibframe/Work">
25 <rdfs:label>Work</rdfs:label>
26 <skos:definition>Resource reflecting a conceptual essence of a cataloging resource.</skos:definition>
27 <dcterms:modified>2016-04-21 (New)</dcterms:modified>
28 </owl:Class>
29 <owl:Class rdf:about="http://id.loc.gov/ontologies/bibframe/Instance">
30 <rdfs:label>Instance</rdfs:label>
31 <skos:definition>Resource reflecting an individual, material embodiment of a Work.</skos:definition>
32 <dcterms:modified>2016-04-21 (New)</dcterms:modified>
33 </owl:Class>
34 <owl:Class rdf:about="http://id.loc.gov/ontologies/bibframe/Item">
35 <rdfs:label>Item</rdfs:label>
36 <skos:definition>Single example of an Instance.</skos:definition>
37 <dcterms:modified>2016-04-21 (New)</dcterms:modified>
38 </owl:Class>
39 <owl:Class rdf:about="http://id.loc.gov/ontologies/bibframe/Hub">
40 <rdfs:label>Hub</rdfs:label>
41 <rdfs:subClassOf rdf:resource="http://id.loc.gov/ontologies/bibframe/Work"/>
42 <skos:definition>An abstract resource that functions as a bridge between two Works.</skos:definition>
43 <dcterms:modified>2021-06-09 (New [GH75])</dcterms:modified>
44 </owl:Class>
45 <owl:Class rdf:about="http://id.loc.gov/ontologies/bibframe/Text">
46 <rdfs:label>Text</rdfs:label>
47 <rdfs:subClassOf rdf:resource="http://id.loc.gov/ontologies/bibframe/Work"/>

```

〈그림 42〉 bibframe.rdf 파일의 내용 중 일부

〈그림 42〉는 bibframe.rdf 파일을 열어본 결과 화면 중 일부입니다. rdf 파일 형식은 웹 브라우저 혹은 텍스트 편집기에서 쉽게 열어 내용을 확인할 수 있습니다.

Profile Editor의 핵심은 BIBFRAME 용어를 활용하여 서지정보를 입력한 각 개체에 대한 정보를 미리 설정해 놓는 것입니다. 대표적으로 행위자(Agent) 개체를 입력하고 설정해 보겠습니다.

〈그림 43〉 Bibframe Profile 편집기의 Agent 요소 관리 화면

위 그림은 BIBFRAME 2.0 행위자(Agent)에 대한 프로파일 설정 화면입니다. 먼저 Agent라는 개체에 대한 ID, 명칭(Title), 설명, 생성 및 수정 일자, 만든사람에 대한 정보를 입력하고 Agent에 속하는 세부 요소(Person, Family, Corporate Body 등)에 대한 설정을 수행합니다. Person을 선택하면 아래 그림과 같은 설정 정보를 확인할 수 있습니다.

〈그림 44〉 Bibframe Profile 편집기의 Agent의 하위요소인 Person 관리 화면

위 그림은 Agent의 Person에 대한 설정 화면입니다. Agent와 마찬가지로 ID와 명칭(Resource Label), 입력 지침, 그리고 Resource URI 정보를 입력하도록 되어 있습니다. Person의 URI는 “http://id.loc.gov/ontologies/bibframe/Person”입니다. 해당 URI에

접속하면 Person을 정의한 RDF 문서를 확인할 수 있습니다.

```

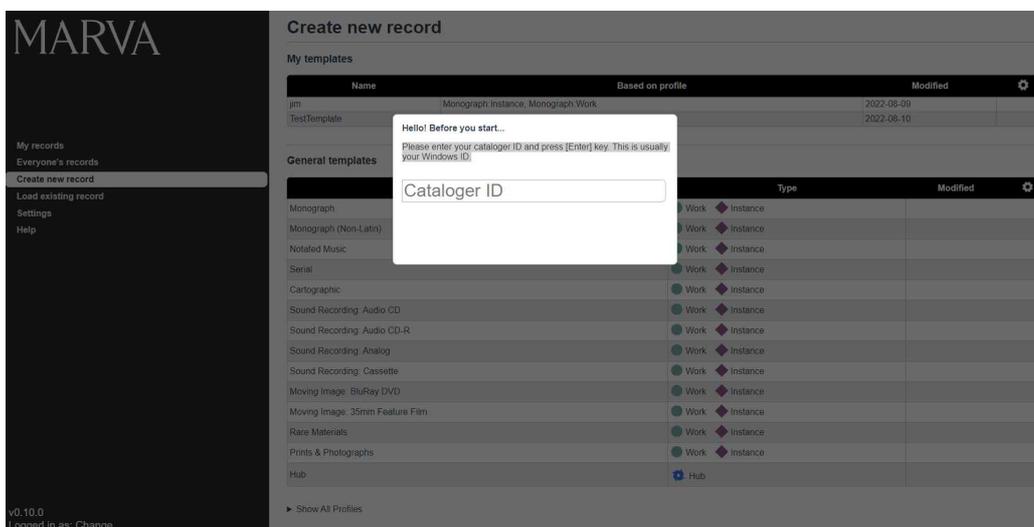
1 <rdf:RDF xml:base="http://id.loc.gov/ontologies/bibframe/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax
2 <owl:Class rdf:about="http://id.loc.gov/ontologies/bibframe/Person">
3   <rdfs:label>Person</rdfs:label>
4   <rdfs:subClassOf rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
5   <rdfs:subClassOf rdf:resource="http://id.loc.gov/ontologies/bibframe/Agent"/>
6   <skos:definition>Individual or identity established by an individual (either alone or in collaboration w
7   <dcterms:modified>2016-04-21 (New)</dcterms:modified>
8 </owl:Class>
9 </rdf:RDF>

```

〈그림 45〉 Person에 대한 RDF 문서

Bibframe Profile Editor는 향후 BIBFRAME 기반의 관리시스템을 적용한다면 구현해야 하는 기능입니다. 단, 위에서 언급한 것처럼 이는 온톨로지 파일을 GUI 관리시스템으로 구현한 형태입니다. 엄격하게 관리 및 통제가 필요한 부분입니다.

이제 Marva Editor에서 실제 서지데이터 입력화면을 살펴보도록 하겠습니다.

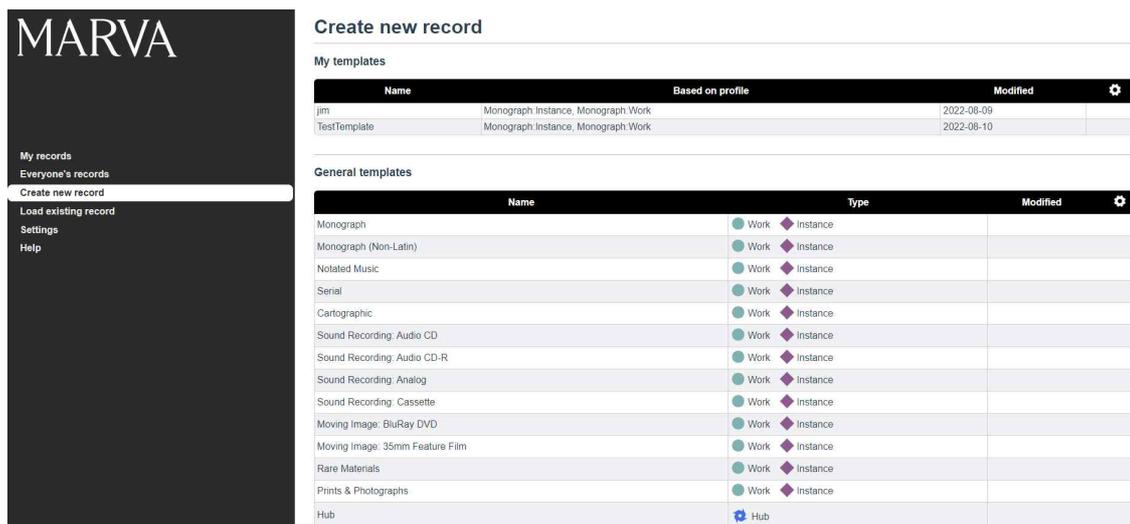


〈그림 46〉 Marva 편집기 초기화면

편집기에 들어가면 목록자 ID를 생성해야 합니다. 실제 도서관 시스템이 아니므로 여러분이 원하는 ID를 생성하면 됩니다. 해설서에서는 'BIB4Kor'라는 명칭으로 아이디를 생성하겠습니다. 아이디를 생성했다고 해서 별도 로그인도 필요하거나 이전 아이디로 나만의 메뉴를 구성하지는 않습니다.

전체 메뉴 구성은 내 코드(My records), 공용 레코드(Everyone's records), 신규 레코드 생성(Create new records), 레코드 불러오기(Load existing record), 환경설정(Settings), 도움말(Help)로 구성되어 있습니다.

여기서는 신규 레코드 생성 메뉴를 활용하여 유형별로 데이터를 입력해보도록 하겠습니다. 다음 그림은 신규 레코드 생성의 초기화면입니다.



〈그림 47〉 MARVA 편집기의 신규레코드 생성 초기화면

신규 레코드 생성을 선택하면 13개 유형의 미리 저장된 템플릿을 확인할 수 있습니다. 이들은 단행본(Monograph), 비라틴어 단행본(Monograph (Non-Latin)), 악보(Notated Music), 연속간행물(Serial), 지도(Cartographic), 녹음자료: 오디오 CD(Sound Recording: Audio CD), 녹음자료: CD-R 오디오(Sound Recording: Audio CD-R), 녹음자료: 아날로그(Sound Recording: Analog), 녹음자료: 카세트(Sound Recording: Cassette), 동영상: 블루레이(Moving Image: BlueRay DVD), 녹음자료: 35mm(Sound Recording: 35mm Feature Film), 희귀본(Rare Materials), 인쇄 사진(Prints & Photographs)으로 자료의 내용 및 매체 유형이 반영되어 있습니다.

단행본 입력을 선택하면 다음과 같은 입력화면을 볼 수 있습니다. 우리는 ‘한강’ 작가의 ‘채식주의자’를 입력해보도록 하겠습니다. 기본 서지정보는 미국의회도서관 자료를 활용합니다 (<https://catalog.loc.gov/vwebv/holdingsInfo?searchId=3576&recCount=25&recPointer=1&bibId=19235547>). 일부 입력 요소 중 일반 텍스트 입력이 가능할 경우에는 국립중앙도서관에서 확인가능한 채식주의자 서지정보도 함께 입력합니다. 이는 여러분의 이해를 돕기 위한 것이라는 점만 주의해주십시오.

BOOK

The vegetarian : a novel

Full Record MARC Tags

Personal name	Han, Kang, 1970- author.
Uniform title	Ch'aesikchuujja. English
Main title	The vegetarian : a novel / Han Kang ; translated from the Korean by Deborah Smith.
Edition	First United States paperback edition.
Published/Produced	New York : Hogarth, [2016] ©2015
Request this Item LC Find It Item Availability	

LCCN Permalink <https://lccn.loc.gov/2016296327>

Description 201 pages ; 21 cm

ISBN 9781101906118 (softcover)

LC classification PL992.26.K36 C4313 2016

<그림 48> 미국의회도서관 vegetarian 서지 검색결과 화면

The screenshot shows the MARVA editor interface. On the left is a navigation menu with categories like 'Work', 'Title Information', 'Form/Gene', etc. The main area contains several input fields: 'Creator of Work' (set to 'Person'), 'Work Title' (with 'Preferred Title for Work' and 'Part name' sub-fields), 'Form/Gene' (with 'Search LOGFF' and 'Thesaurus (for LOGFF)' sub-fields), and 'Government publication'. On the right, there is a sidebar with 'Library of Congress subject headings' and 'Instance' information.

<그림 49> MARVA 편집기의 저작 입력화면 중 일부1

저작(work)에 대한 입력을 먼저 수행하도록 되어 있습니다. 중요한 것은 데이터 입력 시 미국의회도서관의 전거 레코드를 참조하도록 설정되어 있다는 점입니다. 만약 전거레코드에서 입력값이 검색되지 않으면 문자로 입력할 수 있습니다.



〈그림 50〉 MARVA 편집기의 저작 입력화면 중 일부2

Person 항목에 Search LCNAF란을 선택하거나 오른쪽의 돋보기 아이콘을 선택하면 아래와 같은 검색 및 문자 입력 화면을 확인할 수 있습니다.



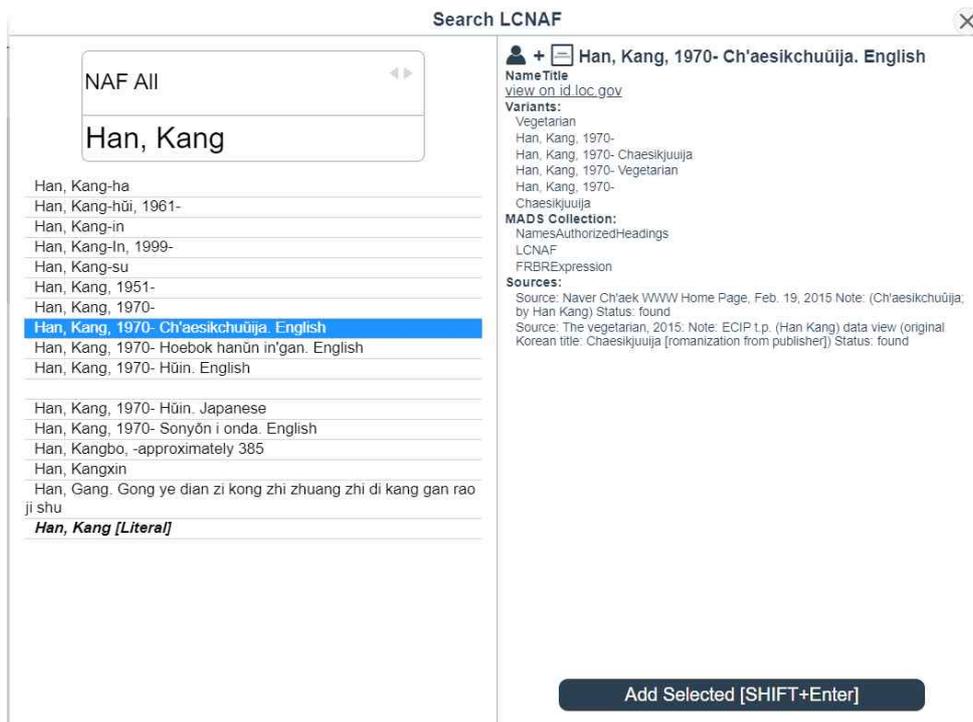
〈그림 51〉 MARVA 편집기의 저작 입력화면 중 저자명 전거 검색 화면1



〈그림 52〉 MARVA 편집기의 저작 입력화면 중 저자명 전거 검색 화면2

한강의 영어 표기 이름 Han, Kang을 입력하여 나온 검색결과입니다. 1970년생, 채식주의자 서지명이 표기된 작가 정보를 확인할 수 있습니다. Marva는 아직 정식버전이 아닙니

다. 그래서 검색에 있어서도 정확한 명칭을 입력하지 않으면 검색이 불가능한 경우가 다수 존재합니다.



〈그림 53〉 MARVA 편집기의 저작 입력화면 중 저자명 선거 검색 화면3

표기된 저작자 중 적합한 인명을 선택하면 해당 상세정보를 확인할 수 있습니다. 추가 버튼을 선택하면 데이터를 입력할 수 있습니다(〈그림 54〉 참조).



〈그림 54〉 MARVA 편집기의 저작 입력화면 중 저자명 입력 후 화면

만약 한글로 '한강'을 입력하려고 할 경우에는 전거를 참조하지 않고 문자열로 입력이 가능합니다. 단, 한글입력은 허용되지 않습니다. 영어로 han kang라고 입력하면 조회된 값이 없음을 표시되고 단순 문자열이라는 의미의 [Literal]이 표시됩니다. 여기서 'Literal' 즉 문자열로 값이 입력된다는 의미는 우리가 앞서 살펴봤던 링크드 데이터 구조에서 마지막

목적어 부분이 문자열로 저장된다는 의미입니다. 앞선 검토내용을 다시 살펴보겠습니다.

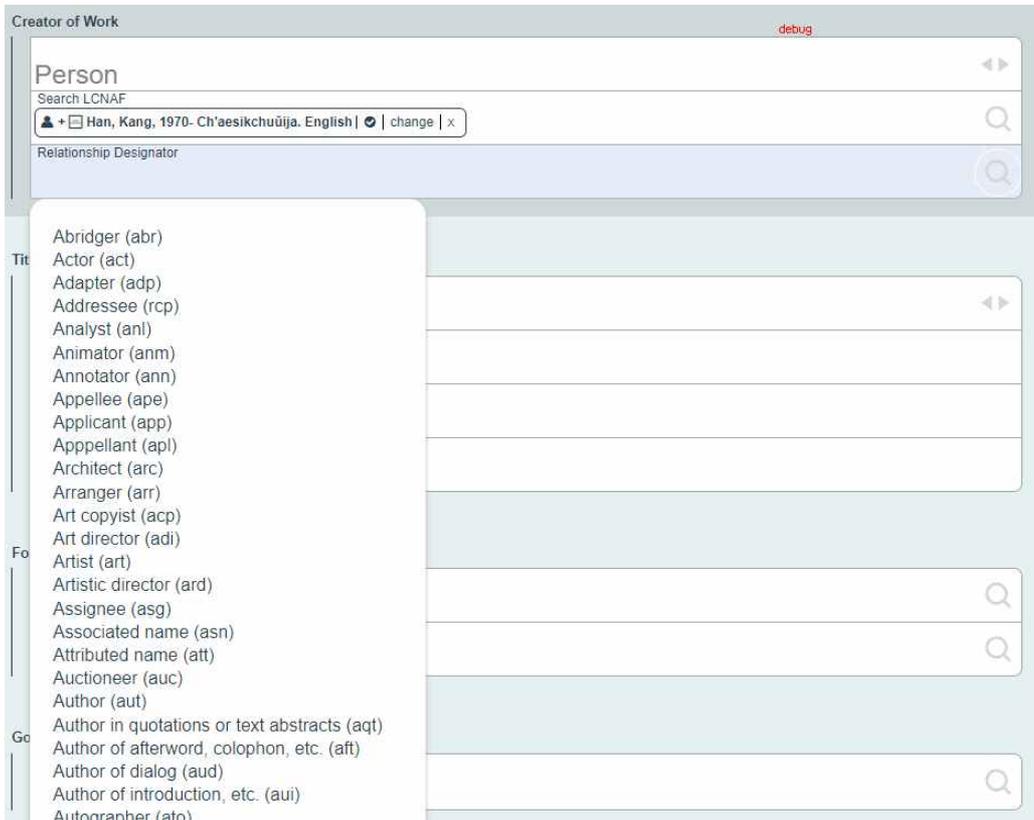
RDF의 구조 =	주어(Subject) -	술어(Predicate) -	목적어(Object)
RDF의 구조 =	URI -	URI -	URI
RDF의 구조 =	URI -	URI -	Literal

문자열로 입력하면 URI 없이 문자열로 값이 입력이 되고, 이는 보다 상세한 정보를 제공하거나 다른 개체로 링크 없이 그대로 입력 작업이 완료됩니다. 이용자는 URI 값이 아니므로 문자로 된 형태의 정보만 취할 수 있고 다른 정보로 이동하거나 상세 정보를 얻을 수 없습니다.



〈그림 55〉 MARVA 편집기의 저작 입력화면 중 저자명을 문자로 입력했을 경우

다음은 해당 저작자의 역할어를 선택해야 합니다. 입력란에서 키보드의 스페이스를 누르면 전체 역할어가 알파벳순으로 나열됩니다. 이미 업무에 능숙한 사람이라면 기존에 알고 있는 역할어명을 입력 및 조회하여 사용하면 됩니다. 여기서는 Author를 선택하겠습니다.



〈그림 56〉 MARVA 편집기의 저자의 역할 검색 화면

입력이 완료된 화면은 다음 그림과 같습니다.



〈그림 57〉 MARVA 편집기의 저자의 역할 입력 완료화면

저작 입력 화면의 오른쪽 상단을 보면 위 그림처럼 식별자라고 할 수 있는 URI가 표현된 것을 확인할 수 있습니다. 우리가 입력하는 저작물의 경우는 “/resources/works/7240c24c-4b2c-4f35-9ad6-32dab860ca23”입니다. 이제 이 저작은 고유한 식별자를 갖게 되었으며, 향후 이 저작의 URI를 기반으로 다양한 인스턴스와 관계를 갖게 됩니다. 위 그림의 debug 버튼을 누르면 실제 RDF/XML로 표현된 서지정보

를 확인할 수 있습니다. 우리가 입력한 정보에 대한 RDF/XML 표현은 아래와 같습니다.

```
-----  
<bf:contribution xmlns:bf="http://id.loc.gov/ontologies/bibframe/">  
  <bflc:PrimaryContribution xmlns:bflc="http://id.loc.gov/ontologies/bflc/">  
    <bf:agent>  
      <bf:Agent xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
        rdf:about="http://id.loc.gov/authorities/names/n2015010884">  
        <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#NameTitle"/>  
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Han, Kang, 1970-  
          Ch'aesikchuŭija. English</rdfs:label>  
      </bf:Agent>  
    </bf:agent>  
  </bflc:PrimaryContribution>  
</bf:contribution>  
-----
```

중점적으로 보아야하는 부분은 굵은 글씨로 표시된 부분입니다. <bf:agent>는 BIBFRAME에서 정의한 agent에 대한 요소를 활용해서 서지정보를 기술하겠다는 의미입니다. 즉 행위자 정보를 표현하겠다는 의미입니다. 그 아랫줄은 실제 agent에 해당하는 값들이 입력됩니다. 우리의 경우는 다음과 같습니다.

```
-----  
<bf:Agent xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  rdf:about="http://id.loc.gov/authorities/names/n2015010884">  
-----
```

우리가 실습하고 있는 위 코드는 한강이라는 작가에 대해서 설명하는 것이고 한강이라는 작가를 미국의회도서관의 전거데이터에서 부여한 URI로 고유하게 식별하였습니다. 즉 한강이라는 작가는 다른 개체와 구별됩니다. 이를 좀 더 이해하기 쉽게 우리가 앞서 살펴봤던 RDF 그래프 형식으로 풀어보겠습니다.



〈그림 58〉 한강 작가의 RDF 그래프 표현1

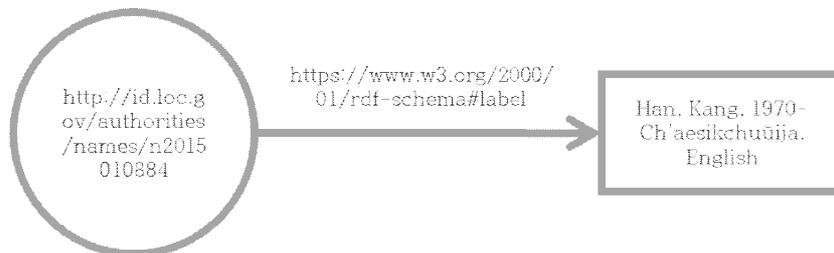
〈그림 58〉에서 표현한 것처럼 한강의 채식주의자는 생성된 저작 URI로 고유하게 구분되고, 이 저작물의 Agent 즉, 행위자는 미국 의회도서관 전거 데이터의 한강 작가에 대한 고유 URI로 구분됩니다. 코드의 이해가 어렵다면 그래프로 도식화하면 이해하기 쉽습니다. 이제 다음 코드를 살펴보겠습니다.

