AIM@SHAPE

AIM@SHAPE

Advanced and Innovative Models And Tools for the development of Semantic-based systems for Handling, Acquiring, and Processing knowledge Embedded in multidimensional digital objects

IST NoE No 506766

Deliverable D1.2.1.1

Ontology for Virtual Humans 2nd version

Circulation: [1]CO

Partner(s): **EPFL, UNIGE, ITI, IMATI, MPII**

Authors: A. Garcia-Rojas, L. Moccozet, G. Vasilakis, M. Pitikakis, M. Spagnuolo, A. Belyaev

Version: 02

Stage: 100% Final

Date: Thursday, September 7, 2006

---

[1] Please indicate the dissemination level using one of the following codes:

PU = Public

PP = Restricted to other programme participants (including the Commission Services).

RE = Restricted to a group specified by the consortium (including the Commission Services).

CO = Confidential, only for members of the consortium (including the Commission Services).

consisting of:

CNR-**IMATI**-GE	C.N.R. – Istituto di Matematica Applicata e Tecnologie Informatiche Dept. of Genova, Italy
DISI	Università di Genova – Dipartimento di Informatica e Scienze dell'Informazione, Italy
EPFL	École Polytechnique Federale de Lausanne, Switzerland
FhG/**IGD**	Fraunhofer Institut für Graphische Datenverarbeitung, Germany
INPG	Institut National Polytechnique de Grenoble, France
INRIA	Institut National de Recherche en Informatique et Automatique, France
ITI-CERTH	Informatics and Telematics Institut – Center for Research and Technology Hellas, Greece
UNIGE	Université de Genève, Switzerland
MPII	Max-Planck-Institut für Informatik, Germany
SINTEF	Stiftelsen for industriell og teknisk forskning ved Norges Tekniske Høgskole, Norway
TECHNION	Technion – Israel Institute of Technology, Israel
UU	Utrecht University, Netherlands
WEIZMANN	Weizmann Institute of Science, Israel

Document History

Vers.	Issue Date	Stage	Content and changes
1	17 July 2006	95%	Second version of the ontology
2	14 august 2006	100%	Explain how we can answer the CQs using the search engine.

Executive Summary

This document contains a description of the deliverable **D1.2.1.1** of the IST NoE AIM@SHAPE. The deliverable ***D1.2.1.1 – Ontology for Virtual Humans 2nd version*** – is intended to provide a second version of the ontology for Virtual Humans which is in development within the network.

The task leader is **EPFL** and is actively supported by all other AIM@SHAPE partners.

For a better comprehension of this document we recommend to read the First version of the Virtual Human ontology (D1.2.1.1).

This document is structured as follows. First, we present an introduction of the ontology's changes between the first version and this one. In Section 1 we detail the mentioned changes. Section 2 describes the current structure of the ontology. Section 3 presents two case scenarios where the ontology helps to make easier the process of design and animation of virtual humans. We conclude with a list of the competency questions that can be answered using this ontology. Finally in section 4 we present some conclusions.

Table of Contents

INTRODUCTION

Ontology development is a continuous process. In the previous version of this document, the first version of the Virtual Human (VH) ontology, we provided a starting point for the formalization of the knowledge of the VH creation. We presented a list of Competency Questions (CQs) that defined the domain and scope of the ontology, and based on those CQs we designed the first version of the ontology.

In this document we present the second version of the VH Ontology. Here, we propose to change the relation among some concepts, and we include new concepts. This evolution was due to the integration of the VH ontology with the Common Shape Ontology (CSO), and the need to provide answers to our CQs.

One major change presented in the section 1, is the relation between the VH and its different shape representations. In this document we define 3 kinds of shape representations: geometrical, 3D animation and structural. We consider that VHs can have one or all these kinds of representations.

Some integration of new metadata is given to the 3D animation concept. We present in the section 1.2 a classification of two types of animations: motion captured animation and key frame animation. However, we still have to provide more definitions of semantics around this concept.

To guarantee the reuse and share of the VH information, we work closely with the standardization activity (Task8.5). The implementation of standards, specifically H-Anim, exists in the coding of the VH representations. This version of the VH Ontology integrates partially the H-Anim standard, presented in the section 1.3. This standard allows that VHs created using authoring tools from one vendor can be animated using tools from another.

Since we are able to describe the structure of the VH in such a level of detail (provided by the standard), we came out with new competency questions (CQs) that are also listed in the section 1.3.

To give a general overview of the ontology, we present in the section 2 a diagram that reflects the mentioned changes and the new metadata. The H-Anim standard is integrated as part of the structural shape representation (which is part of the CSO). Here, we present the metadata that describes the H-Anim standard.

At the end of this document, in the section 3, we present two use case scenarios with competency questions associated. In these scenarios, the users search for specific models and animations. The questions posed by the users were answered using the search engine developed in the network. Finally we conclude this document in the section 4.

1 ONTOLOGY EVOLUTION

In this chapter we will describe the changes made in the virtual human ontology during this period. The first change we will explain is related to the basic structure of the ontology. The following changes are new concepts we needed to integrate in order to answer the competency questions provided in the first version of this document.

