

# 호주의 보건복지관련 통계서비스 및 METeOR(Metadata Online Registry) 운영

*Australian Health and Welfare Statistics Services and METeOR Management*

이연희 한국보건사회연구원 책임전문원

호주보건복지연구소(AIHW; Australian Institute of Health and Welfare)는 국가의 건강상태에 관한 보고를 목적으로 1987년 호주보건연구소법에 의해 설립되었다. 1992년 호주보건연구소의 역할과 기능을 확대하여 복지와 관련된 정보 및 통계를 담당하는 호주보건복지연구소로 개편하였다. 본고는 호주보건복지연구소가 서비스하고 있는 통계 및 메타데이터 관리와 관련된 현황을 살펴봄으로써, 이를 통해 우리나라의 보건복지관련 통계나 관련 메타데이터를 효율적으로 운영하는데 다소 지침이 될 수 있을 것으로 판단된다.

## 1. 통계정보 서비스현황

AIHW는 별도의 통계사이트를 운영하지 않는다. 그러나 AIHW 홈페이지([www.aihw.gov.au](http://www.aihw.gov.au))를 보면, 메인 페이지의 일부 영역(좌측)을 할당하여 이용자가 통계에 접근할 수 있는 이정

표를 표시하고 있다(그림 1 참조).

AIHW는 8개의 커다란 범주로 나눠 통계와 관련된 정보를 제공한다. 물론 그 하부에는 좀 더 구체적인 범주별 통계를 제시하고 있다. 여기서 8개의 범주란 ①노령, 장애 및 간병인(Ageing, disability and carers), ②가족 및 아동(Families & children), ③병원(Hospitals), ④주택 및 노숙(Housing & homelessness), ⑤호주 원주민(Indigenous Australians), ⑥인구집단(Population groups), ⑦위험요인, 질병 및 사망(Risk factors, diseases & death), ⑧서비스, 노동력 및 소비(Services, workforce & spending)로 구분한다.

AIHW는 8개 각 범주에 대한 간단한 개요를 통해 주제의 중요성을 설명하고 있다. 세부적인 통계는 보고서로, 또는 인터넷상으로 제공한다. 예컨대, 노령에 관한 주제는 우선 노령에 대한 호주 사회의 중요성을 간략히 설명한다. 그 다음은 노인인구(65세 이상)의 수와 비율, 여성노인, 85세 이상 인구비율 등을 통계표 형태가 아

그림 1. 호주보건복지연구소 웹사이트 메인페이지 화면



닌 문장형태로 제시하고 있다. 각 범주에서 핵심이 되는 통계는 앞의 사례에서 설명했던 것처럼 서술식의 문장형태로 제공되거나, 또는 설명력을 높이기 위해 차트나 이미지를 이용하여 제시하고 있다. 그 외에 부가적인 통계는 관련된 보고서를 링크시켜 이용자에게 제공하고 있다. 다시 말하면, AIHW는 개별 통계에 대한 구체적인 서비스를 하는 것은 아니며, 주요 통계는 간단한 설명 위주로 하되, 고정된 이미지 형태로 서비스한다. 주요 통계외의 통계들은 통계의 구체적인 설명과 해석 등을 첨가하여 보고서 형태로 인터넷상에 올려놓고 이용자들에게 서비스된다. 즉, 이용자의 요구에 맞춰 통계항목을 선택할 수 있는 유연성을 제공하지는 않고 있다.

AIHW의 통계서비스는 8개 범주가 어떤 정

해진 패턴과 구조에 의해 구성된 것이 아니라 각 범주별로 상이하게 통계를 제공하고 있다. 즉 어떤 주제는 한 단계 구체화된 세부 주제로 나눠 추가적인 통계정보를 서비스하기도 하지만 그렇지 않은 경우도 있다.

AIHW의 통계서비스 현황을 분석한 바에 따르면, 정보기술을 최대한 활용한 이용자 편의 중심의 우리나라 통계사이트들과는 다른 형태로 통계가 서비스되고 있음을 알 수 있다.

