Proceedings of the 2013 Federated Conference on
Computer Science and Information Systems pp. 549–556

Microformat and Microdata Schemas for Interactive 3D Web Content

Jakub Flotyński, Krzysztof Walczak
Poznań University of Economics, Poland
email: {flotynski, walczak}@kti.ue.poznan.pl

***Abstract*—The paper presents new Microformat and Microdata schemas for creating descriptions of interactive 3D web content. Microformats and Microdata are increasingly popular solutions for creating lightweight attribute-based built-in semantic metadata of web content. However, although Microformats and Microdata enable basic description of media objects, they have not been intended for 3D content. Describing 3D components is more complex than describing standard web pages as the descriptions may relate to different aspects of the 3D content—spatial, temporal, structural, logical and behavioural. The main contribution of this paper are new Microformat and Microdata schemas for describing 3D web components and 3D scenes with metadata and semantic properties. The proposed schemas may be combined with X3D, a well-established 3D content description standard. Thanks to the use of the standardized solutions, the presented approach facilitates widespread dissemination of 3D content for use in a variety of multimedia applications on the web.**

***Index Terms*—3D content, semantic metadata, Microformats, Microdata, X3D**

I. Introduction

INTERACTIVE 3D technologies have enabled significant progress in the quality and functionality of human-computer interfaces. Widespread use of interactive 3D technologies, including virtual reality (VR) and augmented reality (AR), has been recently enabled by increasing hardware performance, availability of versatile input-output devices, as well as rapid growth in the available network bandwidth. However, the potential of 3D/VR/AR technologies in everyday applications can be fully exploited only if accompanied by the development of efficient and easy-to-use methods of creation, publication and sharing of interactive 3D multimedia content.

Building, searching and combining distributed three-dimensional interactive content are much more complex and challenging tasks than in the case of typical web pages. The relationships between components of an interactive three-dimensional virtual scene may include, in addition to its basic meaning and presentation form, also spatial, temporal, structural, logical, and behavioural aspects.

The aforementioned problems may be alleviated by describing 3D content with appropriate metadata and semantic properties. Research on the Semantic Web was initiated by T. Berners-Lee and the W3C (World-Wide Web Consortium) in 2001. This research aims at evolutionary development of the current web towards a distributed semantic database linking structured content and documents. Semantic description of web content makes it understandable for both humans and computers achieving a new quality in building web applications that can ”understand” the meaning of particular components of content and services as well as their relationships, leading to much better methods of searching, reasoning, combining and presenting web content.

On the basis of Semantic Web recommendations such as the Resource Description Framework (RDF) [1], the RDF Schema [2] and the Web Ontology Language (OWL) [3], a number of vocabularies, schemas and ontologies have been devised for a variety of application domains, in particular for multimedia systems, e.g., the Multimedia Content Description Interface [4], the Ontology for Media Resources [5] and the Core Ontology for Multimedia [6]. Available approaches to creating semantic descriptions of media content introduce a number of common attributes convenient for the general type of web resources, e.g., identifier, title, description, contributor, etc. In addition, they provide specific classes and properties intended for images, audio and video, but not for complex 3D web components, which may be described by multiple specific properties such as interactivity, animations, illumination, levels of detail, etc. Such metadata properties may be useful for exploration and analysis of 3D content, in particular for multimedia retrieval and optimization of queries for 3D content by providing values of attributes that are relatively constant and whose calculation is time-consuming.

Microformats [7] and Microdata [8] are increasingly used approaches to creating built-in semantic descriptions of web content with schemas defined in common repositories on the web. Embedding metadata directly in web content has a few important advantages in comparison to approaches that decouple resources from their descriptions. First, with embedded metadata, resources are unambiguously and inextricably linked with their descriptions. Second, it enables more concise descriptions and faster and less complicated authoring and analysis of semantically described content. Furthermore, it facilitates combining the semantic descriptions of resources with descriptions of web pages that embed the resources. Finally, it permits storage of content in structurally simpler databases. However, although existing Microformats and Microdata enable basic semantic descriptions of several types of multimedia objects, such as images, audio and video, they do not provide support for describing 3D content.

The main contribution of this paper are new Microformat and Microdata schemas for creating semantic metadata of interactive 3D web components and 3D scenes. The proposed

978-1-4673-4471-5/$25.00

schemas facilitate indexing and retrieval of 3D content that meets specific criteria. The schemas include a number of specific properties that may be useful for contextual 3D content presentation dependent on, e.g., hardware/software platform, user–system interaction paradigms, user preferences, etc. The schemas may be combined with Extensible 3D (X3D) [9]—a well-established 3D content description standard, but they are not limited to this language. Thanks to the use of standardized solutions, the proposed approach enables flexible description and widespread dissemination of 3D content for use in a variety of multimedia web systems, e.g., in cultural heritage, education, simulations, geospatial visualisations, etc.

The remainder of this paper is structured as follows. Section II provides an overview of the state of the art in the domain of semantic and metadata descriptions of web resources, in particular 3D web content. In Section III, new Microformat and Microdata schemas are proposed for creating built-in semantic metadata of interactive 3D web content. Section IV describes a possible application of the proposed schemas in a system for searching 3D models by their metadata properties. Finally, Section V concludes the paper and indicates the possible directions of future research.