1.1 Ontology Structure

In the first version of the ontology we have described the creation of VHs at different levels:

- **Human body modeling and analysis**: morphological analysis, measuring similarity, model editing and reconstruction.
- **Animation of virtual humans**: autonomous or pre-set animation of VH.
- **Interaction of virtual humans with virtual objects**: virtual –smart– objects that contain the semantic information indicating how interactions between virtual humans and objects are to be carried out.

That structure was specifically focused on the virtual human domain. As part of the ontology‘s evolution we needed to change that structure in order to be compatible with other definitions of shapes. These changes were made keeping in mind the objective of answering our competency questions.

Now the VH representation is described by the relationship among four main levels of knowledge:

- **Geometry:** The geometry is the physical visual representation of the Virtual Human and it can be characterized by two main aspects: the first one is the body shape, and the second one deals with the accessories, garments, etc. Therefore, we can create a general description (class) suitable for any kind of geometry shape, and allow the specialized subclasses to inherit its properties, specifically for Virtual Humans or objects. This common class contains the geometric properties: number of vertices, edges, scale, material and texture. The Virtual Human has a Geometrical Representation that provides the kind of geometry it has (e.g. Surface mesh).
- **Animation:** The animation of a Virtual Human has to be formalized both for face and body. This is because the way of animating each one is different. Body animation can be represented as KeyFrame or MoCap, and can be encoded with different formats (e.g. MPEG-4, VRML, etc.). On the other hand, facial animation can be made through Morphing or vertex based control. In order that an animation can take place, the character should have a skeletal structure. This structure is defined in the Structural Descriptor. Here we have based this structure in H-Anim since it is the adopted standard in most of the animation formats. Nevertheless any other structures can be also represented.
- **Morphology:** Morphological Descriptor contains information like: age, weight, height, gender, etc.
- **Behavior:** an individual Descriptor and Behavior controller are used for describing the behavior. The Individual descriptor contains the constant definition of the behavior of the Virtual Individual like his personality, cultural identification, background, etc. The behavior controllers are algorithms that drive the behavior of the character considering the emotional state and its individuality.

These levels are intricately related: the size of a Virtual Human can be computed from the 3D shape of the body; the mental state will influence the gestures and motion that the Virtual Human can perform and the way it will perform them.

This new structure allows us to access directly to the different levels of knowledge, and also to generalize concepts within other ontology domains. We can also extend concepts of each field of knowledge. In the next subsection we provide new concepts of specific domains of the creation of virtual humans.

1.2 New concepts

To be able to answer specific CQs, we need to integrate specific semantics of the knowledge domains. During this period we have focused on the Animation. This process is very long and expensive, and results in a huge amount of files. For animators it is required that a solution allows sharing and reusing their work. Here, we have stated to define semantics to classify and annotate animations for its future storage and retrieval. Therefore, we have considered including concepts of standards to share animations more easily.

1.2.1 Animation

We have extended the Animation definition to be able to answer specific CQs:

- How many key frames does the animation have?
- What is the kind of interpolator used?
- What is the morphology (size, gender) of the actor who performed the animation?
- Is it raw data? (has it been mapped onto a skeleton or is it only sensor data)
- Does the data have any filtering?

To produce animations there exist several techniques [1], which can be divided into motion capture and key frame:

1) Motion capture: This animation technique consists in recording movements of a real person using sensors on his body. This technique provides realistic movements, but it is not easy to modify it in order to create new postures. The results are attached to a specific anatomy (actor's anatomy), and a lot of retargeting work has to be made for reusing it in different anatomies.

2) Keyframe animation: It is mainly defined as important frames that record the values of articulations, and the computer calculates with an interpolator the frames in between to produce the animation. This kind of animation can use other techniques as Inverse Kinematics or Dynamics. Its advantage is the high control of the animation but they are not realistic.

These two definitions are the starting point for the classification of animations. The integration of these concepts inside the ontology is presented in section 2.

1.2.2 Standards and File Formats

In computer graphics, there exists the H-Anim[2] standard that defines the skeletal structure of human-like characters. This structure mainly represents a simplified human skeleton with some specifications of the name, number and components of human articulations (as presented in Figure 1-1). It provides an articulated structure with standard nomenclature and best practice. A typical example of the benefit of such standards is that an animation sequence can be shared among different 3D characters if they use H-Anim as animation control structure.

The H-Anim standard is implicitly defined in the ontology. It is necessary for the level of information we want to achieve when searching for an animation or a VH, e.g. to answer the questions:

- What are the joints affected by this animation sequence?
- Does this VH model have the needed joints to play a given animation?
- Does the VH have landmarks? (of a VH shape having landmarks is known through its skeletal structure)
- What feature points has a given face? (these feature points are special animation landmarks located on the face)
- Does a joint have segments that have defined the given landmarks?
- What is the parent joint of a given body part (e.g. a leg)?

There are some file formats that implement this standard. They can be mapped inside the ontology. This means also that the structure of the ontology is in some way represented inside the files (geometry, skinning, skeleton, etc.). This standard has been used or could be used in many 3D exchange formats like MPEG-4 [3], VRML/X3D[4], Collada [5], CAL3D [6], BVH [7], etc.