## 2. METeOR(Metadata Online Registry) 소개

AIHW는 호주 국민의 건강과 안녕에 관한 국

가정보의 비교가능성, 일관성, 그리고 관련성을 개선할 목적으로 메타데이터 개발을 위한 국가의 인프라를 제공한다. 특히 AIHW는 호주의 국민건강과 지역사회서비스, 주택 등에 관한 메타데이터 항목들과 표준을 관리한다. 연구소는 특수 프로그램이나 정책분야 그리고 정부의 개입이 요구되는 영역에서의 메타데이터 표준을 개발하는데 참여하고 있다. 이것은 국가 정보로서의 요건을 식별하고, 또한 데이터에 관한 상세서를 개발하는 일을 하는 National Minimum Data Sets 같은 클라이언트들과 함께 작업을 한다. 뿐만 아니라 AIHW는 특수 데이터 사전의 개발, 가이드라인 개발, 그리고 그 외에 메타데이터와 관련된 자료를 문서화 하는 것도 지원하고 있다.

AIHW는 METeOR(Metadata Online Registry)라는 호주의 건강과 지역사회 서비스, 주택지원 부문 등에 관한 국가 메타데이터 표준저장소를 운영하고 있다. 이 시스템은 국가적으로 이미 인정한 다양한 데이터 정의(definitions)들과 기존에 명시한 구성요소들을 바탕으로 신규로 데이터 정의를 생성하는 틀들을 온라인상에서 접근하도록 하고 있다. 따라서 METeOR은 메타데이터가 저장되어 있고, 이것을 관리하고, 또한 보급하는 역할을 수행한다.

실제로 웹사이트의 METeOR를 통해 2,600개 이상의 데이터 표준을 다운받을 수 있다. 이들 표준은 유사한 표준을 생성하는데 들어가는 자원의 낭비를 막을 수 있으며, 국가적으로 인정된 표준을 기초로 이용자들의 정보시스템을 만들 수 있다. 그 외에도 METeOR의 데이터 표준을 이용하면, 다른 많은 수집된 데이터들을 비

교할 수 있는 기준 데이터를 얻을 수 있다.

## 1) METeOR의 메타데이터 구조

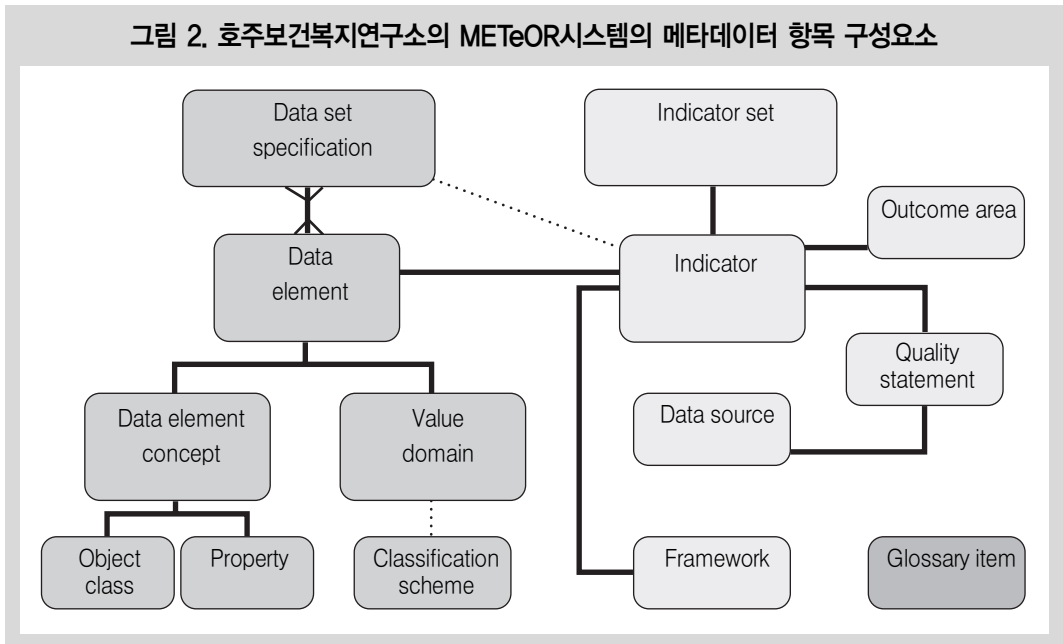
METeOR은 건강과 주택, 그리고 지역사회 서비스와 관련된 통계 및 정보에 관한 호주의 국가메타데이터 표준을 위한 저장소이다. 여기서 메타데이터란 데이터에 대한 데이터로 불린다. 즉, 데이터의 내용이나 품질, 조건, 또는 그 외 특성에 대한 기본 정의나 구조화된 표현을 말한다.

