Usability Testing of a Multi-Level Modeling Framework for Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health

Sebastian GRUBER<sup>a,b,c,1</sup>, Bernd NEUMAYR<sup>c</sup>, Daniela WURHOFER<sup>a</sup> and Jan David SMEDDINCK<sup>a</sup>

<sup>a</sup>Ludwig Boltzmann Institute for Digital Health and Prevention, Salzburg, Austria
<sup>b</sup>Salzburg Research Forschungsgesellschaft, Salzburg, Austria
<sup>c</sup>Johannes Kepler University Linz, Linz, Austria

Abstract. The JITAI is an intervention design to support health behavior change. We designed a multi-level modeling framework for JITAIs and developed a proof-of-concept prototype (POC). This study aimed at investigating the usability of the POC by conducting two usability tests with students. We assessed the usability and the students’ workload and success in completing tasks. In the second usability test, however, they faced difficulties in completing the tasks. We will work on hiding the complexity of the framework as well as improving the frontend and the instructions.

Keywords. Semantic Web, Artificial Intelligence, Data Analysis, Data Aggregation

1. Introduction

The just-in-time adaptive intervention (JITAI) is a design for a digital health intervention with great potential to contribute to sustainable behavior change for individuals [1]. It combines real-time data and knowledge about an individual to provide personalized interventions using artificial intelligence. For example, a particular JITAI may provide exercise suggestions while waiting for a bus and considering individual preferences.

In JITAI research, different systems for implementing JITAIs store data according to different meta-models. Therefore, data analysis across studies is difficult and requires prior data integration. In previous work [2], we proposed the design of a framework that combines multi-level modeling (MLM), composition hierarchies and specialization hierarchies to eliminate the need for data integration. However, we also pointed out the complexity of the framework and the size of models created with the framework.

The aim of this study was to investigate the usability of a proof-of-concept prototype (POC). The POC builds on JavaFX and standardized semantic web technologies, such as the Resource Description Framework (RDF), the SPARQL Query Language and the Shapes Constraint Language (SHACL). The POC does not hide the complexity of the framework but provides means to navigate through the structure of a model (e.g., providing a tree view) and to support working with a model (e.g., providing input forms to create new model elements).

<sup>1</sup> Corresponding Author: Sebastian Gruber, Ludwig Boltzmann Institute for Digital Health and Prevention, Salzburg, Austria, E-Mail: sebastian.gruber@dhp.lbg.ac.at
2. Methods

We conducted two usability tests with students from a lecture course on “Semantic Artificial Intelligence” held at the Johannes Kepler University Linz in 2022.

In the first usability test, we introduced JITAIs and MLM with a basic version of the POC by asking students to complete eight tasks (e.g., writing SPARQL queries over a given model) within a week using a provided tutorial.

In the second usability test, we presented the framework and the full version of the POC with a 30-minute presentation and asked students to complete 10 tasks (e.g., creating a model from scratch and writing SPARQL queries) within 60 minutes.

We invited the students to complete a survey after each usability test. They rated the items of the Usability Metric for User Experience (UMUX) [3] and the subscales of the NASA Task Load Index (TLX) [4] and submitted their task solutions. The full set of materials used is available at: https://github.com/LBI-DHP/dHealth23.

3. Results

We assessed the overall scores of the UMUX (0 – very low, 100 – very high) and unweighted/raw TLX (0 – very low, 20 – very high). We evaluated the submissions of the students to assess the relative number of tasks the average student completed successfully, failed to complete or did not even attempt. Table 1 summarizes these results.

Table 1. Summary of the survey responses for both usability tests. M = mean, SD = standard deviation.

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>UMUX [0, 100]</th>
<th>Raw TLX [0, 20]</th>
<th>Tasks [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
<td>76.0</td>
<td>19.8</td>
<td>9.6</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>67.4</td>
<td>13.9</td>
<td>10.1</td>
</tr>
</tbody>
</table>

4. Discussion

The first usability test only served as a preparation to reduce the cognitive burden on the students in the second usability test. In the second usability test, however, the students faced difficulties in completing tasks and we had to provide additional help. These difficulties may be due to the inherent complexity of the framework, accidental complexity of the frontend, inadequate instruction on the POC, and/or lack of understanding of the framework. We will therefore work on hiding the complexity of the framework and improving the frontend and the instructions (e.g., developing guidelines).

References