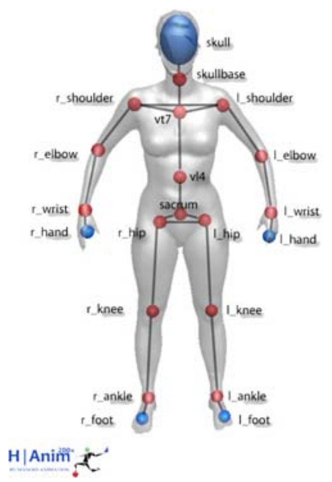

Figure 1-1 H-Anim standard scheme

Once we have identified the importance of the metadata that the standards can provide, we need now to implement the extraction of the information that each of these files contains in order to populate the ontology. This issue has to be solved in the future; as well we need to encourage the use of standards, because the process of sharing models among VH creators is limited by the lack of the use of standards and semantics.

In the section 4, we present some examples where, thanks to the standardization of models and animations, it is possible to reuse them avoiding extra work.

2 VIRTUAL HUMAN ONTOLOGY

As mentioned in the section 1.1, we have made changes in the structure of the ontology that define the VH creation. These changes are mainly reflected in the way of referring to the VH descriptors. In the previous version of the ontology we considered 3 descriptors: structural, morphological and general. We have removed some of those descriptors in order to integrate them in the common shape ontology, and to be coherent with the definition of shape of the other ontology domains. Moreover, the new structure allows the user to access directly to the concerned resources instead of having intermediate classes.

We also change the definition of the structural descriptor to be able to describe the H-Anim standard. For this integration we added to the common shape ontology a structural representation of the shape. Under this representation the H-Anim and other structure representations can be described.

2.1 Ontology Diagram

In the Figure 2-1, we present how the concepts are connected to each other in the ontology. The concepts that have the prefix **ShapeOntology** are part of the Common Shape Ontology, while concepts that are in a circle are subclasses of the concept **Resource** (in the bottom of the diagram). As a consequence, the circled concepts have inherited the properties of the Resource: author, version, file, etc. The resource concept was created to avoid repetition of properties in the classes created (Behavior Controller, Geometry Representation, Structural Descriptor and Animation 3D).