METeOR의 메타데이터 저장소는 13개의 메타데이터 항목을 가진다. 13개 항목은 객체클래스, 속성, 데이터요소 개념, 데이터요소, 값 도메인, 분류체계, 데이터세트 사양, 용어사전 항목, 지표, 지표 세트, 성과 분야, 품질서, 데이터소스 등이다(그림 2 참조).

데이터 요소 및 데이터 요소 개념, 분류체계, 값 도메인, 용어사전 항목, 지표, 지표 세트, 성과분야, 그리고 품질서 등은 알파벳으로 보여준다. 데이터세트 사양은 등록기관으로 보여준다. 객체클래스들은 객체클래스 상세서의 트리형태로 부여되어 있다. 상세서는 트리구조로 접근할 수 있으며, 이탤릭체로 되어 있다. 속성들은 속성그룹으로 보여준다. 데이터 소스는 등록항목들을 관리하지는 않지만 알파벳순으로 보여준다.

METeOR는 2가지 형태의 메타데이터 즉, 메타데이터 항목(metadata items)과 탐색적 항목(navigational items)을 사용하는데, 메타데이터 항목은 관련된 등록기관으로부터 표준으로 인정될 때까지의 일련의 등록 과정을 통과하여 처

그림 2. 호주보건복지연구소의 METeOR시스템의 메타데이터 항목 구성요소



리된 항목들을 말한다. 즉, 이것은 공식데이터 표준이 되는 메타데이터 항목들을 의미한다.

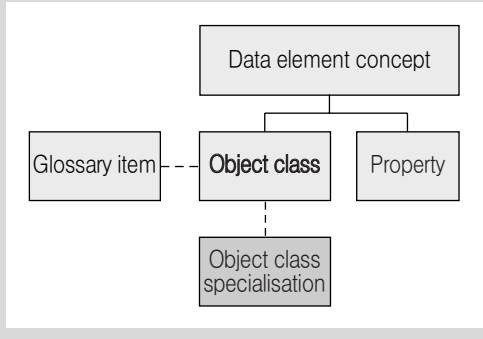
현재 METeOR에서 사용되는 메타데이터 항목은 객체클래스, 속성, 데이터요소 개념, 데이터요소, 값 도메인, 분류체계, 데이터세트 사양, 용어사전 항목 등 8개가 있다. 객체클래스는 설명할 필요가 있고 흥미가 있는 어떤 사람이나, 조직, 구조 또는 이벤트를 표현한다. 예를 들어 Person, Dwelling, Family는 METeOR의 객체클래스에 포함되는 예이다. 특별한 속성을 가지는 특정 객체클래스 조합은 어떤 데이터 요소 개념을 생성한다. 앞에서 보았던 객체클래스 중 Person은 데이터 요소 개념을 생성하기 위해 종교 가입(Religious affiliation) 속성과 함께 결합될 수 있다. 객체클래스들은 다른 객체클래스들을 특수화(specialisations)시킬 수 있다(그림 3 참

조). 예컨대, Adult는 Person의 연령집단과 관계 있는 특수화이다.

속성(Property)은 객체클래스의 특성을 말한다. 예를 들어 객체클래스 Person은 성별과 생일 같은 특성을 가질 수 있으며, 이들 특성은 속성으로 참조된다. 특수 객체클래스를 가지는 특수속성 조합은 어떤 데이터 요소 개념을 생성하게 된다. 예를 들자면 Person은 Person-성별, Person-생년월일 이라는 데이터 요소 개념들을 생성한다. 속성들은 라이프스타일 특성, 재정 특성 등과 같은 유사한 속성들을 묶는 속성 집단들이 될 수 있다.

