Persona Modeling Process
From Microdata-based Templates to Specific Web Ontologies

Stefan Negru and Sabin Buraga
Faculty of Computer Science, Alexandru Ioan Cuza University of Iasi, General Berthelot, 16, Iasi 700483, Romania

Keywords: Personas, Knowledge Modeling, Semantic Web, Ontology, Microdata, Human-Computer Interaction.

Abstract: The use of Personas method for communicating user requirements in Human-Computer Interaction is well established and widely used. However, little research has been conducted regarding linking personas data with other pieces of data concerning the product development and design process or even creating a personas model. We consider that using semantic web technologies such as microdata and RDFa to annotate personas data is an adequate step towards solving these issues. This paper aims to provide a conceptual model (a HTML5 microdata schema and an OWL specification) that includes concepts and properties used to model personas. In order to exemplify our model and extract data, we made use of a semantically annotated persona template.

1 INTRODUCTION

The Personas method was proposed by Cooper (Cooper, 1999; Cooper et al., 2007) and has gained increasing attention in the Human-Computer Interaction (HCI) field, as it facilitates the communication of information about target users among the project team members (Pruitt and Adlin, 2010), but also provides a greater understanding of the main audience. Considering these advantages, many companies such as Microsoft, SAP, FedEx have adopted personas as a part of their design process (Blomquist and Arvola, 2002; Chapman and Milham, 2006; Drego and Dorsey, 2010; Drego et al., 2010; Browne, 2011).

A persona is regarded as a user archetype which can be used to “help guide decisions about product features, navigation, interactions, and even visual design” (Goodwin, 2005). Being an archetype, a persona represents a group of users who share common behavioral or physical characteristics, goals, frustrations, preferences and other similar characteristics. Even though a persona represents a group of users, a fictional individual is created, based on user interviews, to represent specific aspects of that group (Pruitt and Adlin, 2010).

Despite the popularity of this method, there are several debates (Long, 2009) regarding the implementation of this technique as many authors advocate distinct differences in the way it can be used. Also “there have been no adequate studies addressing the reliability, validity, or utility of the method” (Chapman and Milham, 2006), and – even if methodological problems were overcome –, this technique suffers certain practical limitations: “how personas are reconciled with other information, and who is responsible for interpreting them.” (Chapman and Milham, 2006). Additionally, validating the obtained information can be time consuming and troublesome, as the personas profiles are validated manually against online user profiles or through the customer service team (Caddick and Cable, 2011).

Ontologies have proven to be a very useful tool in modeling data concepts and relationships between them, by providing semantics for humans and formalism for machine processing and reasoning. Also, microformats1, RDFa2 and microdata3 enable the publication and consumption of the data available on the Web – particularly, using the HTML54 vocabulary. In this paper we take into consideration such semantic web technologies, as a first step towards solving the issues presented above, and opening new possibilities such as automatically validation and generation of personas based on online user profiles. We propose PersonasOnto ontology as a way of modeling personas related concepts and their relationships.

1http://microformats.org/
2http://www.w3.org/TR/rdfa-core/
3http://www.w3.org/TR/microdata/
4http://www.w3.org/TR/html5/
The next section of the paper offers a review of related work. We continue with a discussion regarding several aspects of the personas method. Section 4 presents the personas knowledge modeling, and section 5 illustrates several use cases. The article ends with conclusions and future directions of research.

2 RELATED WORK

The interest on describing persons and user profiles has a wide spread as several vocabularies such as FOAF\(^5\), SIOC\(^6\) and Person\(^7\) offer facile manner to publish information in a machine-readable format. Also, as the need for software system (Razmerita et al., 2003; Negru, 2010; Cheng et al., 2011) to automatically adapt to their users has increased, a number of domain specific ontologies concerning user profiles have been developed.

Such ontologies – as (Felden and Linden, 2007; Cena et al., 2011) – explore the implications of user models in the context of recommender systems. Also, (Razmerita, 2007) presents a generic ontology-based user modeling architecture applied in Knowledge Management Systems. Other related work such as (Golemati et al., 2007) provides a standard ontology for modeling user profiles with the aim of facilitating communications between applications. A related ontology is AEGIS ontology\(^8\) which helps the mapping between accessibility concepts, and how they can be mapped within accessibility scenarios.