```

<rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#NameTitle"/>
<rdfs:label      xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Han,      Kang,      1970-
Ch'aesikchuüija. English</rdfs:label>

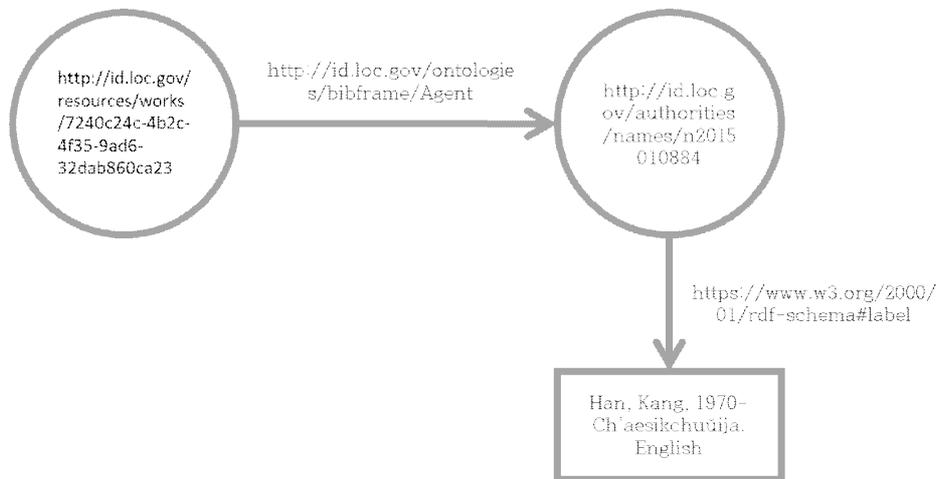
```

사실 이미 저작자에 대한 정보는 URI로 고유하게 식별되었습니다. 그런데 코드에서 “Han, Kang, 1970- Ch'aesikchuija. English”라고 표현된 부분은 rdf 용어집 중 “label” 속성을 활용해서 문자로 표현한 것입니다. 이를 다시 이해하기 쉽게 그래프로 표현하면 다음과 같습니다.



〈그림 59〉 한강 작가의 RDF 그래프 표현2

자, 이제 이 두 가지 그림을 다시 합쳐서 검토해 보겠습니다.



〈그림 60〉 한강 작가의 RDF 그래프 표현3

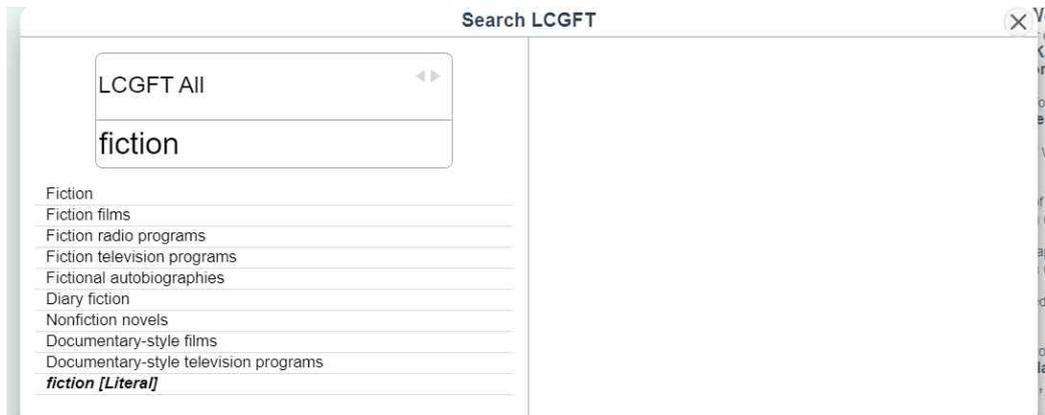
우리가 간단히 편집기로 입력한 한강이라는 작가에 대한 정보는 RDF 형식으로 변환하면 위의 그림과 같이 두 개의 트리플이 서로 연결된 상태임을 확인할 수 있습니다.

이제 저작에 다른 정보들을 순차적으로 입력해보도록 하겠습니다. GUI 환경하에 입력하는 정보들로 특별한 배경지식이 필요하지는 않습니다. 저자 외에 저작의 제목을 입력하면 다음과 같습니다.

The screenshot shows a web interface titled "Title Information" with a "debug" button in the top right. Below the title is a large text input field containing "The vegetarian". To the right of this field is a double-headed arrow icon. Below the main input field are two smaller input fields: "Part number" and "Part name", both currently empty.

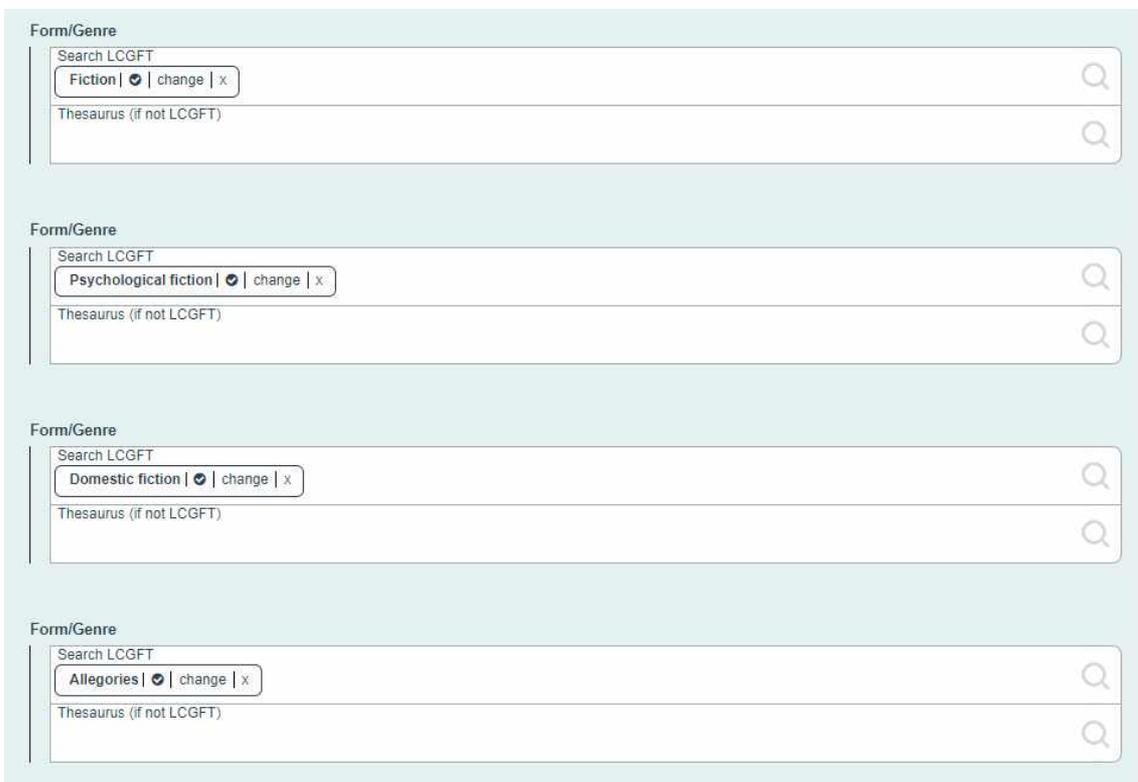
〈그림 61〉 MARVA의 저작 중 제목 입력 화면

다음은 장르와 형식(Genre & Form)에 대한 정보입력입니다. 이 정보 역시 저자명 전거 처럼 이미 등록된 항목 중 선택입력이 기본입니다. 우리는 미국의회도서관의 목록 정보에서 확인한 'fiction', 'Psychological fiction', 'Domestic fiction' 'Allegories'를 입력하겠습니다. 해당 용어는 모두 미국의회도서관 링크드 데이터 서비스 시스템의 'https://id.loc.gov/authorities/genreForms.html'에 등록된 내용을 참조합니다.



〈그림 62〉 MARVA의 저작의 장르, 형식 입력화면(LCGFT 검색 후 선택)

아래 그림은 모든 장르와 형식정보를 입력한 후의 화면입니다.



〈그림 63〉 MARVA의 저작의 장르, 형식 입력 완료 후 화면

장르와 형식의 경우 복수의 요소를 입력해야합니다. MARVA에서 복수의 요소를 입력하기 위해서는 외쪽 주 메뉴에서 ‘+’ 버튼을 눌러 해당 요소 입력창을 추가해주면됩니다. ‘Form/Genre’의 경우 추가 버튼 예시는 아래 그림과 같습니다.



〈그림 64〉 MARVA의 장르/형식 추가 버튼

다음은 출판년도와 원생산지를 입력하는 화면으로 출판년도는 문자형식으로 직접 입력하고, 생산지의 경우 기존 용어집을 참조하도록 되어 있습니다. 여기서는 미국의회도서관의 국가용어집(<http://id.loc.gov/vocabulary/countries/>)을 검색하여 입력합니다. 해당 정보 입력 후 화면은 아래와 같습니다.

〈그림 65〉 MARVA에서 저작의 출판년도, 원생산지 입력 후 화면

다음은 저자 외 해당 저작물에 기여한 사람에 대한 정보를 입력합니다. 우리의 경우 미국의회도서관에서 소장하고 있는 번역본으로 번역자가 존재합니다. 저자를 입력할 때와 마찬가지로 저자전거를 조회하여 입력 가능하며, 역할 역시 번역자로 이미 등록된 역할어를 선택하여 입력합니다. 입력 결과는 아래와 같습니다.

〈그림 66〉 MARVA에서 저작의 다른 기여자 입력 후 화면

다음은 저작물의 주제명을 입력하는 화면입니다. 마찬가지로 주제명 전거 데이터를 조

회 후 입력하도록 되어 있습니다. 우리는 미국의회도서관에서 부여한 주제명을 선택하여 입력합니다. Genre/Form과 마찬가지로 복수 입력을 수행합니다. 입력 대상은 ‘Self-actualization (Psychology) in women—Fiction’, ‘Self-realization in women—Fiction’, ‘Vegetarianism—Fiction’입니다. 입력 후 결과 화면은 아래와 같습니다.

The screenshot displays the MARVA interface with three main sections:

- Subject of the Work** (top): Contains a search bar for LCSH/LCNAF with the selected term 'Self-realization in women—Fiction' and a search bar for the Thesaurus with the selected term 'Library of Congress subject headings'.
- Subject of the Work** (middle): Contains a search bar for LCSH/LCNAF with the selected term 'Vegetarianism in literature' and a search bar for the Thesaurus with the selected term 'Library of Congress subject headings'.
- Classification numbers** (bottom): Contains a search bar for 'Library of Congress Classification' with the entered classification number 'PL992.26.K36 C4313 2016'.

〈그림 67〉 MARVA에서 저작의 다른 기여자 입력 후 화면

다음은 도서관에서 관리하는 식별자(식별체계)를 입력합니다. 우리가 잘 알고 있는 도서관의 분류번호를 합니다. MARVA의 저작에서는 미국의회도서관 분류번호와 듀이십진분류번호를 입력하도록 되어있습니다. 입력값은 모두 문자형태로 직접 입력합니다. 단, 듀이십진분류의 경우 판번호를 조회 후 입력하도록 구성되어 있으며, 해당 분류번호를 입력한 기관을 통제어에서 선택하도록 되어 있습니다. 다음은 해당 정보를 입력한 결과 화면입니다.

Classification numbers

Library of Congress Classification

Classification number
PL992.26.K36 C4313 2016

Cutter number

Classification numbers debug

Dewey Decimal Classification

Classification number
895.734

Cutter number

Dewey Edition number
23 | x

Assigner
DLC | x

〈그림 68〉 MARVA에서 저작의 식별번호 입력 후 화면

다음은 저작의 입력사항은 콘텐츠 유형과 언어정보입니다. 두 가지 모두 미국의회도서관의 통제어휘를 따르도록 구성되어 있습니다. 콘텐츠 유형은 콘텐츠 타입(<http://id.loc.gov/vocabulary/contentTypes>)을 조회하고, 언어정보는 언어용어집(<http://id.loc.gov/vocabulary/languages/>)을 참조합니다. 아래는 입력 후 화면입니다.

Content Type

text | x

Language debug

Language code
Korean | x

Language note

〈그림 69〉 MARVA에서 저작의 콘텐츠 유형과 언어 정보 입력 후 화면

저작 입력의 마지막 단계는 관리 메타데이터(Admin Metadata)입력입니다. 입력 정보는 목록자 ID, 인코딩 수준, 상태, 목록입력 날짜, 변경날짜, 목록입력기관, 메타데이터 입력시 기준이되는 표준, 기술 언어, 로컬 등록번호와 주기사항입니다. 앞서와 마찬가지로 대부분의 정보는 미국의회도서관에 사전에 등록된 용어집이나 코드를 조회하도록 되어 있습니다. 전체 입력 후 화면은 아래와 같습니다.

Admin Metadata		debug
Your cataloger ID (Windows ID)	BIB4Kor	
Encoding level	full ● x	🔍
Status	new ● x	🔍
Creation date	220421	
Change date	20220531130544	
Assigning agency	krsnlk (uncontrolled) x	🔍
Description authentication	pcc ● x	🔍
Description conventions	rda ● x	🔍
Description language	Korean ● x	🔍
Description modifier	krsnlk (uncontrolled) x	🔍
Profile		
Local system number	KMO202219035	
Assigner	krsnlk	
Label		
Process information	create work	
Note	맨부커 인터내셔널상 수상, 2016 산클레멘테 문학상, 2018	

〈그림 70〉 MARVA에서 저작의 콘텐츠 유형과 언어 정보 입력 후 화면

지금까지 다양한 형태의 인스턴스를 갖는 한강 작가의 채식주의자에 대한 저작내용을 입력해보았습니다. 편집기에서는 저작(work) 정보를 입력한 후에는 실제 인스턴스 정보를 입력하도록 되어 있습니다. MARVA의 경우는 저작과 인스턴스가 하나의 화면에서 동시에 입력되도록 구성되어 있습니다. 인스턴스 정보를 입력할 때 기본은 해당 인스턴스가 어떤 저작의 것인지를 URI로 식별하는 것입니다. 우리의 경우 저작 정보를 입력할 때 해당 URI가 자동 할당된 것을 확인하였습니다. 이를 입력한 화면은 아래와 같습니다.

Instance	/resources/instances/7240c24c-4b2c-4f35-9ad6-32dab860ca23
Instance Of	<input type="text" value="http://id.loc.gov/resources/works/7240c24c-4b2c-4f35-9ad6-32dab860ca23"/>

〈그림 71〉 MARVA에서 인스턴스의 해당 저작 URI 입력화면

해당 저작의 URI를 입력한 후에는 제목정보와 책임사항, 판사항을 기술하도록 되어 있

습니다. 모두 일반 문자형으로 정보입력이 가능하며 입력 후 화면은 아래와 같습니다.

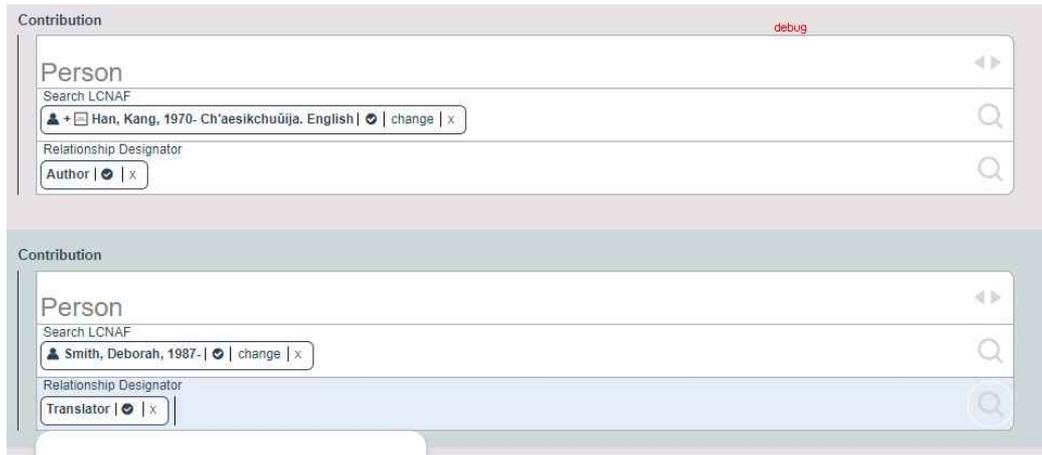
〈그림 72〉 MARVA에서 인스턴스의 제목, 책임사항, 판사항 입력화면

다음은 제공자 활동(Provision activity) 정보를 입력합니다. 주로 출판사와 관련된 정보로 장소, 출판사이름, 출판연도와 주기사항을 입력합니다. 장소는 미국의회도서관의 용어집을 참조하여 선택하며, 나머지 정보는 직접 텍스트 입력이 가능합니다. 아래는 입력 후 결과화면입니다. 출판사 정보의 경우 일반 텍스트 입력이 가능하며, 우리나라 출판사인 창비를 문자로 입력하였습니다.

〈그림 73〉 MARVA에서 인스턴스의 제공자 활동 정보 입력화면

다음은 저자를 포함한 다양한 기여자 정보를 입력하도록 되어 있습니다. 입력방식은 저작에서와 마찬가지로 저자명전거 데이터를 조회하여 참조하도록 되어 있습니다. 아래는 저

자와 번역자 정보를 조회하여 입력한 결과화면입니다.



〈그림 74〉 MARVA에서 인스턴스의 저자 및 기여자 정보 입력화면

다음은 간행방식(Mode of issuance)을 입력합니다. 간행방식도 이미 정의된 용어집 (<http://id.loc.gov/vocabulary/issuance/>)에서 적합한 방식을 선택하도록 구성되어 있습니다. 아래는 입력 후 화면입니다.



〈그림 75〉 MARVA에서 인스턴스의 간행방식 입력화면

다음은 식별자 정보를 입력합니다. 앞서와 마찬가지로 복수의 식별자 입력이 가능합니다. 인스턴스에 해당하는 식별자를 입력하는 화면으로 청구기호와 같은 유형의 식별자 정보를 입력합니다. 우리는 미국의회도서관과 국립중앙도서관의 청구기호를 입력해보도록 하겠습니다. 예시자료로 두 개 기관의 청구기호를 넣었다는 점만 유의해주시오.

The image shows two sections of the MARVA interface for entering identifiers. The top section is titled 'Identifiers' and contains a text input field with 'LCCN' and a value of '2016296327'. Below the input field is a search icon and a label 'Invalid/Canceled?'. The bottom section is also titled 'Identifiers' and contains a text input field with 'Other identifier' and a value of '813.62'. Below this is a 'Qualifier' field with the text '국립중앙도서관 청구기호'. A small red 'debug' label is visible in the top right corner of the bottom section.

〈그림 76〉 MARVA에서 인스턴스의 식별자 입력화면

다음은 인스턴스에 대한 미디어 유형, 형태사항, 수록매체 정보를 입력합니다. 미디어 유형과 수록매체는 각각 해당용어집(<http://id.loc.gov/vocabulary/mediaTypes>), [〈http://id.loc.gov/vocabulary/carriers〉](http://id.loc.gov/vocabulary/carriers)을 참조하며, 나머지는 일반 텍스트로 입력 합니다.

The image shows a form with four sections: 'Media type', 'Extent', 'Dimensions', and 'Carrier type'. The 'Media type' section has a dropdown menu with 'unmediated' selected and a search icon. The 'Extent' section has a text input field with '201 pages', a 'Note' field, and an 'Applies to' field. The 'Dimensions' section has a text input field with '21 cm'. The 'Carrier type' section has a dropdown menu with 'volume' selected and a search icon. A small red 'debug' label is visible in the top right corner of the 'Extent' section.

〈그림 77〉 MARVA에서 인스턴스의 미디어유형, 형태사항, 수록매체 입력화면

인스턴스와 직접적으로 연관되는 마지막 정보는 인스턴스의 URI를 입력하는 것입니다. 여기서는 국립중앙도서관 LOD의 URI를 활용하였으며, 다음과 같습니다.

URL of Instance	
URL	http://lod.nl.go.kr/page/KMO202219035
Note	국립중앙도서관 LOD

〈그림 78〉 MARVA에서 인스턴스의 URI 정보 입력화면

마지막은 인스턴스의 관리 메타데이터 입력합니다. 입력 정보는 목록자 ID, 인코딩 수준, 상태, 목록입력 날짜, 변경날짜, 목록입력기관, 메타데이터 입력 시 기준이되는 표준, 기술 언어, 로컬 등록번호와 주기사항입니다. 앞서와 마찬가지로 대부분의 정보는 미국의회도서관에 사전에 등록된 용어집이나 코드를 조회하도록 되어 있습니다. 전체 입력 후 화면은 아래와 같습니다.

Admin Metadata		debug
Your cataloger ID (Windows ID)	BIB4Kor	
Encoding level	full <input checked="" type="radio"/> x	🔍
Status	new <input checked="" type="radio"/> x	🔍
Creation date	2022-09-01	
Change date	20220531130544	
Assigning agency	krsnlk (uncontrolled) x	🔍
Description authentication	pcc <input checked="" type="radio"/> x	🔍
Description conventions	rda <input checked="" type="radio"/> x	🔍
Description language	English <input checked="" type="radio"/> x Korean <input checked="" type="radio"/> x	🔍
Description modifier	krsnlk (uncontrolled) x	🔍
Profile		
Local system number	KMO202219035	
Assigner	krsnlk	
Label		
Process Information	create work	
Note	맨부커 인터내셔널상 수상, 2016 산클레멘테 문학상, 2018	

〈그림 79〉 MARVA에서 인스턴스의 관리 메타데이터 정보 입력화면

지금까지 MARVA 편집기를 기준으로 BIBFRAME 용어집 기반의 서지 정보 입력을 수행

했습니다. 서두에 언급한 것처럼 사실 GUI 환경에서의 입력화면이기 때문에 기존 목록작업과 큰 차이점을 알 수 있는 부분은 적습니다. 편집기 상단의 코드보기(</>) 버튼을 눌러서 최종결과물을 확인할 수 있는데 우리의 경우 아래와 같습니다.



〈그림 80〉 MARVA에서 입력 완료 후 코드보기 화면

우리는 여기서 MARC와는 다른 형식의 결과물을 확인할 수 있습니다. 조금이나마 URI를 중심으로 개체를 기술하고 연결하는 것이 무엇인지 알 수 있었기를 바랍니다. 또, 데이터 입력, 저장형식이 어떻게 바뀌더라도 우리가 일하는 방식에는 큰 차이가 없을 것이라는 안도감을 갖기를 바랍니다.

05 KORMARC의 BIBFRAME 변환

BIBFRAME은 MARC 데이터의 변환만을 목적으로 하지는 않습니다. 그럼에도 도서관의 원천데이터인 MARC의 변환도 가능하도록 설계되어 있습니다. 대상 MARC은 MARC 21입니다. 때문에 KORMARC와는 완벽하게 매핑되지 않습니다. BIBFRAME이 목적인 바를 이루려면 단순히 MARC를 BIBFRAME으로 변환하는 것은 큰 의미가 없습니다. 도서관 데이터에 웹에 친숙한 형태로 변환되어 웹에서 다른 데이터들과 연결될 때 목적인 바를 이루었다고 할 수 있으며, 그 효과가 증명된다고 할 수 있습니다. 즉, 단순히 MARC이 BIBFRAME으로 변환되었다는 것은 큰 의미가 있다고 보기는 어렵습니다. 하지만 아직까지 도서관 원천데이터의 핵심인 MARC의 변환에 대한 이해와 변환은 필요합니다. 여기서는 미국의회도서관이 제공하는 MARC 21과 BIBFRAME 변환에 관한 정보를 기반으로 검토해보도록 하겠습니다.

미국의회도서관은 BIBFRAME 2.0 사용을 지원하기 위해 MARC 21 서지 레코드 필드의 변환 과정을 BIBFRAME의 허브(Hub), 저작(Work), 인스턴스(Instance), 개별자료(Item) 항목으로 매핑하기 위한 변환 명세서(Conversion Specification)를 제공²⁾하고 있습니다. 이 명세서는 MARC 21 관점에서 작성되었는데, 변환 절차에 대한 문서 자료(MS Word 형식)와 MARC 21 필드와 BIBFRAME 요소 간 매핑 파일(MS Excel 형식)로 구성되어 있습니다.

본 해설서에서는 MARC21 관점에서 작성된 미국의회도서관의 변환 명세서를 토대로 국립중앙도서관에서 제공하고 있는 한국문헌자동화목록형식(KORMARC)과 일대일 확인 작업을 통해 필드별 주요 차이점과 특징들을 정리하였습니다. 변환 명세에 대해 각 파일별로 정리한 내용은 분량이 많은 관계로 부록으로 별도 첨부하였습니다. 여기서는 변환 절차에 대한 주요 내용(Process Note)과 필드별로 차이가 있는 항목들을 중심으로 정리한 내용만 기술하였습니다.

2) <https://www.loc.gov/bibframe/mtbf/>

5.1 변환 명세서에 대한 참고사항

앞서 언급한 것처럼 KORMARC, MARC21과 BIBFRAME 2.0 간의 매핑정보는 별도의 파일(MS Excel)형식으로 제공하며, 현재 보고서에는 분량관계로 직접 첨부하지 않았사브니다. 변환 명세서에 대한 참고사항은 해당 파일을 참고할 경우 사전에 알고 있어야하는 주의사항, 참고사항에 대해 기술하였다. KORMARC의 경우는 아직 공식적인 변환 명세서가 존재하지 않기 때문에 MARC21을 기준으로 명세서에 제공된 참고사항이 주요 내용이라고 생각하면 됩니다. 엑셀 형식의 변환 명세서를 읽고 해석해야 할 경우 이 참고사항들을 미리 숙지하거나 참고하면 쉽게 이해할 수 있습니다.

변환 명세서는 MARC 21에서 BIBFRAME 2.0으로의 매핑이 핵심입니다. 이때, 해당 항목을 거의 사용하지 않거나 전혀 사용하지 않는 경우에는 “nac(no attempt to convert, 변환 시도 없음)”로 표시합니다. 최근에 MARC 21 형식으로 승인된 필드나 하위필드는 대부분 “nac”로 표기되어 있습니다. 그러나 이는 현재 기준으로 향후 해당 필드의 활용성이 높아지면 수정될 수도 있습니다. MARC 외부 환경과 관련이 없는 일부 MARC 레코드 요소들은 “ignore(무시함)”으로 표시되어 있습니다.

명세서에서는 많은 약칭들을 사용합니다. 저작(Work)은 W, 인스턴스(Instance)는 I를 사용하지만, 개별자료는 Item, 허브는 Hub로 표시합니다. Hub는 말 그대로 일종의 가교 역할을 수행합니다. 저작(Work)을 연결하기 위한 추상적인 자원인데 MARC는 물론 RDA, LRM 저작물과 BIBFRAME 저작물을 동등하게 연결하기 위한 노력의 일환입니다.

007 및 008 필드 등 변환 명세서에서 사용하는 이름, 주제, 장르 용어 및 개별 요소에 대해서는 대부분 <http://id.loc.gov>에 정의된 어휘(vocabularies)를 사용합니다.

명세서는 MARC에서 자원을 기술할 때 사용하는 요소들을 BIBFRAME 모델을 준수하도록 일치시키거나 병합하는데 중점을 두고 있습니다. 이때 MADS(Metadata Authority Description Schema, 메타데이터 전거기술 스키마) 이름(names)과 주제 파일(subject files)에서 사용하는 대상 객체의 레이블(labels)과 RWO(Real World Object, 실세계 객체)를 식별하는 작업이 필요합니다. 이를 위해 사용할 데이터에 대한 문자열(strings)과 함께 로컬 속성(properties)을 사용합니다. “bflc”네임스페이스의 이러한 로컬 요소는 이름과 제목을 변환하고자 할 때 사용합니다.

이상의 참고사항들은 MARC21을 BIBFRAME으로 전환하고자 할 때의 기준입니다. 따라서 KORMARC에서만 사용하거나 표현되는 대상에 대해서는 미국의회도서관에서 정의하고 있는 어휘나 규칙 외에 별도의 내용들을 생성해 변환 · 매핑해야 할 수도 있습니다. 하지만 KORMARC to BIBFRAME을 위해 필요한 어휘 및 규칙 생성과 매핑, 변환 등에 대한

내용은 본 연구의 범위를 넘어서는 내용이므로 후속 연구가 요구되며, 부록의 변환 파일에서 KORMARC에만 해당하거나 MARC 21과 다른 내용에 대해서는 별도 표시(셀 혹은 글자색 구분)하였습니다.

5.2 KORMARC 서지 변환 명세를 위한 프로세스 노트(Process Note)³⁾

프로세스 노트에서 다루는 범위 중 프로세스 0 ~ 3, 5는 서지 변환을 위한 것이며, 4는 전거 변환에 필요한 것입니다.

■ 프로세스 0 - 이름(name) 및 제목(title) 처리

0.1) 이름 및 이름/제목 필드 구성요소

- 이름(명칭) 부분 : X11 필드에서 ▼e, ▼4, ▼h, ▼i, ▼j 를 제외한 ▼t 이전의 모든 식별기호
- 역할 부분 : X11 필드에서 ▼e, ▼4, ▼j
- 이름(명칭)(표제) 부분 : ▼h,v,x,y,z,w,0-8을 제외한 ▼t 이후의 모든 식별기호
- 총서 부분 : 8XX 필드의 ▼v, 7XX와 8XX의 ▼x
- 장르 부분 : ▼h
- 관계 부분 : 7XX 필드의 ▼i

0.2) 관계(Relationship, 역할어) 구성요소

- 1XX 필드의 경우, 이름(명칭)/제목(표제)가 설명하고 있는 자원임
- 6XX 필드의 경우, 관계(역할어)는 bf:subject임(ConvSpec-600-662-v1.7.1-KORMARC⁴⁾ 파일 참조)
- 700-730 필드에서 식별기호 ▼i가 없는 경우,
 - 제2지시기호가 2(분출표목)이라면, 관계(역할어)는 bf:hasPart이며
 - 그 외의 경우 관계(역할어)는 bf:relatedTo로 변환하고,
 - 식별기호 ▼i가 있는 경우, bf:relation 속성을 사용해 ▼i의 콘텐츠를 가져감(프로세스 0.3.3 참조)
- 8XX 필드의 경우, 관계(역할어)는 bf:hasSeries임(프로세스 6참조)
- 760-788 필드의 경우,
 - 태그 및 제1지시기호에 의해 관계(역할어)를 결정하며(ConvSpec-760-788-Links-v1.7-KORMARC 파일 참조)
 - 식별기호 ▼i도 있는 경우, bf:relation 속성을 사용해 ▼i의 콘텐츠를 가져감(프로세스 0.3.3 참조)

3) <https://www.loc.gov/bibframe/mtbf/ConvSpec-Process0-7-v1.7.docx> 파일을 번역하였으며, KORMARC 필드와의 일대일 확인을 통해 MARC21과 차이가 있을 수 있는 부분에 대해서는 별도 주석으로 처리하였습니다.

4) MARC21 to BIBFRAME에서는 참고할 변환 명세 엑셀 파일에 대한 설명을 미국의회도서관에서 정리한 파일명(예: ConvSpec-660-662)으로 표시하고 있으나, 본 해설서에서는 이러한 설명에 포함되는 파일명을 부록에 포함된 KORMARC 필드와 일대일 매핑 내용을 확인한 파일명(예: ConvSpec-600-662-v.1.7.1-KORMARC) 기준으로 정리하였습니다.

0.3) 이름, 제목 및 관계(역할어)에 대한 기본 RDF 패턴

0.3.1) 이름에 대한 RDF

이름(명칭)이 1XX 필드에서 오는 경우, bf:Contribution 대신 bf:PrimaryContribution을 사용

