IST FP7 231507

# D3.5 Integration in Virtual Environments

| | |
|---|---|
| *File name* | PuppyIR-D3.5-Integration-in-Virtual-Environments.doc |
| *Author(s)* | Franciska de Jong (UT), Marie-Francine Moens (KUL), Job Zwiers (UT), Ian Ruthven (UoS) |
| *Work package/task* | WP3 |
| *Document status* | Final |
| *Contractual delivery date* | M34 |
| *Confidentiality* | CO |
| *Keywords* | virtual reality, mixed reality, virtual guide, vizualisation, interactive narrative |
| *Abstract* | In this report we describe how some the interaction models and information services from the PuppyIR Framework could be integrated into Virtual Environments that are geared towards use in the context of cultural objects and/or younger audiences. |

## Table of Contents

# Executive Summary

In this report we present an inventory of some of the options and obstacles for how to integrate the interaction models and information services from the PuppyIR Framework developed in WP2 and WP3 into virtual environments that are geared towards use by younger audiences. The report is inspired by some of the recently proposed initiatives at KUL, UoS, Museon and UT, such as MUSE, CETINGS and `meSch`.

# 1 Introduction

Within PuppyIR workplan no plans were included for developing interaction models and/or services that were geared towards the integration in virtual environments. But given the long-term interest in modelling virtual works and interaction tasks involving virtual agents or characters at several of the partner institutes within the PuppyIR consortium it was anticipated that the results of the project could be fit for integration in environments involving interaction with virtual agents and virtual worlds in general. This report is an inventory of the some of the options and obstacles for uptake of the PuppyIR results in virtual environments from a technical perspective, inspired by some of the recently proposed initiatives at KUL, UoS, Mueson and UT, such as MUSE, CETINGS and `meSch` that will be introduced in more detail below.

In the PuppyIR Framework several novel interaction models have been designed and implemented. The novelty is mainly in the support for collaborative search tasks in combination with touch tables (e.g. D7.2 Museum Demonstrator), and on the realization of interfaces for individual search scenarios with children as user and specific collections as starting point. For the latter category several metaphors (Show and Tell and Beat the Witch[1]) turned out to be suitable basis for the enhancement of search sessions, while the design and development of EmSe (aka the Hospital Demonstrator, described in D7.4) is centered around the Body Browser as interface. The concept of the body itself is not a metaphor, but due the cartoonish design of the Body Browser it shares with some of the other of interfaces the role of stimulating the imagination and activating the cognitive and conceptual background of children while engaging with the PuppyIR interfaces and information services.
The PuppyIR work in WP3 allowed the coupling of novel interaction designs to the processing of the content and to adapt it to the needs of children by filtering for topics and /or by adjusting the language to the abilities of children from specific age groups.

In this report the question is addressed if and how these approaches can be taken further to function as a good basis for integration in virtual environments. In Section 2.we will describe two recent initiatives aiming at the realization of virtual environments in which interaction with information is a salient feature. In section 3 an overview the research issues that need to be addressed before a fruitful coupling of the PuppyIR Framework to virtual environments could be realized. Section 4 (Conclusion) summarizes the main findings.

[1] Show and Tell helps children manage the complexity of searching by removing difficult decisions, such as creating queries, and supporting organization of information. Beat the Witch encourages children to assess multiple information objects and make decisions about relevance. Cf. [1].

# 2 Interaction with Information and Virtuality

Some of the earliest interactive systems in IR were based on conversations or dialogues modelling reference interviews between librarians and library patrons. The aim was to use a familiar process – obtaining help from an expert intermediary such a librarian – to elicit useful information on a user's background knowledge, preferences, and information on what functions they may wish to use. Another conceptual metaphorical device for interacting with information was to exploit existing understandings of information spaces such as libraries or bookshops. As also observed in [1] and [7] many metaphorical interfaces only serve as display interfaces rather than full interactive metaphors.
One specific application of visual interface design techniques is the virtual or 3D interface. In such a system the user is immersed in the interface environment; he or she can explore that environment by moving around, examining or collecting objects, and so on. The best current examples of such virtual interfaces are to be found in the world of video games.