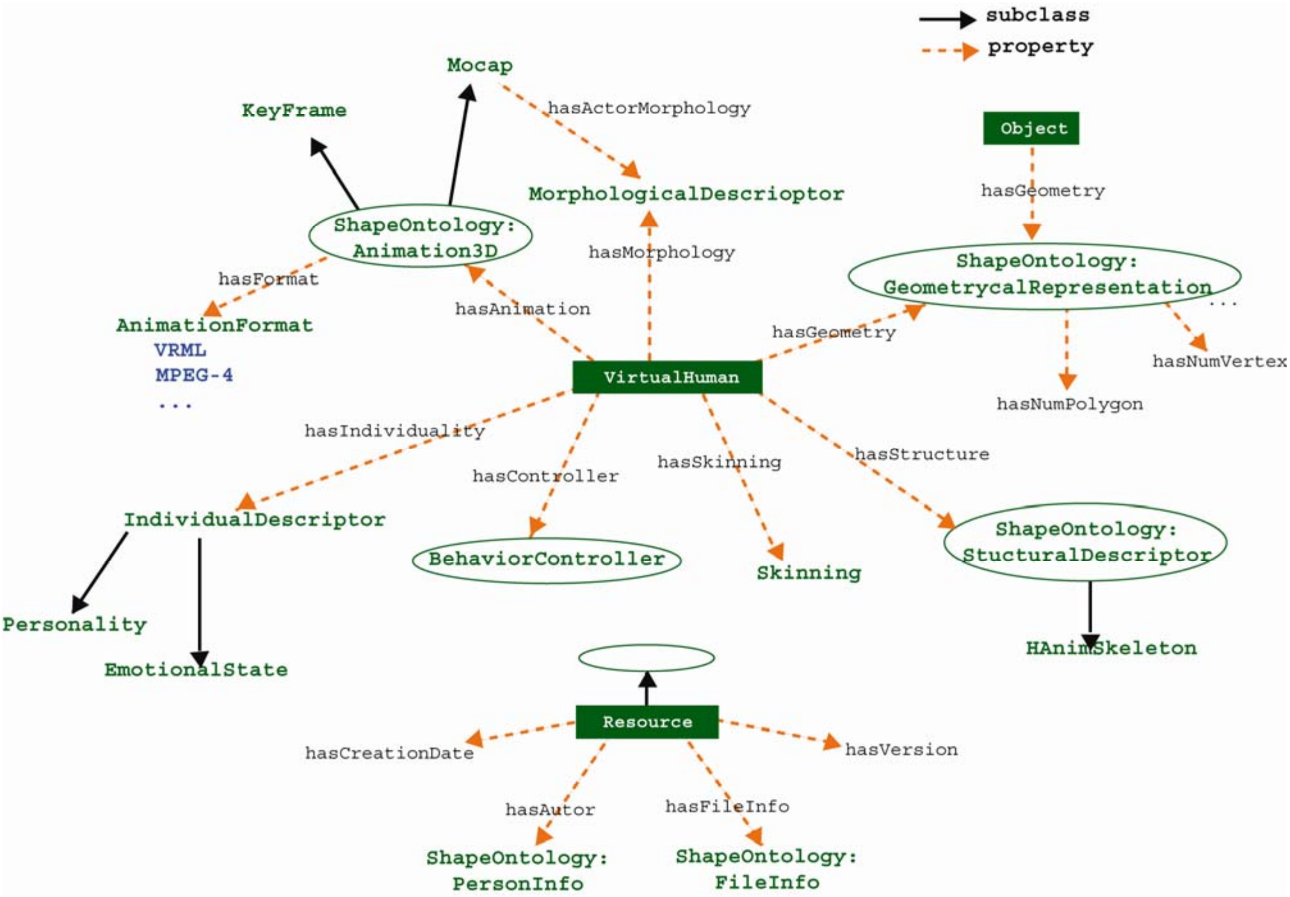

Figure 2-1 Ontology Structure

The Resources are the content that the user gets when searching for something. Some of the resources can be found in the same file or in separated files. In the case of behavior controller, the user can get an algorithm.

2.2 Structural Descriptor

The conceptualization of the structural shape representation is a new feature of the common shape ontology. It will be presented in the "D1.5.1 Report of an integrated view of the domain ontologies - 2nd version". The Virtual Human Ontology specializes this feature with the goal of conceptualizing the H-Anim structure and other kind of structures like topological graphs, etc.

As shown in Figure 2-2 the structural descriptors have been grouped into two categories: Graph and Multidimensional. The class shapeOntology:Graph represents the descriptors that encode the shape of an object as a graph; the descriptors represented by shapeOntology:Multidimensional encode the shape by using multidimensional elements like surfaces and volumes beyond segments or curves.

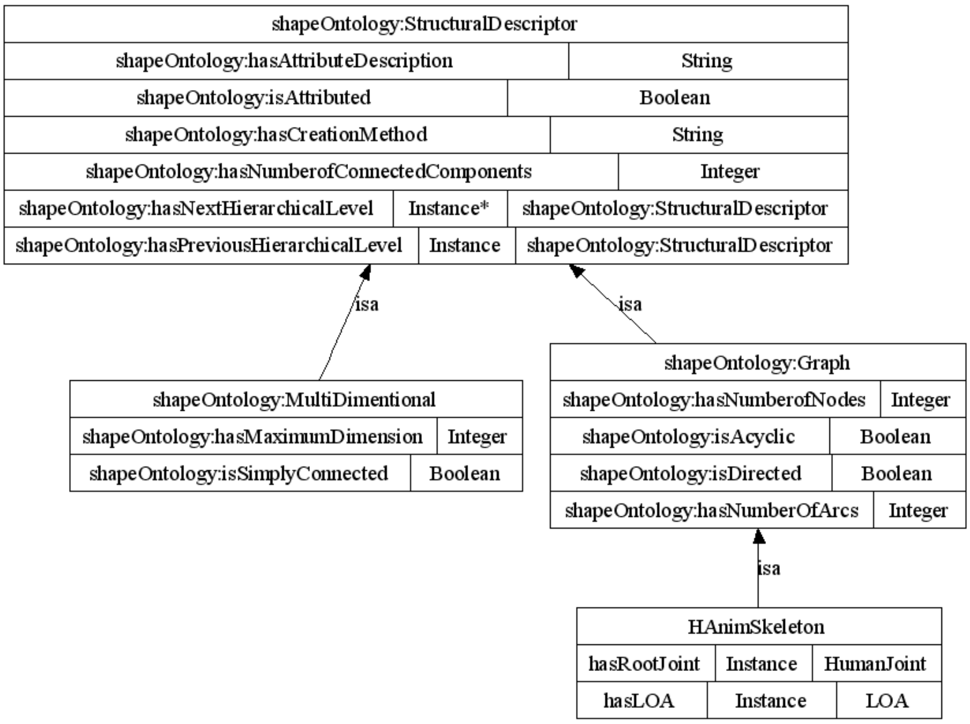

Figure 2-2 Structural descriptor properties and hierarchy

The class Graph is endowed with the minimal set of attributes needed to describe a graph: number of nodes and arcs, two flags that specify if the graph is directed and acyclic. The class Multidimensional has two attributes: isSimplyConnected, a boolean flag set to true if all the elements that constitute the descriptor are simply connected; and hasMaximumDimension, an integer that specifies the maximum dimension of the descriptor (e.g. a multidimensional descriptor consisting of volumes and surfaces has 3 for Maximum Dimension).

Beyond the attributes strictly related to graphs and multidimensional descriptors, the two classes share other attributes whose purpose is to provide general information about the descriptors. These attributes have been grouped in the class shapeOntology:StructuralDescriptor. The Boolean attribute isAttributed is set to true if the elements of the descriptor are endowed with attributes, such as describing the geometry of the object; in this case hasAttributeDescription provides a short text describing the meaning of those attributes. The attribute hasCreationMethod is a text field that provides information about the methodology used to extract the structural descriptor (e.g.

volumetric thinning or Reeb graph). The number of connected components of the descriptor is provided by the attribute hasNumberofConnectedComponents. Finally since a structural descriptor may be represented as a hierarchy of descriptors, hasNextHierarchicalLevel and hasPrevioustHierarchicalLevel implements the relations among the level of the hierarchy.