```
<resource> bf:contribution [ a bf:Contribution ;
    bf:agent [ a bf: Person, Organization, Meeting, Family or Jurisdiction
        rdfs:label "label from Process 1.3";
        identifiedBy [ a Identifier ... ]; see Subfield ▼0 spec
        bf:nameXXMatchKey "string from Process 1.1";
        bf:nameXXMarcKey "string from Process 1.2" ] ;
    bf:role [ a bf:Role
        [rdfs:label "..."] ; see Process 1.4
        bf:code "..."] ]. see Process 1.4
    If URI from ID for role, then instead:
    bf:role URI for role
```

0.3.2) 제목에 대한 RDF

제목 식별기호에서 제목 클래스를 구성함; 제목의 하위 속성들을 필드와 동일한 순서로 유지

```
bf:Work bf:title [ a bf:Title
    bf:mainTitle "content of ▼a (X30, 240) or ▼t (X00, X10, X11)" ;
    bf:partnumber "content of ▼n" ;
    bf:partName "content of ▼p" ;
    bf:titleXXMatchKey "see Process 2.1" ;
    bf:titleXXMarcKey "see Process 2.2" ;
    bf:titleSortKey "see Process 2.3" ] .
    bf:identifiedBy [ a Identifier ... ] see Subfield ▼0 spec
```

0.3.3) 관계(역할어)에 대한 RDF

```
<resource>bf:relatedTo** URI
```

** bf:relatedTo는 bf:hasPart를 대신할 수도 있으며, bf:hasSeries, 혹은 또 다른 특정 관계 속성 중 하나일 수 있음

특정 관계도 표현해야 하는 경우

Bib 7XX ▼나 Auth 5XX ▼에 오직 관계 레이블만 있는 경우:

```
<resource>bflc:relationship [ a bflc:Relationship;
    bflc:relation [ a bflc:Relation;
        [rdfs:label "name of relationship" ] ] ] .
```

관계 레이블 및/또는 관계 URI가 있는 경우:

```
<resource>bflc:relationship [ a bflc:Relationship;
    bflc:relation [ URI for relation;
        [rdfs:label "name of relationship" ] ] ] .
```

```
URI a bf:Work or bf:Instance;
    identifiedBy [ a Identifier ... ]; see Subfield ▼0 spec
    bflc:titleXXMatchKey "See Process 1.1";
    bflc:titleXXMarcKey "See Process 1.2".
```

■ 프로세스 1 - 이름(name) 처리

X00, X10, X11 이름 변환

이름 키에 대한 참고사항: 필드가 이름/제목 필드인 경우 식별기호 ▼t 이전의 식별기호들만 이름의 일부로 포함함. 제목과 이름에는 몇 개의 식별기호가 있을 수 있으며, ▼t 이후에 오는 경우는 제목의 일부임

1.1) 매칭 키 생성

모든 내용을 대상으로: 모든 지시기호 및 식별기호 코드를 삭제함 - 필드 순서는 유지

```
X00 - abcdjq - bflc:name00MatchKey
X10 - abcdng - bflc:name10MatchKey
X11 - acdengq - bflc:name11MatchKey
```

1.2) MARC 키 생성

모든 내용을 대상으로: 모든 지시기호 및 식별기호 코드를 유지함 - 시작 부분에 태그를 고정 - 제목(표제)가 있는 경우에도 전체 필드는 유지함 - 필드 순서는 유지: tagii⁵⁾▼atext▼btext▼gtext)

5) tagii: 각 필드번호(tag)와 제1, 제2지시기호(indicators)를 표시하는 기호입니다. 해당 필드번호-지시기호(1, 2)-각 식별번호 별 내용 유지 형식을 의미합니다.

```
X00 - bf:lc:name00MarcKey
X10 - bf:lc:name10MarcKey
X11 - bf:lc:name11MarcKey
```

1.3) rdfs:label 생성

모든 내용을 대상으로: 각 식별기호 코드를 공백으로 대체함 - 필드 순서는 유지

```
X00 - abcdjq - rdfs:label
X10 - abcdng - rdfs:label
X11 - acdengq - rdfs:label
```

1.4) 역할 설정

MARC 서지 레코드가:

- X00, X10 필드에 ▼e, X11 필드에 ▼j 또는 ▼4가 없는 경우, 역할은 기여자 (contributor)'이며, ID:<<https://id.loc.gov/vocabulary/relators/ctb>>에서 URI를 추가함
- X11 필드에 ▼e 또는 ▼j가 있는 경우,

```
bf:role  bf:Role  rdfs:label "content of ▼e (X00, X10) or ▼j (X11)"
```

참고사항: 식별기호 콘텐츠에 "and", &, 혹은 ","가 있는 경우 식별기호에 여러 가지 역할들이 있음. 이 경우 각각을 bf:role로 분리해 처리함.

- ▼4(각각의 ▼4)에 3자리 코드가 있는 경우,

```
          bf:role  a  bf:Role          bf:code  "content of ▼4"
or        bf:role  a  bf:Role          URI for role from ID
```

- 4(각각의 ▼4)에 URI(<http://...>)가 포함된 경우,

```
bf:role  URI from ▼4
```

- 필드 태그가 1XX인 경우, 이름(명칭) 정보에 대해 bf:PrimaryContribution 클래스를 사용함(프로세스 0.3 참조)

■ 프로세스 2 - 제목(title) 처리

X00, X10, X11, X30, 240 필드의 제목(표제) 변환

참고 : 아래의 't'로 시작하는 식별기호 문자열의 경우, ▼t 이후의 제목에서 나타나는 하위필드만 포함함. 몇 개의 하위 필드는 ▼t 전후로 나타날 수 있으며, 이전에 있으면 제목이 아니라 이름의 일부임

2.1) 매칭 키 생성

모든 내용을 대상으로: 모든 식별기호 코드를 삭제함 - 필드 순서는 유지

```
X00 - tfgklmnoprs - bflc:title00MatchKey
X10 - tdfgklmnoprs - bflc:title10MatchKey
X11 - tfgklmps - bflc:title11MatchKey
X30 - adfgklmnoprs - bflc:title30MatchKey
240 - adfgklmnoprs - bflc:title40MatchKey
```

2.2) MARC 키 생성

모든 내용을 대상으로: : 모든 지시기호 및 식별기호 코드를 유지함 - 시작 부분에 태그를 고정 -이름(명칭)이 있는 경우에도 전체 필드는 유지함 - 필드 순서는 유지 - 구분기호를 ▼ 기호로 변환: tagii\$atext\$btext\$gtext

```
X00 - bflc:title00MarcKey
X10 - bflc:title10MarcKey
X11 - bflc:title11MarcKey
X30 - bflc:title30MarcKey
240 - bflc:title40MarcKey
```

2.3) (필요한 경우) 정렬 문자열 만들기

제2 지시기호가 공백이거나 0이 아닌 경우, 제2지시기호에 지정된 문자와 후행 "/" 구두점을 제거해 정렬 문자열을 만들음. 새 문자열은 bflc:titleSortKey로 명명함.

정렬 문자열에 포함하는 식별기호는 다음과 같이 정리함(모든 내용을 대상으로: 각 식별기호 코드를 공백으로 대체-필드 순서는 유지)

```
X00 - tfgklmnoprs
X10 - tdfgklmnoprs
X11 - tfgklmps
X30 - adfgklmnoprs
240 - adfgklmnoprs
```

■ 프로세스 3 - 856필드(전자적 위치 및 접속)

3.1) 필드에 ▼u가 없으면, 856필드는 변환하지 않음

3.2) 856이 ▼3=목차 식별기호를 갖는 경우, 저작물(Work)의 bf:tableOfContents에 ▼u를 추가함(ConvSpec-5XX-v1.7.1-KORMARC 파일을 참조)

```
<bf:Work> bf:tableOfContents [ a bf:TableOfContents ;  
    rdf:value "URL from ▼u"^^xs:anyURI ] ;
```

3.3) 856필드의 제2지시기호가 2(관련자료)인 경우

```
Instance - supplementaryContent - SupplementaryContent  
    rdf:value "URL from ▼u"^^xs:anyURI  
    note - Note - rdfs:label "content of ▼z"  
    note - Note - rdfs:label "content of ▼y"  
    note - Note - rdfs:label "content of ▼3"  
  
<bf:Instance> bf:supplementaryContent [ a bf:SupplementaryContent ;  
    rdf:value "URL from ▼u"^^xs:anyURI ;  
    bf:note [ a bf:Note ;  
        rdfs:label "note text" ] ] ;
```

3.4) 856필드의 제2지시기호가 b (해당 정보 없음), 0(자료자체), 1(자료의 버전) 또는 8(표출어를 생성하지 않음)인 경우, 856 필드 ▼u에서 변환할 URL은 다음과 같으며, 제2지시기호가 b, 0, 1 또는 8인 경우 856필드의 다른 내용들은 변환하지 않 것

```
*.loc.gov  
*.fdlp.gov  
*.gpo.gov  
*.hathitrust.org
```

3.5) 인스턴스가 전자자료인 경우(008/23기호가 o(온라인 전자자료)나 s(전자자료)인 경우)

6) 두 칸 띄어쓰기

```

Instance - hasItem -
  Item - electronicLocator - rdfs:Resource
    rdf:value "URL from ▼u"^^xs:anyURI
    note - Note - "rdfs:label "content of ▼z"
    note - Note - "rdfs:label "content of ▼y"
    note - Note - "rdfs:label "content of ▼3"

```

인스턴스가 전자자료가 아닌 경우, 아날로그 인스턴스의 제목(표제)으로 새 인스턴스를 만들고 저작물(Work)을 가리키도록 함

```

Instance      a      Electronic
  - link to the Work
  - hasItem -
    Item - electronicLocator - rdfs:Resource
      rdf:value "URL from ▼u"^^xs:anyURI
      note - Note - rdfs:label "content of ▼z"
      note - Note - rdfs:label "content of ▼y"
      note - Note - rdfs:label "content of ▼3"

```

```

<bf:Instance> bf:hasItem [ a bf:Item ;
  bf:electronicLocator [ a rdfs:Resource ;
    rdf:value "URL from ▼u"^^xs:anyURI ;
    bf:note [ a bf:Note ;
      rdfs:label "note text" ] ] ];

```

■ 프로세스 4 - 4XX와 5XX태그의 전거(典據)⁷⁾

4.1) 4XX 전거

▼t가 없는 400, 410, 411 태그의 경우(오직 이름만 있는 경우)

- 프로세스 1과 0(이름 부분)을 적용해 매칭키를 만들고, bib 1XX와 같이 처리하고

W - contribution을 만듦(즉, 이름/제목이 아닌, 이름인 bib 7XX처럼 처리함).

만약 MARC 데이터에 역할이 없다면, <<https://id.loc.gov/vocabulary/relators/ctb>>를 참조해 역할을 추가함.

7) 통합서지용 태그가 아니라 KORMARC 전거통제용에 기술된 400, 410, 411(보라부출-개인명, 단체명, 회의명), 500, 510, 511(도보라 부출-개인명, 단체명, 회의명)과 관련된 내용입니다. Process Note의 원문은 Authority 4XX and 5XX tags로 되어 있습니다. MARC to BIBFRAME Specifications는 전체적으로 통합서지용 필드 내용으로 정리되어 있어 해당 내용으로 매칭했습니다. 프로세스 노트의 경우는 통합서지용 외에 내용도 들어가 있어서 해설서에서는 원 파일에 있는 내용 그대로 정리했습니다.

▼t가 있는 400, 410, 411 태그 및 430 태그의 경우

- ▼a 이전에 있는 부분(이름 부분)과 ▼t 및 그 이후 부분을 포함하는 부분(제목 부분)을 구분함; 430의 경우 제목 부분은 모두 ▼a 뒤에 있음
- ▼t 이전의 이름 부분으로 bflc:nameXXMatchKey를 만들(430은 무시함)
- 1XX용으로 만든 bflc:nameXXMatchKey와 비교함(430은 무시함)
- 비교내용이 일치하면 이름 부분을 버림
- 제목 부분을 변형 제목처럼 취급할 때에는: W - title - VariantTitle
- 변형 제목 부분을 처리하려면, ConvSpec-240,X30-UnifTitle-v1.7-KORMARC 파일을 참조
- 이름 부분이 있지만 1XX 이름 부분과 일치하지는 않는 경우, 4XX와 같이 처리하고 <<https://id.loc.gov/vocabulary/relators/ctb>> URI를 추가함

4.1) 5XX 전거

▼t가 없는 500, 510, 511 태그의 경우(오직 이름만 있는 경우)

- 프로세스 1과 0(이름 부분)을 적용해 매칭키를 만들고, bib 1XX와 같이 처리하고 W - contribution을 만들

▼t가 있는 500, 510, 511 태그의 경우

- ▼a와 ▼t 사이의 데이터(이름 부분)와 ▼t와 그 이후 부분을 포함하는 데이터(제목 부분)를 구분함
- 프로세스 1.1, 1.2, 1.3, 1.5를 사용해 ▼a로 시작하는 데이터를 처리해 매칭키를 만들고, ConvSpec-1XX,7XX,8XX-Names-v1.7-KORMARC파일을 참조해 이름(명칭) 부분을 처리함
- 이름 부분이 있는 경우 bib 7XX와 같이 처리하고 <<https://id.loc.gov/vocabulary/relators/ctb>> URI를 추가함
- 프로세스 2를 사용해 ▼t 이후의 데이터를 처리해 매칭키를 만들고, ConvSpec-240,X30-UnifTitle-v1.7-KORMARC 파일을 참조해 제목(표제) 부분을 처리함
 - 관계(역할어)를 결정하려면, 프로세스 0.3.3의 '관계(역할어)에 대한 RDF(RDF for relationships)'를 참조함
 - ▼i의 관계 정보는 제목으로 변환됨. 프로세스 0.3.3의 '관계(역할어)에 대한

RDF(RDF for relationships)’를 참조

- 500, 510, 511 태그에 ▼i가 없으면 ▼w/0을 확인함
- ▶ ▼w/0이 f인 경우, W -derivativeOf - W를 생성
- ▶ 500, 510, 511 태그에 ▼i가 없고, ▼w/0이 f가 아닌 경우, W - relatedTo - W를 생성

530 태그의 경우

- 프로세스 2를 사용해 ▼a 이후 데이터를 처리해 매칭키를 만들고, ConvSpec-240,X30-UnifTitle-v1.7-KORMARC파일을 참조해 제목(표제) 부분을 처리함

- 관계(역할어)를 결정하려면, 프로세스 0.3.3의 ‘관계(역할어)에 대한 RDF(RDF for relationships)’를 참조함
- 530 태그에 ▼i가 있으면, 다음 내용을 생성함

```

bflc:relationship [a bflc:Relationship ;
bf:relatedTo      <related W uri>
bflc:relation     rdfs:label "content of ▼i"

```

- 530 태그에 ▼i가 없으면 ▼w/0을 확인함
- ▶ ▼w/0이 f인 경우, W -derivativeOf - W를 생성
- ▶ 530 태그에 ▼i가 없거나 ▼w/0이 f가 아닌 경우, W - relatedTo - W를 생성

■ 프로세스 5 -이름/제목 및 제목 전거(典據) 선택

5.1) 레코드 선택

이름 파일에서 처리할 레코드를 선택함

- 130태그인 경우는 선택
- 100, 110, 111 태그에 ▼t가 있는 경우 선택

5.2) 누락사항 변환

변환되지 못한 레코드 내 모든 필드는 다음과 같이 숨김 처리함

```

bflc:missingConversionSpec - literal (marc field with tag and indicators and subfields)

```

5.3) 주기사항

연속된 시리즈 자료나 통합된 자원을 식별할 수 있는 명확한 방법은 없으며, LC 전거 파일에 포함된 제목 및 이름/제목의 대략적인 수는 (2017년) 1,256,426, (2016년) 1,178,841, (2012년) 1,055,887임.

선택한 레코드 내에서 연속되는 시리즈 자료에 대한 레코드를 식별하려면 008/16이 a 인 경우를 확인함. 추가적으로 008/12와 008/13이 n이 아닌지 확인할 것.

■ 프로세스 6 -총서사항 처리

6.1) 4900_ 태그의 변환은 ConvSpec-490,510-Links-v1.7-KORMARC 파일을 참조

6.2) 4901_/8XX 그룹이 2개 이상인 경우 그 둘을 쌍을 만들고 동일한 순서로 있는 것으로 가정함. 예를 들어 첫 490 태그는 첫 8XX 태그와 함께 가며, 490 태그의 데이터는 프로세스 6.4 또는 6.5 단계에 필요할 수 있음

6.3) ISSN은 8XX ▼x 또는 490 ▼x에 위치할 수 있음—490 및/또는 8XX 태그에는 ISSN(▼x)이 있을 수 있으므로 둘 중 하나에서 가져옴

6.4) 각 400/410/411800/810/811 태그에 대해 다음과 같은 내용들을 생성함

Work	hasSeries Work		
bf:contribution/bf:Contribution	agent construct	See Process 1.3	
bf:title/bf:Title	bf:maintitle	See Process 2	
rdfs:label	All subfields less ▼v, ▼w, ▼x, and ▼0-8, separated by space, keep order in field		
identifiedBy	Issn	rdf:value	▼x or 490 ▼x (if necessary)
identifiedBy	Lccn	rdf:value	8xx-w if ▼w begins with (DLC)
Also add ## - assigner - Agent - http://id.loc.gov/vocabulary/organizations/dlc			
identifiedBy	Identifier	rdf:value	8xx-w if ▼w does not begin with (DLC)
Also add ## - assigner - Agent - bf:code			
			▼w data in parentheses
If ▼3 exists, use Subfield ▼3 Process			

6.5) 각 440 또는 830 태그에 대해 다음과 같은 내용들을 생성함

Work	hasSeries	Work		
	bf:title/bf>Title	bf:maintitle		See Process 2
	rdfs:label		All subfields less ▼v, ▼w, ▼x, and ▼0-8, separated by space,	
			keep order in field	
	identifiedBy	Issn	rdf:value	▼x or 490 ▼x (if necessary)
	identifiedBy	Lccn	rdf:value	8xx-w if ▼w begins with (DLC ⁸)
	Also add ## - assigner - Agent -	http://id.loc.gov/vocabulary/organizations/dlc		
	identifiedBy	Identifier	rdf:value	8xx-w if ▼w does not begin with (DLC)
	Also add ## - assigner - Agent -	bf:code	▼w data in parentheses	
		If ▼3 exists, use	Subfield ▼3 Process	

6.6) MARC 레코드에 ▼x(ISSN)이 있는 490 태그가 있고, 해당하는 8XX 태그는 없는 경우 6.5단계에 설명된 내용과 같이 Work - hasSeries - Work Structure 구조를 만들

6.7) MARC 레코드에 8XX 태그만 있는 경우, 인스턴스에 다음을 추가함

■ 프로세스 7-구두점 유지 및 제거

목록화된 서지 레코드에 포함된 구두점을 다루는 것은 과학적으로 정밀하지 못함. 일반적으로 rdfs:label 요소의 경우 구두점은 원자료(source)에 있는 그대로 유지됨. 다른 데이터 요소의 경우 끝 구두점이 제거되거나, 제거되지 않을 수 있음. 해당되는 경우, 변환은 PCC Guidelines for Minimally Punctuated MARC Records(MARC레코드의 최소 구두점 사용에 대한 PCC 가이드라인)⁹⁾을 따름. 구체적인 지침은 다음에 설명함.

7.1) 단일 MARC 필드가 여러 BIBFRAME 속성이나 클래스로 분할되는 경우 각 하위 필드 사이의 구두점을 제거함

[예시]

245 10 ▼a Title of resource : ▼b subtitle of resource / ▼c author of resource.
bf:title [a bf>Title ; bf:mainTitle "Title of resource" ; bf:subtitle "subtitle of resource"] . bf:responsibilityStatement "author of resource" ;

8) 440은 ▼w를 사용하지 않고, 830에는 ▼w가 있는데 해당 내용은 '서지레코드 제어번호(Bibliographic record number)'라 LC에서 관련 사례로 해당 코드를 넣어둔 것입니다. 우리나라의 경우 삭제 적용해도 무방합니다.

9) <https://loc.gov/aba/pcc/documents/PCC-Guidelines-for-Minimally-Punctuated-MARC-Data-v.1.2.docx>

264 1 ▼a New York : ▼b HarperCollins, ▼c 2020.

```
bf:provisionActivity [ a bf:Publication ;  
    bf:agent [ a bf:Agent ;  
        rdfs:label "HarperCollins" ] ;  
    bf:date "2020" ;  
    bf:place [ a bf:Place ;  
        rdfs:label "New York" ] ] ;
```

300 ▼a 50 pages ; ▼c 24 cm.

```
bf:extent [ a bf:Extent ;  
    rdfs:label "50 pages" ] ;  
bf:dimensions "24 cm." ;
```

700 10 ▼a Miranda, Lin-Manuel, ▼d 1980-, ▼e composer, ▼e librettist, ▼e performer.

```
bf:agent [ a bf:Agent ;  
    rdfs:label "Miranda, Lin-Manuel, 1980-" ;  
    bf:role [ a bf:Role ;  
        rdfs:label "performer" ],  
    [ a bf:Role ;  
        rdfs:label "librettist" ],  
    [ a bf:Role ;  
        rdfs:label "composer" ],
```

7.2) MARC 필드가 단일 BF 속성¹⁰⁾에 배치되는 경우, 각 하위 필드 사이에 구두점을 유지함

[예시]

264 1 ▼a New York : ▼b HarperCollins, ▼c 2020.

```
bf:provisionActivityStatement "New York : HarperCollins, 2020" ;
```

490 0 ▼a IEEE digital and mobile communications series ; ▼v 1

```
bf:seriesStatement "IEEE digital and mobile communications series ; 1" ;
```

10) 한 필드 내에 식별기호로 구분되는 내용들이 여러 개의 비브프레임(BF) 속성으로 나뉘어 변환되는 경우들(7.2의 바로 위 사례처럼 700 필드 내 데이터를 변환할 때 bf:agent, bf:role 등으로 나뉘어서 변환하지요)도 있으나, 여기에서는 예시에 보는 것처럼 264 및 490의 각 식별기호의 데이터 내용을 단일 비브프레임 속성(bf:provisionActivityStatement / bf:seriesStatement)에 구분해 넣고 있기 때문에 '단일한 비브프레임 속성'으로 번역하였습니다.

7.2.1) MARC 필드 245 ▼c 에서 bf:responsibilityStatement를 생성할 때 텍스트 문자열 내에서 구두점을 유지함

[예시]

245 10 ▼a Salt, fat, acid, heat : ▼b mastering the elements of good cooking / ▼c by Samin Nosrat and art by Wendy MacNaughton ; with a foreword by Michael Pollan.
bf:responsibilityStatement "by Samin Nosrat and art by Wendy MacNaughton ; with a foreword by Michael Pollan " ;

7.2.2) 여러 문장이 포함된 메모¹¹⁾ 필드를 변환할 때 구두점을 유지함

[예시]

351 ▼a Part I arranged in three series. Series 1: 1 Early Family Papers, 1762-1910; Series 2: General Correspondence, 1934-1964; Series 3: Special Correspondence, 1946-1963.
bf:collectionArrangement [a bf:CollectionArrangement ; bf:collectionOrganization "Part I arranged in three series. Series 1: 1 Early Family Papers, 1762-1910; Series 2: General Correspondence, 1934-1964; Series 3: Special Correspondence, 1946-1963."] ;

506 ▼a Use digital image. Original served only by appointment because material requires special handling.
bf:usageAndAccess [a bf:AccessPolicy ; rdfs:label "Use digital image. Original served only by appointment because material requires special handling."] ;

7.2.3) MARC 필드 210 ▼b 및 220 ▼b에서 bf:AbbreviatedTitle 또는 bf:KeyTitle에 대한 bf:qualifier를 만들 때 부가적 식별정보 주위에 괄호를 유지함

[예시]

210 0 ▼a Br. j. hist. math. ▼b (Online)
bf:title [a bf:AbbreviatedTitle ; bf:mainTitle "Br. j. hist. math." ; bf:qualifier "(Online)"] .

11) 원문에는 "note fields"로 표기되어 있습니다. 이를 "주기" 필드라고 하면 실제 주기사항을 나타내는 필드명과 혼란이 있을 수 있어 메모라는 표현을 사용했습니다.

222 0 ▼a Economic education bulletin ▼b (Great Barrington)

```
bf:title [ a bf:KeyTitle ;  
          bf:mainTitle "Economic education bulletin" ;  
          bf:qualifier "(Great Barrington)" ] .
```

7.3) 대괄호는 목록 작성자가 제공한 메타데이터와 기록되지 않는(non-transcribed) 메타데이터를 나타내는 데 사용되므로 유지해야 함

[예시]

245 10 ▼a [Charles and Ray Eames, designers, looking at the chair at their office] / ▼c [edited by] Hanna Garth and Ashanté M. Reese.

```
bf:title [ a bf>Title ;  
          bf:mainTitle "[Charles and Ray Eames, designers, looking at the chair at their office]" ]  
.  
bf:responsibilityStatement "[edited by] Hanna Garth and Ashanté M. Reese" ;
```

250 ▼a [Explicit edition].

```
bf:editionStatement "[Explicit edition]" ;
```

264 2 ▼a [New York] : ▼b [Buena Vista Home Entertainment], ▼c 2019.

```
bf:provisionActivity [ a bf:Distribution ;  
                      bf:agent [ a bf:Agent ;  
                                rdfs:label "[Buena Vista Home Entertainment]" ] ;  
                      bf:date "2019" ;  
                      bf:place [ a bf:Place ;  
                                rdfs:label "[New York]" ] ] ;
```

7.4) 마지막 구두점(보통 마침표)은 대부분의 01X-4XX 및 65X-8XX 필드에서 제거되어야 하지만 필드 210, 222, 300 및 5XX 필드에서는 유지됨

[예시]

250 ▼a First edition.