As most of the existing user profiling ontologies the related work presented above emphasizes the importance of context awareness. This aspect and other relevant user modeling issues are reviewed in (Fischer, 2001).

Although seemingly overlapping concepts personas and user profiles, they are different. Personas represent a group of users who share characteristics and preferences, while user profiles (also known as user models) are a collection of personal data associated to specific users or stereotypes. Despite these differences, certain aspects regarding user modeling techniques can be transposed on the personas method.

In the next section we will consider a few aspects regarding the personas method.

3 PERSONAS METHOD

Unlike the target audience segmentation or user profiling, personas provide greater depth and context to generic target audience groups by focusing on one archetype which embodies the predominant qualities of the larger group. These qualities are often structured by a persona document (specification) (Long, 2009; Pruitt and Adlin, 2010; Caddick and Cable, 2011).

The data included into this document is a result of user interviews and it is distilled into one or multiple fictitious characters which corresponds to a certain persona type (Cooper et al., 2007). In this document, each character is developed in greater detail, along with scenarios which describe how that character might interact with a product. The personas and their associated scenarios form the basis for specifying how users want to experience and interact with a certain product or application.

3.1 Personas Document Template

As an important part of this method is the persona document and the information it contains, we developed our own HTML5 template\(^9\) (Figure 1), following existing guidelines (Long, 2009; Pruitt and Adlin, 2010; Caddick and Cable, 2011).

Figure 1: Overview of the personas HTML5 template.
2010; Caddick and Cable, 2011). Being a HTML document which contains information semantically annotated by using microdata/RDFa, it facilitates publication, consumption, and reuse of information.

The template is structured as follows:

1. Type and Background Information – contains details regarding the persona type and fictional background information such as birthday, name, gender, location and other information of interest;
2. Main Details – consists of a detailed description regarding this persona along with some characteristics (main points), goals and frustrations;
3. Scenarios – includes either a description of scenario tasks or images depicting scenarios;
4. Other Details – contains other relevant notes regarding the persona.

In the next section we will discuss how we can make use of this HTML template and annotate relevant data using existing vocabularies and our proposed vocabulary.

4 PERSONAS ONTOLOGY ENGINEERING

As we previously mentioned, several existing widely used vocabularies like as FOAF and schema.org/Person could be utilized, thus we applied such vocabularies to annotate information in our HTML template.

In a small scale experiment, we provided our template to several teams working on different projects to fill up with information. Using existing tools – such as RDFa Distiller\(^\text{10}\) and Microdata to RDF Distiller\(^\text{11}\) – we extracted the data presented below, from a project’s personas document. This data consists of RDF triples expressed in Turtle format.

```turtle
[ a schema:Person;
  schema:givenName "Jane"@en-us;
  schema:familyName "Doe"@en-us;
  schema:gender "Female"@en-us;
  schema:birthDate "1989-05-11"^^xsd:date;
  schema:image <student.jpg>;
  schema:location [ a schema:Place;
    schema:name "City,Country"@en-us ];
  schema:worksFor [ a schema:Organization;
    schema:name "Univ. City,Country"@en-us;
    schema:name "Univ. City,Country"@en-us ];
  schema:alumniOf [ a schema:EducationalOrganization
    schema:name "Univ. City,Country"@en-us ]
] 
```

The data obtained is unsatisfactory as it omits several important pieces of data from the persona document. In order to solve this issue, we modeled our own Persona microdata schema.

4.1 Persona Schema

One advantage of the HTML5 microdata is that it is designed such that each piece of information in a document has assigned types from a single vocabulary, though each entity may have several types and properties from other vocabularies (W3C, Interest Group, 2012). Additionally, the Person schema offers properties such as affiliation, interactionCount (count of a specific user interactions with an item), performerIn, spouse, awards etc. which facilitates to establish certain knowledge regarding the user background and context.