Under the described general classes we can start to particularize any structure, which is the case of the HAnim. At the same time we can know the characteristics of the structure by querying the ontology. For example, we can say that H-Anim is a structural descriptor that belongs to the Graph classification; it has the number of nodes equal to the number of joints; it is not acyclic and directed.

Another representation of a virtual human topology that has changed is the representation of the body parts. Each human part has a root joint where the hierarchy of the part comes. Some joints cannot have a part, but can have a segment. Therefore we can name those joints as a root of a specified part. The following Figure 2-3 of the ontology diagram illustrates this structure.

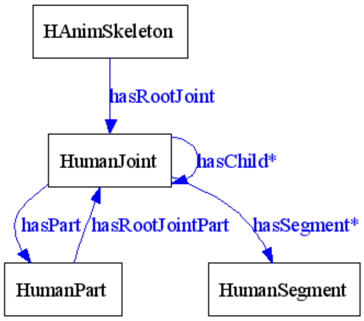

Figure 2-3 Human parts diagram

These presented modifications allow us to describe several features provided by the H-Anim standard. We need now to start populating this structure with the different file formats in order to validate it.

3 USE CASE SCENARIOS

In this section we present two case scenarios where we answer to specific CQs. The first case aims to find an animation that can be reused with a specific human model, while the second one explores the internal structure of a virtual human that will help the user to set up an animation technique (inverse kinematics).

For the first case, let's say that our designer Mireille wants to have an animation of a dancing woman. She looks in the ontology for animations of women that have the keyword dancing (Figure 3-1 search 1). She found one in C3D [8] format which can be opened in Motion Builder [9]. She makes a retargeting of the Mocap animation to the H-Anim skeleton, and she exports the animation to 3D Studio Max [10]. She found that the Mocap animation is made for a woman of 1.70 m. Therefore, she needs to look for a woman with a height of 1.70 m, with H-Anim structure, and also with a number of polygons between 50k and 80k, because that is the range we can load models and animate in our viewer without problems (Figure 3-1 search 2). She found a model in Collada[5] format that can be opened in 3D Studio Max. Finally she can attach the animation to the model and play it. The result of the last search is presented in the Figure 3-2, which is a screenshot of the search engine [11].

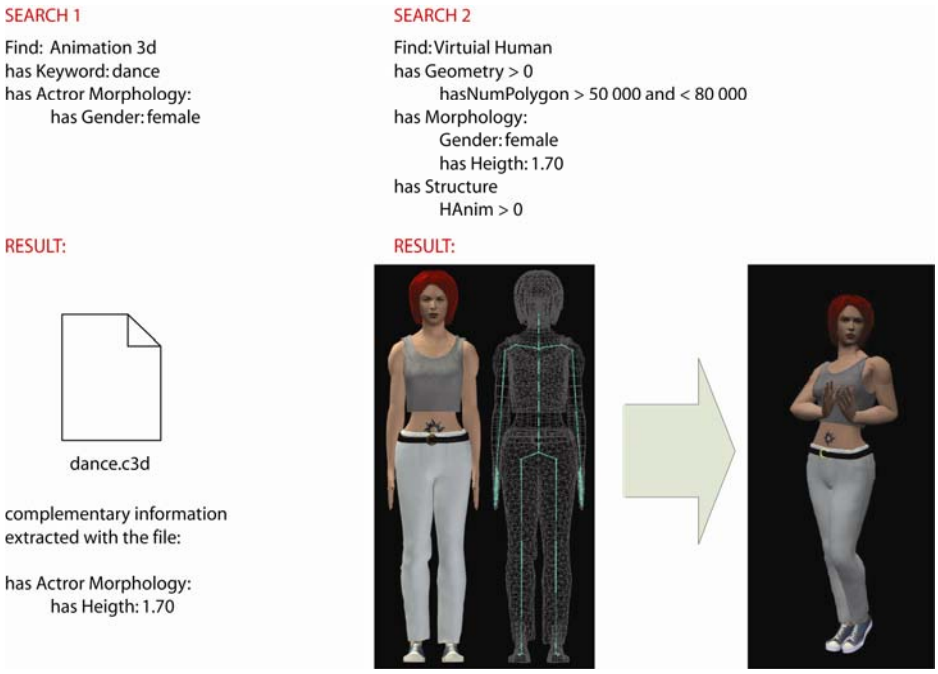

Figure 3-1 Search process to get Virtual Human dancing

This process can take 30 minutes in the best case, and much more depending on how "clean" the files are that the designer found. We can notice the benefit it could imply if we simplify the work of designers using an optimized search like this.

