Digital Culture & Heritage
Patrimoine & Culture Numérique

Haus der Kulturen der Welt, BERLIN
Aug. 31st - Sept. 2nd, 2004
31 Août - 2 septembre 2004

AN OPEN-ENDED TOOL TO COMPOSE MOVIES FOR CROSS-CULTURAL DIGITAL STORYTELLING: TEXTABLE MOVIE

Cati Vaucelle, Glorianna Davenport

Media Lab Europe, Dublin, Ireland

http://www.mle.ie

Abstract

This paper presents *Textable Movie*, an open-ended tool that allows any storyteller to become "video-jockey" able to improvise a media story in real-time drawing from an available collection of annotated images and videos. In the framework of digital storytelling, Textable Movie promotes the idea of maker controlled media and can be contrasted to automatic presentation systems. Its graphical interface takes text as input and allows users to improvise a movie in real-time based on the content of what they are writing. Media segments are selected according to how the users label their personal audio and video database. As the user types in a story, the media segments appear on the screen, connecting writers to their past experiences and inviting further story-telling. Video Jockeys perform using their own or someone else's video database presenting different visual perspectives on the same story. By co-creating movie stories that are first improvised from a personal video database and then projected into someone else's video database as two tellers merge their stories, young adults are challenged in their beliefs about other communities. In this paper we will present our ongoing research on a future tangible version of Textable Movie for a direct control and visualization of multiple point of views.

Keywords: digital stories, storytelling, real time, database, video-jokey, cross-cultural exchange

Zusammenfassung (DE)

Dieser Beitrag präsentiert *Textable Movie*, ein erweiterbares Tool, das jeden Geschichtenerzähler zum Video-Jockey werden lässt. Es kann eine Geschichte in Echtzeit improvisieren, indem auf eine vorhandene Sammlung von kommentierten Bildern und Videos zugegriffen wird. Im Rahmen des digitalen Storytelling fördert *Textable Movie* im Gegensatz zu automatischen Präsentationssystemen die Idee des Nutzergesteuerten Mediums. Das grafische Interface verwendet Text als Input. Es erlaubt dem Nutzer einen Film in Echtzeit zu improvisieren, der darauf basiert, was dieser gerade schreibt. Mediensegmente werden entsprechend der Beschriftung der persönlichen Audio- und Videodatenbank der Nutzer ausgewählt. Während der Nutzer eine Geschichte schreibt, erscheinen die Mediensegmente auf dem Bildschirm und verknüpfen die Schreiber somit mit ihren vergangenen Erfahrungen und laden zum weiteren Geschichtenerzählen ein. Video Jockeys arbeiten indem sie ihre eigene Datenbank oder die eines anderen benutzen und dieselbe Geschichte aus

unterschiedlichen visuellen Perspektiven zeigen. Dadurch, dass die Filmgeschichten in Zusammenarbeit gestaltet werden - zuerst werden sie von einer persönlichen Videodatenbank improvisiert und dann, wenn zwei Geschichtenerzähler ihre Geschichten verbinden, in die Videodatenbank von jemand anderen projiziert - sind die Vorstellungen, die junge Erwachsene von anderen Gemeinschaften haben, in Frage gestellt. In diesem Beitrag werden wir unsere permanente Forschung an einer künftigen Version von *Textable Movie*, die eine direkte Kontrolle und Visualisierung aus mehreren Perspektiven möglich macht, präsentieren.

Schlüsselwörter: digitale Geschichten, Storytelling, Echtzeit, Datenbank, Video-Jockey.

Résumé (FR)

Cette communication présente *Textable Movie*, un outil ouvert qui permet à tout narrateur de devenir un "*video-jockey*" capable d'improviser en temps réel un récit audiovisuel à partir d'une collection d'images et de vidéos annotées. Dans l'exercice de la narration numérique, *Textable Movie* promeut l'idée du contrôle de la création du média, et se démarque des systèmes automatiques de présentation. Son interface graphique considère le texte comme une entrée et permet aux utilisateurs d'improviser en temps réel un film basé sur le contenu de ce qu'ils écrivent. Les séquences sont sélectionnées en fonction de comment les utilisateurs ont annoté leurs données audiovisuelles personnelles. Lorsque une histoire est rédigée, les séquences apparaissent à l'écran, remémorant à l'auteur ses expériences passées et l'invitant à poursuivre sa narration. Les *video-jockeys* utilisent leurs propres bases vidéo ou celles d'autrui afin de présenter différentes mises en perspective visuelle de la même histoire. En co-créant des narrations filmiques qui sont d'abord improvisées d'après une base de données vidéo personnelle, et sont ensuite projetées dans la base d'une autre personne, lorsque deux utilisateurs fusionnent leurs histoires, les jeunes adultes sont interpellés sur leurs croyances et leurs a priori concernant les autres communautés. Dans cet article nous présentons nos recherches en cours sur une future version de *Textable Movie* qui permettra un contrôle direct et la visualisation de plusieurs points de vue.