II. Semantic Descriptions of Multimedia Web Content

In this section, the state of the art in the field of semantic and metadata descriptions of multimedia web content is presented. In particular, metadata and ontologies as well as Microformats and Microdata for describing multimedia content are discussed. Next, methods of creating attribute-based embedded semantic and metadata descriptions of interactive 3D web content are considered.

A. Metadata and Ontologies for Multimedia Content

Several vocabularies, metadata schemas and ontologies have been proposed for describing multimedia content. DIG35 [10] defines metadata schemas for digital images. CableLabs [11] introduces vocabularies for both images and videos. The QuickTime File Format Specification [12] provides a schema for describing movie files. The Multimedia Content Description Interface (MPEG-7) [4] is a standard that defines a set of sophisticated tools for creating metadata—Descriptors, Description Schemes, the Description Definition Language and Coding Schemes. There is a wide range of target multimedia content that may be described with MPEG-7, including images, audio, video and 3D objects [13][14], however the standard is strongly focused on audio-visual data. The standards mentioned above typically include a number of generic properties (e.g., resource identifier, title, description, contributor, etc.), and a number of specific properties for describing images, audio and video, but not for complex interactive 3D web content.

A few ontologies have been proposed for multimedia content. The Ontology for Media Resources [5] has been devised by the W3C on the basis of the Resource Description Framework (RDF) [1], the RDF Schema [2] and the Web Ontology Language (OWL) [3] as a common solution for describing multimedia published on the web. It provides an interoperable core vocabulary that is mapped to a set of metadata formats for media content (e.g., DIG35, CableLabs and MPEG-7). A number of concepts defined in this ontology are common for web content of different types. There is a limited set of attributes typical for multimedia content, e.g., frameSize, compression, duration and samplingRate. This ontology lacks classes and properties typical for interactive 3D content, such as illumination, animations, navigation, levels of detail, etc.

The Core Ontology for Multimedia (COMM) [6][15] is another solution designed for describing media content. COMM is based on MPEG-7, but it represents knowledge with open Semantic Web solutions avoiding some interoperability problems that occur in MPEG-7, e.g., with semantically equivalent descriptors that are processed in different manners [15]. This ontology is convenient for describing images, audio and video, but it contains only a limited set of concepts suitable for interactive 3D web content.

Some other works are devoted to metadata for describing interactivity of 3D objects [16] and their interfaces [17]. Such descriptions may be used for finding 3D components by their properties [18]. In [19], metadata schemas for media objects have been proposed in the context of teaching architecture. The Metadata 3D Initiative [20] is a project in which a number of companies and research centres collaborate on the standardization of schemas for 3D (stereoscopic) content to make interoperable lenses, cameras, rigs, stereoscopic image processors, etc. In [21], the Multimedia Web Ontology Language, an extension to OWL, designed for creating ontologies and models for probabilistic reasoning in multimedia processing is presented.

B. Microformat and Microdata Schemas for Multimedia Content

Microformats [7] and Microdata [8] are solutions that are currently increasingly used for encoding semantic metadata of web content. In contrast to RDF-based approaches, which enable semantic descriptions with ontologies distributed across the web, Microformats and Microdata permit rapid creation of lightweight built-in semantic descriptions of content with schemas defined in common repositories on the web [7][22]. Such descriptions may be understandable for widely-used web search engines, such as Google, Yahoo and Bing. Currently, both Microformats and Microdata provide a variety of schemas for describing different types of web resources, in particular for multimedia content including images, audio and video.

The `hMedia` Microformat [23] is used for describing images, audio and video with a common set of properties. This schema is convenient for creation of general semantic descriptions that do not go into specific details with regard to individual media types. The `hAudio` [24] has been designed for describing audio content. There are no specific Microformats for describing images, video and 3D content.

In comparison to Microformats, Microdata provides a few more compound schemas, which form a hierarchical structure with inheritance of properties. The root of this hierarchy is the `MediaObject` [25] schema, which defines a set of properties common for different types of media objects (like the `hMedia` Microformat)—images, audio and video. In addition, more extensive semantic and metadata descriptions of the particular media types may be created with descendant schemas of the `MediaObject`—`ImageObject` [26], `AudioObject` [27] and `VideoObject` [28]. Like the approaches mentioned above, Microdata does not provide schemas and properties sufficient for creating metadata and semantic descriptions of interactive 3D web content.

C. Attribute-based Embedded Metadata for 3D Web Content

In [29][30], a method of creating lightweight attribute-based semantic descriptions built into interactive 3D web content has been presented. The method enables metadata and semantic descriptions of both real objects and their virtual 3D counterparts by putting metadata into individual X3D metadata nodes. The resulting metadata and semantic descriptions of 3D web content are equivalent to Microformat and Microdata descriptions of typical web pages in terms of expressiveness. Moreover, such descriptions may use the same schemas, thus the method permits bidirectional transformation between descriptions that are built into web pages and descriptions embedded in 3D web content. Due to the use of the standard syntax and structure of X3D documents, the compatibility of the proposed approach with available X3D browsers is preserved.

