

구글 클라우드 기반 FHIR 정보 관리

조하람, 김민주, 이정훈
제주대학교 데이터사이언스학과
{baekhyunexoh, full070}@naver.com, jhlee@jejunu.ac.kr

FHIR information management on Google Cloud

Haram Cho, Minju Kim, Junghoon Lee
Dept. of Data Science, Jeju National University

요약

본 논문에서는 구글 클라우드의 Healthcare API의 의료정보 표준에 따라 정의된 객체들을 저장하는 과정을 수행하고 활용방안을 설계한다. 클라우드에서 프로젝트 생성, API 활용 설정, 인증, 데이터셋 생성 등의 과정을 거쳐 cloud shell을 통해 curl 명령을 수행함으로써 Json으로 표현된 의료자원들에 대해 CRUD 연산을 수행할 수 있다. 무결성 체크 과정에서 기본 객체의 id 필드 명세가 중요하며 데이터 저장시 객체에 구글 클라우드의 특정한 request 속성추가가 필요하다.

1. 서론

최근 표준화된 의료정보에 의한 의료기록의 공유와 이에 따른 비용 절감에 대한 관심들이 많아지면서 구글, 아마존, 마이크로 소프트 등에서도 이에 관련된 솔루션들을 제공하고 있다. 이 의료정보 공유를 위해서는 복잡한 의료 자원의 모델링, 데이터의 저장, 그리고 효율적인 교환 등이 필요하다. 먼저 의료 정보에 대한 표준화된 모델을 만들기 위해 다년간 많은 노력이 있었으나 통일하기 어려운 의료정보의 특성상 뚜렷하게 통용되는 표준이 없었다. 그러나 HL7의 FHIR(Fast Healthcare Interoperability Resource)를 중심으로 의료정보 표준이 많은 이해당사자들에 의해 수용될 가능성이 높아지고 있으며 특히 미국에서의 도입이 활발하다 [1].

FHIR에서는 환자(Patient), 의료진(Practitioner), 의학 진료(Encounter), 검사(Observation), 상태(Condition) 등을 정의하고 있으며 이를 각각은 필요한 속성들을 포함하고 있다. 정보 모델이 구축되면 이들이 안전하고 효율적인 통신 메커니즘을 통해 인가된 환자 개인, 병원, 보험회사 등 다양한 당사자들에게 공유될 수 있는데 이 과정에서는 웹 기술을 기반으로 하여 REST API를 통해 데이터의 교환이 이루어진다. FHIR 모델은 의료정보를 교환하는 방식이며 정보의 저장은 각 기관의 저장방식에 의해

결정된다. 별적인 서버나 클라우드의 관계형 데이터베이스일수도 있고 FHIR 모델의 데이터 모델 정의를 따르는 저장 시스템일 수도 있다. 전자의 경우는 저장과 통신 사이에서 변환하는 과정이 필요하며 각 의료 관련 기관마다 저장방식이 다르기 때문에 고비용의 변환 과정이 필요할 것으로 보인다.

이와 아울러 이미 오픈소스나 상용화된 FHIR 객체 저장 서버가 제공되고 있는데 HAPI 서버의 경우는 웹을 통해 FHIR 객체를 생성하고 검색해볼 수 있는 테스트 기능을 제공하고 있다 [2]. FHIR 객체들은 기본적으로 JSON 형태가 가장 널리 사용되고 있으며 객체 생성을 지원하는 Fred와 같은 도구들도 있다. 또 Logica Sandbox는 의료관련 응용을 개발할 수 있는 플랫폼을 제공하고 있으며 데이터의 무결성 지원과 데이터 접근에 있어서의 Oauth 기반 인증을 제공한다. 클라우드들도 의료정보 저장과 활용에 많은 관심을 가지고 있다 [3]. 특히 구글 클라우드는 의료정보를 FHIR 기준에 따라 저장할 수 있는 API를 제공하고 있는데 이 저장된 의료정보에 기존의 Machine Learning과 Big Query 기능을 제공한다.

2. 의료정보의 구글 클라우드 저장

구글 클라우드에서 FHIR 객체를 저장하려면 <그림 1>과 같이 먼저 클라우드 프로젝트를 생성하고 Healthcare API 사용 설정을 하는데 이 과정은

별도의 크레딧이 필요하다. 다음 과정에서는 Oauth 설정으로 데이터에 대한 접근 범위를 결정하며 이후 데이터 세트와 데이터 저장소를 생성하며 이 스토리지 객체를 선택함으로써 현재 저장된 FHIR 객체를 조회할 수 있다. 저장된 객체들에 대해서는 몇 가지 접근 방식이 제공되고 있으나 Cloud shell을 이용하면 Unix 계열의 인터페이스에 의해 데이터 조작이 가능하다.