```
bf:editionStatement "First edition" ;
```

264 1 ▼a New York : ▼b HarperCollins, ▼c 2020.

bf:provisionActivityStatement "New York : HarperCollins, 2020" ;

500 ▼a Includes index.

bf:note [a bf:Note ;
rdfs:label "Includes index."] ;

650 0 ▼a Baseball players ▼v Drama.

bf:subject [a bf:Topic ;
rdfs:label "Baseball players--Drama"] ;

655 7 ▼a Motion pictures.

bf:genreForm [a bf:GenreForm ;
rdfs:label "Motion pictures"] ;

5.3 숫자로 된 식별기호의 변환¹²⁾

■ 식별기호 ▼0

1.1) ▼0이 URI를 포함하지 않는 경우

12) <https://www.loc.gov/bibframe/mtbf/ConvSpec-NumericSubfields-v1.7.docx>파일을 번역하였으며, KORMARC 필드와의 일대일 확인을 통해 MARC21과 차이가 있을 수 있는 부분에 대해서는 별도 주석으로 처리함

```

▼0(DE-101c)DE-9406763
▼0(isni)1234567899999799

    bf:identifiedBy      [      a      Identifier ;
      rdf:value          "DE-9406763" ;
      bf:source          [bf:code      "DE-101c" ] ]

    bf:identifiedBy      [      a      Identifier ;
      rdf:value          "1234567899999799" ;
      bf:source          [bf:code      "isni" ] ]

혹은 일반적으로 다음과 같이 표시:

    bf:identifiedBy      [      a      Identifier ;
      rdf:value          "alpha/number string" ;
      bf:source          [bf:code      "content of (...) before identifier" ] ]

```

1.2) ▼0이 URI를 포함하는 경우

```

▼0http://id.loc.gov/authorities/names/n82070361
▼0(uri)http://id.loc.gov/authorities/names/n82070361 (drop "(uri)")
▼0http://rdaregistry.info/termList/RDAproductionMethod/1017

    bf:<property as defined in spec> <URI> ] ;
    <URI> a bf:<class as defined in spec>
    rdfs:label "text from field or subfield"

bf:agent    <http://id.loc.gov/authorities/names/n82070361> ] ;
    <http://id.loc.gov/authorities/names/n82070361>
    a bf:Person ;
    rdfs:label "Villa, Susie Hoogasian" .

    bf:baseMaterial <http://ii.loc.gov/vocabulary/material/pap> ] ;
    <http://id.loc.gov/vocabulary/mmmaterial/pap> a bf:BaseMaterial ;
    rdfs:label "paper" .

```

1.3) 한 필드에 여러 개의 ▼0이 있는 경우, ▼0이 식별하는 하위 필드의 바로 다음에 오는 것으로 가정하고 1.2 단계의 예시를 따름:

```

▼a Term 1 ▼0 URI for Term 1 ▼b Term 2 ▼0 URI for Term 2 ▼c Term 3

```

필드의 끝이나 시작 부분에 여러 개의 ▼0이 함께 나타나는 경우, 1.1.단계의 예시를 따름:

▼a Term 1 ▼b Term 2 ▼c Term 3 ▼0 URI ▼0 URI
 ▼0 URI ▼0 URI ▼a Term 1 ▼b Term 2 ▼c Term 3

■ 식별기호 ▼2

	bf:source	a	Source	[bf:code "content of ▼2"]
or	bf:source	<http://id.loc.gov/vocabulary/.../<content of▼2>>		

■ 식별기호 ▼3

KORMARC의 식별기호 ▼3은 자료 범위를 지정함

- 자료의 속성을 지정하지만, 해당 속성은 자료의 일부분에만 적용되며, “지정된 자료”가 어느 부분에 속하는지를 알려줌
- 단어, 구, 낱자 등을 나열함. 예를 들면: 서신, 일기, 말(馬), 사본(dup., duplicate), 음화(neg., negative), 사건 파일 등)

다음과 같이 사용할 수 있음

- 일기, 기록물, 노트, 연례보고서 등의 아카이브 컬렉션의 부분을 나타내기 위해 사용될 수 있음
- 키트(kits), 조각상, 연구자료를 비롯해 퍼즐, 말(馬), 등록된 연구 데이터 및 3차 저작물(3rd work)과 같이 여러 다른 작업을 포함하고 있는 저작물 등 복합 자료의 일부를 표시하기 위해 사용될 수 있음
- 특히 연속간행물 자료에 적용되는쇄(刷, imprint)와 같은 속성이 적용될 때 여러 낱자나 열거형(예: 1980, 1927-42, 1-19, (1930-49))을 나타내기 위해 사용될 수 있음
- 탈산성(deacidified) 복사자료, 사본, 음화와 같은 보존 정보의 식별을 위해 사용될 수 있음

bflc:appliesTo 속성과 bflc:AppliesTo 클래스를 사용함

적용예시:

3.1) R1, R2, R3 부분들로 구성된 자원 R이 있음. 각 부분 모두 개별 URI로 주소를 지정할 수 있고, R2 및 R3에만 적용되는 주석이 있는 경우 다음과 같이 변환함.

```

<resource> bf:note [ a bf:Note ;
                    rdfs:label "text of note" ;
                    bflc:appliesTo URI or [rdfs:label ...] for R2 ;
                    bflc:appliesTo URI or [rdfs:label ...] for R3 ] .

```

3.2) 각 부분들에 대한 URI는 없지만, 머리말(foreward) 부분에만 주석이 있는 경우 다음과 같이 변환함.

이 사항은 (주석과 같은) 객체 속성에 대해서만 작용함.

```

<resource> bf:note [ a bf:Note ;
                    rdfs:label "foreword by Thelma Plews" ;
                    bflc:appliesTo [rdfs:label "foreword" ] .

```

3.3) bf:duration(예, bf:duration "10분")과 같은 '비공식적인' 주석이 있는 경우, 해당 내용을 주석에 기재함.

```

<resource> bf:note [ a bf:Note ;
                    bf:duration "10 minutes" ;
                    bflc:appliesTo [rdfs:label "first movement" ] .

```

MARC 적용 사례:

[037 - 입수처(SOURCE OF ACQUISITION)]

```

037 ##▼3-2013▼bPortico
037 ##▼32014▼bOxford University Press

bf:acquisitionSource [
    a bf:AcquisitionSource ;
    rdfs:label "Portico" ;
    bflc:appliesTo [rdfs:label "-2013" ] ] .

bf:acquisitionSource [
    a bf:AcquisitionSource ;
    rdfs:label "Oxford University press" ;
    bflc:appliesTo [rdfs:label "2014" ] ] .

```

※ MARC21 기준 사례에는 제1지시기호가 있는 037 3#의 사례로 명시되어 있으나, KORMARC의 037은 제1, 2지시기호 모두를 정의하지 않아 037 ##로 수정해 정리함

[260 - 발행, 배포, 간사사항 (PUBLICATION, DISTRIBUTION, ETC. (IMPRINT))]

해당 필드가 적용된 자료의 여러 서술 내용들을 구별하기 위한 정보임.

```

260 2#▼31980-May 1993▼aLondon : ▼bVogue
260 3#▼3June 1993-▼aLondon : ▼bElle
260 2#▼32000?-2002▼aNew York : ▼bColumbia University Press

bf:provisionActivity [
    a      bf:Distribution ;
    bf:agent      [rdfs:label "Vogue" ] ;
    bf:place      [rdfs:label "London" ] ;
    bf:appliesTo [rdfs:label "1980-May 1993" ] ] .

bf:provisionActivity [
    a      bf:Distribution ;
    bf:agent      [rdfs:label "Elle" ] ;
    bf:place      [rdfs:label "London" ] ;
    bf:appliesTo [rdfs:label "3June 1993-" ] ] .

bf:provisionActivity [
    a      bf:Distribution ;
    bf:agent      [rdfs:label "Columbia University Press" ] ;
    bf:place      [rdfs:label "New York" ] ;
    bf:appliesTo [rdfs:label "2000?-2002" ] ] .

```

[300 - 형태사항(PHYSICAL DESCRIPTION)]

```

300 ##▼3poems▼a1▼fpage ; ▼c108 cm. x 34.5 cm

<instance> bf:extent [
    a      bf:Extent ;
    rdf:value      "1" ;
    bf:unit      [rdfs:label "page" ] ;
    bf:dimensions      "108 cm. x 34.5 cm" ;
    bf:appliesTo      [rdfs:label "poems" ] ] .

```

※ 주어진 예시에서의 자료 수량은 1, 단위는 page, 크기는 c108 cm. x 34.5 cm이며, 시(poems)에 적용됨

[337 - 매체유형(MEDIA TYPE)]

```

337 ##▼aunmediated▼2rdamedia▼3liner notes

<resource> bf:media [
    a      bf:Media ;
    rdfs:label      "unmediated" ;
    bf:appliesTo      [rdfs:label "liner notes" ] ] .

```

[351 - 자료의 구조와 배열(ORGANIZATION AND ARRANGEMENT OF MATERIALS)]

351	##▼3Diaries and notebooks▼aOrganized into four series: I. Youth, 1846-1852. II. Early Career,1853-1865. III. Political Life, 1866-1895. IV. Retirement, 1896-1903;▼bChronological arrangement.
bf:arrangement	[
a	bf:Arrangement ;
bf:pattern	"Chronological arrangement" ;
bf:organization	"Organized into four series: I. Youth, 1846-1852. II. Early Career,1853-1865. III. Political Life, 1866-1895. IV. Retirement, 1896-1903" ;
bflc:appliesTo	[rdfs:label "Diaries and notebooks"]] .

[490 - 총서사항(SERIES STATEMENT)]

490	1#▼31972/73-1975-76: ▼aResearch report / National Education Association Research
<resource> bf:note	[
a	bf:Note ;
bf:seriesStatement	"Research report / National Education Association Research" ;
bflc:appliesTo	[rdfs:label "1972/73-1975-76"]] .

■ 식별기호 ▼4

프로세스 노트 0.2와 1.4를 참조할 것

■ 식별기호 ▼5

식별기호 ▼5는 형식 전체에서 동일한 정의로 사용됨

▼5 - 필드 적용 기관

해당 필드의 데이터가 적용되는 사본을 보유한 기관 또는 조직의 MARC 코드. 필드 데이터는 개별자료(item)에 대한 보편적인 설명에 적용되지 않거나, 개별자료에 보편적으로 적용될 수는 있지만 인용된 위치에만 관련이 있을 수 있음. MARC21 기록에 사용되는 기관 목록은 MARC Code List for Organizations를 참조하며, KORMARC의 경우는 '부록 A. 제어 식별기호(https://librarian.nl.go.kr/kormarc/kormarc_2014/sub/appendix_a.html)'를 기준으로 함

변환 작업 중 ▼5의 기관 코드가 DLC가 아닌 경우에는 필드를 변환하지 말 것¹³⁾

▼5가 DLC인 경우, 예 설명된 대로 변환하되 050 00 필드(미국국회도서관 청구기호)를 변환할 때 생성된 BIBFRAME 항목에 배치해 생성함. BIBFRAME 항목이 없는 경우, 새로 생성함. 두 경우 모두 하단의 설명에 나와 있는 것과 같이 bflc:applicableInstitution/bf:Agent/URI 정보를 추가함.

MARC 적용 사례:

[246 - 여러 형태의 표제(VARYING FORM OF TITLE)]

246	3#▼a<title of reissue>;▼5DLC
<pre> <resource> bf:title [a bf:VariantTitle ; rdfs:label "title of reissue" ; bflc: applicableInstitution [a bf:Agent ; bf:code "DLC"]]. </pre>	

※ 자료의 재발행에 따라 나타나는 변형된 제목

[655 - 색인어--장르/형식(INDEX TERM--GENRE/FORM)]

655	#7▼aAnnotations (Provenance)▼zSweden▼y18th century.▼2rbprov▼5DLC
<pre> <resource> bf:genreForm [a bf:GenreForm ; rdfs:label "Annotations (Provenance)" ; bf:place [rdfs:label "Sweden"] ; bf:date "18th century" ; bflc:applicableInstitution [a bf:Agent ; bf:code "DLC"]] . </pre>	

※ 희귀도서 규칙에 따라 자료에 부여된 형식/장르 용어

[700 - 부출표목--개인명(ADDED ENTRY--PERSONAL NAME)]

700	1#▼aJefferson, Thomas, ▼d1743-1826, ▼eformer owner.▼5DLC
<pre> <resource> bf:contribution [a bf:Contribution; bf:agent [a bf:Person; rdfs:label "Jefferson, Thomas, 1743-1826"] ; bf:role [a Role rdfs:label "former owner"] ; bflc:applicableInstitution [bf:code "DLC"]]. </pre>	

※ 도서관의 도서자료 사본이 이전에 Thomas Jefferson의 소유였기 때문에 기록에 해당 항목이 추가됨

13) KORMARC의 경우는 향후에 국립중앙도서관 관련 코드 변환에 대한 내용으로 적용할지 여부 검토가 필요한 항목입니다.

■ 식별기호 ▼6

MARC 태그 010-856 데이터는 880 태그와 쌍을 이루어 다른 문자 집합의 메타데이터를 나타낼 수 있음. 각각의 태그들은 각 태그 내 ▼6 코딩을 통해 함께 연결됨.

```
245 10 ▼6880-01▼aMinguo shi qi Guangdong xue ren yu Zhongguo xi nan yan jiu
880 10 ▼6245-01/(B▼a民国时期广东学人与中国西南研究
```

880 필드는 일반적으로 MARC 레코드의 일반 데이터 필드의 지시기호 번호로 연결되지만 '00' 번호가 적용되는 경우 '그 쌍이 해제(unpaired)'될 수 있음. 예를 들면:

```
880 ▼6 500-00/▼1 ▼a 北宋刻本 一捲一木匣 行二十七字.
```

6.1) 일부 MARC 태그(주로 식별자나 승인된 접근점)의 경우, MARC 태그 880의 연결 데이터는 변환하지 않으며, 각각의 개별 변환 에서 '무시함(ignore)'으로 표시됨.

6.2) 하단에 나열된 필드 이외의 MARC 태그의 경우, MARC 태그 880의 데이터만 변환되고 MARC 태그 010-856의 연결된 쌍 데이터(paired data)는 무시됨. 이는 개별 변환 에 “식별기호 ▼6.2를 보라(See Subfield ▼6.2 Spec)’라는 표기로 표시됨. 이러한 필드의 경우 각 에 나타나는 표시 및 해당 식별기호에 대한 변환 지침을 따르며, ▼6에 지정된 스크립트 표시를 포함함.

```
bf:tableOfContents [ a bf:TableOfContents ;
  rdfs:label "1. 대한민국 임시 정부 수립 및 통합의 역사와 그 현재적 함의 / 오 일환 -- 2. 기독교 문명과 상해 프랑스 조계 대한민국 임시 정부 / 김 명섭 -- 3. 대한민국 임시 정부 수립 과정 및 이후 중국 내 기독교 독립 운동가의 활동 / 김 명배 -- 4. 대한민국 임시 정부와 만주, 연해주 지역 독립 운동 / 박 환 -- 5. 대한민국 임시 정부와 국내 독립 운동 / 윤 정란 -- 6. 해방 직전 이 승만과 기독교인 친한회(The Christian Friends of Korea)의 대한민국 임시 정부 승인 / 박 명수 -- 7. 북음통일을 위한 대한민국 임시 정부사의 기독교적 함의 / 허 문영."@ko-hang ] ;
```

6.3) 필드 245, 250, 26X 및 490에 대한 특별 지침

필드 245, 250, 26X 및 490의 경우, 쌍을 이루는 MARC 태그는 두 문자 집합의 메타데이터에 대해 단일한 트리플로 결합되어야 함. 이는 개별 변환 에 “식별기호 ▼6.3을 보라(See Subfield ▼6.3 Spec)’라는 표기로 표시됨. 예를 들면:

```

<bf:title>
  <bf:Title>
    <bf:mainTitle>국내 증권업 자본 규제 평가 및 시사점</bf:mainTitle>
    <bf:mainTitle xml:lang="ko-Latn">Kungnae chüngkwönöp chabon kyuje p'yöngka mit sisachö
    m</bf:mainTitle>
    <rdfs:label>국내 증권업 자본 규제 평가 및 시사점</rdfs:label>
  </bf:Title>
</bf:title>

<bf:editionStatement>Shohan</bf:editionStatement>
<bf:editionStatement xml:lang="ja-jpan">初版</bf:editionStatement>

<bf:provisionActivity>
  <bf:ProvisionActivity>
    <bf:place>
      <bf:Place>
        <rdfs:label xml:lang="ja-jpan">東京都新宿区</rdfs:label>
      </bf:Place>
    </bf:place>
    <bf:date>2013</bf:date>
    <bf:agent>
      <bf:Agent>
        <rdfs:label xml:lang="ja-jpan">藤原書店</rdfs:label>
      </bf:Agent>
    </bf:agent>
  </bf:ProvisionActivity>
</bf:provisionActivity>