Another important aspect is that schema.org allows the extension of existing schema, and we made use of it in this section by developing the Personas schema\(^\text{12}\). As represented in Table 1, we proposed a set of new properties such as personaType, tagLine, mainPoint, frustrationPoint, endGoal, scenario, context and other. Some properties like myersBriggs and topicInterest were inspired from the FOAF vocabulary.

Properties like minHeight, maxWeight, minFeetSize refer to a persona body measurement such as height, weight, bust size, waist size, and feet size. We used min and max values instead of a range, due to the fact that they are more precise than a range.

Along with some of the new properties, we proposed several additional schemas like Disability, Personality, Emotion, Scenario and Context, in order to provide a more detailed vocabulary which fits our purposes. These schemas themselves come along with new properties and a range which specifies expected data types.

The Scenario schema has the following new properties and expected data types:

- **product**: A product (subject of a given usability test); for example, a software application or even a physical product such as mobile phone or car. Expected Type: Product or Application;
- **scenarioName**: Scenario name. Expected Type: Text;
- **userTask**: Tasks to be performed by the user. Expected Type: Text;

\(^\text{10}\)http://www.w3.org/2012/pyRdfa/
\(^\text{11}\)http://www.w3.org/2012/pyMicrodata/
\(^\text{12}\)publicly available at: http://blankdots.com/open/schema/
Table 1: Persona Schema Properties and Description.

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>personaType</td>
<td>Text</td>
<td>The type of Persona: Primary, Secondary, Negative, Suplemental, Served or Customer.</td>
</tr>
<tr>
<td>tagline</td>
<td>URL or Text</td>
<td>A tagline specific to a persona.</td>
</tr>
<tr>
<td>mainPoint</td>
<td>Text</td>
<td>Main points specific to the persona category.</td>
</tr>
<tr>
<td>minHeight, maxHeight</td>
<td>Float</td>
<td>Minimum and maximum height of a persona.</td>
</tr>
<tr>
<td>minWeight, maxWeight</td>
<td>Float</td>
<td>Minimum and maximum body weight of a persona.</td>
</tr>
<tr>
<td>minBustSize, maxBustSize</td>
<td>Float</td>
<td>Minimum and maximum bust size of a persona.</td>
</tr>
<tr>
<td>minWaistSize, maxWaistSize</td>
<td>Float</td>
<td>Minimum and maximum waist size of a persona.</td>
</tr>
<tr>
<td>minFeetSize, maxFeetSize</td>
<td>Float</td>
<td>Minimum and maximum feet size of a persona.</td>
</tr>
<tr>
<td>experienceGoal</td>
<td>Text</td>
<td>Experience goals are simple, universal, and personal.</td>
</tr>
<tr>
<td>endGoal</td>
<td>Text</td>
<td>End goals represent the user’s motivation for performing the tasks associated with using a specific product.</td>
</tr>
<tr>
<td>lifeGoal</td>
<td>Text</td>
<td>Life goals represent personal aspirations of the user that typically go beyond the context of the product being designed.</td>
</tr>
<tr>
<td>businessGoal</td>
<td>URL or Text</td>
<td>Business goals represent the goals of the organization the persona works for.</td>
</tr>
<tr>
<td>technicalGoal</td>
<td>Text</td>
<td>Technical goals reflect technical aspects regarding an application/product for example: run in a variety of browser, data privacy etc.</td>
</tr>
<tr>
<td>experienceLevel</td>
<td>Text</td>
<td>Reflects the experience level of a persona with the application or product; proposed levels: Beginner, Intermediate, Advanced.</td>
</tr>
<tr>
<td>technicalLevel</td>
<td>Text</td>
<td>Reflects the technical level of a persona.</td>
</tr>
<tr>
<td>disability</td>
<td>Disability or Text</td>
<td>A persona disability/disabilities relevant to accessibility aspects of the application/product.</td>
</tr>
<tr>
<td>myersBriggs</td>
<td>Personality or Text</td>
<td>Inspired by FOAF Myers Briggs personality classification which includes 16 4-letter textual codes (Myers et al., 1998).</td>
</tr>
<tr>
<td>topicInterest</td>
<td>Thing</td>
<td>A thing of interest for a certain persona, inspired by FOAF topic interest.</td>
</tr>
<tr>
<td>affectiveState</td>
<td>Emotion or Text</td>
<td>The affective state of the user at a certain moment, if a proper schema is not used, please provide a certain emotion using text format.</td>
</tr>
<tr>
<td>frustrationPoint</td>
<td>Text</td>
<td>Elements of a UI/UX or certain characteristics that frustrate the user or (s)he sees as pain points. These elements will help in usability evaluation.</td>
</tr>
<tr>
<td>userRole</td>
<td>Text</td>
<td>The role of the user in the application/product (if necessary - for example: admin, basic user etc.).</td>
</tr>
<tr>
<td>scenario</td>
<td>Scenario or URL</td>
<td>The scenario where the users represented by this persona will be used to test the usability.</td>
</tr>
<tr>
<td>context</td>
<td>Context</td>
<td>Description of the context that best fits a scenario recommended values: Physical/Virtual.</td>
</tr>
<tr>
<td>otherNotes</td>
<td>Text</td>
<td>Other notes and observations regarding a persona.</td>
</tr>
</tbody>
</table>