<그림 1> FHIR 저장소 활성화 과정

<그림 2>는 cloud shell에서 curl 명령을 수행하여 Healthcare 레코드를 조회한 결과이다. curl 명령어로 CRUD 기능을 모두 수행할 수 있는데 URL을 구글 클라우드에서 정의된 양식으로 명시하여야 한다. <https://healthcare.googleapis.com/v1/projects/myproject/locations/asia-northeast3/datasets/mydataset/fhirStores/mystorage/fhir/targetresource> 와 같은 URL에서 이탤릭 부분은 이전 과정에서 생성한 데이터 세트, 데이터 저장소 등을 지정하고 있다. 또 -H 옵션 뒤에 나오는 헤더 부분에서는 인증 키와 콘텐츠 타입 등을 명시하며 Json 등으로 의료자원들을 기술하고 있는 파일을 명세할 수 있다. 또, 이 파일은 클라우드에 접근하는 일반 PC 등에서 작성하여 cloud shell로 업로드할 수 있다.

```

[hleetwins@cloudshell:~ (fhirfhir)$ curl -O https://cloud.google.com/healthcare-
% Total    % Received % Xferd  Average Speed   Time   Time  Current
          Dload Upload Total Spent   Left Speed
100  203 100  203    0     0  835      0 --:--:-- 0:00:00  835
[hleetwins@cloudshell:~ (fhirfhir)$ ls
Patient.json README-cloudshell.txt
[hleetwins@cloudshell:~ (fhirfhir)$ more Patient.json
{
  "name": [
    {
      "use": "official",
      "family": "Smith",
      "given": [
        "Darcy"
      ]
    }
  ],
  "gender": "female",
  "birthdate": "1970-01-01",
  "resourceType": "Patient"
}
  
```

<그림 2> cloud shell과 FHIR 자원 접근

FHIR에서 가장 기본적인 자원인 Patient는 환자 개인에 대한 정보를 가지고 있는데 이 Patient 객체

를 Healthcare 클라우드에 저장하면 자동적으로 022e654d-c9b0-4407-b8df-dc9968027bcf와 같은 형태의 고유 ID가 부여된다. 이후 Observation 등의 다른 자원에서 환자를 reference 할 때는 이 ID를 사용해야 하며 데이터의 무결성을 강하게 체크하기 위해 존재하지 않는 Patient ID를 reference에 포함하고 있는 경우는 데이터 저장이 거부된다. 또 자원의 속성 필드에 유효하지 않은 값이 지정되어 있어도 데이터 저장이 거부된다.

<그림 3>은 하나의 자원이 아닌 복수의 자원을 동시에 입력하는 Bundle을 한 번에 저장한 결과를 보이고 있는데 여기서 객체의 resourceType, type, id 필드의 값을 각각 "Bundle", "transaction", "bundle-transaction" 등으로 변환하고 request 속성을 생성하여 method와 URL 속성에 POST와 자원의 타입을 명세해야 한다.

ID	First Name	Last Name	Gender	Birth Date
MR-44240387	Westley	Kennedy	male	1919-08-23
MR-59064677	Isiah	Byrd	unknown	1960-07-21
MR-44240387	Westley	Kennedy	male	1919-08-23
MR-25212979	Lori	Kelley	female	1901-10-04
MR-59064677	Isiah	Byrd	unknown	1960-07-21
MR-057316140	Eleanor	Vaughn	male	1956-01-29
MR-775941672	Roger	Hunt	unknown	1969-12-28
MR-44240387	Westley	Kennedy	male	1919-08-23
MR-59064677	Isiah	Byrd	unknown	1960-07-21
MR-775941672	Roger	Hunt	unknown	1969-12-28

<그림 3> Patient 객체의 저장 결과

3. 결론 및 추후 과제

FHIR 정의에 의한 의료정보 공유는 지속적으로 활성화되고 클라우드에 저장 및 인공지능을 이용한 분석 등이 확대될 것으로 예상된다. 추후 구글 클라우드에 심전도와 관련된 많은 양의 데이터를 저장하고 Big Query 등을 통해 환자의 상태 예측과 사전 이상 발견 등의 응용을 개발할 예정이다.

Acknowledgment

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학지원사업의 연구결과로 수행되었음”(2018-0-01863)

참고문헌

- [1] <https://www.hl7.org/fhir/>
- [2] <https://hapi.fhir.org/>
- [3] <https://cloud.google.com/healthcare-api?hl=ko>