The proposed approach is depicted in Fig. 1. The primary entity of the semantic description of 3D content is an X3D `MetadataSet` node. Since the method enables semantic descriptions of both real objects and their virtual 3D counterparts, both types of resources are referenced in the same manner—by their URIs. If a particular 3D component is to be described, it has to be assigned a URI in the X3D `DEF` attribute. The `name` attribute of the `MetadataSet` indicates a list of types of the described item, each of which may determine a set of semantic metadata properties. New item properties may be added to a `MetadataSet` independently of the schemas used. The optional `value` attribute specifies the URI (navigable or non-navigable) of the described object. The `reference` attribute of the `MetadataSet` contains a list of references to attributes that have been specified in other semantic descriptions and need to be shared with the primary one.

In addition to specifying the type and the URI of an item, the `MetadataSet` serves also as a container for item properties and relationships with other resources, which are reflected by nested typed metadata nodes (integer, float, double, string). The name and value of a property/relationship are given by the `name` and `value` attributes, respectively. The `reference` attribute is used to distinguish a property from a relationship. Metadata describing a property (e.g., data type) or a relationship may be contained in an additional metadata node of a desirable type, which is nested in the property/relationship typed element.

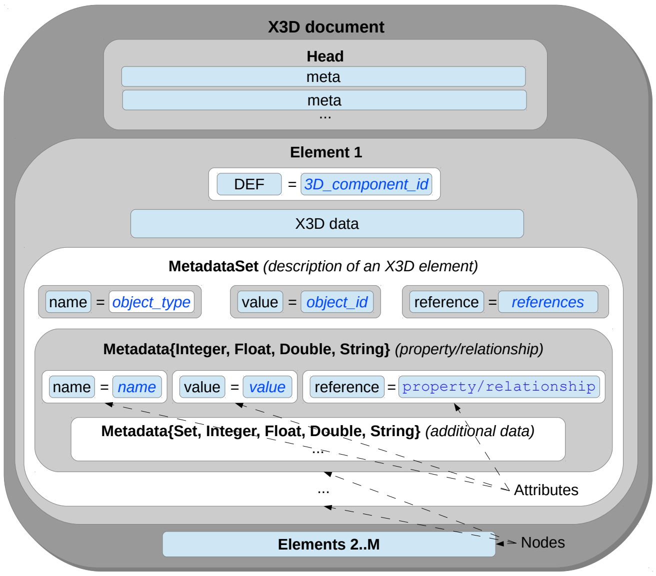

Fig. 1. Attribute-based embedded semantic descriptions of 3D content

In [31], a novel method of harvesting embedded attribute-based semantic metadata descriptions from distributed X3D web content has been proposed. The presented solution is an XSLT-based equivalent to the GRDDL [32] approach (originally intended for typical web pages), which has been designed for selection and processing of semantic descriptions of complex 3D/VR/AR scenes and components distributed across the web. The selection is based on the actual format, type and structure of the components, which are determined by their syntax. Filtering of the metadata to be extracted is necessary for reducing large semantic descriptions of complex 3D/VR/AR content to excerpts relevant to a particular application. The harvesting of metadata is a preliminary stage of the semantic analysis of the content, and it may precede the following activities, such as loading the generated semantic descriptions into a database and querying the system for semantically described 3D components.

III. Metadata Schemas for 3D Content

The aforementioned approaches address different aspects of creating metadata and semantic descriptions of web content, but they do not provide metadata schemas convenient for describing interactive 3D web content. To enable semantic descriptions of interactive 3D web components and complex 3D scenes, new metadata schemas are proposed in this section. They are intended to facilitate indexing, exploration and analysis of 3D content, and searching for 3D components and scenes described with embedded metadata attributes. Furthermore, the schemas include a number of specific properties that may be useful for contextual 3D content presentation dependent on, e.g., hardware/software client platform, user–system interaction paradigms, user preferences.

First, a classification of semantic metadata properties of interactive 3D web content is introduced. Then, new Microformat and Microdata schemas are proposed for describing 3D content with metadata and semantics.

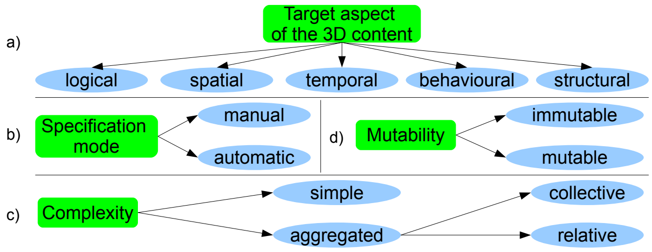

Fig. 2. The classification of metadata properties of interactive 3D web content in terms of the target aspect of the described 3D content (a), the specification mode of the property (b), the complexity of the property (c), and the mutability of the property (d)

A. Classification of Metadata Properties of Interactive 3D Web Content

The classification of metadata properties of interactive 3D web content is depicted in Fig. 2. The proposed classification is discussed in the four following aspects.

1) Target aspect of the 3D content: The metadata properties may be divided into different groups in terms of the described target aspect of the 3D content—*logical* (e.g., content description, presented object), *spatial* (e.g., dimensions, levels of detail), *temporal* (e.g., duration of an animation), *behavioural* (e.g., interactivity, animations) or *structural* (e.g., 3D sub-components).

2) Specification mode: The metadata properties may be distinguished in terms of the mode of specification, which may be *manual* or *automatic*. The first group incorporates properties that cannot be derived from the described 3D content and must be specified manually, in particular, by the author of the 3D content, e.g., the URI of the prototype of the described 3D component, the description of the component, and the semantic roles of its sub-components. The second group includes properties that can be automatically derived from the described 3D content, e.g., dimensions, animations, interactions.