데이터 요소 개념(Data element concept)은 객체클래스와 속성의 조합으로 구성되어 데이터 요소를 정의할 목적으로 생성된 개념으로 정의된다. 하나의 객체클래스와 하나의 속성만이 각

그림 3. 객체클래스와 관련된 메타데이터 항목 유형간 관계



데이터 요소 개념에 결합될 수 있다. 데이터 요소 개념은 데이터를 수집하는 방법에 대한 개념이 아니라 데이터를 구하는(seek) 방법에 대한 개념을 표현한다. 데이터 수집에 대한 설명은 값 도메인을 가지는 조합을 통해 얻어진다. 특수한 값 도메인을 가지는 특수 데이터 요소 개념의 조합은 하나의 데이터 요소를 생성한다. 예를 들어 Person-date of birth는 데이터 요소를 생성하는데 날짜 DDMMYYYY 값 도메인과 결합되어 DDMMYYYY 데이터 요소 개념을 가진다.

데이터 요소는 데이터 요소 개념과 값 도메인을 결합하여 생성되는 것으로서, 식별 가능하면서 정의 가능한 정보의 기본단위이다.

값 도메인(Value domain)은 데이터 요소가 실행될 때 허용 가능한 값들의 집합을 제공한다. 여기서 값 도메인은 숫자이거나 숫자이외 문자일 수도 있다.

분류체계(Classification scheme)는 국가나 국제기구로부터 인정되고 승인된 공식용어시스템이며, 데이터를 분류하는데 사용되고 있다.

METeOR에서 사용하는 분류체계의 사례로는 호주 직업 표준 분류(Australian Standards Classification of Occupations)와 질병 및 관련 건강문제 국제 통계분류 등이 있다. 분류체계는 하나 이상의 값 도메인에 의해 수행될 수도 있는데, 예를 들어 호주직업표준분류는 직업코드(Occupation code(ASCO 2nd edn)) 값 도메인에 의해 분류된다.

데이터셋 상세서(Data set specification; DSS)는 데이터 요소들로 구성된 그룹을 열거하며 그 그룹 내에 있는 조건들이 수집된다. DSS는 데이터 요소들이 포함되는 순서나 그것들이 필수적인지, 그리고 지원해야 할 검증규칙과 수집 특성은 무엇인지 등을 정의할 수 있다.

용어사전 항목(Glossary items)은 특수한 환경에 있는 용어의 의미를 정의한다. 일례로, 여성의 임신횟수에 대한 데이터 수집을 위해서는 출산(live birth)이라는 것을 이용하는데, 데이터 요소 Female-number of previous pregnancies (live birth), total NN에 적합한 계산규칙을 적용하기 위해 출산이라는 용어의 의미를 이해해야만 한다.

METeOR의 메타데이터 중 탐색적 항목(Navigational items)은 메타데이터 항목을 좀 더 쉽고, 좀 더 유용하게 탐색 및 관리하기 위해 등록담당자들이 METeOR에 추가로 부여한 항목들이다. 그러므로 탐색적 항목은 공식적인 데이터 표준은 아니다. METeOR에서 사용되는 탐색적 항목에는 객체클래스 상세서와 속성그룹이 있다. 객체클래스 상세서는 객체클래스의 계층도를 만드는데 기초가 되며, 또한 계층도를 이용해서 탐색하는데도 기본이 된다. 속성그룹은

유사한 속성을 묶는 것과, 유사한 특성을 가진 속성들의 검색을 허용한다.

## 2) 메타데이터 표준

메타데이터 표준을 개발하는 것은 호주의 보건복지에 관한 국가정보의 품질과, 적합성, 일관성, 가용성을 개선하는데 목적이 있다. 표준을 개발하는 동인은 더 좋은 정보의 욕구로부터 발생한다.

메타데이터 표준은 정의된 환경에서 사용하기 위해 데이터의 의미와 수용 가능한 표현을 설명한다. 의미의 일관성에 대한 요구는 데이터를 사용하는 이용자들 간에 정보 공유를 촉진하는데 매우 중요하다. 메타데이터 표준은 내용과 정의의 일관성에서도 중요하다. 공유할 데이터가 전혀 없다면 표준화할 필요성도 전혀 없다. 그러나 공유할 데이터가 있다면 해당 데이터를 사용할 필요가 있는 모든 사람들은 데이터가 어떻게 수집되든, 또는 어떻게 저장되든 관계없이 데이터의 그 의미를 명확히 이해할 수 있도록 보장해야 할 필요가 있다.