<bf:seriesStatement>Zhongguo li dai hui hua da xi</bf:seriesStatement>
<bf:seriesStatement xml:lang="zh-hani">中國歷代繪畫大系</bf:seriesStatement>

```

▼6의 구조는 MARC 서지 형식 (<http://www.loc.gov/marc/bibliographic/ecbdcntf.html>)의 부록 A에 설명되어 있음. MARC 레코드에서 발견된 스크립트 코드는 해당 부분에서 식별됨.

KORMARC의 경우 ▼6에 사용하는 기호는 KORMARC 부록 A.제어 식별기호 (https://librarian.nl.go.kr/kormarc/kormarc_2014/sub/appendix_a.html)를 기준으로 함.

6.4) 데이터 손실을 방지하기 위해서는 ▼6에 나열된 MARC 태그에 대한 지침에 따라 쌍으로 연결되지 않은(unpaired) 880 태그를 처리해야 함. 짝을 이루지 않는 880

필드 데이터는 항상 MARC에서 BIBFRAME으로 변환되어야 함.

5.4 KORMARC 서지 변환 - MARC21 필드와의 차이점들을 중심으로¹⁴⁾

■ 리더부(Leader)

MARC21의 리더와 비교했을 때 KORMARC과 차이가 있는 세부 항목은 다음 표와 같으며, 각각의 세부 내용에 대해 BIBFRAME으로 변환하는 내용에 대한 기준 마련이 필요함. 일부 필드의 세부 사항 차이에 대해서는 별첨한 부록의 해당 파일(ConvSpec-LDR-v1.7.1-KORMARC)을 참조할 것.

〈표 3〉 리더부 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
06-레코드 유형(Type of record) w - 고서	<ul style="list-style-type: none"> MARC21에 없는 항목으로 '고서' 유형에 대한 표기 추가가 필요함 	
08 - 제어유형(Type of control) b - KS X 1001	<ul style="list-style-type: none"> MARC21에서는 MARC-8으로 되어 있으므로, KORMARC의 해당 항목 변환 시 기호 의미 차이에 대해 고려해야 함 	MARC21에서는 '08-제어유형'에 대해 BIBFRAME 변환 시 '무시함(ignore)'
08 - 제어유형(Type of control) z - KS X 1001, USC/Unicode 이외의 문자부호	<ul style="list-style-type: none"> MARC21에 없는 항목임 	
18 - 목록기술형식(Descriptive cataloging form) c - KCR3판 이상	<ul style="list-style-type: none"> MARC21은 ISBD punctuation omitted(ISBD 구두점 생략 형식)을 이 기호로 사용하므로, KORMARC의 해당 항목 변환 시 기호 의미 차이에 대해 고려해야 함 KCR3판 이상을 표현할 수 있는 어휘 URI 생성 · 추가 필요 	
18 - 목록기술형식(Descriptive cataloging form) k - KORMARC 기술규칙	<ul style="list-style-type: none"> MARC21에 없는 항목임 KORMARC 기술규칙을 표현할 수 있는 어휘 URI 생성 · 추가 필요 	

■ 001-005, 007(제어, 형태기술 필드)

001-005, 007 필드는 KORMARC에서 '미정의'로 추가한 항목 외의 대부분의 세부 항목이 MARC21과 동일한 형태로 존재하고 있으나, '007-고서(Antiquarian)'의 경우는 MARC21에는 없고 KORMARC에만 나타나는 항목으로 해당 내용을 표현할 수 있는 기술규칙과 관련 어휘 URI의 생성과 추가가 필요함. 일부 필드의 세부 사항 차이에 대해서는 별첨한 부록의 해당 파일(ConvSpec-001-007-v1.7.1-KORMARC...)을 참조할 것.

14) <https://www.loc.gov/bibframe/mtbf/>에서 제공하고 있는 MS Excel 변환 파일을 바탕으로 KORMARC 각 필드와의 비교를 통해 정리하였으며, 각 파일의 세부 내용을 일대일로 비교해 정리한 파일들은 부록으로 별첨하였습니다. MARC21에만 있고 KORMARC에는 없는 내용은 해설서 내의 차이점 설명에 포함하지 않았고, 부록 파일에 취소선 등으로 구분할 수 있도록 정리하였습니다.

〈표 4〉 001-005, 007(제어, 형태기술 필드) 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
007--고서(ANTIQUARIAN) 00 - 자료범주표시(Category of material) o - 고서(Antiquarian) 01 - 특정자료종별(Specific material designation) a - 권속장 b - 절첩장 c - 호접장 d - 포배장 e - 선장 f - 족자 g - 접포 h - 모장 i - 날장 u - 특정화하지 않음(Unspecified) z - 기타 - 부호화하지 않음	<ul style="list-style-type: none"> MARC21에 없는 항목임 007/00 부호 o일 때 고서임을 식별하는 물리적 특성에 관한 특정의 부호화된 정보를 기술해야 함 고서 필드는 KORMARC에만 있는 필드이므로, 이를 표현할 수 있는 기술규칙 및 어휘 URI 생성·추가 필요 	

■ 006, 008(고정길이 필드, 부호화정보 필드(일반, 부가적 자료 특성 등))

006(고정길이 데이터 요소-부호화정보 필드-부가적 자료 특성)의 경우, MARC21에서는 006/01-17은 008/18-34처럼 변환한다고 명시되어 있으나, KORMARC의 경우에는 006 필드가 00-13의 14자리로 구성되어 있어, KORMARC 006 통합서지용 안내에 적시된 설명으로 수정 적용 해야 함. 또한, 부가적 자료 특성 기호 중 00-자료형태(Form of material)에 있어 MARC21에는 명시되어 있지 않은 'w-고서'에 대한 항목 추가가 필요함.

〈표 5〉 006 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
006 - 고정길이 데이터 요소(부호화정보필드-부가적 자료 특성(반복, 재량))	<ul style="list-style-type: none"> MARC21에서 변환에 명시한 006-01-17 자리수와 KORMARC의 006/01-13 자리수가 달라 KORMARC의 해설서 내용 기준으로 정리해야 함 006/01-13 자리에는 008/18-25, 29-31, 33-34 자리의 내용을 표시하도록 함 	
006 - 고정길이 데이터 요소(부호화정보필드-부가적 자료 특성(반복, 재량)) 00 - 자료형태(Form of material) w - 고서	<ul style="list-style-type: none"> MARC21에 없는 항목임 고서 필드는 KORMARC에만 있는 필드이므로, 이를 표현할 수 있는 기술규칙 및 어휘 URI 생성·추가 필요 	MARC21 to BIBFRAME에서는 '00-자료형태' 내용은 nac (변환시도 없음)으로 처리함

008(고정길이 데이터 요소-부호화정보필드-일반 정보)의 경우, 일부 내용에 있어

KORMARC에서만 정의해 사용하는 항목 및 MARC21의 기호와 다른 의미를 부여한 항목들이 존재하고 있음. 이와 관련해 BIBFRAME으로 변환해 표현할 수 있는 기술규칙 및 어휘 URI 생성 등이 필요함. 일부 필드의 세부 사항 차이에 대해서는 별첨한 부록의 해당 파일(ConvSpec-006-008-v1.7.1-KORMARC)을 참조할 것.

〈표 6〉 008 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
008—공통항목(ALL MATERIALS) 18-25 특정자료종별	<ul style="list-style-type: none"> 008/18-25 부문 중 해당부분 참조 	
008—공통항목(ALL MATERIALS) 26-27 - 한국대학부호(고서는 '미정의')	<ul style="list-style-type: none"> MARC21에 없는 항목임 한국대학부호는 2자리로 표시하며 대학에서 간행된 자료에 대해 기술하고, 대학간행물이 아닌 경우 빈칸(b)으로 채움 (https://librarian.nl.go.kr/kormarc/kormarc_2014/sub/annex_2.html) 한국대학부호와 관련한 내용을 표현할 수 있는 어휘 URI 생성 · 추가 필요 	
008—공통항목(ALL MATERIALS) 28 - 수정레코드(Modified record) b - 변형없음(Not modified) b - 문자를 빈칸(b)으로 입력한 경우 h - 한자를 한글로 입력한 경우 r - 로마자 이외의 문자를 로마자로 입력한 경우 s - 레코드의 단축 x - 한자 이외의 문자를 한글로 변형 입력한 경우' - 부호화하지 않음(No attempt to code)	<ul style="list-style-type: none"> MARC21에 없는 항목임 수정레코드는 1자리로 표시하며, 레코드 생성 시 서지정보 일부가 단축되거나 입력 불가능한 기호 등이 발생한 경우 등의 유형을 나타냄 	MARC21에서는 이 항목에 대해 BIBFRAME 변환 시 '무시함(ignore)'
008—공통항목(ALL MATERIALS) 32 - 목록진거(Cataloging source) b - 국가서지작성 기관(National bibliographic agency) a - 자관 c - 협력기관(Cooperative cataloging program) d - 기타 기관(Other) u - 미상(Unknown) - 부호화하지 않음(No attempt to code)	<ul style="list-style-type: none"> MARC21에 없는 항목임 	MARC21 to BIBFRAME에서는 nac(변환시도 없음)으로 처리함
008—공통항목(ALL MATERIALS) 38-39 - 한국정부기관 부호(고서는 '미정의')	<ul style="list-style-type: none"> MARC21에 없는 항목임 한국정부기관부호는 2자리로 표시하며 한국정부기관에서 발행한 자료에 대해 기술하고, 해당 사항이 없으면 빈칸(b)으로 채움 (https://librarian.nl.go.kr/kormarc/kormarc_2014/sub/annex_4.html) 한국정부기관부호와 관련한 내용을 표현할 수 있는 어휘 URI 생성 · 추가 필요 	
008—도서(BOOKS), 전자자료(COMPUTER FILES), 음악/녹음자료(MUSIC), 시청각자료(VISUAL MATERIALS) 22 - 이용대상자 수준(Target audience) b - 일반이용자용 (General)	<ul style="list-style-type: none"> MARC21과 이용대상자 수준의 분류가 다소 차이가 있음 일반이용자, 중학생, 고등학생에 대한 이용대상자 수준의 분류와 의미가 달라 차이가 있는 이용대상자를 표시할 수 있는 어휘 URI와 rdfs:label이 필요 	일반이용자의 경우 KORMARC에서는 b, MARC21에서는 g 기호로 사용)

c - 중학생 d - 고등학생 z - 미상 또는 세분하지 않는 경우(Unknown or not specified)		
008—도서(BOOKS), 계속자료 (CONTINUING RESOURCES) 24-25 - 내용형식(Nature of contents) j - 족보	<ul style="list-style-type: none"> • 006/07-08 • MARC21은 24-27로 되어 있음 • j기호의 경우 MARC21은 특허자료로 되어 있어, 장르 유형 중 '족보'를 표현할 수 있는 기술규칙 및 어휘 URI 생성 · 추가 필요 	
008—도서(BOOKS) 33 - 문학형식(Literary form) k - 추리소설 t - 논설문 u - 평론 v - 문집 w - 향가, 시조, 가사, 구비문학	<ul style="list-style-type: none"> • 문학형식 중 추리소설, 논설문, 평론, 문집, 향가 및 시조 등의 기호는 MARC21에 없는 항목임 • KORMARC에만 있는 문학형식들의 내용을 표현할 수 있는 기술규칙 및 어휘 URI 생성 · 추가 필요 	KORMARC에서 평론을 의미하는 u 기호는 MARC21에서는 알려지지 않은 형태(unknown)를 표시하는 기호임
008—음악/녹음자료(MUSIC) 18-19 - 작곡형식(Form of composition) ka - 제레악 kb - 연례악 kc - 거동음악 kd - 가사 ke - 시조 kf - 판소리 kg - 단가 kh - 병창 ki - 통속민요 kj - 잡가 kk - 토속민요 kl - 산조 km - 시나위 kn - 봉장취 ko - 농악 kp - 향제 대풍류 kq - 삼현육각 kr - 불교음악 ks - 무속음악 kt - 연희 ku - 한국가곡 kv - 한국대중음악	<ul style="list-style-type: none"> • 음악/녹음자료에 대한 작곡형식 중 MARC21에 없고 KORMARC에만 존재하는 작곡형식들이 다수 존재함 • 음악/녹음자료에 대한 작곡형식 유형 중 각각의 형식에 해당하는 내용을 표시할 수 있는 URI와 rdfs:label 내용 추가 필요 	
008—음악/녹음자료(MUSIC) 20 - 악보형식(Format of music) h - 국악보 9 - 합창보(Chorus score)	<ul style="list-style-type: none"> • '국악보'는 MARC21에 없는 내용임 • 음악/녹음자료에 대한 악보형식 중 국악보에 해당하는 내용을 표시할 수 있는 URI와 rdfs:label 내용 추가 필요 • MARC21에서는 h가 합창보로 되어 있으나, KORMARC에서는 h를 국악보로 부여하고, 합창보에는 9 기호를 부여함 	
008—고서 18 - 삽화표시 b - 해당 없음 a - 삽화 b - 지도 c - 초상화(또는 인물사진) d - 표 e - 설계도 f - 도판 g - 악보 h - 영인물 i - 문장(紋章)	<ul style="list-style-type: none"> • 008-고서 자료에 대한 내용은 MARC21에 없고 KORMARC에만 존재함 • 각 세부 항목에 대한 내용을 표시할 수 있는 기술규칙 및 어휘 URI 생성 · 추가 필요 	

<p>j - 계보 k - 서식 및 양식 l - 견본(실물견본) o - 사진 p - 채색장식(彩色裝飾) s - 변상도(變相圖) t - 산도(山圖), 묘도(墓圖) - 부호화하지 않음</p> <p>19 - 광곽 b - 적용 안함 a - 사주단변(四周單邊) b - 사주쌍변(四周雙邊) c - 상하단변(上下單邊) 좌우쌍변(左右雙邊) d - 상하쌍변(上下雙邊) 좌우단변(左右單邊) e - 상하단변(上下單邊) f - 상하쌍변(上下雙邊) g - 사주무변(四周無邊) - 부호화하지 않음</p> <p>20 - 사경 및 사본의 계선 b - 적용 안함 a - 오사란(烏絲欄) b - 주사란(朱絲欄) c - 남사란(藍絲欄) d - 금계(金界) e - 은계(銀界) f - 백계(白界) g - 기타 - 부호화하지 않음</p> <p>21 - 판구 b - 적용 안함 a - 대흑구(大黑口) b - 소흑구(小黑口) c - 백구(白口) d - 화구(花口) e - 기타 - 부호화하지 않음</p> <p>22 - 어미 b - 적용 안함 a - 上 또는 下 흑어미(黑魚尾) b - 上 또는 下 백어미(白魚尾) c - 上 또는 下 2엽화문어미(葉花紋魚尾) d - 上 또는 下 3엽화문어미(葉花紋魚尾) e - 상하흑어미 (上下黑魚尾) f - 상하백어미 (上下白魚尾) g - 상하2엽화문 어미(上下2葉花紋魚尾) h - 상하3엽화문 어미(上下3葉花紋魚尾) i - 혼엽화문 어미(混葉花紋魚尾) j - 흑혼입화문 어미(黑混入花紋魚尾) k - 기타 - 부호화하지 않음</p> <p>23 - 개별자료 형태 b - 해당없음</p>		
---	--	--

<p>a - 마이크로필름 (Microfilm) b - 마이크로피치 (Microfiche) c - 마이크로오패크 (Microopaque) d - 큰활자인쇄자료(Large print) o - 온라인 전자자료 (Online) q - 직접 접근 전자 자료(Direct electronic) r - 복제자료(Regular print reproduction) s - 전자자료(Electronic) - 부호화하지 않음(No attempt to code) 24-25 - 내용형식(Nature of contents) b - 해당 없음(No specified nature of contents) b - 서지 d - 운서(韻書) e - 법전(法典) g - 서화 h - 금석문(金石文) i - 방목(榜目), 관안(官案) j - 족보(族譜), 계보(系譜), 세보(世譜) m - 불전(佛典) o - 주석(註釋), 평주(評註), 해제 p - 조약집(條約集), 외교문서 s - 통계 t - 경전(經典) v - 지지(地誌) w - 지도(地圖) x - 고문서(古文書) - 부호화하지 않음(No attempt to code) 29-31 - 판종(版種) A00-Z99 판종유형 및 내용 b - 적용안함 - 부호화하지 않음(No attempt to code) 33 - 문학형식 b - 해당없음 e - 文(산문, 수필, 잡문) f - 소설 i - 서간 m - 일기, 기행 v - 문집(文集) w - 향가, 시조, 가사, 구비문학 - 부호화하지 않음(No attempt to code) 34 - 전기 b - 해당 없음 a - 자서전 b - 개인전기서 c - 전기물의 합저서 d - 전기적 정보가 포함된 자료 - 부호화하지 않음(No attempt to code)</p>		
--	--	--

■ 010-049(식별자 등)

010-049(식별자 등)는 도서관 제어번호, 서지번호, 각종 등록번호 등의 식별자와 관련된 정보를 포함하는 항목으로, KORMARC에는 MARC21에는 사용하지 않는 012(국립중앙도서관 제어번호), 023(출판예정도서목록제어번호(CIP)), 049(소장사항) 필드가 추가되며, 일부 필드의 세부 사항 차이에 대해서는 별첨한 부록의 해당 파일 (ConvSpec-010-048-v.1.7.1-KORMARC)을 참조할 것.

〈표 7〉 010-049 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
012-국립중앙도서관 제어번호 (The National Library of Korea Control Number) 식별기호 코드 ▼a - 국립중앙도서관 제어번호(반복불가) ▼z - 취소/사용하지 않는 국립중앙도서관 제어번호(반복) ▼8 - 필드 링크와 일련번호(반복)	<ul style="list-style-type: none"> 012는 국립중앙도서관에서 부여하는 필드로, MARC21에 없는 항목임 미국국회도서관 제어번호(Lccn)과 같이 국립중앙도서관제어번호를 표시할 수 있는 기호 및 변환규칙 필요 	
023-출판예정도서목록제어번호(CIP system control number) 식별기호 코드 ▼a - 출판예정도서목록(CIP) 제어번호(반복불가) ▼z - 취소/사용하지 않는 출판예정도서목록(CIP)제어번호(반복) ▼8 - 필드 링크와 일련번호(반복)	<ul style="list-style-type: none"> MARC21에는 없는 항목임 CIP번호를 표시할 수 있는 변환규칙 필요 	
049-소장사항(Local Holdings) 제1지시기호 - 배가위치 0 - 해당자료 전체가 동일한 곳에 배가(별치)되는 경우 1 - 해당자료 전체가 동일한 곳에 배가(별치)되지 않는 경우 제2지시기호 - 미정의(b) 식별기호 코드 ▼a - 소장기관부호(반복불가) ▼l - 등록번호(반복) ▼v - 권·연차기호(반복) ▼c - 복본기호(반복) ▼f - 별치기호(반복)	<ul style="list-style-type: none"> 이 필드는 MARC21이나 다른 나라 MARC와는 다른 형식임 이 필드에는 각 로컬도서관에서 도서관 부호와 소장자료의 등록번호, 권·연차기호, 복본기호, 별치기호 등을 기술함 도서관에 따라 소장사항은 049필드나 090 필드가 아닌 다른 로컬필드(092, 093 등) 또는 소장 MARC의 필드(852 등)를 사용할 수 있음 각 로컬도서관의 소장사항에 대한 내용을 표시할 수 있는 변환규칙 필요 	

■ 050-088, 090(분류 및 청구기호 등)

050-088 및 090은 도서관 자료의 분류 및 청구기호에 대한 사항을 포함하는 항목으로, KORMARC에는 MARC21에는 사용하지 않는 052(국립중앙도서관 청구기호), 056(한국십진분류기호), 066(사용문자세트) 필드가 추가되며, 일부 필드의 경우 MARC21과 다른 기호를 부여해 정의하고 있음. 세부 사항 차이에 대해서는 별첨한 부록의 해당 파일

(ConvSpec-050-088-v.1.7-KORMARC)을 참조할 것.

〈표 8〉 050-088, 090 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
<p>052-국립중앙도서관 청구기호 (National Library of Korea Call Number) 제1지시기호 - 자료의 국립중앙 도서관 소장여부 0 - 국립중앙도서관 소장자료 1 - 국립중앙도서관 미소장자료 2 - 국립중앙도서관 분관 소장자료 3 - 국립중앙도서관 매체변환자료 (CD-ROM) 4 - 국립중앙도서관 매체변환자료(마이크로자료) 제2지시기호 - 적용분류표 0 - KDCP로 분류된 자료 1 - KDC로 분류된 자료 2 - DDC로 분류된 자료 3 - 조선총독부 신서부분류표로 분류된 자료 4 - 조선총독부 양서부분류표로 분류된 자료 5 - 조선총독부 고서부분류표로 분류된 자료 6 - 미국정부문서 분류표로 분류된 자료</p> <p>식별기호 코드 ▼a - 분류기호(반복불가) ▼b - 도서기호(반복불가) ▼c - 권·연차기호(반복불가) ▼6 - 대체문자 연결(반복불가) ▼8 - 필드 링크와 일련번호(반복)</p>	<ul style="list-style-type: none"> 052는 국립중앙도서관에서 부여하는 필드로, MARC21에 없는 항목임 MARC21에서 052는 지리분류기호를 식별하는 필드이며, KORMARC은 지리분류기호 필드를 055로 정의하고 있음 W-classification-Classification 부분에 국립중앙도서관 청구기호 표시에 대한 항목을 추가해야 함 	
<p>055 - 지리 분류기호(GEOGRAPHIC CLASSIFICATION)</p>	<ul style="list-style-type: none"> MARC21에서는 052가 지리분류기호이며, 055는 캐나다에서 부여한 분류기호를 표시하는 필드임 KORMARC에서는 052는 국립중앙도서관 청구기호를, 055에 지리분류기호를 정의하고 있음 W-geographicCoverage-GeographicCoverage로 변환 규칙 적용 필요 	
<p>056 - 한국십진분류기호(Korean Decimal Classification Number) 식별기호 코드 ▼a - 한국십진분류기호(반복) ▼b - 도서기호(반복불가) ▼2 - 판표시(반복불가) ▼8 - 필드 링크와 일련번호(반복)</p>	<ul style="list-style-type: none"> 056은 KORMARC에서만 부여하는 필드로, MARC21에 없는 항목임 분류기호 검색을 위해 청구기호(090 필드)의 분류기호(식별기호 ▼a)로 KDC를 사용하는 경우에도, 청구기호에 기술되는 분류기호를 포함하여 부출되는 KDC 분류기호를 모두 이 부분에 기술함 MARC21에서 DDC 분류기호의 BIBFRAME 변환 규칙에 맞춰 W-classification-ClassificationKdc의 형태로 국립중앙도서관 청구기호 표시에 대한 항목을 추가해야 함 	
<p>066 - 사용문자세트(Character Set Present)(반복불가, 해당시필수) 식별기호 코드 ▼a - 도서관부호(반복불가)</p>	<ul style="list-style-type: none"> MARC21에는 066 필드가 없으며, KORMARC에서 066 필드는 레코드에 사용한 문자세트가 KS X 1001 또는 ISO 10646 (유니코드) 이외의 문자 세트인 경우, 한국도서관 부호표 (https://librarian.nl.go.kr/kormarc/kormarc_2014/sub/annex_7.html)를 참조해 레코드를 작성한 도서관부호를 기술함 KORMARC to BIBFRAME에서 변환이 필요한 경우 표시할 수 있는 변환규칙 필요 	

<p>085 - 기타 분류기호(Other Classification Number)</p> <p>제1지시기호 - 적용분류표</p> <p>b - 그 외의 분류표</p> <p>0 - KDCP (한국십진분류법-박봉석)</p> <p>1 - NDC(일본십진분류법)</p> <p>2 - 조선총독부 신서부분류표</p> <p>3 - 조선총독부 양서부분류표</p> <p>4 - 조선총독부 고서부분류표</p> <p>5 - 사부분류</p> <p>제2지시기호 - 미정의(b)</p> <p>식별기호 코드</p> <p>▼a - 분류기호(반복)</p> <p>▼b - 도서기호(반복불가)</p> <p>▼q - 부여기관(반복불가)</p> <p>▼2 - 분류기호의 정보원(반복불가)</p> <p>▼6 - 대체문자 연결(반복불가)</p> <p>▼8 - 필드 링크와 일련번호(반복)</p>	<ul style="list-style-type: none"> MARC21은 기타 분류기호 필드로 084를 사용하며, 085에는 합성분류번호 요소(SYNTHESIZED CLASSIFICATION NUMBER COMPONENTS)를 배정 W-classification-Classification 부분에 기타 분류기호 표시에 대한 항목을 추가해야 함 	
<p>090 - 자관 청구기호(Local Call Number)</p> <p>▼a - 분류기호(반복불가)</p> <p>▼b - 도서기호(반복불가)</p> <p>▼2 - 권·연차기호(반복불가)</p> <p>▼8 - 필드 링크와 일련번호(반복)</p>	<ul style="list-style-type: none"> 해당 기관에서 부여한 청구기호상의 분류기호를 기술하며, 여기 기술된 분류기호는 분류기호 검색을 위해 적용분류표에 해당하는 분류기호 필드에 다시 기술함 BIBFRAME으로의 변환이 필요한 경우, W-classification-Classification 부분에 기타 분류기호 표시에 대한 항목을 추가해야 함 	

■ 1XX, 7XX, 8XX(명칭)

1XX, 7XX, 8XX는 개인명, 단체명, 회의명 등의 명칭(이름)을 표시하는 내용과 관련된 항목으로, 세부 필드에서 MARC21에는 적용하지 않지만 KORMARC에만 적용되는 식별기호들은 하단의 표에 정리함.

400, 410, 411, 440의 경우 8XX(800, 810, 811, 830)와 같이 변환 처리하며, 세부 사항 차이에 대해서는 별첨한 부록의 해당 파일 (ConvSpec-1XX,7XX,8XX-Names-v1.7-KORMARC)을 참조할 것.

〈표 9〉 1XX, 7XX, 8XX 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
<p>100, 700, 800 - 개인명(PERSONAL NAME)</p> <p>제2지시기호 - 부출표목의 유형(Second - Type of added entry)</p> <p>식별기호 코드</p> <p>▼f - 저작 연도(반복불가)</p> <p>▼g - 기타 정보(반복불가)</p> <p>▼h - 자료유형표시(반복불가)</p> <p>▼i - 관계정보(반복)</p> <p>▼k - 형식부표목(반복)</p> <p>▼l - 저작의 언어(반복불가)</p> <p>▼m - 연주수단(반복)</p> <p>▼n - 책·권차, 편차(반복)</p> <p>▼o - 편곡(반복불가)</p> <p>▼p - 저작의 편제(반복)</p>	<ul style="list-style-type: none"> MARC21에는 없고 KORMARC에서만 정의되는 식별기호들의 변환 관련 내용은 별첨한 부록의 관련 파일 (ConvSpec-240,X30-UnifTitle-v1.7- KORMARC...)을 참고할 것 제2지시기호(부출표목의 유형)의 경우 100에는 적용하지 않으며, 700, 800에만 적용함 100에는 적용하지 않으며, 700, 800 필드에만 적용하는 식별기호는 ▼h, ▼m, ▼o, ▼r, ▼s, ▼x, ▼3, ▼5가 있음 100, 700에는 적용하지 않으며, 800 필드에만 적용하는 식별기호는 ▼i가 있음 100, 800에는 적용하지 않으며, 700 필드에만 적용하는 식별기호는 ▼v, ▼w가 있음 	

<p>▼r - 음악자료의 장·단조(반복불가) ▼s - 판(version)(반복불가) ▼t - 저작의 표제(반복불가) ▼v - 권차/순차표시(반복불가) ▼w - 서지레코드 제어번호(반복) ▼x - 국제표준연속간행물 번호(반복불가) ▼3 - 자료 범위 지정 (반복불가) ▼5 - 필드 적용 기관 (반복불가) ▼8 - 필드 링크와 일련번호 (반복)</p>		
<p>110, 710, 810 - 단체명 (CORPORATE NAME) 제1지시기호 - 미정의(b) 식별기호 코드 ▼f - 저작 연도(반복불가) ▼g - 기타 정보(반복불가) ▼h - 자료유형표시(반복불가) ▼i - 관계정보(반복) ▼k - 형식부표목(반복) ▼l - 저작의 언어(반복불가) ▼m - 연주수단(반복) ▼n - 분과 및 부회 회의 회차(반복) ▼o - 편곡(반복불가) ▼p - 저작의 편제(반복) ▼r - 음악자료의 장·단조(반복불가) ▼s - 판(version)(반복불가) ▼t - 저작의 표제(반복불가) ▼v - 권차/순차표시(반복불가) ▼w - 서지레코드 제어번호(반복) ▼x - 국제표준연속간행물 번호(반복불가) ▼3 - 자료 범위 지정 (반복불가) ▼5 - 필드 적용 기관 (반복불가) ▼8 - 필드 링크와 일련번호 (반복)</p>	<ul style="list-style-type: none"> MARC21에는 없고 KORMARC에서만 정의되는 식별기호들의 변환 관련 내용은 별첨한 부록의 관련 파일 (ConvSpec-240,X30-UnifTitle-v1.7- KORMARC...)을 참고할 것 MARC21에서는 110, 710, 810의 제1지시기호를 정의하고 있으나 KORMARC에서는 정의하지 않음(미정의(b)) 110, 810에는 제2지시기호를 적용하지 않으며(미정의(b)), 710 필드에만 제2지시기호를 적용함 110에는 적용하지 않으며, 710, 810 필드에만 적용하는 식별기호는 ▼h, ▼m, ▼o, ▼r, ▼s, ▼x, ▼3, ▼5가 있음 110, 710에는 적용하지 않으며, 810 필드에만 적용하는 식별기호는 ▼v, ▼w가 있음 110, 810에는 적용하지 않으며, 710 필드에만 적용하는 식별기호는 ▼i가 있음 	
<p>111, 711, 811 - 회의명(MEETING NAME) 제1지시기호 - 미정의(b) 식별기호 코드 ▼f - 저작 연도(반복불가) ▼h - 자료유형표시(반복불가) ▼i - 관계정보(반복) ▼k - 형식부표목(반복) ▼l - 저작의 언어(반복불가) ▼n - 분과 및 부회 회의 회차(반복) ▼p - 회차표제(반복) ▼q - 관할 구역명 다음에 오는 회의명(반복불가) ▼s - 판(반복불가) ▼t - 저작의 표제(반복불가) ▼v - 권차/순차표시(반복불가) ▼w - 서지레코드 제어번호(반복) ▼x - 국제표준연속간행물 번호(반복불가) ▼3 - 자료 범위 지정 (반복불가) ▼5 - 필드 적용 기관 (반복불가) ▼8 - 필드 링크와 일련번호 (반복)</p>	<ul style="list-style-type: none"> MARC21에는 없고 KORMARC에서만 정의되는 식별기호들의 변환 관련 내용은 별첨한 부록의 관련 파일 (ConvSpec-240,X30-UnifTitle-v1.7- KORMARC...)을 참고할 것 MARC21에서는 111, 711, 811의 제1지시기호를 정의하고 있으나 KORMARC에서는 정의하지 않음(미정의(b)) 111, 811에는 제2지시기호를 적용하지 않으며(미정의(b)), 711 필드에만 제2지시기호를 적용함 111에는 적용하지 않으며, 711, 811 필드에만 적용하는 식별기호는 ▼h, ▼s, ▼x, ▼3, ▼5가 있음 111, 711에는 적용하지 않으며, 811 필드에만 적용하는 식별기호는 ▼v, ▼w가 있음 111, 811에는 적용하지 않으며, 711 필드에만 적용하는 식별기호는 ▼i가 있음 	

■ 200-247(표제와 표제 관련 필드)

200-247 부분은 240(통일표제) 항목을 제외한 나머지 표제 및 표제 관련 필드에 대한 내용을 정리한 항목으로 KORMARC이 MARC21과 상이한 부분은 필드별 표제 부출 여부 및 출력 형태에 대한 지시기호, 일부 상이한 기호 표시 사용에 대한 부분임. MARC21 대비 KORMARC의 200-247 필드의 구체적인 차이와 관련해서는 별첨한 부록의 해당 파일 (ConvSpec-200-247-Titles-v.1.7-KORMARC...)을 참조할 것.

■ 240, X30(통일표제)

240과 X30은 통일표제와 관련된 타이틀로, KORMARC이 MARC21과 상이한 부분은 표제와 관련된 200-247 필드와 마찬가지로 일부 필드의 표제 부출 여부 및 출력 형태에 대한 지시기호 사용 등에 대한 부분임. MARC21 대비 KORMARC의 240, X30 필드의 구체적인 차이와 관련해서는 별첨한 부록의 해당 파일 (ConvSpec-240,X30-UnifTitle-v.1.7-KORMARC)을 참조할 것.

■ 250-270(판차, 발행 등 필드)

250-270은 자료의 판사항, 자료별 특성, 발행 및 배포 사항 등을 표시하는 항목으로 구성되며, MARC21에는 없으나 KORMARC의 모든 필드에 정의되고 있는 ▼8(필드 링크와 일련번호) 기호 외에 특이사항은 없음. 세부 내용에 대해서는 별첨한 부록파일 (ConvSpec-250-270-v.1.7-KORMARC)을 참조할 것.

■ 3XX(형태사항 등)

3XX는 다양한 정보자료에 대한 세부 형태 사항들을 표현하는 항목으로 구성되며, KORMARC의 모든 필드에 정의되고 있는 ▼8(필드 링크와 일련번호) 기호 외에, KORMARC에는 384(음조) 필드 내 음조 유형을 표시하는 제1지시기호가 정의된다는 차이점이 있음. 세부 내용에 대해서는 별첨한 부록파일(ConvSpec-3XX-v.1.7.1-KORMARC)을 참조할 것.

〈표 10〉 3XX 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
384 - 음조(KEY) 제1지시기호 - 음조 유형 b - Original key와의 관계를 알 수 없는 경우 0 - Original key 1 - 조옮김된 음조	<ul style="list-style-type: none"> MARC21에서는 384필드가 반복(R)으로 되어 있으나, KORMARC에서는 반복불가(NR)로 되어 있음 MARC21에서는 제1, 2지시기호를 정의하고 있지 않아 이에 대한 변환 규칙이 필요한 경우 신규 생성 필요 MARC21에서는 식별기호 ▼0, ▼1, ▼3을 정의하고 있으나 KORMARC에는 정의하고 있지 않음 	

제2지시기호 - 미정의(b) 식별기호 코드 ▼a - 음조(반복불가) ▼6 - 대체문자 연결(반복불가) ▼8 - 필드 링크와 일련번호(반복)		
---	--	--

■ 490(총서사항), 510(인용/참고주기)

490(총서사항) 및 510(인용/참고주기) 필드는 총서명의 관계 및 관사 출력을 표시하는 제2지시기호 및 ▼8(필드 링크와 일련번호) 기호 외에 MARC21 대비 신규 정의되는 항목은 없음. 세부 내용에 대해서는 별첨한 부록파일 (ConvSpec-490,510-Links-v.1.7-KORMARC)을 참조할 것.

■ 5XX(주기사항)

5XX(주기사항)은 여러 정보자료에 대한 다양한 설명을 표시하는 항목으로 KORMARC에서는 MARC21에 없는 590(소장본 주기) 필드가 추가로 정의하고 있음. MARC21에는 적용하지 않지만 KORMARC에만 적용되는 일부 필드의 식별기호들은 하단의 표에 정리함. 각 필드별 차이에 대해서는 별첨한 부록의 해당 파일 (ConvSpec-5XX-v1.7.1-KORMARC)을 참조할 것.

〈표 11〉 5XX 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
501 - 합철주기(WITH NOTE) 식별기호 코드 ▼a - 합철주기(반복불가) ▼c - 표제(반복) ▼d - 첫 번째 책임표시(반복) ▼e - 두 번째 이하 책임표시(반복) ▼g - 기타 정보(반복) ▼n - 판사항(반복) ▼p - 발행사항(반복) ▼q - 형태사항(반복) ▼t - 총서사항(반복) ▼x - 대등표제(반복) ▼5 - 필드 적용 기관(반복) ▼6 - 대체문자 연결(반복불가) ▼8 - 필드 링크와 일련번호(반복)	<ul style="list-style-type: none"> 501 필드 식별기호 중 KORMARC에만 정의하고 있는 기호는 ▼c, ▼d, ▼e, ▼g, ▼n, ▼p, ▼q, ▼t, ▼x, ▼8¹⁵⁾임 다른 필드의 유사 식별기호 변환 규칙을 참고해, 각 식별기호의 내용을 BIBFRAME으로 변환하기 위한 세부 규칙이 필요함 	
502 - 학위논문주기(DISSERTATION NOTE) 제1지시기호 - 학위논문의 종류 0 - 석사학위논문 1 - 박사학위논문 제2지시기호 - 미정의(b) 식별기호 코드 ▼a - 학위논문의 종류(반복불가)	<ul style="list-style-type: none"> 502 필드 식별기호 중 KORMARC와 다른 의미로 기호를 사용하고 있는 항목은 ▼a, ▼b, ▼c이며, MARC21에는 없지만 KORMARC에서만 정의하고 있는 기호는 ▼8임 다른 필드의 유사 식별기호 변환 규칙을 참고해, 각 식별기호의 내용을 BIBFRAME으로 변환하기 위한 세부 규칙이 필요함 	

<ul style="list-style-type: none"> ▼b - 학위수여기관(반복불가) ▼c - 학과 및 전공(반복불가) ▼d - 학위수여연도(반복불가) ▼g - 기타정보(반복) ▼o - 학위논문 식별자(반복불가) ▼6 - 대체문자 연결(반복불가) ▼8 - 필드 링크와 일련번호(반복) 		
<p>505 - 내용주기(FORMATTED CONTENTS NOTE)</p> <p>제1지시기호 - 표출어 제어</p> <ul style="list-style-type: none"> 0 - 완전한 내용주기(Contents) 1 - 불완전한 내용주기(Incomplete contents) 2 - 부분 내용주기(Partial contents) 8 - 표출어를 생성하지 않음(No display constant generated) <p>제2지시기호 - 내용표시의 수준</p> <ul style="list-style-type: none"> b - 기본형(Basic) 0 - 확장형(Enhanced) <p>식별기호 코드</p> <ul style="list-style-type: none"> ▼a - 형식화된 내용주기(반복불가) ▼d - 첫 번째 책임표시(반복) ▼e - 두 번째 이하 책임표시(반복) ▼n - 권차(반복) ▼g - 기타 정보(반복) ▼t - 표제(반복) ▼u - URI(반복) ▼6 - 대체문자 연결(반복불가) ▼8 - 필드 링크와 일련번호(반복) 	<ul style="list-style-type: none"> • 505 필드 식별기호 중 KORMARC에서만 정의하고 있는 식별기호는 ▼d, ▼e, ▼n, ▼8임 • 다른 필드의 유사 식별기호 변환 규칙을 참고해, 각 식별기호의 내용을 BIBFRAME으로 변환하기 위한 세부 규칙이 필요함 	
<p>590 - 소장본 주기(Holding Item Note)</p> <p>식별기호 코드</p> <ul style="list-style-type: none"> ▼a - 낙장(落張), 파손(破損), 배접(背接), 보사(補寫), 포갑(包匣)(반복불가) ▼b - 인문(印文)(반복불가) ▼c - 장서기(藏書記), 수증기(受贈記), 수령기(受領記), 수권기(受券記)(반복불가) ▼d - 지어(識語), 묵서(墨書)(반복불가) ▼w - 소장 원본, 복제본의 청구기호(반복불가) ▼x - 소장 원본, 복제본의 형태사항(반복불가) ▼y - 열람용 소장본(반복불가) ▼z - 소장 관련 관리부호(반복불가) ▼6 - 대체문자 연결(반복불가) ▼8 - 필드 링크와 일련번호(반복) 	<ul style="list-style-type: none"> • 이 필드는 MARC21에는 없고 KORMARC에서만 정의하고 있음 • 소장본 주기사항의 내용을 표시할 수 있는 변환규칙 필요 	

■ 600-662(주제명부출표목)

600-662는 각 자료의 주제명 부출표목을 표시하는 필드로 각 필드의 기술 대상 데이

15) \$8은 MARC21에서는 정의하고는 있으나, 변환 명세 엑셀파일(ConvSpec-5XX-v.1.7.1) 기준으로는 LC에서 명시를 하지 않고 있습니다. 여기서는 일단 추가하여 표기하였습니다. 이후 표에 등장하는 \$8 모두 동일합니다.

터에 따라 지시기호 표시에 있어 MARC21과 KORMARC에 차이가 있음. 특히 주제명표/시소러스 종류를 구분하는 제2지시기호의 차이가 큼. 세부 필드별 지시기호의 차이에 대해서는 별첨한 부록파일(ConvSpec-600-662-v1.7.1-KORMARC...)을 참조할 것.

■ 720, 740-758(부출표목)

720, 740-458은 통제되지 않은 명칭, 분출표제, 지명, 분류학명 등의 부출표목을 표시하는 필드로, ▼8(필드 링크와 일련번호) 기호 외에 MARC21 대비 신규 정의되는 항목은 없음. 각 필드별 차이에 대해서는 별첨한 부록의 해당 파일(ConvSpec-720,740-758-v1.7-KORMARC...)을 참조할 것.

■ 760-788(연관저록)

760-788 필드는 상위 및 하위 총서, 원저서, 특별호를 비롯해 자료의 세부 구성 내용 등을 표시하는 항목임. MARC21에서는 이 필드의 세부 지시기호 및 식별기호를 상세히 정의하고 있지 않아, KORMARC에 정의된 각 해당 필드의 지시기호, 식별기호와 함께 각각의 내용을 BIBFRAME으로 변환할 때 필요한 세부 규칙의 참고 사항도 Conversion의 해당 부분에 함께 정리하였음. 각각의 세부 내용에 대한 정리 사항은 부록의 해당 파일(ConvSpec-760-788-Links-v1.7-KORMARC)을 참고할 것.

■ 841-887(소장, 위치, 변형문자 등)

841-887은 기술 대상 정보자료에 대한 소장사항과 위치정보를 비롯한 기타사항을 표시하는 필드임. MARC21 대비 KORMARC에만 정의된 필드는 890(미입력문자표시)가 있으며, 이에 대한 사항은 하단의 표에 정리함.

'KORMARC-통합서지용'에서는 841-845, 853-855, 863-878 필드를 서지데이터용 형식에서 열거만 하였고, 필드별 세부 식별기호 등에 대한 완전한 기술내용에 대해서는 'K O R M A R C - 소 장 정 보 용 (https://librarian.nl.go.kr/kormarc/kormarc_holding_data/index.html)'을 참고하도록 되어 있음. 이에 따라 841-845, 853-855, 863-878 필드는 해당 내용을 기준으로 지시기호 및 식별기호 등을 정리하였음. 각 필드별 차이에 대해서는 별첨한 부록의 해당 파일(ConvSpec-841-887-v1.7-KORMARC)을 참조할 것.

〈표 12〉 841-887 변환 고려사항

구분	BIBFRAME으로의 변환에 있어 고려사항	비고
890 - 미입력문자표시(반복불가, 해당시필수) 식별기호 코드 ▼b - 문자를 빈칸(b)으로 입력한 경우(반복) ▼h - 한자를 한글로 입력한 경우(반복) ▼r - 로마자 이외의 문자를 로마자로 입력한 경우(반복) ▼s - 레코드의 단축(반복불가) ▼x - 한자 이외의 문자를 한글로 변형 입력한 경우(반복)	<ul style="list-style-type: none"> • 이 필드는 MARC21에는 없고 KORMARC에서만 정의하고 있음 • 이 필드 내용에 대해 BIBFRAME으로 변환이 필요한 경우 신규 변환 규칙 정의가 필요함 	

5.5 국립중앙도서관 서지데이터 변환

지금까지 검토한 것처럼 MARC 21, KORMARC 모두 기존 구조를 유지하고, 모든 기술 요소들을 완벽하게 BIBFRAME으로 전환하는 것은 어렵습니다. 특히 KORMARC의 경우는 URI 기반으로 데이터를 기술하기 위한 데이터가 없습니다. 그럼에도 다양한 서지데이터들이 어떻게 변환되고 어떻게 보여지는지를 확인할 필요는 있습니다. 여기서는 국립중앙도서관 서지 데이터를 BIBFRAME으로 변환한 결과를 중심으로 검토해보도록 하겠습니다. 변환 대상은 단행본 하나를 대상으로 하며, 지금까지 검토한 한강의 채식주의자를 중심으로 합니다. 앞서 언급한 것처럼 BIBFRAME으로 변환하기 위해서는 해당 자원의 주어부, 술어부, 목적어부에 대한 URI가 필요합니다. 물론 우리가 사용하는 모든 요소들은 이미 BIBFRAME 용어집을 따르고 있기 때문에 별도로 URI가 필요하지 않지만 해당 자원을 구분할 수 있는 주어부의 URI와 전거데이터와 같이 MARC 21과 다르게 KORMARC에서 사용하는 형식들을 구분하기 위한 URI와 해당 값들이 필요합니다. 국립중앙도서관의 경우는 이미 링크드 데이터 서비스를 수행하고 있고, 일부 필요한 정보들은 이를 참조할 수 있습니다. 그러나 본 해설서에서 이 URI를 사용하면 전체적인 이해에 어려움이 따를 수 있어 여기서는 이 URI를 사용하지 않겠습니다. 대신 국립중앙도서관의 대표 도메인 'nl.go.kr'을 사용하여 가상으로 표현하도록 하겠습니다.

변환 예제에서 사용할 (가상)URI 체계는 다음과 같습니다.

data.nl.go.kr/works/[제어번호]

data.nl.go.kr/instances/[제어번호]

data.nl.go.kr/authorities/[전거종류 혹은 명칭]/[식별번호]

data.nl.go.kr/vocabulary/[해당용어집대표명 혹은 코드명]/[식별번호]

도서 변환 대상은 국립중앙도서관 소장자료인 채식주의자입니다. <그림 81>에서 보는 것처럼 국립중앙도서관에서 검색 시 MARC보기/다운로드와 LOD보기/다운로드가 가능합니다. 여기서 변환할 대상은 MARC입니다. LOD의 경우 BIBFRAME과 마찬가지로 링크드 데이터 형태로 변환된 데이터를 제공해줍니다. 차이점이라고하면 우리의 경우 BIBFRAME 용어집에서 정의한 표준용어를 기준으로 변환한다는 점만 다릅니다. RDF 형식으로 어떻게 변환되고 연결되는지 실제 사례를 보려면 현재 제공하는 LOD 보기/다운로드 혹은 국립중앙도서관 LOD(<http://lod.nl.go.kr/home/>) 서비스를 활용하면 됩니다.

The screenshot shows the search results for the book '채식주의자: 한강 장편소설'. The metadata includes:

- 제목/저자사항: 채식주의자 : 한강 장편소설 / 지은이: 한강
- 발행사항: 2022
- 판사항: 개정판
- 발행사항: 파주 : 함비, 2022
- 형태사항: 276 p. ; 20 cm
- 주기사항: 내용: 채식주의자 - 종교변형 - 나무 불꽃
- 표준번호/부호: ISBN 9788936434595 038100 415000
- 분류기호: 한국십진분류법 > 813.62 유아십진분류법 > 895.734
- 주제명: 한국 현대 소설 [韓國現代小說]

 The holdings table below shows one record:

소장정보	권/권자	판권	저작자	발행년도	ISBN	청구기호	자료이용장소	자료상태
<input type="checkbox"/>			지은이: 한강	2022	9788936434595	813.62-22-688-2	2층 문학실(리뷰실-해피)	자료실내비치

<그림 81> 국립중앙도서관 홈페이지의 채식주의자 검색결과 화면

변환을 위한 원천 MARC 데이터와 BIBFRAME 요소를 간략히 요약하면 <표 13>과 같습니다.

<표 13> 채식주의자 KORMARC와 BIBFRAME 대응요소 간략 정리본

TAG	IND	내용	대응요소	비고
001		KMO202219035		URI identifier로 사용
005		20220531130544	changeDate	
007		ta ¹⁶⁾	해당사항없음	mediaType(Media) carriers(Carrier type) 사용가능하나 구성상 별도 체계 필

				요
008		220421s2022 ggk ¹⁷⁾ 000 f korb n	creationDate originPlace genreForm language IntendedAudience Assigning agency	발행국 코드 집합 필요 장르 코드 집합 필요 장소 코드 집합 필요 언어 코드 집합 필요 이용대상 코드 집합 필요 목록작성기관 코드 집합 필요
020		▼a9788936434595▼g03810:▼c\15000	identifiers acquisitionTerms	
049	0	▼IEM8235535▼IEM8235536▼c2▼JE87508▼c3	identifiers AccessionNumber	국립중앙도서관 등록번호
052	01	▼a813.62▼b22-68	identifiers	BIBFRAME에 요소 없음 nlk call number 신규 구성 필요 identifiers에 기술
056		▼a813.62▼26		BIBFRAME에 요소 없음
082	01	▼a895.734▼223	classification ClassificationDdc	듀이십진분류기호(국립중앙도서관부여)
245	00	▼a채식주의자 :▼b한강 장편소설 /▼d지은이: 한강	mainTitle subtitle	
250		▼a개정판	editionStatement	instance
260		▼a파주 :▼b창비,▼c2022	originPlace originDate Place of publication Name of publisher Date of publication	
300		▼a276 p. :▼c20 cm	Extent Dimensions	
505	00	▼t채식주의자 --▼t몽고반점 --▼t나무 불꽃	note	relatedTo

16) 한국문헌자동화목록형식-통합서지용의 007 형태기술필드(physical Description Fixed field)가 코드화 되었다는 가정하에

586		▼a맨부커 인터내셔널상 수상, 2016	awards	
586		▼a산클레멘테 문학상, 2018	awards	
650	8	▼a한국 현대 소설[韓國現代小說]	subject	
700	1	▼a한강.▼g韓江.▼d1970-▼4aut	agent Relationship Designator	
740	02	▼a몽고반점	note	
740	02	▼a나무 불꽃	note	
950	0	▼b\15000	acquisitionTerms	

해당 MARC 데이터를 RDF 형식으로 변환하면 다음과 같이 표현가능합니다. 아래는 코드 값으로 직관적으로 이해하기 어렵습니다. 여기서 사용한 코드값은 앞서 검토했던 Marva 편집기를 활용한 것입니다. 이 편집기에서 입력한 내용을 코드화한 후 가상으로 설정한 국립중앙도서관 URI를 활용하여 일부 편집하였습니다. MARC21을 기준으로 한 예제는 Marva 메뉴 중 Everyone's records에서 확인할 수 있습니다.