3) Complexity: Simple and *aggregated* properties may be distinguished. *Simple* properties may be automatically retrieved from the described 3D component without appealing to its sub-components, e.g., the background of the scene or navigation modes. The main advantage of extracting *simple* properties is a possibility to reference them in queries that are built with a query language.

In contrast to *simple* properties, *aggregated* metadata properties are set with regard to the sub-components of the described 3D content, as opposed to other media types that do not incorporate a complex hierarchy of nested objects. *Aggregated* properties are determined by recursive processing and analysis of the content, which may be difficult and time-consuming, e.g., determining the number of levels of detail, animation types, light sources for the described content and all its sub-components. The main advantage of using *aggregated* metadata properties is that the calculation of them is performed once, before the described 3D content is loaded into a system. The presented approach may be preferred for 3D content that is rarely modified, for which the calculated properties are valid for a relatively long time. In such cases, the presented approach accelerates queries sent to the system, which do not need to initiate time-consuming processing of the content.

Two types of *aggregated* automatically determined metadata properties may be distinguished—*collective* and *relative* properties. *Collective* properties are specified by the analysis of sub-components of the described 3D component and determining a single aggregated value or a list of aggregated values. The results are related to the described 3D content as a whole, e.g., the dimensions, mass or volume of the described component are calculated as the sums of the dimensions, masses and volumes of its particular sub-components.

Like *collective* properties, *relative* properties are determined with regard to the sub-components of the described component, but the value of a *relative* property always describes either a relationship between the described component and its particular sub-components, or a relationship between particular sub-components of the described component. Such properties are convenient for expressing relative physical quantities, e.g., collision detection, velocity, acceleration, angular velocity, angular acceleration that are determined only by relationships between particular objects. In contrast to *collective* properties, *relative* properties are not aggregated.

4) Mutability: Metadata properties may be categorized in terms of the mutability into *immutable* and *mutable* properties. The first group comprises properties that do not change during the content presentation, e.g., the URI of the object presented by the described component or the names of additional packages required for correct content presentation. The second group contains properties that potentially can (but do not have to) change, e.g., dimensions, mass, illumination may change because of disappearing of some sub-components, switching off some light sources, etc. In the presented approach, it is not assumed that *mutable* properties are up to date for the whole period of the 3D content presentation. Instead, an author of the described 3D content may arbitrarily select a condition or a point in time for which the property is specified. It is recommended to select a specific condition/point in time, e.g., the start or the end of an animation that modifies the described property, and to indicate the selected point by the metadata of this property.

B. Metadata Properties of the Proposed Schemas

The list of metadata properties of the proposed `Microformat3D` and `Microdata3D` schemas are presented in Table 1. In the table, the aspect of 3D content (logical, spatial, temporal, behavioural and structural), the name and data types are provided together with the complexity, mutability and description for each property. In the *Type* column, the first data type is specified according to the proposed Microformat, the second one—to the proposed Microdata schema.

Target aspect of the 3D content	Property	Type	Complexity	Mutability	Description
logical	presentedObject	string	simple	immutable	A uniform identifier of the object (prototype) that is presented by the described 3D content. The presented object may be either real or virtual with a URI specified, e.g., as an HTTP address.
	packages	string[]	collective	immutable	The list of the names or URIs of additional packages necessary for the presentation of the whole described 3D content including all its sub-components. It is specific to the particular 3D content description standard used. For instance, X3D content may require such additional packages as Geospatial, NURBS, Human Animation, Distributed Interactive Simulation or CAD.
spatial	dimensions	double[] / float[]	collective	mutable	The width, height and length of the described 3D component including all its sub-components.
	mass	double / float	collective	mutable	The mass of the described 3D component including all its sub-components.
	volume	double / float	collective	mutable	The volume of the described 3D component including all its sub-components.
	fog	double / float	collective	mutable	The minimal visibility range for the described 3D scene and all its sub-components.
	background	string	simple	mutable	The name/value of the color or a URI of the image that is used as the background of the described 3D scene.
	illumination	(string, long)[]	collective	mutable	The list of types and numbers of light sources of the particular type, which are used in the scene and its sub-components. The light types may be, e.g., `point`, `spot`, `directional`, `area`, `model`, `ambient`.
	levelsOfDetail	(long, double/float)[]	collective	mutable	The list of the levels of detail used for the described 3D scene. Each level is specified by the number of polygons and a range. If multiple sub-components are combined within the described complex 3D scene, this parameter should be a combination of the levels of detail with regard to these sub-components.
	collisions	string[][]	relative	mutable	The list of the sets of URIs of 3D sub-components of the described 3D scene, for which collision detection is enabled.
temporal and behavioural	interactivity	string[]	collective	mutable	The list of user interactions allowed for the described 3D component and its sub-components, e.g., `selection`, `manipulation`, `navigation`, `system control`, `symbolic input`, etc.
	navigation	string[]	simple	mutable	The list of navigation modes allowed for the described 3D scene, e.g., `any`, `fly`, `walk`, `examine`, `lookat`, `slide`, `rotate`, `pan`, `game`, `jump`, `none`.
	animations	string[]	collective	mutable	The list of animation types used in the described 3D scene and its sub-components, e.g., `position`, `orientation`, `scale`, `structure`, `shape`, `appearance`, etc. For complex 3D scenes, the list should include animations in relationships between particular sub-components of the scene as well as between the scene and its sub-components, e.g., `position` animation.
structural	imageComponents	(string, string)[]	collective	immutable	The list of image components (textures) that are linked to the described component and its sub-components, with their URIs and semantic roles in the described 3D content, e.g., a texture of a dish, a sculpture, an exhibit, etc.
	audioComponents	(string, string)[]	collective	immutable	The list of audio components that are linked to the described component and its sub-components, with their URIs and semantic roles in the described 3D content, e.g., a background sound, a piano sound, etc.
	videoComponents	(string, string)[]	collective	immutable	The list of video components that are linked to the described component and its sub-components, with their URIs and semantic roles in the described 3D content, e.g., a projection, a movie, etc.
	3DComponents	(string, string)[]	collective with relative properties	immutable	The list of 3D sub-components with their URIs and semantic roles in the described 3D component, e.g., artefact, exhibition stand, wall, floor, furniture, etc.