For PuppyIR a step towards the further enhancement of the interaction models for children would be to integrate metaphors and interactive storylines in 3D search environments. The approach towards human – content interaction envisaged in MUSE, CETINGS and `meSch` could be a stepping stone.

## *2.1 MUSE (Machine Understanding for Interactive StorytElling)*

MUSE (Machine Understanding for interactive StorytElling) is a recently started EU project in which KUL participates. The main motivation of is to develop a new methodology, building blocks, and prototype that can be reused in the development of 3D interactive environments from free text materials. MUSE will introduce a new way of exploring and understanding text information for "bringing text to life" through 3D interactive storytelling, achieved by transforming textual documents into virtual worlds where the user can interact with the content in a game-like environment.

Taking as input natural language texts like children's stories or medical patient education materials, MUSE will process the natural language, translate it into formal knowledge that represents the actions, actors, plots and surrounding world, and then render these as virtual 3D worlds in which the user can explore the text through interaction, re-enactment and guided game play.

## *2.2 CETINGS*

CETINGS is a project proposal (still under review) with participation of UT that aims to develop the concept of *interactive mediated narrativity* into an instrument for presentation, contextualisation, interpretation and exploration of cultural artefacts. CETINGS will create a technological framework that will support the creation and deployment of richer, more natural and more rewarding forms of interactive narration. It will do this by *generalising* and *organically integrating*, in a concept-denoted *interactive narrative world*, three distinct forms of mediated narrativity: non-linear interactive narrativity carried by recorded content (video, voice recordings, etc.); interactive narrativity expressed through generated content (animation, synthesised voice, etc.); and explore-able 3D spaces that adapt dynamically to the context of each particular exploration. The research will address both the refinement of such experiences as well as the creation of appropriate technological solutions. CETINGS aims to develop the concept of *interactive mediated narrativity* into an instrument for presentation, contextualisation, interpretation and exploration of cultural artefacts that are representative of content curated at heritage institutions and educational collections.

For illustration, a user experience in an interactive narrative world could include listening to the computer generated voice of a historical figure “reading” fragments from existing texts, selected such that they respond to the interest of the visitor expressed as unstructured questions. This could appear more or less as a “conversation” with the historical figure, and could be automatically initiated when, situated in a 3D virtual exhibition room with scanned images of some important manuscripts, the visitor approaches a manuscript authored by the respective personality. After the conversation is initiated, the 3D manuscript room could be replaced by a video projection room where visual cues, such as changing static pictures (key frames from the video footage), invite the visitor to explore video footage related to the immediate topic of the conversation. Accepting an invitation, the visitor enters, in parallel with the conversation, an interactive video narration that is automatically generated in real-time (during viewing) from fragments of the video archive. The video narrative follows the topics dictated by the visitor’s input, but is continuous, appearing as a traditional documentary, but custom made to the visitor’s interest. Choosing an object that appears in the video footage, say an invention belonging to the respective personality, takes the visitor to a 3D exhibition room that shows digital replicas of that and/or related objects. This room is inhabited by a virtual guide, with expertise in these devices, who takes

From the point of view of the experience of the child end-user, an interactive narrative world can be a place for

- interactive narrative viewing and listening: narratives that in real time can adapt to the child’s information need
- conversation and play – realised through superimposing narrative structures over 3D spaces populated by (autonomous) agents and narrative fragments
- guided discovery and exploration – realised through wandering, guided, in nested virtual 3D spaces inhabited by interactive narratives that are visually represented, or through digital replicas of physical artefacts