- **productTask**: Task performed by the product in response to the user. Expected Type: Text;
- **interactionMedium**: The medium of interaction; for example, Touch, Gestural, Mouse+keyboard etc. Expected Type: Text;
- **description**: Description of the scenario. Expected Type: URL or Text;
- **context**: Scenario context. Expected Type: Context;
- **participant**: Participants to this scenario. Expected Type: Person;
• **usabilityTest**: Usability test. Expected Type: UsabilityTest or Text.
• **evaluatedElements**: Focus/key elements in this scenario, tested later on, for example notifications element from the user interface. Expected Type: Text.

The Context schema has the following properties:
• **contextType**: Context type – recommended values: Physical/Virtual or Tangible/Intangible;
• **location**: Useful information about the location;
• **event**: An event has a location and a time;
• **geo**: The geographical coordinates of the context;
• **sensorData**: Sensors context data;
• **photos**: Photographs regarding the context;
• **reviews**: A set of reviews regarding the context.

Disability schema contains the following properties: visualImpairment, hearingImpairment, gustatoryImpairment, somatosensoryImpairment, intellectualImpairment, mentalEmotionalDisorder, developmentalDisability.

After modeling these schemas, we annotated the information in the HTML template using the above vocabularies and we extracted the data below.

```xml
[a <http://schema.org/Person/Persona>;
  schema:givenName "Jane"@en-us;
  schema:familyName "Doe"@en-us;
  schema:gender "Female"@en-us;
  schema:birthDate "1989-05-11"^^xsd:date;
  schema:image <student.jpg>;
  schema:location [ a schema:Place;
    schema:name "City,Country"@en-us ];
  schema:worksFor [ a schema:Organization;
    schema:name "Student"@en-us;
    schema:name "Univ. City,Country"@en-us ];
  schema:alumniOf [ a schema:EducationalOrganization;
    schema:name "Univ. City,Country"@en-us ];
  schema:jobTitle "Student"@en-us;
  schema:name "City,Country"@en-us ];
  schema:alumniOf [ a schema:EducationalOrganization;
    schema:name "Univ. City,Country"@en-us ];
  schema:personasType "Primary"@en-us;
  schema:mainPoint [ a rdf:Bag;
    rdf:_1 "Experience with Android OS;" @en-us;
    rdf:_2 "connected socially to most of her colleagues and some teachers."
  ];
  schema:technicalLevel "Medium"@en-us;
  schema:businessGoal "increase knowledge"@en-us;
  schema:lifeGoal "be productive"@en-us;
  schema:scenario [ a <http://schema.org/Scenario>; schema:description "she will check grades and progress online"@en-us ]]
```

The obtained data contains all the relevant information regarding a persona and could be processed and linked with other data in the spirit of Linked Data initiative (Bizer et al., 2009).