Table 1. Semantic metadata properties of the proposed Microformat3D and Microdata3D schemas

The proposed Microformat and Microdata schemas are partially based on metadata properties devised in previous research works, e.g., the ARCO (Augmented Representation of Cultural Objects) 3D virtual museum system [33][34]. The presented list focuses only on properties specific for interactive 3D web content, and it is common for the new proposed Microformat and Microdata schemas. The schemas make use of the aforementioned Microformats and Microdata for media resources (`hMedia`, `hAudio`, `MediaObject`, `ImageObject`, `AudioObject`, and `VideoObject`) extending them with new metadata properties for 3D content. Attributes common for different media types have been omitted in the list, as they are inherited from the parent schemas. The inherited properties (not listed in Table 1) are mainly immutable and manually specified (e.g., `title`, `contributor`, `description`) or they are simple attributes, determined automatically without processing of their sub-components (e.g., `encodingFormat`, `uploadDate`). In Table 1, only the `presentedObject` must be specified manually, e.g., by a content creator. Other properties may be automatically determined—usually as aggregations of values from sub-components.

Microformats and Microdata introduce several equivalent metadata schemas, e.g., the `hAudio` and `hMedia` Microformats have as counterparts the `AudioObject` and the `MediaObject` in Microdata. Although these schemas usually contain common sets of properties, their numbers and types of attributes are frequently different, e.g., the `contributor` attribute indicates a `hcard` in Microformats and a `Person` or an `Organization` in Microdata. Hence, although the new proposed `Microformat3D` and `Microdata3D` schemas have the same properties specific for describing 3D content, they differ in the numbers and types of properties that are inherited from their parent standard-specific schemas.

The proposed schemas do not impose the use of any particular units for individual metadata properties, but it is recommended to conform to the quantities and units specified in the described 3D content, e.g., metres and kilograms may be used for specifying the `dimensions` and the `mass` of the described object. Like any metadata describing a particular semantic property, unit descriptions may be nested into the descriptions of the relevant properties.

Although, overall, it is recommended to embed descriptions directly in the described 3D content, this is not required. The proposed `Microformat3D` and `Microdata3D` schemas are intended to be used directly within 3D content descriptions (in particular in X3D documents), but they may be also used in other types of documents, e.g., web pages embedding the 3D content. Hence, the name of the new Microformat does not start with a letter indicating a particular standard of the parent document—as opposed to the `hMedia` and `hAudio` Microformats that have been originally designed for HTML web pages.

IV. Searching 3D Content by Semantic Metadata Properties

This section explains the architecture of a planned system enabling 3D content retrieval using attribute-based built-in descriptions. The system leverages the proposed `Microformat3D` and `Microdata3D` schemas. First, the client–system interaction is discussed, second, an illustrative example is considered.

A. *Client–system Interaction*

The interaction between a client (a web browser) and the system is discussed in the two following aspects—loading 3D content into the system and querying the system for 3D content.

1) Loading 3D content into the system: Loading of 3D content into the system (Fig. 3a) is performed by the *Web Client*. Any browser can be used as the client, not imposing any specific and difficult to meet software requirements for the client side. A user sends 3D content with a built-in semantic description to the *3D Loader* web service via a web page with a file input component. The *3D Loader* sends the 3D content with its embedded metadata to the *GRDDL Agent* that extracts the metadata and creates a separate RDF document according to the rules presented in [31]. The document contains a semantic description equivalent to the built-in description. No changes are introduced to the primary X3D document. Finally, both the X3D document with built-in metadata and the generated RDF document are stored in the *Database*.

2) Querying the system for 3D content: Fig. 3b presents the consecutive steps performed every time 3D content is requested the system. First, a user utilizes a *Web Client* to build a query that specifies desirable metadata properties of the content. The query is embedded in an HTTP request and sent to the *Query Handler* web service. The *Proxy* is a Java application that mediates in the communication and extends the request with a context description of the client–system interaction that may specify, e.g., the client device and the software platform, user preferences and location, interaction paradigm, etc. Next, the *Query Handler* web service translates the extended query into a SPARQL [35] statement which is delivered to a *SPARQL Query Engine*. The engine retrieves desirable 3D components from the *Database*. The components are sent back to the *Query Handler* that invokes a *Web Page Builder* to create a representative web page. Finally, the web page is delivered to the *Web Client*.