뿐만 아니라 메타데이터 표준은 지나친 다양성이 비효율성을 일으키거나 효율성을 지연시킬 때 필요하다. 즉, 메타데이터 표준은 정보가 서로 다른 그룹 간에 교환되는 다양한 방법을 축소하는 수단을 제공한다. 또한 메타데이터 표준은 여러 개발 노력들에 시너지 효과를 주는 수단을 제공하기도 한다. 한편, 메타데이터 표준은 데이터 개발 비용을 감소시킨다. 즉, 메타데이터 표준은 다른 사람들이 처음부터 시작하지 않고도 사용할 수 있도록 하는, 어떤 문제해

결 방법을 제공한다. 그뿐 아니라 메타데이터 표준은 작업하는 조직에게 일반적이고 일관성 있는 플랫폼을 제공하며, 메타데이터 표준을 사용함으로써 지역 및 국가 수준에서 채택과 실행을 단순화 시킨다

METeOR은 국가 메타데이터 표준을 개발하는 데 있어서 잘 정의된 처리절차를 지원하고 있다. 어떤 표준을 승인 처리하는 과정에는 전문가 그룹과 데이터 위원회, 등록기관 등이 있다. 전문가 그룹은 국가적으로 인정한, 해당 주제의 전문가 그룹들이며, 전문가 그룹이 항목을 동의하면, 심의를 위해 등록담당자에게 제출된다. 제출된 항목은 제안된 항목의 등록 상태를 계속 수정하면서 완성된다. 데이터 위원회는 심의 요청된 항목을 표준으로 승인할 것인가를 결정하는 위원회이다. 데이터 위원회가 항목을 검토하여 만일, 표준으로 채택하겠다고 하면 후보 등록상태로 수정한다. 후보 항목들은 위원회에서 성공적으로 검토된 후, 표준화 중(Standardisation pending)상태를 할당받게 되고, 관련 등록기관을 표시한다. 등록기관의 경우, 적용 가능한 분야의 등록기관은 데이터위원회나 정보위원회 권고를 검토하고, 적절한 항목에 표준상태를 할당하게 된다.

메타데이터 표준은 지역사회 서비스 부문과 보건 부문, 그리고 주택지원 부문으로 나눈다. 메타데이터 표준에 있어서 지역사회 서비스 부문은 국가 단위, 주 단위, 지역단위 정부와 비정부기관에서 제공되는 서비스들을 포함한다. 최근 몇 년간 호주의 정보 인프라는 국민의 욕구와 제공된 지원정책에 관하여 국가적으로 일관되면서 비교 가능한 데이터를 수집할 수 있도록

지원하는 방법을 개발하고 있다. 여기서 인프라는 국가데이터 표준에 일관된 데이터를 수집하는 산업별 협정(national agreement)과 국가위원회 검토를 바탕으로 하는 데이터표준용 승인과정(approval process), 승인된 데이터표준의 정기 간행물 등이 해당된다.

### 3) 지표(Indicator)

METeOR의 지표 개발은 호주인의 건강과 복지에 관한 국가정보의 품질과 적합성, 일관성, 가용성을 개선하는데 목적이 있다. 또한 METeOR의 지표 개발은 통계에서 사용된 데이터가 국가적으로 그리고 국제적으로 비교 가능하게 하면서 상호운용성을 촉진하는데도 있다.