## 2.3 `meSch` (Material EncounterS with digital Cultural Heritage)

`meSch` is an IP project proposal under review and involving UoS and Museon. It has the goal of designing, developing and deploying tools for the creation of *tangible interactive experiences* that will connect the physical experience of museums and exhibitions with relevant digital cross-media information in novel ways. A wealth of digital cultural heritage content is currently available in on-line repositories and archives, it is however accessed only in a limited way and utilised through rather static modes of delivery. `meSch` will bridge the gap between visitors’ cultural heritage experience on-site and on-line by providing a platform for the creation of *tangible smart exhibits*, that will enable heritage professionals to compose and realise physical artefacts enriched by digital content without the need for specialised technical knowledge: the platform will include an authoring toolkit for the *composition of physical/digital narratives* to be mapped to interactive artefacts, and an embedded multi-sensor digital system platform for the construction of ad-hoc physical smart exhibits. The ultimate goal of the project is to support the creation of an open community of cultural heritage institutions driving and sharing a new generation of physical/digital museum interactives.

The understanding of children’s information interaction, especially with museum artefacts, gained from PuppyIR has led to specific objectives within `meSch` for child end-users, for example:

- The searching, selection and composition of multimedia content into multiple narratives will help children create their own tours through online museum visits;
- The annotation of the narratives with smart objects' behaviour descriptors that will trigger changes in the objects' behaviour according to the type of visitor (e.g. children);
- The automatic recommendation of new objects of interact to reduce the complexity of searching for children;
- The creation of interfaces that help end-users, specifically children, manage complex interactions through metaphorical interfaces (as in PuppyIR Task 2.3)

Annex A describes an elaborate scenario that could be supported by this kind of platforms.

# 3 Research issues

In a future setting the information seeking child could be the actor in a Virtual Reality (VR). A goal could be to create a virtual world made up both from information extracted from the results of the information seeking (stories, fairy tales, but also child blogs) in which the child and his or her friends become the main actor. Eventually we could establish a mixed reality, a mixture of the personal world of the child and the world created by the information sources (e.g., the child meets Napoleon in an appropriate setting, when he or she wants to find information on Napoleon). In PuppyIR we did not want to go that far, because technology is not advanced enough to properly provide this functionality. However, PuppyIR seems a good starting point both from a child user and from content perspectives to take the development of these types of information access further. The fact that it could help to bridge the gap between content classification and access (from the point of view of “finding” objects) and interpretation and contextualisation in the form of rich narratives is fully in line with the driving force behind PuppyIR. [3]

For several topics further investigations and/or explorations are needed, among which:

- Translating text into VR representations
- Character selection/creation for narratives
- Multi-user Environments
- User profiling

## 3.1 Translating text into VR representations

One of the major research challenges is to develop a consistent mapping strategy between the semantics of textual descriptions and the knowledge layer underpinning the virtual world.

The knowledge that we need in order to translate text into another medium might be of a different nature than the knowledge types so far acquired. In addition it has never been studied how to integrate this knowledge into text understanding. Integrating this knowledge will require advances in how information extraction methods are combined with translation techniques.

*Relevant state-of the-art in extraction of semantic representations from natural language*

Real understanding of text requires making inferences based on world and domain knowledge. Recent research has attempted to automatically acquire world knowledge from the vast amount of natural language text available on the Web. Researchers have categorized different types of common-sense relationships that are necessary for understanding, including *cause-and-effect*, *preconditions*, *is-member-of*, *has-parts*, etc. (LoBue and Yates 2011). Some of the more successful approaches to acquiring such relations with minimal human intervention include: searching based on query patterns like *to X and then Y* to identify *happens-before* relations (Chklovski and Pantel 2004), and extraction of relational tuples using part-of-speech tag features (TextRunner) (Banko et al. 2007), Wikipedia and dependency features (Wu and Weld 2010), or syntactic and lexical constraints (Fader et al. 2011).

The INDIGO project[2] develops a robot that interacts with each visitor in a unique way through language and gestures. The project builds very strongly on the work done in the M-PIRO[3] project, in which software was developed to automatically generate descriptions of museum artefacts. The language that is used for the descriptions differs per type of visitor and their personal preferences. The INDIGO robot provides information about the objects within a museum, based on the manner in which visitors respond to it, and shows visitors other objects and information that applies to their preferences.