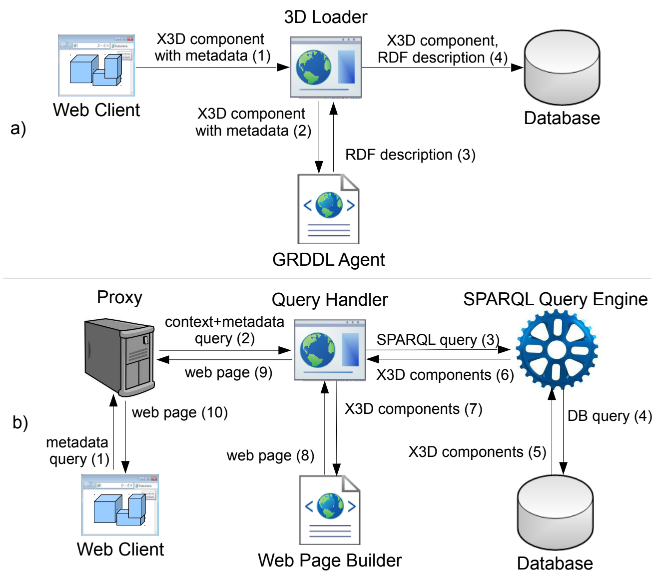

Fig. 3. Loading 3D content into the system (a) and querying the system for 3D content (b)

Listing 1. X3D content described with the proposed `Microformat3D`

```
<Shape DEF='Sculpture'>
 <Appearance>...</Appearance><IndexedFaceSet>...</IndexedFaceSet>
 <MetadataSet name='http://www.kti.ue.poznan.pl/3DContent' value='Sculpture'>
  <MetadataString name='fn' value='Wooden sculpture' reference='property' />
  <MetadataString name='enclosure' value='.../sculpt.x3d' reference='relationship' />
  <MetadataString name='description' value='An example virtual sculpture' reference
        ='property' />
  <MetadataString name='presentedObject' value='http://.../museum/sculpture'
        reference='relationship' />
  <MetadataString name='dimensions' value='0.5 0.3 1' reference='property'>
   <MetadataString name='unit' value='meter' reference='property' />
  </MetadataString>
  <MetadataString name='collisions' value='http://.../dish.x3d http://.../handle.x3d'
        reference='relationship' />
  <MetadataString name='illumination' value='point' reference='property' />
  <MetadataString name='interactivity' value='selection manipulation navigation'
        reference='property' />
  <MetadataString name='navigation' value='fly walk' reference='property' />
  <MetadataSet name='3DComponents' reference='property'>
   <MetadataSet name='http://www.kti.ue.poznan.pl/StringTuple'>
    <MetadataString name='URI' value='http://.../dish.x3d' reference='property' />
    <MetadataString name='role' value='dish' reference='property' />
  </MetadataSet></MetadataSet>
 </MetadataSet>
</Shape>
```

B. *Illustrative Example*

Below, two examples of the client–system interaction are presented with regard to the steps described in the previous subsection. In the examples below, 3D content is described with the proposed Microformat, but an equivalent description could be created as well using the Microdata schema.

1) Loading 3D content into the system: Example X3D content that presents a 3D model of a sculpture is presented in Listing 1. The X3D `head` and several other elements reflecting the geometry and appearance have been omitted as they are not crucial for this example. The method of embedding semantic descriptions into 3D content has been explained in detail in [29][30]. In the presented example, the metadata properties have been specified manually while creating 3D components. However, it is desirable to develop an additional tool automatically calculating and embedding the properties into 3D content.

The 3D content is described with the `Microformat3D` (line 3) with multiple properties inherited from the `hMedia` (`fn`-4, `enclosure`-5, `description`-6), logical (`presentedObject`-7), spatial (`dimensions`-8, `collisions`-11, `illumination`-12), temporal and behavioural (`interactivity`-13, `navigation`-14) as well as structural (`3DComponents` 15-19). Next, the 3D content is sent to the *3D Loader* web service and stored in the *Database*, which is implemented using Oracle XML DB [36].

Listing 2. Example conditions sent to the system by a user (a), and the SPARQL query including contextual requirements inserted by the *Proxy* (b)

```
a) ?component description ?description
      FILTER regex(?description, "artefact").
   ?component material ?material
      FILTER regex(?material, "wood").
   ?component illumination "point".
   ?component animations "position".

b) select ?URI where {
                 ?component enclosure ?URI.
                 ?component interaction "manipulation".
                 ?component navigation "walk".
                 ?component levelsOfDetail ?lod.
                 ?lod numOfPolygons ?polygons.
                 FILTER (?polygons >= 100 000).
                 { select count(?lod) as ?n where
                          { ?component levelsOfDetail ?lod. }}.
                 FILTER (?n >= 3). }
```

2) Querying the system for 3D content: Querying the system for 3D content starts with specifying conditions using the *Web Client* (e.g., a web page). The *Web Client* builds a SPARQL query, which is encoded with the SPARQL Protocol for RDF [37], built into HTTP address and sent to the *Query Handler* web service. In the presented example, a user specifies semantic properties of the desirable objects by requiring 3D models of artefacts made of wood. In addition, the following metadata properties of the 3D objects are specified—the artefacts should be illuminated by `point` light sources and their `positions` should be animated (Listing 2a). The *Proxy* inserts additional contextual requirements into the request that specify 3D components suitable for desktop devices equipped with a keyboard and a high-resolution screen—with the `manipulation` interaction, the `walk` navigation mode enabled, having at least 100k `polygons` and at least 3 `levelsOfDetail` (Listing 2b). The *Query Handler* conveys the extended query to the *SPARQL Query Engine* (implemented with Apache Jena [38]). Next, 3D components that satisfy the given conditions are retrieved from the *Database* and provided to the *Query Handler*. Finally, the *Web Page Builder* creates a web page (Fig. 4), inserting the found 3D components into a web page template. The resulting document is delivered to the *Web Client*.

