

Semi-Automated Method to Generate Simulated Clinical Data from OpenEHR Platform – Think!EHR

Abdul Mateen RAJPUT^{a,1}

^aKöln University Hospital, Köln, Germany

Keywords. OpenEHR, Think!EHR, KNIME, ETL Process

1. Introduction

OpenEHR is openly available and vendor independent technology based on two-level modelling; Reference Model and Archetype Model [1]. Think!EHR is an OpenEHR platform used by many HiGHmed partner sites, it offers clinical data storage, management, querying, retrieval and exchange.

To work with Think!EHR, Templates have to be developed by combining different Archetypes. Examples of Archetypes are Blood pressure and Body temperature, a Template example would be Vital Signs. Think!EHR provides an example of data values, once a Template is uploaded to the platform or the cloud instance i.e. EHRscape.com.

Since our aim was to generate data from all the Templates and of numerous patients, an automated approach was needed. Our tool of choice was KNIME (Konstanz Information Miner) [2], which is a freely available software. It has many dedicated nodes which have predefined functionalities and readily available nodes for different tasks.

2. Methods

Our environment was Think!EHR (Ver. 2.45) running on a Windows machine. In our example we used Localhost because the system was installed locally. The same method can be applied by using Ehrscape REST API URL [3]. However, one needs to change the base URL accordingly as we used localhost:8081 for this paper.

Following steps have to be performed to generate and retrieve data of a patient:

1. First step was to POST Operational Template in OPT format to Think!EHR:
 - Post <https://localhost:8081/rest/v1/template>
2. This Get method retrieves the complete list of Templates available Think!EHR:
 - Get <http://localhost:8081/rest/v1/template>
3. To retrieve example patient data based on template (step 1), following link was used. Actual OPT can also be retrieved, by replacing "example" with "opt":
 - Get [http://localhost:8081/rest/v1/template/\[TemplateId\]/example](http://localhost:8081/rest/v1/template/[TemplateId]/example)

¹ Corresponding Author, Abdul-Mateen Rajput, Universitätsklinikum Köln, Kerpener Str. 62, 50937 Köln, Germany; E-mail: Abdul.mateen@uni-koeln.de.

To avoid manual querying, repeating step 3, for hundreds of patients' data, the following workflow (shown in figure 1) has been developed to automate the process.



Figure 1. This workflow generates data of 10 patients based on the Template Pankreaskarzinom. From left, the first node creates an empty table. Second node starts the loop and numbers of loop can be set in the setting, in our case it was 10. Third node uses the link mentioned in text in step 3 and retrieve the data of a patient. Fourth node checks the conditions whether number of loops have been executed and collects the data. Once all the loops are executed, it ends the loop. Fifth node write the JSON files to given directory.

3. Results

The following link was used to retrieve the data of 10 patients based on "Pankreaskarzinom" Template:

- Get <http://localhost:8081/rest/v1/template/Pankreaskarzinom/example>

Figure 2 shows the output of the workflow, where each line represents a new record based on the Template as discussed earlier. The same can be done with all other Templates with just an additional node which loops over the name of Templates.

```

"laborbefund_pankreaskarzinom/context/bericht_id": "Bericht ID 95", "laborbefund_pankreaskarzinom/context/status_des_auftrages": "Status des Auftrages 31",
"laborbefund_pankreaskarzinom/context/bericht_id": "Bericht ID 96", "laborbefund_pankreaskarzinom/context/status_des_auftrages": "Status des Auftrages 78",
  
```

Figure 2. The output of the workflow, where each row contains a record of a patient with the Template "Laborbefund_Pankreaskarzinom".

4. Discussion and Conclusions

The approach presented in this paper shows the possibility to generate clinical data from Think!EHR. The approach is based on REST API methods and the analytical tool KNIME was used to automated the data retrieval process.

Data generated with this approach is not only compatible to all OpenEHR system but it has also appropriate data types and values representing the real patient data. This would reduce the need of real patient data if one needs to test the health IT systems.

Acknowledgment

The project is funded by the German Federal Ministry of Education and Research (BMBF, grant id: 01ZZ1802U).

References

- [1] Min L, Tian Q, Lu X, Duan H. Modeling EHR with the openEHR approach: an exploratory study in China. BMC Medical Informatics and Decision Making. 2018;18:75.
- [2] KNIME | Open for Innovation. <https://www.knime.com/>. Accessed 4 Mar 2020.
- [3] API Explorer. <https://www.ehrscape.com/api-explorer.html>. Accessed 31 Mar 2020.