Although microdata is an easy way of annotating HTML document, it cannot express two aspects that RDFa supports: datatypes of literals and XML literals (Vestlandsforsking, 2012).

### 4.2 Personas Ontology Description

This subsection presents a description of the PersonasOnto ontology. PersonasOnto provides a mean of annotating XHTML and HTML documents with RDFa, but it can also be mapped in its RDF representation to HTML5 microdata.

Table 2 presents an overview of the proposed ontology classes and subclasses and Figure 2 provides a graphical representation of the ontology classes and properties.

The Person class is the central one in the ontology, as it contains characteristics such as age, gender, name, date of birth, but also the Persona class. As a subclass of Person, the Persona class could be regarded of being a separate identity of a person, although it has the same usage. Also, as we previously mentioned, a persona represents the characteristics of a group of persons, and usually has a context associated.

Classes such as AffectiveState, Personality, Disability, Organization are used to describe several characteristics or a person like an emotion (e.g. anger), a disability (such as partial blindness), his/her personality (e.g. ESTJ – Extraversion, Sensing, Thinking, Judgment) and the organization (s)he works for, or (s)he is in.

Other classes like Context, Scenario, Task, Goals, UsabilityTest and subclasses PersonaType are connected with the Persona subclass. For example, a persona has a certain type (Primary, Secondary, etc.), has the characteristics of a (fictional) person, has certain life or experience Goals, (s)he performs certain tasks in a scenario. The scenario is placed in a certain context and can be also used in a usability test.

Many of the classes, subclasses and properties are

---

13 publicly available at: http://blankdots.com/open/personasonto.owl

14 http://www.w3.org/TR/xhtml1/

15 http://www.w3.org/RDF/
<table>
<thead>
<tr>
<th>Class Name</th>
<th>Subclasses</th>
<th>Class Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AffectiveState</td>
<td>Emotion, Feeling, Mood</td>
<td>The affective state of a person at a certain point.</td>
</tr>
<tr>
<td>Context</td>
<td>Intangible, Tangible, Place, Time</td>
<td>The Context in which a scenario takes place.</td>
</tr>
<tr>
<td>Person</td>
<td>Participant, Persona</td>
<td>Basic information about a person and more precisely about a participant or a persona.</td>
</tr>
<tr>
<td>Personality</td>
<td>–</td>
<td>The MyersBriggs personality of a person.</td>
</tr>
<tr>
<td>Resource</td>
<td>Document, Image, Product</td>
<td>Type of resources available.</td>
</tr>
<tr>
<td>Scenario</td>
<td>–</td>
<td>A series of tasks the user performs in a certain context.</td>
</tr>
<tr>
<td>Task</td>
<td>InteractiveTask, NonInteractiveTask</td>
<td>Tasks performed by a Person in a Scenario. InteractiveTask could be certain action performed by a user, NonInteractiveTask could be a machine response to that action.</td>
</tr>
<tr>
<td>UsabilityTest</td>
<td>–</td>
<td>A usability test evaluates the usability of a certain Product in a Scenario.</td>
</tr>
</tbody>
</table>

the same as in the Persona schema described in the previous subsection. On the other hand, classes like AffectiveState and Personality have certain properties which better define the relationship between them.
Such a property is expressed below in Turtle\textsuperscript{16} format.

```turtle
### personasOnto.owl#hasInfluence
:hasInfluence rdf:type owl:ObjectProperty ;
rdfs:domain :Disability ;
rdfs:range [ rdf:type owl:Class ;
owl:unionOf { :AffectiveState ;Personality }].
```

The Context Class includes a tangible (physical world) or intangible (virtual world) environment, but also could specify spatial and temporal concepts. This inclusion is expressed below.

```turtle
### personasOnto.owl#includes
:includes rdf:type owl:ObjectProperty ;
rdfs:range [ rdf:type owl:Class ;
owl:unionOf { :Place ;:Time } ;
rdfs:domain [ rdf:type owl:Class ;
owl:unionOf { :Intangible ;:Tangible }].
```

Other classes like UsabilityTest can be extended to include domain specific concepts.