Fig. 4. An example web page with the requested 3D components

V. Conclusions and Future Works

In this paper, Microformat3D and Microdata3D schemas have been proposed for describing interactive 3D web content. Although a number of schemas and ontologies have been designed for describing the metadata and semantics of multimedia content on the web, they have been intended mostly for images, audio and video and not for interactive 3D content. The lack of commonly accepted schemas for describing the metadata and semantics of 3D resources is one of the important obstacles to widespread creation, dissemination and reuse of interactive 3D web content.

The proposed Microformat3D and Microdata3D schemas allow for metadata and semantic descriptions of interactive 3D web components and 3D scenes. The proposed schemas make use of the available Microformats and Microdata schemas for flexible semantic descriptions, and they may be combined with X3D, which is the leading standard for describing 3D content on the web. The compatibility with well-established web approaches enables the use of the presented schemas for describing, retrieving and exploring (finding, classifying, clustering, cataloguing, etc.) interactive 3D content in a variety of multimedia web systems, using a number of available tools (editors, validators, parsers, transformers, etc.), with minimal additional effort. The presented approach may be used for query optimization in applications with context-based user–system interaction.

The proposed approach stresses the compatibility of the created descriptions with the current syntax and structure of X3D documents and available X3D browsers. To provide the conformance of the presented solution to popular web search engines, the proposed schemas need to be encoded using the original Microformat and Microdata syntax, which has been intended for web pages and which is not compatible with 3D content standards. Therefore, the metadata should be inserted into web pages embedding the described 3D content.

We plan implementation of the proposed approach as an extension to the ARCO virtual museum system. Such extension will allow for creating built-in metadata and semantic descriptions of virtual museum exhibitions, and will enable evaluation of the schemas in terms of the achieved optimization of queries to digital repositories of 3D exhibits.

Possible directions of future research incorporate several facets. First, the paper describes only an initial set of properties that may be used for describing 3D content. Based on practical experiences from implementation of systems using this kind of descriptions, the presented schemas may be further extended to include more specific properties describing 3D content. Also, possible values of the properties together with the preferred syntax should be specified. Second, the presented metadata model may be implemented using powerful RDF and RDF-based technologies such as OWL and RDFS. This will permit sophisticated exploration of semantically described content, including querying data sources and reasoning. Third, a tool for automatic computation of the proposed metadata properties should be developed and used for new 3D components loaded into the system. Next, the contextual rules managed by the *Proxy* might be described with Semantic Web standards to be accessible and processable with widely-used semantic tools. Finally, an evaluation of the system should be performed to assess the benefits from the query optimization provided by the proposed metadata schemas.

VI. Acknowledgements

This research work has been partially funded by the Polish National Science Centre grant No. DEC-2012/07/B/ST6/01523.

References

[1] W3C. Resource Description Framework (RDF): Concepts and Abstract Syntax. *http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/*. Retrieved April 25, 2013.

[2] W3C. RDF Vocabulary Description Language 1.0: RDF Schema W3C Recommendation 10 February 2004. *http://www.w3.org/TR/rdf-schema/*. Retrieved April 25, 2013.

[3] OWL. OWL Web Ontology Language Reference. W3C Recommendation 10 February 2004. *http://www.w3.org/TR/owl-ref/*. Retrieved April 25, 2013.

[4] ISO/IEC 15938-10:2005. Information technology - Multimedia content description interface – Part 10: Schema definition. *http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm ?csnumber=40527*. Retrieved April 25, 2013.

[5] W3C. Ontology for Media Resources 1.0. W3C Recommendation 09 February 2012. *http://www.w3.org/TR/mediaont-10/*. Retrieved April 25, 2013.

[6] A Core Ontology for Multimedia. *http://comm.semanticweb.org/*. Retrieved April 25, 2013.

[7] Microformats. *http://microformats.org/*. Retrieved April 25, 2013.

[8] W3C. HTML Microdata. Editor's Draft 26 April 2013. *http://www.w3.org/html/wg/drafts/microdata/master/*. Retrieved April 26, 2013.

[9] Web3D. Extensible 3D (X3D) Part 1: Architecture and base components ISO/IEC 19775-1:2008. *http://www.web3d.org/files/specifications/19775-1/V3.2/Part01/Architecture.html*. Retrieved April 25, 2013.

[10] The DIG35 Phase 2 Initiative Group. DIG35 Specification—Metadata for Digital Image. *http://www.bgbm.org/tdwg/acc/Documents/DIG35-v1.1WD-010416.pdf*. Retrieved April 25, 2013.

[11] CableLabs. Video-On-Demand Content Specification Version 2.0. *http://www.cablelabs.com/specifications/MD-SP-VOD-CONTENT2.0-I02-070105.pdf*. Retrieved April 25, 2013.