```
<rdf:RDF xmlns:bflc="http://id.loc.gov/ontologies/bflc/" xmlns:bf="http://id.loc.gov/ontologies/bibframe/"
xmlns:bfsimple="http://id.loc.gov/ontologies/bfsimple/" xmlns:madsrdf="http://www.loc.gov/mads/rdf/v1#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:lclocal="http://id.loc.gov/ontologies/lclocal/"
xmlns:pmo="http://performedmusicontology.org/ontology/" xmlns:datatypes="http://id.loc.gov/datatypes/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:void="http://rdfs.org/ns/void#"
xmlns:mstatus="https://id.loc.gov/vocabulary/mstatus/"
xmlns:mnotetype="http://id.loc.gov/vocabulary/mnotetype/" xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:dcterms="http://purl.org/dc/terms/">
```

```
<bf:Instance xmlns:bf="http://id.loc.gov/ontologies/bibframe/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
rdf:about="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
rdf:about="data.nl.go.kr/instances/KMO202219035">
```

```
<bf:title>
```

```
<bf:Title>
```

작성하였습니다.

17) 한국문헌자동화목록형식-통합서지용의 발행국 부호표가 코드화되었다는 가정하에 작성하였습니다.

<bf:mainTitle>채식주의자</bf:mainTitle>
 <bf:subtitle>한강 장편소설</bf:subtitle>
 </bf:Title>
 </bf:title>
 <bf:editionStatement>개정판</bf:editionStatement>
 <bf:provisionActivity>
 <bf:ProvisionActivity>
 <bf:agent>
 <bf:Agent>
 <rdf:type rdf:resource="http://id.loc.gov/ontologies/bflc/PrimaryContribution"/>
 <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">창비 </rdfs:label>
 </bf:Agent>
 </bf:agent>
 <bf:date>2022</bf:date>
 <bf:place>
 <bf:Place rdf:about="http://data.nl.go.kr/vocabulary/countries/ggk">
 <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">ggk</rdfs:label>
 </bf:Place>
 </bf:place>
 </bf:ProvisionActivity>
 </bf:provisionActivity>
 <bf:contribution>
 <bf:Contribution>
 <bf:agent>
 <bf:Agent rdf:about="data.nl.go.kr/authorities/names/KAC201010897">
 <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#NameTitle"/>
 <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">한강, 韓江, 1970- </rdfs:label>
 </bf:Agent>
 </bf:agent>
 <bf:role>
 <bf:Role rdf:about="http://data.nl.go.kr/vocabulary/relators/aut">

```

    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">aut</rdfs:label>
  </bf:Role>
</bf:role>
</bf:Contribution>
</bf:contribution>
<bf:issuance rdf:resource="http://id.loc.gov/vocabulary/issuance/mono"/>
<bf:identifiedBy>
  <bf:Identifier>
    <bf:status>
      <bf:Status>
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">9788936434595</rdfs:label>
      </bf:Status>
    </bf:status>
  </bf:value>813.62</rdf:value>
  <bf:qualifier>국립중앙도서관 청구기호 </bf:qualifier>
</bf:Identifier>
</bf:identifiedBy>
<bf:note>
  <bf:Note>
    <bf:noteType>
      <rdfs:Literal xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
rdf:about="http://id.loc.gov/vocabulary/mnotetype/datasource">
        <rdfs:label>Data source</rdfs:label>
      </rdfs:Literal>
    </bf:noteType>
  </bf:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">국립중앙도서관 국가서지데이터
</rdfs:label>
  </bf:Note>
</bf:note>
<bf:media rdf:resource="http://id.loc.gov/vocabulary/mediaTypes/n"/>
<bf:extent>

```

```

<bf:Extent>
  <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">201 pages</rdfs:label>
</bf:Extent>
</bf:extent>
<bf:dimensions>20 cm</bf:dimensions>
<bf:carrier rdf:resource="http://id.loc.gov/vocabulary/carriers/nc"/>
<bf:electronicLocator>
  <rdfs:Resource xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    <rdf:value>http://data.nl.go.kr/instances/KMO202219035</rdf:value>
  </rdfs:Resource>
  <bf:note>
    <bf>Note>
      <rdfs:label>국립중앙도서관 BIBFRAME LOD </rdfs:label>
    </bf>Note>
  </bf:note>
</rdf:Resource>
</bf:electronicLocator>
<bf:adminMetadata>
  <bf:AdminMetadata>
    <bflc:catalogerId xmlns:bflc="http://id.loc.gov/ontologies/bflc/">BIB4Kor</bflc:catalogerId>
  </bf:adminMetadata>
  <bf:assigner>
    <bf:Agent>
      <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">krsnlk</rdfs:label>
    </bf:Agent>
  </bf:assigner>
  <bf:descriptionAuthentication>
    <bf>DescriptionAuthentication rdf:about="http://id.loc.gov/vocabulary/marcauthen/pcc">
      <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">pcc</rdfs:label>
    </bf>DescriptionAuthentication>
  </bf:descriptionAuthentication>
  <bf:descriptionConventions>
    <bf>DescriptionConventions rdf:about="http://id.loc.gov/vocabulary/descriptionConventions/rda">

```

```

    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">rda</rdfs:label>
  </bf:DescriptionConventions>
</bf:descriptionConventions>
<bf:descriptionLanguage>
  <bf:Language rdf:about="http://data.nl.go.kr/vocabulary/languages/kor">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">한국어</rdfs:label>
  </bf:Language>
</bf:descriptionLanguage>
<bf:descriptionModifier>
  <bf:Agent>
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">krsnlk</rdfs:label>
  </bf:Agent>
</bf:descriptionModifier>
<bf:procInfo xmlns:bflc="http://id.loc.gov/ontologies/bflc/">create work</bf:procInfo>
<bf:encodingLevel xmlns:bflc="http://id.loc.gov/ontologies/bflc/">
  <bf:EncodingLevel rdf:about="http://id.loc.gov/vocabulary/menclvl/f">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">full</rdfs:label>
  </bf:EncodingLevel>
</bf:encodingLevel>
<bf:status>
  <bf:Status rdf:about="http://id.loc.gov/vocabulary/mstatus/n">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">new</rdfs:label>
  </bf:Status>
</bf:status>
<bf:creationDate>2022-09-01</bf:creationDate>
<bf:note>
  <bf>Note>
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">맨부커 인터내셔널상 수상, 2016
산클레멘테 문학상, 2018</rdfs:label>
  </bf>Note>
</bf:note>

```

```

<bf:identifiedBy>
  <bf:Identifier>
    <rdf:value>KMO202219035</rdf:value>
  </bf:Identifier>
</bf:identifiedBy>
<bf:changeDate>20220531130544</bf:changeDate>
<bf:acquisitionTerms>
  <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">15000</rdfs:label>
</bf:acquisitionTerms>
</bf:AdminMetadata>
</bf:adminMetadata>
<bf:instanceOf>
  18)<bf:Work xmlns:bf="http://id.loc.gov/ontologies/bibframe/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
rdf:about="http://data.nl.go.kr/resources/works/">
  <bf:contribution>
    <bflc:PrimaryContribution xmlns:bflc="http://id.loc.gov/ontologies/bflc/">
      <bf:agent>
        <bf:Agent rdf:about="http://data.nl.go.kr/authorities/names/KAC201010897">
          <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#NameTitle"/>
          <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">한강, 韓江, 1970-</rdfs:label>
        </bf:Agent>
      </bf:agent>
      <bf:role>
        <bf:Role rdf:about="http://data.nl.go.kr/vocabulary/relators/aut">
          <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">aut</rdfs:label>
        </bf:Role>
      </bf:role>
    </bflc:PrimaryContribution>
  </bf:contribution>

```

18) 여기부터 저작에 대해 기술합니다.

```

<bf:title>
  <bf:Title>
    <bf:mainTitle>채식주의자</bf:mainTitle>
  </bf:Title>
</bf:title>
<bf:genreForm>
  <bf:GenreForm rdf:about="http://id.loc.gov/authorities/genreForms/gf2014026339">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Fiction</rdfs:label>
  </bf:GenreForm>
</bf:genreForm>
<bf:genreForm>
  <bf:GenreForm rdf:about="http://id.loc.gov/authorities/genreForms/gf2014026218">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Allegories</rdfs:label>
  </bf:GenreForm>
</bf:genreForm>
<bf:genreForm>
  <bf:GenreForm rdf:about="http://id.loc.gov/authorities/genreForms/gf2014026295">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Domestic fiction</rdfs:label>
  </bf:GenreForm>
</bf:genreForm>
<bf:genreForm>
  <bf:GenreForm rdf:about="http://id.loc.gov/authorities/genreForms/gf2014026492">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Psychological fiction</rdfs:label>
  </bf:GenreForm>
</bf:genreForm>
<bf:originDate>2007</bf:originDate>
<bf:originPlace>
  <bf:Place rdf:about="http://data.nl.go.kr/vocabulary/countries/ggk">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">파주, 경기도</rdfs:label>
  </bf:Place>
</bf:originPlace>

```

```

<bf:geographicCoverage>
  <bf:GeographicCoverage rdf:about="http://id.loc.gov/vocabulary/geographicAreas/a-ko">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Korea (South)</rdfs:label>
  </bf:GeographicCoverage>
</bf:geographicCoverage>
<bf:intendedAudience>
  <bf:IntendedAudience rdf:about="http://data.nl.go.kr/vocabulary/targetAudience/f">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">f</rdfs:label>
  </bf:IntendedAudience>
</bf:intendedAudience>
<bf:subject>
  <bf:Topic rdf:about="http://data.nl.go.kr/authorities/subjects/KSH2002030019">
    <madsrdf:isMemberOfMADSScheme xmlns:madsrdf="http://www.loc.gov/mads/rdf/v1#"
rdf:resource="http://id.loc.gov/authorities/subjects"/>
    <madsrdf:authoritativeLabel xmlns:madsrdf="http://www.loc.gov/mads/rdf/v1#">한국 현대
소설[韓國現代小說]</madsrdf:authoritativeLabel>
  </bf:Topic>
</bf:subject>
<bf:classification>
  <bf:ClassificationDdc>
    <bf:source>
      <bf:Source rdf:about="http://id.loc.gov/vocabulary/classSchemes/ddc23">
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">23</rdfs:label>
      </bf:Source>
    </bf:source>
    <bf:assigner>
      <bf:Agent rdf:about="http://data.nl.go.kr/vocabulary/library/011001">
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">국립중앙도서관</rdfs:label>
      </bf:Agent>
    </bf:assigner>
    <bf:classificationPortion>895.734</bf:classificationPortion>
  </bf:ClassificationDdc>
</bf:classification>

```

```

</bf:ClassificationDdc>
</bf:classification>
<bf:content rdf:resource="http://id.loc.gov/vocabulary/contentTypes/txt"/>
<bf:language>
  <bf:Language rdf:about="http://data.nl.go.kr/vocabulary/languages/kor">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Korean</rdfs:label>
  </bf:Language>
</bf:language>
<bf:adminMetadata>
  <bf:AdminMetadata>
    <bflc:catalogerId xmlns:bflc="http://id.loc.gov/ontologies/bflc/">BIB4Kor</bflc:catalogerId>
    <bf:descriptionAuthentication>
      <bf:DescriptionAuthentication rdf:about="http://id.loc.gov/vocabulary/marcauthen/pcc">
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">pcc</rdfs:label>
      </bf:DescriptionAuthentication>
    </bf:descriptionAuthentication>
    <bf:descriptionConventions>
      <bf:DescriptionConventions rdf:about="http://id.loc.gov/vocabulary/descriptionConventions/rda">
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">rda</rdfs:label>
      </bf:DescriptionConventions>
    </bf:descriptionConventions>
    <bf:descriptionLanguage>
      <bf:Language rdf:about="http://id.loc.gov/vocabulary/languages/kor">
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">Korean</rdfs:label>
      </bf:Language>
    </bf:descriptionLanguage>
    <bf:descriptionModifier>
      <bf:Agent>
        <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">krsnlk</rdfs:label>
      </bf:Agent>
    </bf:descriptionModifier>

```

```

<bf:procInfo xmlns:bf:="http://id.loc.gov/ontologies/bf:/">create work</bf:procInfo>
<bf:creationDate>220421</bf:creationDate>
<bf:status>
  <bf:Status rdf:about="http://id.loc.gov/vocabulary/mstatus/n">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">new</rdfs:label>
  </bf:Status>
</bf:status>
<bf:encodingLevel xmlns:bf:="http://id.loc.gov/ontologies/bf:/">
  <bf:EncodingLevel rdf:about="http://id.loc.gov/vocabulary/menclvl/f">
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">full</rdfs:label>
  </bf:EncodingLevel>
</bf:encodingLevel>
<bf:assigner>
  <bf:Agent>
    <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">krsnlk</rdfs:label>
  </bf:Agent>
</bf:assigner>
<bf:identifiedBy>
  <bf:Identifier>
    <rdf:value>KMO202219035</rdf:value>
  </bf:Identifier>
</bf:identifiedBy>
<bf:changeDate>20220531130544</bf:changeDate>
<bf:awards>
  <rdfs:label xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">맨부커 인터내셔널상 수상, 2016
산클레멘테 문학상, 2018</rdfs:label>
</bf:awards>
</bf:AdminMetadata>
</bf:adminMetadata>
</bf:Work>
</bf:instanceOf>