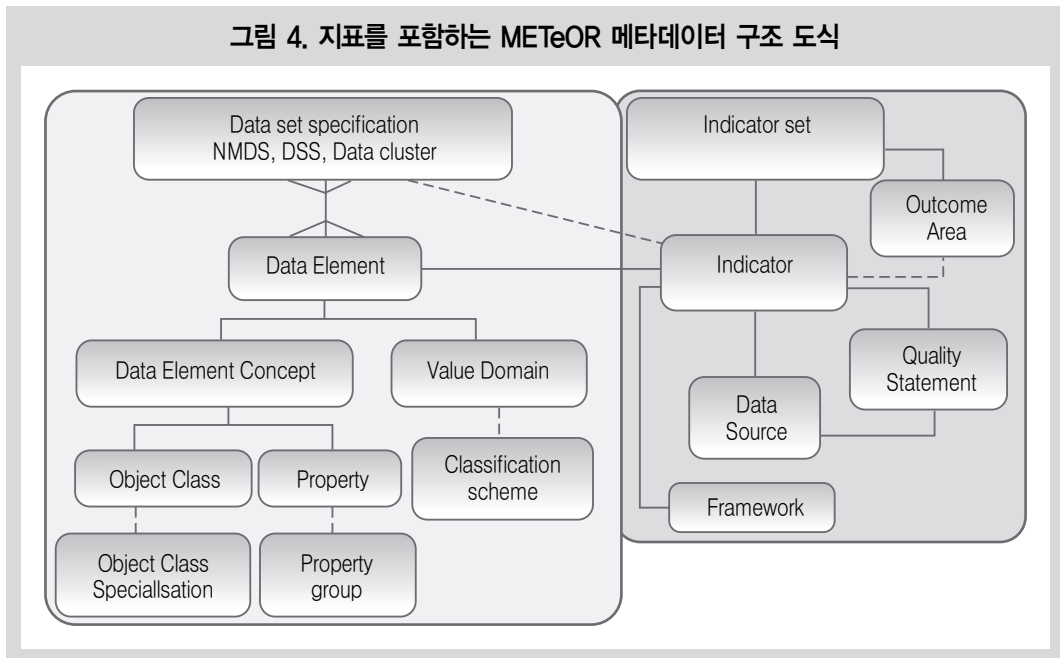
METeOR에는 지표용 메타데이터, 성과 영

역, 지표, 데이터 소스, 품질서, 프레임워크 등 6가지 형태의 지표 항목이 있다. 여기서 지표들은 분모와 분자처럼 데이터 요소와 데이터 세트에 연결되어 있다(그림 4 참조).

지표들은 관련된 등록기관에 의해 표준으로 승인된다. 지표들은 지표서식 모음을 구성하는 즉, 지표 세트(Indicator set), 성과 분야(Outcome area), 지표(Indicator), 품질서(Quality statement), 데이터 소스(Data source) 등 5가지의 형태로 이루어진다.

여기서 지표 세트는 성과지표나, 또는 성능 기준에 대응하는 것들을 사용해서 측정된 전체 목표들과 결과들을 정의하며, 특정 서비스부문이나 또는 명확히 정의된 지배 구조가 가지는 펀딩모델에 기초한다. 지표 세트의 목적은 서비스 제공의 개선을 측정하고 달성하는 것이

그림 4. 지표를 포함하는 METeOR 메타데이터 구조 도식



다. 지표 세트들은 3 가지 속성을 가지는데, 지표 세트 타입, 지표세트와 연결된 성과 영역들, 그리고 이 지표세트와 연결된 지표들이다. 성과 영역에서 성과는 전략적이면서 높은 수준이고 관측 가능해야 하며, 명확히 표현되어야 하고, 측정 및 달성 가능한 용어들이어야 한다. 성과영역은 METeOR에서 2 가지 속성을 가진다. 속성 중 하나는 성과영역과 연결된 지표 세트들과 그리고 다른 하나는 성과 영역과 연결된 지표들이다.

지표(Indicator)는 어떤 사회 관점의 상태를 주장하는, 예를 들어 건강지표, 지역사회에서 발생하는 질병 지표가 되는 특정질병에 대한 백신 접종자수와 같은 통계들과 통계에 대한 설명을 말한다. 지표는 30개의 속성을 가진다. 여기에는 지표 유형(Indicator Type), 지표 이름(Indicator common name), 설명(Description), 근거(Rationale), 지표 세트(Indicator set), 성과 영역(Outcome area), 인구집단연령(Population group age from), Population group age to, 계산 방식(Computation description), 공식(Formulae), 분자(Numerator), 분자 항목(Numerator items), 분모(Denominator), 분모 항목(Denominator items), 세분화 항목(Disaggregation items), 데이터 요소(Data element/Data set), 데이터소스(Data source), NMDS/DSS, 수집방법/사용방법(Collection Methods/Guide for use), 프레임워크 및 차수(Framework and dimensions), 방법론(Methodology), 보고요구사항(Reporting requirements), 데이터 제공 담당 조직(Organisation responsible for providing data), 책임(Accountability), 벤치마크(Benchmark), 국제

비교(International comparisons), 필요한 추가 데이터 개발/수집(Further data development/collection required), 다른 이슈 주의사항(Other issues caveats), 공개날짜(Release date) 등이 포함된다.