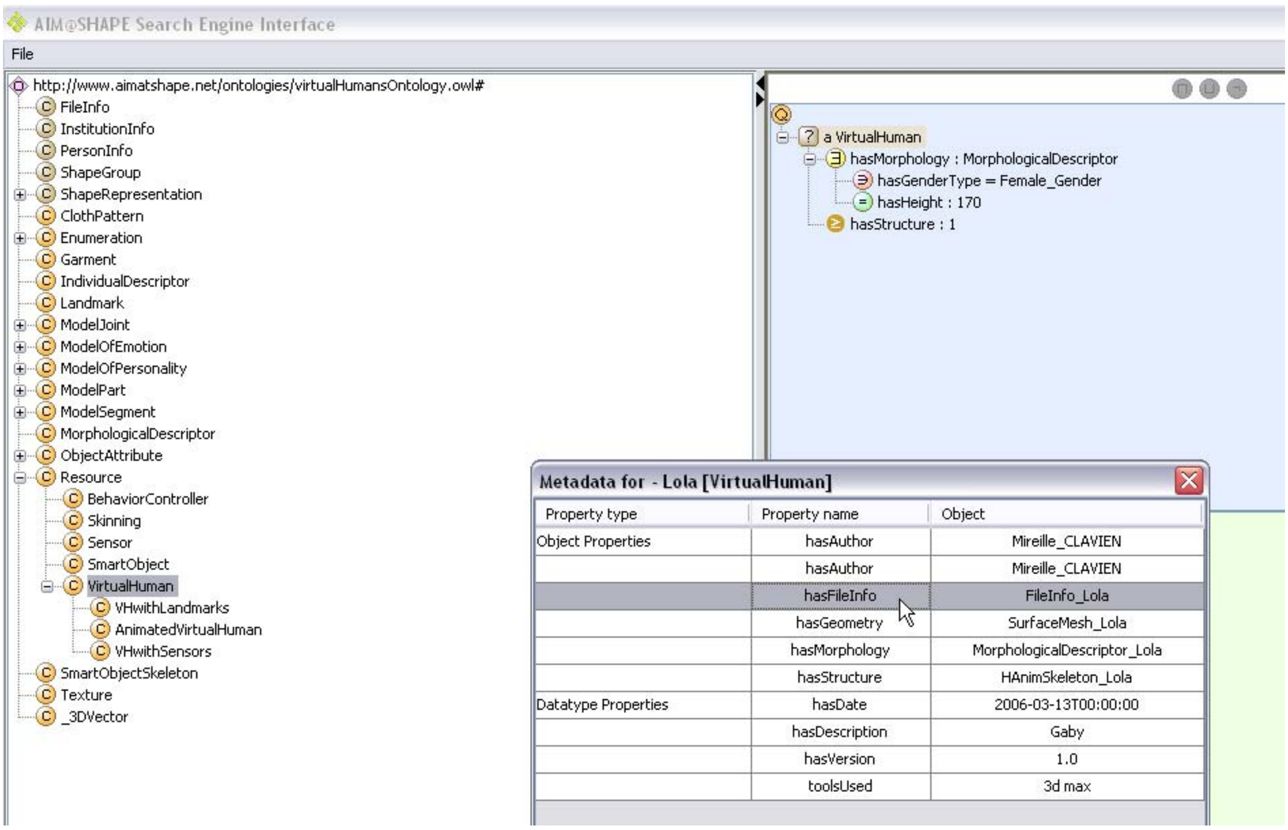

Figure 3-2 Search results of a VH woman using the search engine.

Currently, a designer would first need to identify the dancing animations by playing each one; then she has to check if the dancing sequence fits the selected human model. They may have to do a lot of retargeting because the scale of the model is not the same as the one used for the animation.

In the second scenario, let's consider that a programmer Neo wants to create a game with a male character. This character should be an animatable character, which means that it should have a skeleton. Therefore, he performs the search in the Figure 3-3.