### USE CASES

In section 2 we mentioned a small scale experiment in which we provided our template to several teams working on different projects in order to fill up with information. The purpose of such a project was to build a web-based solution along with a mobile client, for the management of the academic information (e.g. grades, schedule, courses, assignments etc.).

The team responsible for this project identified in a first phase two primary personas: professors and students – each of them having their own goals, frustrations and set of characteristics. We presented the student persona in section 2 and the professor persona is exemplified below.

```turtle
[ a <http://schema.org/Person/Persona>;
schema:givenName "John"@en-us;
schema:familyName "Doe"@en-us;
schema:gender "Male"@en-us;
schema:birthDate "1980-10-23"ˆˆxsd:date;
schema:image <proffesor.jpg>;
schema:location [ a schema:Place;
schema:name "City, Country"@en-us ];
schema:worksFor [ a schema:Organization;
schema:name "City,Country"@en-us ];
schema:jobTitle "Profesor" @en-us;
schema:alumniOf [ a schema:EducationalOrganization;
schema:name "Univ. City,Country" @en-us];
schema:personaType *Primary*@en-us;
schema:mainPoint [ a rdf:Bag;
 rdf:_1 "dislikes applications that are not well structured"@en-us,
 rdf:_2 "has programming knowledge"@en-us,
 rdf:_3 "uses applications to share courses data with students and manage their progress"@en-us;
 rdf:_4 "Working environment: Desktop Mac OS, Mobile: IOS"@en-us. ];
schema:technicalLevel "Advanced"@en-us;
schema:businessGoal "increase student enrollment"@en-us;
schema:endGoal "share courses"@en-us;
schema:experienceGoal "organize work information and tasks"@en-us;
schema:frustrationPoint [ a rdf:Bag;
 rdf:_1 "dislikes weak and bad structured applications"@en-us,
 rdf:_2 "slowness, hard to use and no feedback provided"@en-us,
 rdf:_3 "Unreliability"@en-us. ]]
```

The first challenge the team faced was to combine all the data from the interviews to obtain the personas document. But, as this team followed the method and created several personas by interviewing users, some teams created their personas by imagining possible characteristics and frustrations of their target audience. This latter method requires validation against real user data.

Both of these challenges could be solved by a software system which incorporates in its knowledge base a TBox component (Baader et al., 2007) based on the PersonasOnto and the ABox populated with data extracted from the personas document. The system would be able to generate the personas document based on a selection on online user profiles or validate the created persona against an online user profile.

![Figure 3: Components of a Scenario.](http://www.w3.org/TR/turtle/)

As we see in Figure 3, a scenario contains several components such as participants, tasks, and con-
text. From these components the context is the most problematic as the changes that take place in both the tangible (physical world) and intangible (e.g. virtual world, software simulation) contexts have an impact on the user, which performs a certain task.

In such a case, we need to link the data gathered in the persona document with the data obtained in the usability test, because some inconsistencies may arise. For example, certain context conditions may have an impact on the number of tasks a user has to do, in order to reach the end point of a scenario, thus generating new frustrations.

6 CONCLUSIONS AND FUTURE WORK

This paper presented the PersonasOnto ontology and Persona schema which incorporate concepts and properties used to model personas. We also explored the possibility of annotating information in a personas document by using HTML5 microdata, thus making such a web page machine-readable.

Existing vocabularies and research regarding user profiling and personas have been taken into account, in order to create an accurate personas model.

These vocabularies open new possibilities regarding automatically validation and generation of personas based on online user profiles, but also linking personas data with other data (e.g. data from a usability test document or context data).

It is our future aim to study these issues, with the purpose of incorporating them in a knowledge-driven interactive system.

ACKNOWLEDGEMENTS

This work was partially supported by the European Social Fund in Romania, under the responsibility of the Managing Authority for the Sectorial Operational Program for Human Resources Development 2007-2013 [grant POSDRU/107/1.5/S/78342].

REFERENCES