품질서는 지표나 데이터소스를 보고함에 있어서 이것들의 데이터 품질을 평가할 목적을 가진 다중의 품질 범위 문서이다. 이 항목에 들어 있는 필드들은 ABS Quality 신고서에 기초하며, 또한 지표나 데이터소스와 연결될 수 있다. 품질서는 데이터소스나 지표와 함께 사용하는데 의무적이지는 않지만 상당히 권장하고 있다. 즉, 품질서가 연결된 지표는 METeOR 사용자에게 더 확실한 정보를 제공하기 때문이다.

마지막으로 데이터소스는 데이터의 원천이 되는 곳의 특수한 데이터 집합과 데이터베이스, 그리고 레퍼런스이다. 데이터소스는 METeOR 담당자만이 생성하고 수정할 수 있다.

### 3. 시사점

호주 정부는 국가정보의 철저한 관리와 운영 지침을 유지하기 위해 많은 노력을 기울이고 있다. 특히, 전문가 위원회, 데이터 위원회, 등록기관 등을 두고, 체계적으로 역할을 구분 짓고, 정해진 프로세스에 따라 국가통계와 관련된 메타데이터 관리업무가 추진된다. 뿐만 아니라 통계의 기초가 되는 항목, 즉 메타데이터를 체계적으로 관리하는 METeOR도 운영하고 있다. 호주 보건복지연구소는 METeOR의 전반적인 운영 주체이다. 데이터 표준은 항목에 대한 모호성을



배제함으로써 사용자에게 관련 정보를 제공하여 중복업무의 수고스러움을 감소시킬 뿐 아니라 국가적으로 통일된 정보를 생산하는데 기여한다. 호주보건복지연구소는 이러한 데이터 표준 업무를 주도적으로 수행하고 있고, 이용자에게 관련된 정보도 공개적으로 서비스하고 있다. 특히 메타데이터 표준 및 지표와 관련하여 정교한 아키텍처를 설계해 구조화된 세부 항목으로 관리하고 있어, 과학적, 체계적인 운영이 이루어지고 있음을 다시금 확인할 수 있다.

우리나라는 동일한 조사항목 또는 통계항목이라도 기준과 규격, 정의에 따라 약간의 차이를 보이는 사례가 상당히 있다. 이렇다 보니, 국내통계에서도 유사통계를 비교하는 것이 쉽지 않다. 예컨대 사회보장기본법 개정에 따라 2013년부터 사회보장통계의 생산관리가 시행되고 있다. 사회보장통계는 다부처에 걸쳐서 생

산되고 있고, 해당 부처 및 지자체는 통계의 표준화된 규격 없이 필요에 따라 통계생산자의 역량에 좌우되어 생산·관리되었다고 봐도 과언이 아니다. 이런 현상으로 인해 통계간 연계나 통합이 쉽지 않고, 이용자 측면에서는 기존의 모호성으로 인해 통계이해의 혼란을 야기시킬 뿐 아니라 관리측면에서도 유사 통계임에도 불구하고 이중삼중의 수고를 해야 하는 문제가 발생한다.

따라서 우리나라도 정교하고 정확한 통계의 생산을 위해서는 보건복지관련 메타데이터의 체계적 관리와 더불어 METeOR 같은 시스템의 운영이 절실하다. 또한 이런 역할을 담당할 기관을 두는 것이 우선적으로 필요한데, 무엇보다 보건복지분야를 정확히 아우르는 기관이 그와 관련된 메타데이터 표준을 정의하고 개념화, 구체화하는 역할을 담당함이 바람직하다. 