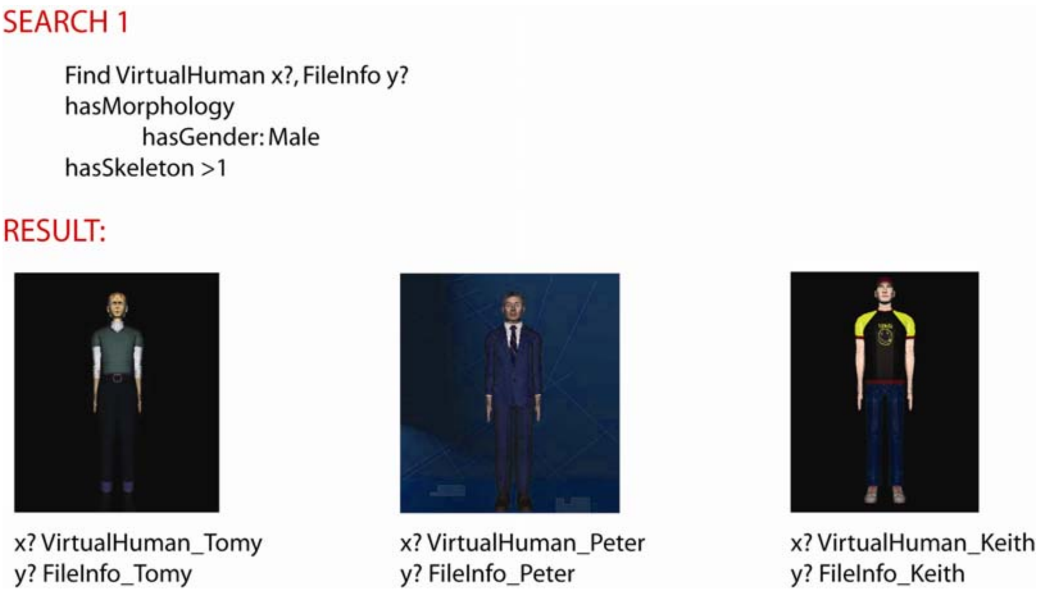

Figure 3-3 Search of a Male VH with H-Anim skeleton

From that search of male characters with skeletal structure, he chooses Keith. To make the animation of Keith, he wants to use Inverse Kinematics. He needs to know the end effectors of Keith. The end-effectors are Human Joints that are localized at the end of the skeletal structure, and do not have children. He performs in the ontology the query shown in the Figure 3-4.

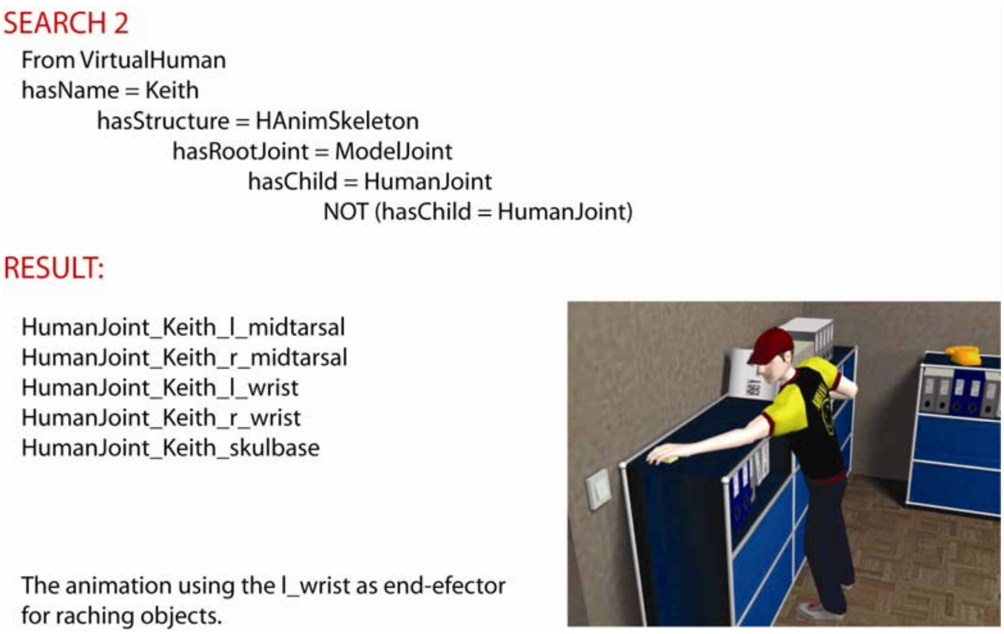

Figure 3-4 Search of end-effectors of a VH

In the last query we searched in Keith's skeletal structure, for all the joints that are children of Root Joint and do not have children. The results are the joints that will be used in the inverse kinematics as end effectors.

In the last scenario, we aimed to prove that the previous knowledge of the metadata of the structure of a virtual human can help to find specific information used to implement an animation technique.

The mentioned queries can be executed using the search engine [11]. The result of the second scenario in the search engine is presented in the figure 3-5.