## *3.2 Character selection/creation for narratives*

In order to be able to design an environment with virtual guides that could support children in the engagement with information, suitable characters should be selected and turned into a virtual agents with a sufficiently rich profile. In the context of heritage collections this could be an historical figure such as Napoleon, or a fictitious figure (talking human, animal or object). The character could be a guide that can help in raising and answering all kinds of questions on canonical historical facts, everyday historical life, etc.

*Relevant state-of the-art in designing virtual agents*

Interaction can take place through high quality computer animations and man-machine interaction. Natural interaction can be established by incorporating detection of gaze and gestures to keep track of the user's needs/wishes/plans.
Often virtual applications require a custom-built virtual human body. At the University of Twente Elckerlyc is developed, a BML[4] realizer for generating multimodal verbal and non-verbal behavior for virtual humans [5]. Elckerlyc[5] includes a virtual agent coming with a set of body and facial animations, and it comes with a default render environment that can import meshes of characters in appropriate XML formats such as COLLADA, the industry standard format designed to allow exchange of assets between different applications in the 3D industry. Elckerlyc has been applied for several settings, and characters, including Griet (the girl that figures in the novel/movie Girl with a Pearl Earring, on the life of the girl on the famous Vermeer painting with the same title). The latter 3D character guides users through an art exhibition, but this kind of agents can also be applied for the explanation of health information by virtual agents. [6]

## *3.3 Multi-user Environments*

In the Museon setting the multi-touch table supports collaboration, but the post-visit application is based on a single user model. Multiplayer games in virtual settings could be an interesting follow-up for the collaborative scenarios for which WP2 has designed interaction models. But they would require attention for a number of issues: spatial simulations that support real-time interaction of concurrent users. Other issues related to the challenges of real-time delivery are: consistency, persistence, security and privacy, interoperability, adaption to nomadic systems, streaming for multimedia content.

[2] http://www.ics.forth.gr/indigo/workpackages.html
[3] http://www.ltg.ed.ac.uk/mpiro/
[4] BML - Behavior Markup Language
[5] "Elckerlyc" is the protagonist of a Dutch morality play with the same name, written at the end of the Middle Ages. It is widely assumed to be the source of the English morality play "Everyman". The protagonist represents every person, as they make the journey towards the end of their life.

*Relevant state-of the-art in multi-user environments*

An important aspect for virtual characters that operate within an interactive narrative context is the ability to perform *continuous* interaction with the user. For example, it cannot be assumed that interaction between characters and users is turn-based, as is often assumed for more classical dialogue systems. Systems like the Elckerlyc framework explicitly address the continuous interaction paradigm and the impact on multimodal behaviour generation.

The technical aspects of designing, developing, and deploying multiplayer online games them is highly interdisciplinary and involves experts from many domains, e.g., graphics, networking, protocol and architecture designs. For educational applications possibly a lot can be learned from game environments such as World of Warcraft and Second Life.

## 3.4 User Profiling

For many scenarios of use the concept of a series of interaction sessions for online enrichment of the on-site museum experience needs to be elaborated, in particular the stimulation of active input from users in automatic recommendation and guidance to enable dynamic contextualisation of information.
In a virtual reality setting the requirements for ensuring consistent and persistent experience across sessions are even less trivial than in more traditional platforms. These requirements impose critical constraints on the approach towards user profiling. The integration of the PuppyIR filtering mechanisms developed in WP3 and virtual platforms are also highly dependent on adequate profiling instruments.