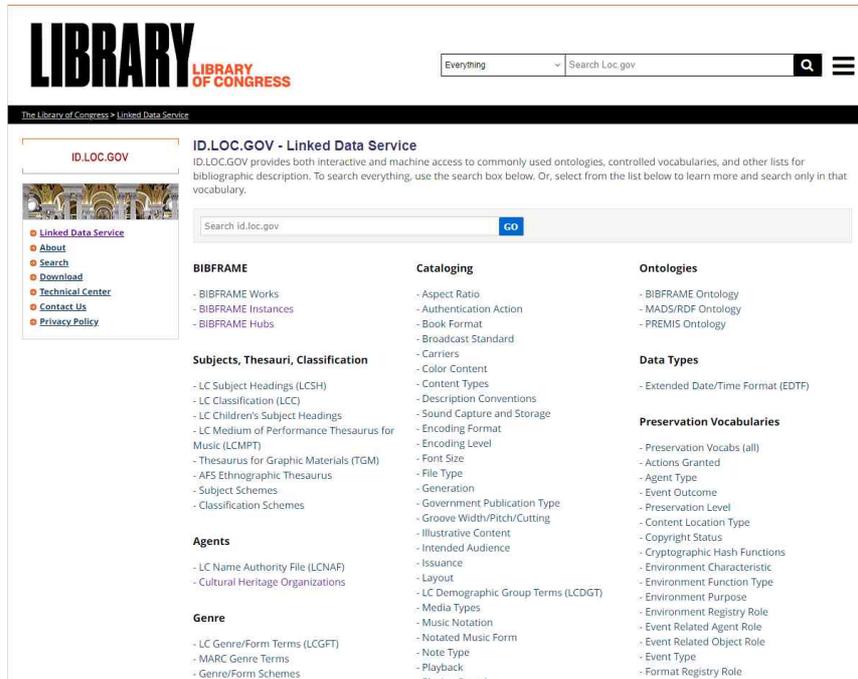
```

</bf:Instance>

</rdf:RDF>

우리는 변환 과정을 통해서 MARC를 BIBFRAME으로 변환하기 위해 준비해야하는 사항이 매핑 규칙을 만드는 것만이 아니라는 점을 알 수 있습니다. 도서관이 만들고 관리하는 전거, 국가명 코드, 장르코드 등은 도서관뿐만 아니라 다양한 데이터들이 참조할 수 있는 유용한 정보들입니다. 또 이런 통제어휘들은 서로 다른 정보자원들을 연결시킬 수 있는 접점 역할을 수행합니다. MARC 데이터를 BIBFRAME으로 변환하더라도 그 데이터가 단순히 서지 정보만을 표현하고 있다면 당초 목표했던 웹에서의 활용도는 높지 않을 것입니다. 그런데 서지 정보가 포함하고 있는 다양한 정보들(주제, 분류체계, 장르, 언어 등)이 함께 공개되고 공유되면 도서관 데이터의 활용도를 더 높아질 수 있습니다. 도서관이 구축하고 관리하는 데이터는 서지 데이터뿐만이 아닙니다. 오히려 서지 정보가 아니더라도 정보자원을 구축 관리하는 기관들이 필요한 다양한 형태의 통제 어휘가 존재하는 곳이 도서관이기도 합니다. 우리가 변환한 단행본의 경우를 다시 살펴보겠습니다. 대상이 단행본이 아니라 연속간행물, 고서, 디지털자료 등으로 변경된다고 하더라도 기본적으로 기술에 필요한 요소는 거의 유사합니다. 단, 대상 저작물의 형태나 장르를 구분하는 요소들에서는 차이를 보이게 됩니다. 이런 차이를 만들어내는 것 또한 중요한 데이터이고 활용도가 높다할 수 있습니다.

미국의회도서관의 경우도 BIBFRAME 표준화 활동과 별개로 링크드 데이터 형태의 도서관데이터를 구축 및 공유하기 위한 별도의 서비스가 존재합니다. 이미 잘 알려진 ID.LOC.GOV(<https://id.loc.gov/>)가 그것입니다.



〈그림 82〉 미국의회도서관 링크드 데이터 서비스 페이지

‘ID.LOC.GOV’는 미국의회도서관이 구축한 온톨로지, 통제어휘, 서지기술에 필요한 다양한 데이터를 제공합니다. 제공하는 데이터들을 정리하면 〈표 14〉와 같습니다.

〈표 14〉 미국의회도서관 링크드 데이터 서비스에서 제공하는 데이터세트 현황

구분	제공 데이터세트
BIBFRAME	<ul style="list-style-type: none"> BIBFRAME 저작, 인스턴스, 허브 데이터 세트 제공 BIBFRAME 저작, 인스턴스, 허브에 해당하는 실제 데이터세트 검색 가능 BIBFRAME으로 표현된 다양한 서지정보를 실제 확인할 수 있음¹⁹⁾
주제, 시소러스, 분류체계 (Subjects, Thesauri, Classification)	<ul style="list-style-type: none"> 주제명 표목(LC Subject Headings) 분류(LC Classification) 어린이 주제명 표목(LC Children's Subject Headings) 음악 연주매체 시소러스(LC Medium of Performance Thesaurus for Music) 그래픽 자료 시소러스(Thesaurus for Graphic Materials) 미국민속학회 민족지학 시소러스(AFS Ethnographic Thesaurus) 주제 체계(Subject Schemes) 분류체계(Classification Schemes)
행위자 (Agents)	<ul style="list-style-type: none"> 이름 전거 파일(LC Name Authority File (LCNAF)) 문화유산 기관(Cultural Heritage Organizations)
장르 (Genre)	<ul style="list-style-type: none"> 장르/형식 용어(LC Genre/Form Terms) MARC 장르 용어(MARC Genre Terms) 장르/형태 체계(Genre/Form Schemes)
언어 (Languages)	<ul style="list-style-type: none"> MARC 언어 코드(MARC Languages) ISO639-1 언어 코드(ISO639-1 Languages) ISO639-2 언어 코드(ISO639-2 Languages) ISO639-5 언어 코드(ISO639-5 Languages)
지리 (Geographic)	<ul style="list-style-type: none"> MARC 국가 코드(MARC Countries) MARC 지리 영역 구분 코드(Geographic Areas)

<p>목록 (Cataloging)</p>	<ul style="list-style-type: none"> • 종횡비(Aspect Ratio) • 인증조치(Authentication Action) • 도서 형식(Book Format) • 방송표준(Broadcast Standard) • 캐리어(Carriers) • 색상(Color Content) • 콘텐츠 유형(Content Types) • 기술 규칙(Description Conventions) • 음성 캡처 및 저장(Sound Capture and Storage) • 인코딩 형식(Encoding Format) • 인코딩 수준(Encoding Level) • 글꼴 크기(Font Size) • 파일 유형(File Type) • 세대(Generation) • 정부 간행물 유형(Government Publication Type) • 홈 폭/피치/커팅(Groove Width/Pitch/Cutting) • 삽화 내용(Illustrative Content) • 대상(Intended Audience) • 간행(Issuance) • 레이아웃(Layout) • 인구통계용어(LC Demographic Group Terms) • 미디어 유형(Media Types) • 악보(Music Notation) • 악보 형식(Notated Music Form) • 주기 유형(Note Type) • 재생(Playback) • 재생속도(Playing Speed) • 극성(Polarity) • 프레젠테이션 형식(Presentation Format) • 출판 방식(Production Method) • 사출(Projection) • 발행주기(Publication Frequencies) • 녹음 매체(Recording Medium) • 녹음 유형(Recording Type) • 감소율(Reduction Ratio) • 지역 인코딩(Regional Encoding) • 관계(Relators) • 부조(Relief) • 자원 구성요소(Resource Components) • 축척(Scale) • 스크립트(Script) • 연속간행물 유형(Serial Publication Type) • 사운드 콘텐츠(Sound Content) • 특수 재생 특징(Special Playback Characteristics) • 상태 코드(Status codes) • 보충 콘텐츠(Supplementary Content) • 지원 자료(Support Material) • 촉각 표기법(Tactile Notation) • 테이프 구성(Tape Configuration) • 기술(Technique) • 비디오 형식(Video Format)
<p>온톨로지 (ontologies)</p>	<ul style="list-style-type: none"> • BIBFRAME 온톨로지(BIBFRAME Ontology) • MADS/RDF 온톨로지(MADS/RDF Ontology) • PREMIS 온톨로지(PREMIS Ontology)
<p>데이터 유형 (Data Types)</p>	<ul style="list-style-type: none"> • 확장 날짜/시간 형식(Extended Date/Time Format)
<p>보존 어휘 (Preservation Vocabulaires)</p>	<ul style="list-style-type: none"> • 보존 어휘 전체(Preservation Vocab (all)) • 부여된 조치(Actions Granted)

	<ul style="list-style-type: none"> • 행위자 유형(Agent Type) • 이벤트 결과(Event Outcome) • 보존 수준(Preservation Level) • 콘텐츠 위치 유형(Content Location Type) • 저작권 상태(Copyright Status) • 암호화 해시 함수(Cryptographic Hash Functions) • 환경 특성(Environment Characteristic) • 환경 기능 유형(Environment Function Type) • 환경 목적(Environment Purpose) • 환경 레지스트리 역할(Environment Registry Role) • 이벤트 관련 행위자 역할(Event Related Agent Role) • 이벤트 관련 객체 역할(Event Related Object Role) • 이벤트 유형(Event Type) • 형식 레지스트리 역할(Format Registry Role) • 하드웨어 유형(Hardware Type) • 금지 대상(Inhibitor Target) • 금지 유형(Inhibitor Type) • 연결 행위자 역할 이벤트(Linking Agent Role Event) • 연결 환경 역할(Linking Environment Role) • 개체 범주(Object Category) • 보존 수준 역할(Preservation Level Role) • 관계 하위 유형(Relationship SubType) • 관계 유형(Relationship Type) • 권리 근거(Rights Basis) • 권리 행위자 역할(Rights Related Agent Role) • 서명 인코딩(Signature Encoding) • 서명 방법(Signature Method) • 소프트웨어 유형(Software Type) • 저장 매체(Storage Medium)
기타 (Other)	<ul style="list-style-type: none"> • 핑거프린트 구성 정보원 부호(Fingerprint Scheme Source Codes) • 언어코드 및 용어 정보원 부호(Language Code and Term Source Codes) • 작곡 양식 코드 정보원 부호(Musical Composition Form Code Source Codes) • 악기 및 음성 코드 정보원 부호(Musical Instrumentation and Voice Code Source Codes) • 이름 및 표제 전거 정보원 부호(Name and Title Authority Source Codes) • 국가 서지 번호 정보원 부호(National Bibliography Number Source Codes) • 자원 유형 체계(Resource Types Scheme) • 표준 식별자 체계(Standard Identifier Schemes)

KORMARC의 경우도 필드별로 다양한 형식의 코드표와 용어집이 존재합니다. <표>처럼 BIBFRAME에서 제공하는 형식들도 호환을 위해 참조하고 있으며, 이외에도 부속서와 부록에 대표적으로 다음 13개의 부호표 및 어휘집이 존재합니다.

<표 15> KORMARC의 부호표와 어휘집 목록

구분	제공 데이터세트
발행국 부호표	<ul style="list-style-type: none"> • 미국 의회도서관의 MARC Code List for Countries를 기초로 2자리 국가 단위 사용 • 우리나라의 경우 국어의 로마자표기법(2014년)을 따랐으며, 특별시·광역시

19) 실제 다양한 BIBFRAME 변환 서지 정보 확인이 가능합니다. 현재 RDF 형식의 데이터 다운로드 링크정보가 존재하나 작동하지 않고 있습니다. 추후 정상적인 다운로드가 가능해지면 다양한 정보의 서지 RDF 다운로드가 가능할 것으로 보입니다.

	<ul style="list-style-type: none"> 시·도별로 상세하게 전개
한국대학부호표	<ul style="list-style-type: none"> 교육법에 의하여 설립된 대학 및 다른 법률에 규정된 대학교육과정 이상의 학력인정 교육기관 대상 영문 두 자리로 구성 대학구분 외에 대학의 상태(신설, 승격, 변경 등) 등 다양한 별도 부호 사용
언어구분부호표	<ul style="list-style-type: none"> 미국 의회도서관의 MARC Code List for Language를 기초로 활용
한국정부기관부호표	<ul style="list-style-type: none"> 중앙행정기관, 지방자치단체, 교육기관, 헌법기관 및 정부투자기관 등 영문 두자리로 구성 이외 변동사항 표기를 위한 별도 부호 사용
한국지역구분부호표	<ul style="list-style-type: none"> KS X 1505 지역코드의 대분류 및 중분류코드와 행정전산망 공통행정코드-지역코드 참조
외국지역구분부호표	<ul style="list-style-type: none"> 미국 의회도서관 MARC Code List for Geographic Areas를 기초로 사용
한국도서관부호표	<ul style="list-style-type: none"> 공공, 대학, 전문도서관, 초중고 학교도서관 대상 코드 외 변동사항 표기를 위한 별도 부호 사용
국가부호표	<ul style="list-style-type: none"> ISO 3166-11997에 근거한 KS X 1510-1:2008의 3문자 국가 코드 사용
표기호별적용수준	<ul style="list-style-type: none"> 리더/17자리의 입력 수준 중 완전수준 및 최소수준의 적용 기준 기술
제어식별번호	<ul style="list-style-type: none"> ▼ w 서지레코드 제어번호 ▼ 0 전거레코드 제어번호 또는 표준부호 ▼ 5 필드 적용 기관 ▼ 6 대체문자 연결 ▼ 8 필드링크와 일련번호
역할어, 정보원 기술부호	<ul style="list-style-type: none"> 교과과정 목표 정보원 부호 국가서지번호 정보원 부호 기능 정보원 부호 날짜와 시간표시 정보원 부호 악기 및 성악 정보원 부호 언어 정보원 부호 역할어 부호 이름 및 표제 전거 정보원 부호 인용한 자료에 관한 정보원 부호 작곡 형식 정보원 부호 장르/형식 정보원 부호 전자적 접근방법 부호 주제 색인 정보원 부호 주제어 정보원 부호 주제범주 정보원 부호 지도 데이터 정보원 부호 지리분류기호 처리 용어 정보원 부호 축약표제 정보원 부호 표준 식별자 정보원 부호 MARC 정보원 부호
판단이 모호한 표목	<ul style="list-style-type: none"> 개인명으로 처리해야 할 표목 단체명으로 처리해야 할 표목 회의명으로 처리해야 할 표목 생략되는 관사
다문자 레코드	<ul style="list-style-type: none"> A모델 - 자국어와 문자변환 B모델 - 단순 다문자 레코드

이렇듯 BIBFRAM RDF 형식으로 KORMARC 데이터를 변환하기 위해서는 다양한 선행 작업이 필요합니다. 앞서 우리는 미국의회도서관이 제공하는 MARC21 to BIBFRAME 명세서를 기준으로 KORMARC로 변환할 경우의 대비해야하는 요소들을 검토하였으며, 상세

내용은 별첨으로 제시한 바 있습니다. 우리의 경우 KORMARC를 기준으로 주요 요소들의 변환을 대비해야하기 때문에 향후 본격적인 BIBFRAME 변환을 수행하거나, BIBFRAME 기반의 새로운 형태의 데이터베이스 구성이 필요할 경우 KORMARC에만 존재하는 기술 요소들을 포함한 온톨로지 구성이 필요합니다. 또한 동일하지만 다른 의미로 사용하고 있는 입력규칙들이 존재한다면 이 역시 별도의 온톨로지로 구성하고 해당 요소들에 대한 명세를 공개해야 KORMARC를 사용하는 기관들의 데이터 표준화를 달성할 수 있습니다. 나아가 KORMARC를 참조하거나 연결하고자 하는 다른 국가, 기관들이 우리의 데이터를 이해하고 활용할 수 있습니다.

국립중앙도서관은 2021년 한국문헌자동화목록형식- 통합서지용, 전자통제용 개정(안) 연구를 수행하면서 MARC21과 KORMARC를 비교하여 차이점을 제시한 바 있습니다(국립중앙도서관, 2021). 우선 본연구에서 제시한 MARC21에는 있으나 KORMARC 통합서지에는 없는 필드는 <표 16>과 같습니다.

<표 16> MARC21에만 존재하는 필드(출처: 국립중앙도서관, 2021)

구분	필드
2014년 KORMARC 개정 이후 MARC21에 신설된 필드	<ul style="list-style-type: none"> • 031(Musical Incipits Information) • 251(Version Information) • 334(Mode of Issuance) • 335(Extension Plan) • 341(Accessibility Content) • 348(Notated MusicCharacteristics) • 353(SupplementaryContent Characteristics) • 363(Normalized Date and Sequential Designation) • 370(Associated Place) • 385(Audience Characteristics) • 386(Creator/Contributor Characteristics) • 388(Time Period of Creation) • 532(Accessibility Note) • 647(SubjectAdded Entry - Named Event) • 688(Subject Added Entry - Type of Entity Unspecified) • 69X(Local Subject Access Fields) • 751(Added Entry - Geographic Name) • 758(Resource Identifier)
복미권 사용을 전제로 하는 필드	<ul style="list-style-type: none"> • 025(Overseas Acquisition Number) • 061(National Library of Medicine Copy Statement) • 071(National Agricultural Library Copy Statement)
국내 적용은 가능하지만 사용 용도가 낮아 배제된 필드	<ul style="list-style-type: none"> • 026(Fingerprint Identifier) • 038(RecordContent Licensor) • 083(Additional Dewey Decimal Classification Number) • 085(Synthesized Classification NumberComponents) • 258(Philatelic IssueData) • 365(Trade Price) • 366(Trade Availability Information) • 881(Manifestation Statements) • 882(Replacement Record Information) • 883(Metadata Provenance) • 884(Description Conversion Information)

〈표 16〉의 필드들은 BIBFRAME에는 존재하는 요소들로 향후 KORMARC의 변환에서는 이 필드들을 대체할 수 있는 방안을 고려하거나 입력 불필요로 처리할 필요가 있습니다. 그럼에도 이 필드들과 BIBFRAME 요소들에 대한 이해가 필요한 이유는 국외 BIBFRAME 도서관 데이터를 연결하거나 이 필드를 사용하는 국내/외 다른 기관과의 연결에 필수적으로 필요하기 때문입니다. 반면 KORMARC에만 존재하고 MARC 21에는 존재하는 않는 필드도 있는데 이는 아래와 같습니다.

- 012 (국립중앙도서관 제어번호) 필드
- 023 (출판예정도서목록제어번호) 필드
- 049 (소장사항) 필드
- 056 (한국십진분류기호) 필드
- 440 (총서사항/부출표목-표제) 필드
- 900 (로컬표목 - 개인명) 필드
- 910 (로컬표목 - 단체명) 필드
- 911 (로컬표목 - 회의명) 필드
- 930 (로컬표목 - 통일표제) 필드
- 940 (로컬표목 - 표제) 필드
- 949 (로컬표목 - 총서표제) 필드
- 950 (로컬표목 - 가격) 필드
- 980 (로컬표목 - 소장표시) 필드

KORMARC에만 존재하는 이 필드들은 향후 BIBFRAME 변환 온톨로지 구성 시 국립중앙도서관의 URI를 기반으로 새롭게 정의하여 발행해야하는 요소들입니다. 단, 056 한국십진분류기호나 023 출판예정도서목록제어번호 등의 경우는 속성보다는 클래스로 구성하여 관리하는 것이 필요합니다. Ddc의 경우 ClassificationDdc 클래스 ‘<http://id.loc.gov/ontologies/bibframe/ClassificationDdc>’와 ShelfMark(<http://id.loc.gov/ontologies/bibframe/ShelfMark>)의 하위요소 ShelfMarkDdc(https://id.loc.gov/ontologies/bibframe.html#c_ShelfMarkDdc)와 같

은 체계를 유지하는 것이 필요합니다.

다음은 동일 필드지만 KORMARC과 MARC21에서 상이한 표시기호를 가지는 경우로 이를 정리하면 <표 17>과 같습니다.

<표 17> MARC21과 KORMARC의 동일 필드 내 다른 표시기호(출처: 국립중앙도서관, 2021)

필드	KORMARC	MARC21
052	• 국립중앙도서관 청구기호	• 지리 분류기호
055	• 지리 분류기호	• 캐나다 청구기호(ClassificationNumbers Assigned in Canada)
084	•	• 기타 분류기호
085	• 기타 분류기호	• 합성된 분류기호 요소 (Synthesized Classification Number Components)

사실 위의 요소들은 BIBFRAME으로 변환 시 BIBFRAME에서 사용하는 요소명을 사용하면 되기 때문에 큰 문제를 발생시키지는 않습니다. 다만, 만약 KORMARC를 BIBFRAME으로 변환하기 위해 MARC21 기반의 솔루션이나 규칙을 활용한다면 이의 조정 시 참고해야 할 필요가 있습니다.

이 외에도 검토가 필요한 부분 중 하나는 형식 차이입니다. <표 18>은 MARC21 식별기호 중 KORMARC에서 적용하고 있지 않는 것들을 정리한 것입니다.

<표 18> MARC21 식별기호 중 KORMARC 미적용 식별기호 : 서지(출처: 국립중앙도서관, 2021)

식별기호	MARC21 적용 필드
▼q (부가적 식별정보)	• 015, 020, 024, 027, 028 필드
▼0(전거레코드 제어번호 또는 표준번호)	• 022, 050, 060, 070, 080, 085, 086, 257, 310, 321, 336, 337, 338, 377, 381, 384, 753 필드
▼1(Real World Object URI)	• 022, 033, 034, 043, 050, 060, 070, 080, 085, 086, 100, 110, 111, 130, 240, 257, 310, 321, 336, 337, 338, 340, 344, 345, 346, 347, 377, 380, 381, 382, 384, 518, 600, 610, 611, 630, 648, 650, 651, 653, 655, 656, 657, 658, 662, 700, 710, 711, 730, 753, 754, 800, 810, 811, 830 필드
▼2 (정보원)	• 017, 043, 044, 100, 110, 111, 130, 240, 310, 321, 506, 520, 700, 710, 711, 730, 752, 753, 800, 810, 811, 830 필드
▼3(자료 범위지정)	• 752 필드
▼4(관계)	• 037 필드

또 MARC21에서 식별기호가 세분되어 있는 필드는 11개로 다음과 같습니다.

- 041(언어부호) 필드

- 043(지역부호) 필드
- 044(발행/제작국명부호) 필드
- 046(특별한 연도 부호) 필드
- 085(기타 분류기호) 필드
- 340(물리적 매체 유형) 필드
- 345(영상자원의 영사 특성) 필드
- 382(연주 매체) 필드
- 540(이용과 복제 제한에 관한 주기) 필드
- 567(방법 주기) 필드
- 650(주제명부출표목 - 일반주제명) 필드

반면 KORMARC에서 식별기호가 세분되어 있는 필드는 다음과 같습니다.

- 501(합철주기) 필드
- 505(내용주기) 필드

마지막으로 동일 필드이지만 식별기호의 의미가 상이한 경우는 066(사용문자세트)입니다.

이 상의 필드들의 특징은 다른 정보자원과 연결하거나 해당 필드의 코드 값으로 동일한 자원간의 관계설정이 가능하다는 점입니다. 즉 이를 활용하면 링크드 데이터들 간의 연결을 통해 보다 풍부한 데이터 구성과 연계가 가능하다는 점입니다. 또한 이들은 도서관을 넘어서 다양한 도메인에서 함께 공유하고 사용할 수 있다는 점을 함께 고려해야 합니다. 즉 향후 BIBFRAME으로 변환 시 우리가 미적용한 필드들은 추가를 고려할 필요가 있으며, 우리가 더 세분화한 내용들은 국제표준에 맞추어 변경하거나 별도의 코드집을 구성하여 개발할 필요가 있습니다.

지금까지는 서지에 대한 부분이며 전거통제 관점에서 KORMARC과 MARC21을 비교해보면 우선 필드를 기준으로 <표 19>와 같이 요약할 수 있습니다.

〈표 19〉 전거통제 관점 KORMARC과 MARC21 필드 비교(출처: 국립중앙도서관, 2021)

구분	필드
2016년 KORMARC 개정 이후 MARC21에 신설된 필드	<ul style="list-style-type: none"> • 031(Musical Incipits Information) • 075(Type of Entity) • 080(UniversalDecimal Classification Number) • 082(Dewey DecimalCallNumber) • 087(Government Document Classification Number) • 147(Heading -Named Event) • 335(Extension Plan) • 348(Format of Notated Music) • 388(Time Period of Creation) • 447(See From Tracing - Named Event) • 547(See Also From Tracing-Named Event) • 677(Definition) • 747(Established Heading Linking Entry-Named Event) • 885(Matching Information)
복미권 사용을 전제로 하는 필드	<ul style="list-style-type: none"> • 050(Library ofCongressCallNumber) • 053(LC Classification Number) • 055(Library and ArchivesCanadaCallNumber) • 060(National Library of Medicine Call Number) • 070(NationalAgricultural LibraryCallNumber)
국내 적용은 가능하지만 사용 용도가 낮아 배제된 필드	<ul style="list-style-type: none"> • 066(Character Sets Present) • 072(Subject Category Code) • 073(Subdivision Usage) • 086(Government Document Call Number) • 884(Description Conversion Information)

반면 KORMARC 전거통제에는 존재하지만 MARC21에 없는 필드는 3개로 아래와 같습니다.

012(국립중앙도서관 제어번호) 필드

052(국립중앙도서관 청구기호) 필드

056(한국십진분류기호) 필드

전거통제용의 경우 KORMARC과 MARC21이 서로 다른 의미로 사용하는 요소는 없는 것으로 확인되었습니다. 전거통제에서 KORMARC에 미적용된 식별기호는 〈표 20〉과 같습니다.