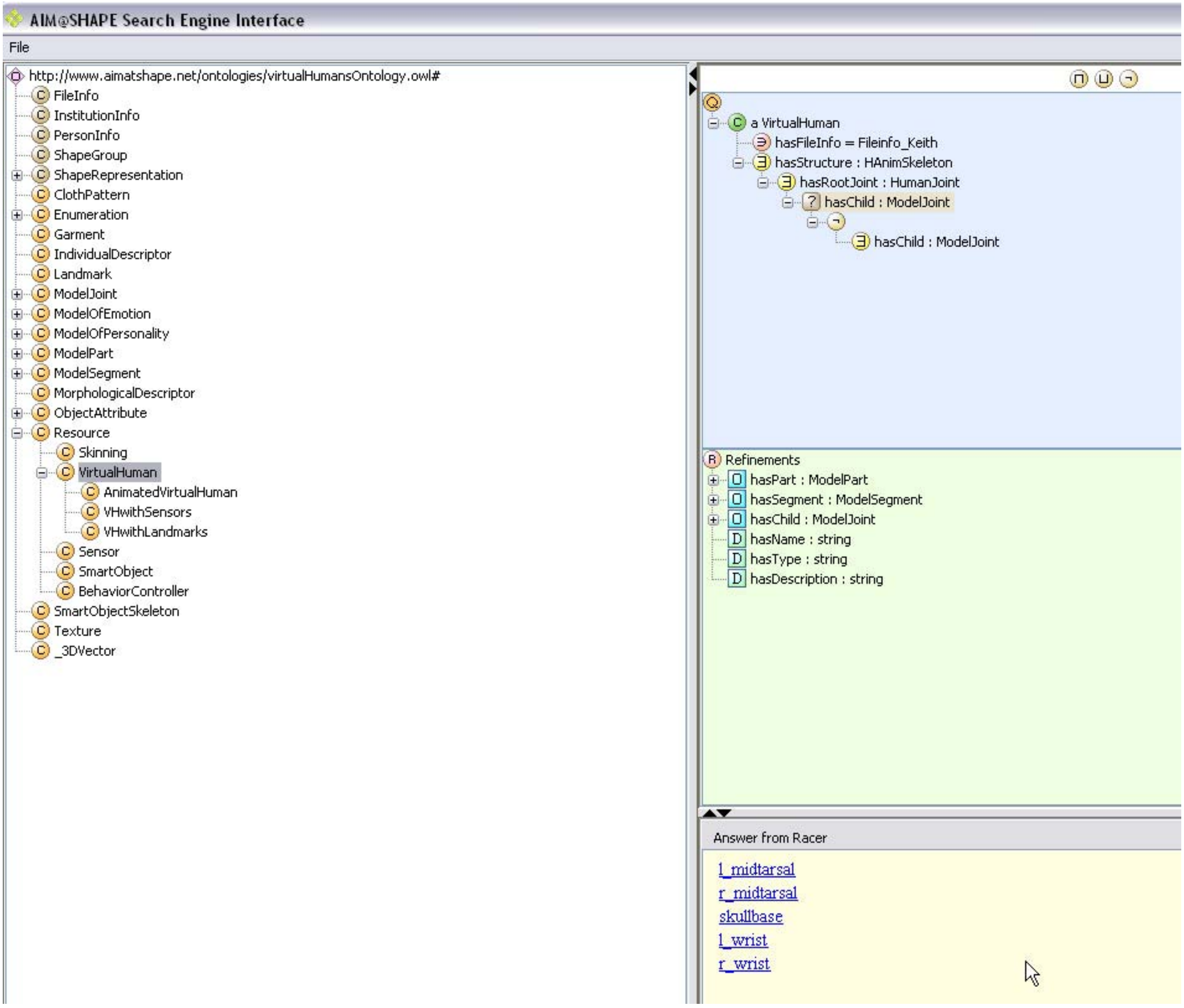

Figure 3-5 Search for the end-joints of Keith using the search engine

The questions presented in both scenarios have been answered by using the virtual human ontology; however more population of the ontology is needed to be able to test all possible cases.

Finally, we present the list of competency questions that we have validated during this period in the same way.

Model history

- Is this model obtained by editing another model?
- Which tools were involved in the synthesis/modification of this VH?

Features listing

- What is the height of the model? Is the model male or female? Does the model have cultural identification?
- Is this model obtained artificially or does it represent a real person?
- Which the VH have a landmark description?
- Which are the available structural descriptors for a particular VH?
- Which are the VHs standing (seating, walking, etc.)?
- How is the body model represented? (a mesh, a point set, etc.)

Animation sequences

- Which model does this animation use?
- Which joints are affected by this animation sequence?
- Are there any animation sequences lasting more than 10 sec that are suitable for this VH? The animation requires X Joints and the model has X Joints.
- Are there any running/football playing animation sequences for this kind of VH? Animation hasKeyword, the animation requires joints and the model has joints.

4 Conclusions and Future Work

For this second version of the ontology we have matured many concepts involved in the creation of VH. We have proposed a new structure that is compatible with the common shape ontology created within the work package.

In the next version we aim to be able to answer all the CQ postulated in the first version of this document. For that we need to make validations and probably small changes in the actual structure. This validation will help us to include other components like behavioral algorithms that are linked to a specific VH definition.

To accomplish this task we need to propose the right representation of the standards by populating the ontology with the different file formats. This requires an automatic extraction of metadata from files at the time of populating the ontology.

5 REFERENCES

1. R. Parent, *Computer Animation: Algorithms and Techniques*, ed. Morgan-Kaufmann. 2001, San Francisco.
2. H-anim., *The humanoid animation working group*. http://www.h-anim.org.
3. MPEG-4, *Overview of the MPEG-4 standard*. http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm.
4. Web 3D consortium. http://www.web3d.org/.
5. COLLADA, *Asset Exchange Schema for Interactive 3D*. http://www.khronos.org/collada/.
6. CAL3D. http://home.gna.org/cal3d/.
7. Biovision, *BVH*. http://www.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/BVH.html.
8. C3D, *Coordinate 3D*. http://www.c3d.org/.
9. Autodesk, *Motion Builder Software*. http://usa.autodesk.com/adsk/servlet/index?id=6837710&siteID=123112.
10. Autodesk 3ds Max. http://usa.autodesk.com/adsk/servlet/index?id=5659302&siteID=123112.
11. George Vasilakis, Marios Pitikakis, Manolis Vavalis, and Catherine Houstis. *A Semantic Based Search Engine for 3D Shapes: Design and Early Prototype Implementation*. in *2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies (EWIMT2005)*. 2005. http://www.iee.org/oncomms/pn/visualinformation/ewimt2005_0067.pdf.