*Relevant state-of the-art in user profiling*

Several approaches do include active input from visitors. An example system that enables users to influence its behaviour was built in the CHIP[6] project. In the CHIP environment, visitors can create a profile and a virtual tour on the museum website based on their interests, before they visit the museum – they prepare the visit in the virtual space. Subsequently, they use this tour as a starting point when visiting the museum. CHIP developed a content-based recommender system, based on semantic web technologies (SKOS). The guidance and recommendations given during the tour are based on the recorded data of the visitor's behaviour and their profile on the website. Visitors can alter the number of objects they want to see and their object preferences during their visit. The guidance also adjusts itself to the time visitors want to spend in the museum and the time they are spending with each object. The online profile of the visitor is directly updated during the museum visit and each visitor can access this data afterwards on the museum website.
A similar approach is taken in the PEACH project[7]. Users can give feedback on the output of the system. The result of the harvested preferences and the data recorded from their visit are used for subsequent visits made online, for example by compiling suggestions about other interesting objects and links to relevant Internet pages [4].
The Dutch Institute for Sound and Vision includes this approach in its repertoire of end-user experiences. RFID chips (in a ring) are used as a means of tagging content for storage. On site, visitors can, for instance, simulate and record their own version of

[6] http://chip-project.org
[7] http://peach.itc.it/home.html

television and radio programs, create short video clips and store their favourite television and radio programs on a ring with an RFID chip. After the visit, visitors can access a website and download the content they created.[8] Technisches Museum Wien takes a similar approach in the medien.welten exhibition[9]. Visitors are given an RFID card which they can scan beside the items of interest. Subsequently, they can access an automatically generated web album that contains structured information about all their selected objects on the museum's site, via their unique card number.

More advanced forms of recommendation and guidance have been explored through more sophisticated input forms. The MiRA project[10], for example, creates mixed reality agents for the exhibition space. The MiRA agent consists of a physical robot and an animated avatar. Visitors are wearing a special headset that enables them to view the conjoint avatar, robot and animation overlapped.

[8] http://www.beeldengeluid.nl
[9] http://www.tmw.at/
[10] http://mira.ucd.ie/

# 4 Conclusions

The work conducted in PuppyIR, especially the collaboration around the Museum demonstrator and interfaces, has led to new findings on how childrens' engagement and learning can be increased in mixed offline and online museum environments. Through new research, such as the meSch proposal, these research findings will be useful as the springboard for new museum demonstrators that use the same concepts as PuppyIR's demonstrators (such as information quests) but incorporate novel tangible interactive devices. The findings, including those gained from the experiences of developing post-visit museum demonstrators, will also facilitate new research into making virtual museum visits more meaningful for children by use of personalized narratives and VR.

# References


[1] D2.3 – Metaphorical interaction models and interfaces (PuppyIR deliverable)

[2] D5.4 – Evaluation Toolkit (PuppyIR deliverable)

[3] Mason R. (2005) Cultural Theory and Museum Studies. In: Macdonald, S, (ed.), A Companion to Museum Studies. Oxford: Blackwell, pp. 17-32.

[4] Stock O., Zancanaro M. (Eds.) (2007) PEACH: Intelligent Interfaces for Museum Visits. In: Cognitive Technologies Series, Springer Verlag.

[5] H. van Welbergen, D. Reidsma, Z.M. Ruttkay and J. Zwiers, (2010). Elckerlyc - A BML Realizer for continuous, multimodal interaction with a Virtual Human, In: Journal on Multimodal User Interfaces, 3(4):271-284.

[6] T. W. Bickmore, L. M. Pfeifer, and M. K. Paasche-Orlow. (2007) Health Document Explanation by Virtual Agents. In: Proceedings of the 7th international conference on Intelligent Virtual Agents. Lecture Notes in Computer Science, Springer-Verlag, vol. 4722, pages 183–196.

[7] Jamshid Beheshti, Leanne Bowler, Andrew Large and Valerie Nesset, (2009) New Directions in Cognitive Information Retrieval Towards an Alternative Information Retrieval System for Children.

# Annex A - Scenario

This Annex contains a use scenario for VR platforms tuned towards children that integrate functionality for interaction with information.