〈표 20〉 MARC21 식별기호 중 KORMARC 미적용 전거용 식별기호(출처: 국립중앙도서관, 2021)

식별기호	MARC21 적용 필드
▼ 0(전거레코드 제어번호 또는 표준번호)	<ul style="list-style-type: none"> • 022, 024, 065, 336, 375, 377, 384 필드
▼ 1(Real World Object URI)	<ul style="list-style-type: none"> • 022, 024, 034, 043, 065, 336, 368, 370, 372, 373, 374, 375, 376, 377, 380, 381, 382, 384, 385, 386, 260, 360, 500, 510, 511, 530, 548, 550, 551, 555, 562, 580, 581, 582, 585, 672, 673, 700, 710, 711, 730, 748,

	750, 751, 755, 762, 780, 781, 782, 785, 883 필드
▼ 2 (정보원)	• 043 필드
▼ 3(자료 범위지정)	• 043, 370 필드
▼ 4(관계)	• 370, 386, 700, 710, 711, 730, 748, 750, 751, 755, 762, 780, 781, 782, 785, 788 필드

전거통제의 경우 그 목적과 성격상 다른 정보, 데이터와의 정확한 식별과 연결에 핵심적인 역할을 담당하게 됩니다. 서지에서와 마찬가지로 전거에서 보이는 이런 차이들은 가급적 수용하여 표기하고, 우리만 사용하는 경우는 공식적인 URI 값을 갖는 요소로 발행이 필수적입니다.

5.6 BIBFRAME 국가도서관 사례

1. 영국 국립도서관

영국 국립도서관(British Library)은 국가서지를 링크드 데이터로 구축하였는데, 단행본을 시작으로 연속간행물, 다작저작 및 통합자원 등으로 확대하였습니다.

영국 국립도서관 링크드 데이터의 스키마는 최대한 기존 기술 스키마를 재사용하였으며, 네임스페이스는 <http://www.bl.uk/schemas/bibliographic/blterms#>입니다. 영국 국립도서관 링크드 데이터에서 영국 국가서지(BNB)의 자원을 기술하는데 사용되는 어휘는 다음과 같습니다:

- Bibliographic Ontology (Bibo)
- Bio: a Vocabulary for Bibliographical Information
- Dublin Core: 기본 메타데이터 용어어휘
- International Standard Bibliographic Description (ISBD): 국제서지기술 표준에서의 어휘
- Org: 조직정보의 링크드 데이터 출판을 위한 온톨로지
- SKOS: 정보조직 시스템의 RDF 출판을 지원하는 어휘집
- RDF Schema: RDF 스키마

- OWL: 웹 온톨로지 언어
- FOAF: 사람 및 사람 사이의 관계를 기술하는 언어
- The Event Ontology: 이벤트 및 이벤트 참가자를 기술하는 언어
- WGS84 Geo Positioning and the preliminary version of RDA: 위도, 경도 등 위치를 기술하는 언어

영국 국립도서관 링크드 데이터는 VIAF, LCSH, Lexvo, GeoNames, MARC 국가 및 언어 코드, Dewey.info 및 RDF Book Mashup과 같은 다양한 링크드 데이터셋으로 연결되어 있습니다. 서비스 및 데이터 다운로드를 위한 URL은 다음과 같습니다. 링크드 데이터 파일은 N-Triples, RDF/XML 형식으로 다운로드 받을 수 있으며, 데이터 검색 및 데이터 셋 다운로드를 위한 URL은 다음과 같습니다.

- 접근 URL : <http://bnb.data.bl.uk>
- 스키마 정보 URL : <http://www.bl.uk/schemas/>
- 전체 파일 다운로드(Bulk downloads, RDF/XML과 N-Triples) URL : <http://www.bl.uk/bibliographic/download.htm>

현재 Share-VDE 가상환경에서의 검색을 위한 최초의 국가서지를 준비하고 있습니다. Share-VDE 인터페이스는 링크드 데이터의 잠재력을 활용하여 쉽고 직관적인 사용자 경험을 제공하고 더욱 광범위하고 상세한 검색 결과를 도서관 이용자에게 제공하는 것을 목적으로 합니다. 참여 기관의 도서관 목록은 BIBFRAME 어휘 및 기타 온톨로지를 사용하여 MARC에서 RDF로 변환되며, 개체 식별, 조정, 데이터 보강 작업을 통해 개체 클러스터를 형성합니다. 클러스터 지식 기반(Sapientia)은 시맨틱 웹 모델을 사용하는 동시에 참여 도서관이 가능한 한 독립적으로 자체 데이터를 계속 처리할 수 있게 하도록 합니다. 또한 데이터는 Share-VDE 데이터베이스에 포함되는 것은 물론, 개별 도서관은 링크드 데이터로 변환된 원본 레코드를 반입할 수 있습니다.

Share-VDE의 주요 초점 영역은 URI를 사용한 MARC 레코드 강화, BIBFRAME 어휘를 사용한 MARC에서 RDF로의 변환, 링크된 데이터 검색 옵션을 제공하도록 개발된 BIBFRAME 데이터 모델을 개조한 가상 검색 플랫폼 생성, RDF에서 접근할 수 있는 엔티티 클러스터(Sapientia Cluster Knowledge Base) 및 관계 데이터베이스 생성, 데이터와의 직접적인 상호 작용을 위한 도구 구현, 배치/자동 데이터 배포 및 업데이트 절차 등입니다.

니다.

2. 독일 국가도서관

독일 국가도서관(Deutsche Nationalbibliothek)은 국가 서지인 DNB(Deutsche Nationalbibliografie)에 대한 링크드 데이터 서비스를 제공하고 있으며, 전거데이터를 포함합니다. 2010년부터 전거데이터의 변환을 시작으로, 2012년 서지데이터에 대한 링크드 데이터로 확장하였습니다.

DNB의 Linked Data 서비스를 통해 제공되는 데이터는 RDF/XML 및 RDF/turtle이며, 링크드 데이터 모델링에 사용된 어휘는 다음과 같은 널리 사용가능한 어휘를 채택하였습니다. 또한 독일 국가도서관의 링크드 데이터는 Dewey.info, MARC-Code 등의 데이터셋과 연결되어 있습니다.

- 서지 온톨로지 (Bibo)
- ISBD
- Dublin Core
- RDF Schema
- OWL
- FOAF

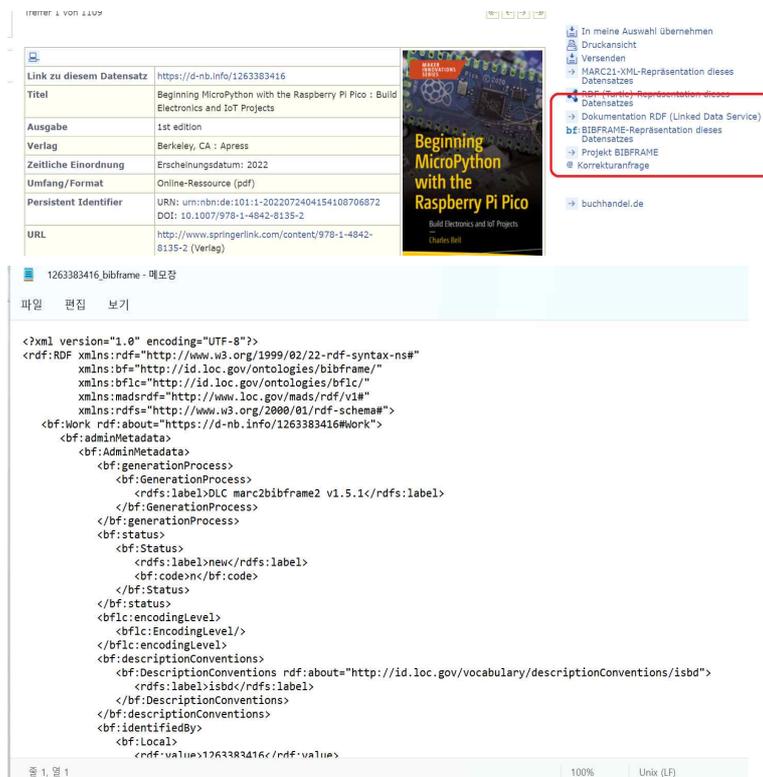
링크드 데이터 서비스 접근 및 데이터 다운로드를 위한 URL은 다음과 같으며, 모든 데이터는 CC0 라이선스 하에서 재사용이 가능합니다.

- 접근 URL : <http://www.dnb.de/EN/lds>
- 파일 다운로드(Dumps) URL : <http://datendienst.dnb.de/cgi-bin/mabit.pl?userID=opendata&pass=opendata&cmd=login>

독일 국가도서관은 BIBFRAME에도 적극적으로 참여하고 있습니다. 1단계 BIBFRAME 프로젝트는 독일 국가도서관 데이터를 BIBFRAM 형식으로 변환하는 데 중점을 두었으며, 2단계는 데이터의 프로토타입 매칭을 업데이트하고 확장하는 것을

목적으로 합니다. 현재 온라인 목록에서 반출형식은 MARC XML, RDF는 물론 BIBFRAME 어휘를 적용한 RDF 파일로도 제공됩니다. 독일 국가도서관의 BIBFRAME 프로젝트에 대한 설명은 다음 URL에서 찾아볼 수 있습니다.

https://www.dnb.de/DE/Professionell/ProjekteKooperationen/Projekte/BIBFRAME/bibframe_node.html



〈그림 83〉 독일 국가도서관 목록검색화면

3. 스웨덴 국립도서관

스웨덴 국립도서관(Kungliga Bibliotek, KB)은 국가도서관으로서 처음으로 국가 서지(Nationalbibliografin)를 BIBFRAME 2.0을 이용해 링크드 데이터로 변환하였습니다. 스웨덴 국립도서관은 Libris XL 프로젝트(새로운 Libris)를 시작으로 기존의 MARC21 시스템을 BIBFRAME을 기반으로 링크드 데이터로 변환하기 위한 도서관

시스템 개발 프로젝트에 착수하였습니다. 스웨덴 국가도서관이 이를 위하여 “KBL” 응용언어를 개발하여 BIBFRAME 및 기타 유사용어를 수용하였습니다. 궁극적으로 국가의 모든 도서관의 도서, 신문, 잡지 등 Libris가 새로운 방식으로 목록 레코드를 생성하고 탐색하는 등 가시화하는 것을 목적으로 합니다.

LIBRIS XL 프로젝트는 MARC21에서 데이터를 반입하고 반출하기 위한 변환 매핑을 구성한 결과, MARC 데이터에서 BIBFRAME으로의 변환이 매우 노동집약적임을 파악하였습니다. 이로 인해 MARC21 데이터를 LIBRIS XL로 마이그레이션 및 변환하는 것에 대한 기본 지침을 특정 메타데이터는 변환에 포함하지 않을 수 있으며, 마이그레이션 및 변환은 기능을 개선하는 것을 최우선으로 한다고 설정하였습니다. 또한 XL BIBFRAME 형식, 매핑, 통합을 수정하기 위하여 노력하였습니다. 그 결과 Libris 목록데이터의 BIBFRAME 기반 검색이 가능합니다. Libris XL를 통해 서지데이터의 저작, 인스턴스, 개념, 에이전트의 키워드 검색이 가능하며, 검색된 목록레코드는 N-Triples, RDF/XML 등으로 다운로드 가능합니다. 아래 이미지는 BIBFRAME 인스턴스 키워드 검색을 통해 발행연도, 수록매체유형, 장르, 소장기관 등의 패킷으로 탐색할 수 있는 Libris XL 검색사이트의 화면입니다.

The screenshot displays the Libris XL search results page. On the left, there is a sidebar with filters for 'HAR BESTÅND' (7k, 4k, 4k, 4k, 3k), 'VERKSTYP' (15k, 4k, 3k, 501, 116), and 'UTGIVNINGSSÄTT' (22k, 885, 184, 123, 16, 1). The main content area shows three search results, each with a title, author, publication details, and a 'BESTÅND' count. The top result is 'Alvesta SOK 30 år' by Svenska, published in 2003, with 24 pages. The second result is 'Mullsjö SOK 1995-2004' by Svenska, published in 2004, with 23 pages. The third result is 'Mullsjö SOK 2005-2014 • Jubileum 60' by Svenska. The interface includes a search bar at the top with 'START', 'SÖK', and 'HJÄLP' buttons, and a navigation bar at the bottom with 'Om Libris och XL', 'Libris nyheter', and 'Introduktionsmaterial' sections.

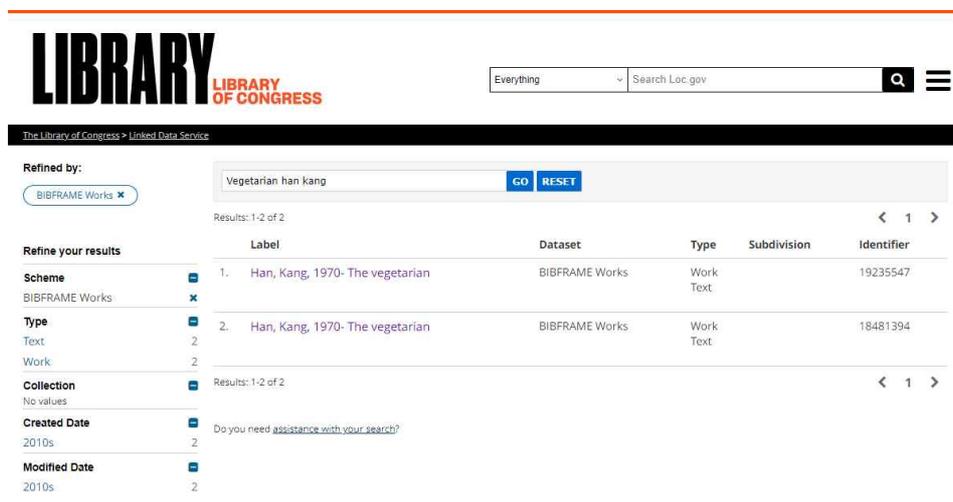
〈그림 84〉 Libris XL 검색화면(<https://libris.kb.se/katalogisering>)

4. 미국 의회도서관

미국 의회도서관은 BIBFRAME 연구와 개발 보급의 중심에 있는 기관으로 해당 용어집을 활용한 변환결과를 보여주기 위한 서비스가 존재합니다. 앞서 잠시 소개했던 “id.loc.gov”가 그것입니다. 이는 미국 의회도서관 전체의 링크드 데이터 서비스로 BIBFRAME이라는 별도의 섹션이 존재합니다. 여기에는 BIBFRAME 저작, BIBFRAME 인스턴스, BIBFRAME 허브 세 가지로 구분된 데이터가 존재하며 각각 검색이 가능한 체계를 갖추고 있습니다. 저작과 인스턴스 외 허브는 조금 독특한 존재입니다. 단어의 본래 뜻 그대로 중심적인 역할을 담당합니다. 링크드 데이터의 가장 큰 장점이라고 하면 말 그대로 연결에 있습니다. 연결을 위해서는 연결하고자 하는 자원, 데이터들이 명확하게 식별되어야 합니다. 도서관은 전통적으로 이의 중요성을 알고 식별자를 활용했습니다. 링크드 데이터의

경우는 URI를 사용합니다. 그런데 아직까지 MARC를 변환하는 링크드 데이터 시스템의 경우 해당 값들이 URI가 아닌 문자를 기반으로 하고 있어 해당 개체를 칭하는 레이블에 크게 의존하고 있습니다. 즉 허브는 관계 설정을 위한 중심적인 역할을 담당한다고 볼 수 있습니다. 허브는 MARC 서지 레코드와 표제 전거레코드를 변환하여 생성합니다. 허브는 다른 허브 혹은 BIBFRAME 저작으로 연결을 만들어냅니다.

BIBFRAME 저작에서 우리가 지금까지 예제로 활용했던 채식주의자를 검색하면 다음과 같은 화면을 확인할 수 있습니다.



〈그림 85〉 미국의회도서관 BIBFRAME 저작에서 채식주의자 검색결과 화면

‘vegetarian’이라는 키워드로 검색하면 검색결과가 너무 많습니다. 전통적인 도서관의 표제 기술방식에 따라 작가이름을 함께 검토하면 해당 저작물 검색에 용이합니다. 해당 저작물을 선택하면 다음과 같은 결과화면을 확인할 수 있습니다.

Bibframe Work

Title
The vegetarian

Type
Text

Contribution
- Han, Kang, 1970- author
- Smith, Deborah, 1987- translator

Subject
- Self-actualization (Psychology) in women--Fiction.
- Self-realization in women--Fiction.
- Vegetarianism--Fiction.

Genre Form
- fiction
- Psychological fiction.
- Domestic fiction.
- Allegories.

Language
- English
- English
- original (kor)
- English

Classification
- LCC: PL992.26.K36 C4313 2016 (Assigner: dlc)
- DDC: 895.73/5 full (Assigner: dlc)

Content
- text
- text (Source: rdacontent)

Expression Of
Han, Kang, 1970- Ch'aesikchuüija. English
Has Instance
The vegetarian

〈그림 86〉 BIBFRAME 채식주의자 검색결과 상세 화면

해당 검색결과 화면을 보면 Work에 해당하는 기술요소들로 해당 저작저작물이 기술되어 있는 것을 확인할 수 있습니다. 오른쪽에 보면 해당 저작에 대한 표현형과 인스턴스 링크를 확인할 수 있습니다. 해당 인스턴스로 이동하면 다음과 같은 화면을 확인할 수 있습니다.

Bibframe Instance

Title
The vegetarian a novel

Type
Print

Identified By
- Lccn: 2016296327
- Isbn: 9781101906118 (softcover)
- Isbn: 9780553448191 (ebook) (invalid)

Dimensions
21 cm

Extent
201 pages

Provision Activity
- Publication: New York (State) 2016
- Publication: New York Hogarth [2016]

Copyright Date
2015

Responsibility Statement
Han Kang ; translated from the Korean by Deborah Smith

Media
unmediated (source)

Issuance
single unit

Carrier
volume (source)

Instance Of
Han, Kang, 1970- The vegetarian
Has Item
PL992.26.K36 C4313 2016

〈그림 87〉 BIBFRAME 채식주의자 인스턴스 상세 화면

〈그림 87〉은 'vegetarian' 소설을 나타내는 인스턴스로 마찬가지로 인스턴스를 표현하는 표준 용어에 맞추어 기술되어 있습니다. 역시 오른쪽 상단에서 해당 인스턴스의 저작과 개

별자료로 이동할 수 있는 링크를 제공합니다. 개별자료는 해당 저작물에 대한 미국의회도서관의 소장정보로 <그림 89>와 같습니다.

Bibframe Item

Shelf Mark
PL992.Z6.K36 C4313 2016
Could not render: bf:assigner

Admin Metadata
Generation Process: DLC:marc2bibframe2 v1.7.0
Status: changed
Encoding Level: full
Description Conventions:
Identified By: bf:Local, 19235547
Change Date: 2017-01-11T17:32:53
Creation Date: 2016-08-18
Description Language: English
Description Modifier:
Description Authentication: pcc
Non-rendered Predicates: iclocal:d906 iclocal:d925 iclocal:d955 bf:assigner

Alternate Formats
- RDF/XML (Verbose)
- N-Triples (Verbose)
- JSON (Verbose)
- RDF/XML (Compact)
- N-Triples (Compact)
- JSON (Compact)

Item Of View

<그림 88> BIBFRAME 채식주의자 개별자료 상세 화면

미국의회도서관의 예를 보면 저작, 인스턴스, 개별자료 간의 관계와 어떻게 보여지는지를 확인할 수 있습니다. 하지만 이런 관계에 기반한 결과화면으로는 링크드 데이터의 장점이 무엇인지 잘 파악하기는 어렵습니다. 이는 다른 사례들에서도 마찬가지입니다. 링크드 데이터가 그 위력을 발휘하기 위해서는 도서관에서의 활용과 연결뿐만 아니라 외부 데이터와 도서관 데이터의 연계 혹은 외부 데이터가 우리 데이터를 활용한 사례들의 확인이 필요합니다. 사례에서 검토한 결과화면으로만 BIBFRAME에 대한 효과를 한정짓지 않기를 바랍니다.

06 마치는 글

본 해설서는 2022년 국립중앙도서관 연구용역인 “차세대 서지 언어 BIBFRAME 명세서 번역 및 해설서 제작”의 일환으로 제작되었습니다. 일반적으로 새로운 서지기술 표준이 등장했고, 이 표준이 기존 MARC를 기반으로 하고 있다면 표준에서 기술하고 있는 용어의 의미와 활용방법에만 초점을 맞추기 쉽습니다. MARC는 아주 오랜 시간 우리에게 친숙한 정보관리, 기술방식이기 때문입니다.

MARC를 기반으로 한 전 세계 도서관의 소통방식은 사실 큰 문제가 있다고 보기 어렵습니다. 사실 우리가 MARC를 기반으로 작업을 한다고 하더라도 우리의 작업환경은 GUI 환경에 맞추어져 있습니다. 실제 해당 레코드를 열어서 편집하는 일들은 거의 일어나지 않습니다. 때문에 우리는 MARC의 구조경직성이나 데이터베이스에서 구성의 문제에 대해서 인지하고 있었지만 크게 위기로 느끼지는 않았을지도 모릅니다. 그 간 정보생태계의 개방성이 점점 강해지면서 도서관 역시 MARCXML과 같은 형식변환, 더블린 코어, MODS와 같은 간단하고 다양한 도메인에서 함께 활용할 수 있는 용어집의 개발과 사용 등을 위한 노력을 해왔습니다. 또한 OAI와 같은 도서관 친화적인 프로토콜을 넘어 OpenAPI 기반의 데이터 교환 등 서비스 영역을 확장해왔습니다. 개방을 넘어서 연결이 중요한 시대가 등장하고 정보의 맥락, 데이터의 맥락이 중요해지면서 도서관은 FRBR, RDA 등 새로운 모형 개발에도 노력했습니다.

BIBFRAME은 새로운 서지 모델을 만들기 위한 노력이면서 이 모형을 표현하기 위한 새로운 용어 개발의 결과물이기도 합니다. 하지만 이 새로운 표준모형을 이해하고 제대로 활용하기 위해서는 왜 이런 표준이 등장했는지 배경에 대한 이해 선행이 필요합니다. 그래야만 왜 이런 모형이 필요하고 이렇게 표현하는지를 이해할 수 있기 때문입니다. 필요성이 납득이 되면 이 후 표준용어를 대하거나 관리시스템이나 서비스 시스템을 어떻게 구성해야 할지 쉽게 상상할 수 있습니다.

본 해설서에서는 이런 점에 중점을 두었습니다. BIBFRAME 전체 모델에 대한 이해와 표준 용어에 대한 이해는 명세서 번역본을 참조하도록 하고, 해설서에서는 이를 이해하기 위한 다양한 관점을 제공하고자 하였습니다. 먼저 이미 도서관계에서는 FRBR와 같은 모형 연구가 있어왔지만 우리는 해당 개념 이해를 위해 많은 노력을 기울여야했습니다. FRBR가 등장한 이 후에 이를 기반으로 한 시스템 구축 연구들도 있었지만 해당 표준을 기반으로 성공한 서비스가 존재한다고 보기는 어려울 것 같습니다. 무엇보다 우리는 해당 모형을 이해하는데 많은 노력을 기울였고 배경에 대한 이해도 높고, 결과물의 긍정적인 면도 알고 있지만 과연 이렇게 시스템을 구축하고 운영하는 것이 가능한가에 대한 의문을 늘 가지고 있었습니다. 사실 이런 의구심은 BIBFRAME 1.0 모델에서도 마찬가지였을 겁니다. BIBFRAME 1.0은 FRBR 모형과 유사한 면을 가지고 있었고, 해당 기술요소에는 RDA에 대한 이해와 적용도 필요했습니다.

해설서에서는 이런 점을 가장 많이 감안했습니다. BIBFRAME 표준 명세는 노력의 결과이고, 이 결과를 제대로 활용하고 이해하는데 필요한 정보들이 필요하다는 점에 최우선 하였습니다. 그리고 새로운 개념이 등장할 때마다 이 전의 것들을 모두 버려야하는 것과 같은 오해를 풀기 위해 노력했습니다. 도서관이 그 동안 정보자원 관리, 서비스, 보존을 위해 지켜왔던 노력들은 링크드 데이터나 BIBFRAME 같은 새로운 개념이 등장했다고 해서 유효하지 않은 것은 아닙니다. 오히려 현재의 변화들은 우리가 일하는 방식과 많이 닮아 있습니다. 해설서의 들어가는 글은 새로운 것에 대한 거부감을 줄이고 우리가 일하는 방식이 틀리지도, 새로운 환경에서 다르지도 않다는 점을 강조하였습니다.

BIBFRAME의 이해에 필요한 가장 핵심적인 요소는 링크드 데이터입니다. 링크드 데이터의 플랫폼은 웹입니다. 즉, BIBFRAME을 제대로 이해하기 위해서는 웹의 변화를 이해하는 것이 무엇보다 필요합니다. 해설서의 2장 웹이 변화는 어떻게 웹이 데이터 중심으로 변화하고 있는지와 그 필요성에 대해서 언급하였습니다. 여러분이 필요성에 대해 공감할 수 있다면 무엇이 필요한지 어떻게 구현할지 어렵פות이 상상할 수 있습니다. 3장은 여러분의 짐작을 구체적으로 알아보기 위한 부분으로 링크드 데이터와 RDF에 대한 내용을 담았습니다. 해설서를 읽고 나서 직접 RDF 형식의 데이터를 생산하거나 모델링하는 수준까지 도달하기는 어렵다고 하더라도 적어도 RDF 형식의 데이터를 읽고 해석할 수 있는 수준까지는 도달할 수 있을 것입니다.

이제 배경과 핵심기술에 대한 이해가 완료되었으므로 BIBFRAME에 대한 이해가 좀 더 수월할 것입니다. 4장은 BIBFRAME의 등장배경과, 모델, 1.0과 2.0의 차이 그리고 BIBFRAME 편집기를 소개하였습니다.

마지막은 우리에게 가장 친숙하고 우리가 가장 잘 알고 있으며, 향후 BIBFRAME을 적용

한다면 직접 목표가 되는 데이터인 KORMARC의 BIBFRAME 변환에 대해 다루었습니다. 대부분 이런 연구들은 필드, 즉 기술요소에 대한 매핑 작업이 주를 이룹니다. 해설서의 경우도 매핑에 초점을 맞추었지만 여기에만 그치지 않았습니다. BIBFRAME은 MARC21의 변환에 초점을 두고 있어 KORMARC와는 이질적인 부분이 존재합니다. 때문에 KORMARC의 각 필드들과 BIBFRAME을 매핑한 결과를 제시하였습니다. 분량의 문제가 있어 해당 정리내용은 별도 파일로 제공하였으며, 해설서에서는 그 중 핵심적인 내용만을 요약하여 제시하였습니다. 더불어 BIBFRAME 데이터를 기반으로 한 구현사례들을 검토하였습니다. 하지만 BIBFRAME의 경우 아직 완료된 프로젝트라고 보기 어렵고 완벽한 서비스 구축 시스템을 찾기는 어려웠습니다. 그래서 국가도서관을 중심으로 사례를 검토하였고 한계점을 함께 제시하였습니다.

본 해설서가 BIBFRAME에 대한 이해와 적용을 위한 완벽한 정보를 제공한다고 보기는 어렵을 것입니다. 그럼에도 이전에 어휘나 모델을 중심으로 이해하려고 했던 시도의 어려움을 조금이나마 덜어줄 수 있는 방안으로는 적합하리라 생각합니다. 새로운 무엇인가의 등장은 늘 어려울 것이라는 선입관이 개입되지만, 적어도 정보나 데이터를 기반으로 하는 새로운 것들은 우리의 그 간 노력이 헛되지 않았음을 상기하면서 자신감을 갖고 BIBFRAME에 도전하기를 바랍니다.

〈부록〉

KORMARC to BIBFRAME 필드별 변환 명세
(별도 파일 참조)

〈참고자료〉

- 국립중앙도서관. 2021. 한국문헌장도화목록형식(KORMARC)-통합서지용, 전거통제용 개정(안)연구. 발간등록번호: 11-1371029-000225-01. 서울: 국립중앙도서관
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2), 199-220.
- Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services(<https://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>)
- Miller, E., Ogbuji, U., Mueller, V., & MacDougall, K. (2012). Bibliographic framework as a web of data: Linked data model and supporting services. In Washington, DC: Library of Congress.
- Casey, D. D., Chapman, J., Harcourt, K., Mering, M., Miller, D., & Smith, S. (2008). On the Record: the Report of the Library of Congress Working Group on the Future of Bibliographic Control-Ten Actions for ALCTS.
- Sanderson, R. (2015). Analysis of the BIBFRAME ontology for Linked Data best practices. Prepared for Office of Network Standards. In The Library of Congress (pp. 04-17).
- Schreur, P. (2017). The evolution of BIBFRAME: from MARC surrogate to web conformant data model.

**차세대 서지 언어 BIBFRAME 명세서 번역 및 해설서 제작 -
BIBFRAME 해설서**

발행일 | 2022. 09. 30.

발행처 | 국립중앙도서관

| 06579 서울특별시 서초구 반포대로 201(반포동)

| 문의전화 02-590-0500 팩스 02-590-0530

| 홈페이지: <https://www.nl.go.kr/>

I S B N | 979-11-6513-290-3

| 979-11-6513-287-3 (세트)

인쇄처 | 위드디피에스

| 03961 서울특별시 마포구 방울내로 11길 37

비매품/무료

94020



9 791165 132903

ISBN 979-11-6513-290-3

ISBN 979-11-6513-287-3 (세트)