[12] Apple. QuickTime File Format Specification. *https://developer .apple.com/standards/qtff-2001.pdf*. Retrieved April 25, 2013.

[13] Bilasco I. M., Gensel J., Villanova-Oliver M., Martin H., On Indexing of 3D Scenes Using MPEG-7, In: *Proceedings of the 13th annual ACM international conference on Multimedia*, Singapore, November 06-12, 2005, pp. 471-474.

[14] Bilasco I. M., Gensel J., Villanova-Oliver M., Martin H., An MPEG-7 framework enhancing the reuse of 3D models , In: *Proceedings of the Web3D Symposium 2006*, Columbia, Maryland, USA, 2006, pp. 65-74.

[15] Arndt R., Troncy R., Staab S., Hardman L., Vacura M., COMM: designing a well-founded multimedia ontology for the web. In: *Proc. of the 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference*, Busan, Korea, November 11-15, 2007, pp. 30-43.

[16] Chmielewski J., Describing Interactivity of 3D Content. In: *Interactive 3D Multimedia Content—Models for Creation, Management, Search and Presentation*, ed. Cellary, W., Walczak K., Springer, London, Dordrecht, Heidelberg, New York, 2012, pp. 195-221.

[17] Chmielewski J., Metadata Model for Interaction of 3D Object, In: *The 1st International IEEE Conference on Information Technology*, ed. Stepnowski, A., M. Moszyński, T. Kochański, J. Dabrowski, Gdańsk, May 18, 2008, Gdańsk University of Technology, 2008, pp. 313-316.

[18] Chmielewski J., Finding interactive 3D objects by their interaction properties, In: *Multimedia Tools and Applications*, Springer, Netherlands, 2012, ISSN: 1380-7501.

[19] Boeykens S., Bogani E., Metadata for 3D Models. How to search in 3D Model repositories? In: *Proceedings of the International Conference of Education, Research and Innovation*, Madrid, Spain, November 17-19, 2008.

[20] Metadata 3D Initiative. *http://www.m3di.org/*. Retrieved April 25, 2013.

[21] Ghosh H., Chaudhury S., Ontology for semantic multimedia web. *http://www.w3cindia.in/conf-site/Hiranmay%20Ghosh%20–%20mowl%20%28ontology %20for%20semantic%20multimedia%20web.pdf*. Retrieved April 25, 2013.

[22] Schema.org. *http://schema.org/*. Retrieved April 25, 2013.

[23] hMedia. *http://microformats.org/wiki/hmedia*. Retrieved May 20, 2013.

[24] hAudio. *http://microformats.org/wiki/hAudio*. Retrieved May 20, 2013.

[25] MediaObject. *http://schema.org/MediaObject*. Retrieved May 20, 2013.

[26] ImageObject. *http://schema.org/ImageObject*. Retrieved May 20, 2013.

[27] AudioObject . *http://schema.org/AudioObject*. Retrieved May 20, 2013.

[28] VideoObject. *http://schema.org/VideoObject*. Retrieved May 20, 2013.

[29] Flotyński J., Walczak K., Describing Semantics of 3D Web Content with RDFa. In: *Proceedings of the First International Conference on Building and Exploring Web Based Environments*, Seville, Spain, January 27-February 1, 2013, pp. 63-68, ISBN 978-1-61208-248-6.

[30] Flotyński J., Walczak K., Attribute-based semantic descriptions of interactive 3D web content. In: *Information Technologies in Organizations : Management and Applications of Multimedia - Toruń*, Wydawnictwa Towarzystwa Naukowego Organizacji i Kierownictwa—Dom Organizatora, Czestochowa, Poland, 2013, pp. 111-138, ISBN 978-83-7285-691-3.

[31] Flotyński J., Harvesting of Semantic Metadata from Distributed 3D Web Content. In: *Proc. of the 6th Int. Conf. on Human System Interaction*, Sopot, Poland, June 6-8, 2013, ISBN 978-1-4673-5636-7.

[32] W3C. GRDDL Primer. W3C Working Group Note 28 June 2007. *http://www.w3.org/TR/grddl-primer/*. Retrieved April 25, 2013.

[33] Mourkoussis N., White M., Patel M., Chmielewski J., Walczak K., AMS Metadata for Cultural Exhibitions using Virtual Reality, In: *Dublin Core International Conference DC 2003*, Seattle, Washington, USA, September 28 - October 2, 2003, pp. 193-201.

[34] Patel M., White M., Mourkoussis N., Walczak K., Chmielewski J., Wojciechowski R., Metadata Requirements for Digital Museum Environments, In: *Int. Journal on Digital Libraries, Special Issue on Digital Museum*, vol 5, no 3, May, 2005, Springer Verlag, pp.179-192.

[35] W3C. SPARQL Query Language for RDF. W3C Recommendation 15 January 2008. *http://www.w3.org/TR/rdf-sparql-query/*. Retrieved April 25, 2013.

[36] Oracle XML DB. *http://www.oracle.com/technetwork/database-features/xmldb/overview/index.html*. Retrieved April 25, 2013.

[37] W3C. SPARQL 1.1 Protocol. W3C Recommendation 21 March 2013. *http://www.w3.org/TR/sparql11-protocol/*. Retrieved April 25, 2013.

[38] Apache Software Foundation. Apache Jena. *http://jena.apache.org/*. Retrieved April 25, 2013.