John is responsible for the educational unit of a lively museum with a natural science collection. Local schools often visit to complement their curriculum and enrich the children's learning. Yesterday John discussed with a group of teachers the need to introduce the topic of sustainability into the museum portfolio of lessons from which visiting schools can choose themes for their visits. He sits at the computer and connects to the `meSch` portal to create a new set of adaptive smart objects to bring the topic of sustainability to life in the museum. The digital content repositories of his institution have been already connected with the part of the `meSch` platform that acts as a brokering system searching for information broadly on both local and remote repositories. Though the portal, he can search seamlessly through the museum's own material, public resources such as Europeana and Web2.0 repositories such as Flickr, YouTube and Qwiki. After the discussion with the teachers, he now knows what is covered in class on the topic: therefore John has a well-formed idea on the type of story that the smart objects will tell and has also created some specific content for it. He loads the files in the `meSch` portal and starts looking for additional material, particularly pictures and videos to illustrate the points.

A recommender system is running behind the `meSch` portal and uses the metadata of what has been uploaded by John to find relevant material among the millions of pieces of information made available through the brokering component. The recommender uses different algorithms, e.g. it searches for similar content and similar media (more text or images), or similar content but different media (if the loaded material is text, it looks for images or videos), or complementary content (if the piece is about nature the recommender looks for the same topic but from a different perspective, for example cultural understandings of nature). The suggested content is displayed on a side bar of the authoring tool John is using.

He begins the iterative process of looking for the right content, and as he goes he saves those items that look promising for composing a nice narrative, or multiple narratives, or different perspectives on the same topic. John is now satisfied with the range of material he has collected. Indeed some of the suggestions presented by `meSch` triggered in him some new ideas on how to contextualize difficult items: while searching for "rain forest" he has found amazing pictures on Flickr from everywhere in the world showing how different rain forests can be, i.e. fern forests in New Zealand and tropical forests in Amazonia, as well as some appalling images of destruction in Congo, Brazil and Sumatra and an animation provided by Greenpeace on YouTube on the topic. He selects an activity template he used before and proved very effective with children: it is a quest where kids must find the smart objects that compose the answer to a question within a given time. He starts composing those pieces of information into the quest template, creating narratives that cover several rooms/objects/topics. A narrative template is independent from the actual content and is used to structure the content, for example, in terms of what comes before and after, and which conditions release what portion of information or interactive behaviour. The narrative is represented as a network: each node contains a snippet of content and is controlled through a set of activation rules that, at run time, will determine if and how the content should be delivered. The set of activation on the nodes represents in abstract terms the behaviour of the smart object: by annotating the nodes with the behaviour (the activation rules) John specifies how each object will behave when the visitors encounter it. In the meSch portal John can now test (as a simulation) what will happen when the adaptive smart object is in use. All works well: the information is delivered in the right order and hints are given at the right time; data of the interaction are collected and in the right format.

Children like competitions so John decides to implement the same narrative templates with two different behaviours and to try them out with the classes: he is thinking of a box that will ask the question and the child have to find which smart object fulfil the quest; in the alternative version, the children have a light that will extinguish unless they find the smart objects that contains the right information to revive it. The narrative template and network is the same, what changes are the activation rules. The box and the light are always connected and both record the interaction so that the two teams can compare the result and see who won. The log of what they have done is used to create a data souvenir of the experience and is used as starting point to explore new things online.

John is now ready to go and make the smart objects. He downloads the templates into a memory sticks and goes to the nearby FabLab in his city to make them. In the FabLab he can design the physical form of the object. He loads the files on the meSch embedded multisensor digital system platform available at the FabLab and start connecting the abstract description of the smart object behaviour with physical elements, e.g. sensors, screens, loudspeakers. Now John can physically try the electronics and see if all works as he expected. Everything works fine: he is now at the last step, the making of the physical form. Every physical component has a predefined 3D encasing shell and both the box and the light models are already available in the system so John selects each and 3D prints them out. Finally, with the help of the FabLAb staff, he polishes and completes the smart objects ready for the first class due to come to the museum next week.

The class arrives and the children are split in two groups competing with each other: one group has the box, the other the light and they start running around to find the smart objects that would help them answer the quest first and more accurately. When they have finished they go to the touchscreen that shows the stories they have composed by finding the smart objects in the museums. Back at school the teacher wants to reinforce the learning and asks the children to go online, look the recording of their visit (reconstructed out of the logs captured by the smart objects) and find three more images through the system that can complement their story.