Mots-clés : Histoires numériques, Narration, Temps réel, Bases de données audiovisuelles, *Video-joke*y, Echanges inter-culturels.

I. Introduction

We imagine a world in which children play, create and exchange visual narratives with ease and transparency. Textable Movie explores a graphical language and interface that can invite storytellers of any age to compose and visualize movies, images and sound environments while writing a story; the system self-selects and edits images in real time based on textual input of the teller. While Textable Movie was first inspired by an application for annotating images (Lieberman, H. & Liu H., 2002), we have found, during pilot studies with teenagers, that its power is maximized when it becomes a presentation system. In this approach, makers first create their own media, (e.g. pictures, sounds, movie segments) based on stories they wish to recall; they then use the text driven presentation mode which allows the maker to fluidly shape a visual story by writing annotation words into the text-based story. By creating a movie-editing paradigm in which text leads and image follows, Textable Movie realizes a very natural, fun and immediate interface to story-making. As this approach creates a symbiotic relationship between the author's imagination and the stories that she wishes to tell, Texable Movie is able to support a diverse set of activities that foster narrative co-construction.

II. Scenario

Imagine a group of teenagers who have access to a video camera. As they fool around with the camera they decide to make a documentary and select as their setting Temple Bar, a well-known recently renovated cultural area of Dublin that acts as a magnet to visitors, locals, artists, filmmakers, and skateboarding youth alike. For each group of teenagers, one of them is voted to be the director; another takes up the role of cameraperson; the rest play the actors; they choose the journalistic approach with a minimum of preplanning. They shoot actions, established places, locations, environment, transportation, fashion, and people. They use different shot types, and camera motions. Back in the studio, the authors divide up tasks: one is the editor and decides how the footage should be annotated; one is the assistant editor and captures, cuts and exports the movie elements; the remaining authors act as narrators and projectionists. They watch their video, and decide how they will cut it for the interactive projection. Then, the editor says : *"Oh, this is you here! We could make it so that when "Tom Cruise" is mentioned during the projection, then you appear on the screen!"* They all laugh.

The clip is annotated as "Tom Cruise" so that the assistant editor, actor during the shooting, will become Tom Cruise during the projection.

III. Application Domain

In proposing a text-led interaction and presentation system, we redress a limitation of traditional editing and draw on observation and previous research that relates verbal and image exploration (Davis, M., 1995; Lieberman, H. & Liu H., 2002). Within the Textable Movie system, we would like users to build their own understanding of how the inference is being made, essentially a straight-forward mapping between the text and the video selected. We believe that this level of transparency allows the system to promote a sense of story magic while at the same time providing the storytellers with a maximum sense of agency.

This process of creating an image collection and projecting different stories using a text based approach to story telling can be readily contrasted to traditional "random access" video editing systems where the interaction paradigm is a timeline. In the latter case, the objective is to make an immutable "final" cut of a movie. The author can only see the whole once she renders the time line. The audience sees only the final version of the object. What happens when we turn this paradigm on its head, when we bring the imagination of the projectionist or the audience in closer synchronicity with the "movie" output? Ironically, the very nature of digital architecture begs us to make a shift from final object to process. Why should we treat a collection of digital clips stored in a database with the rules that shaped the continuous material substrate of celluloid? Why not use the text based nature of annotation, and call chunks from the database in real time?

Previous research and observation together lead to the idea that text could drive a "projection paradigm" where the projection is sequenced by the projectionist/viewer/maker in viewing time through verbal input. Early testing with this paradigm using footage of the application author uncovered a basic limitation: how would the projectionist/viewer know what words to use? This resulted in the idea that the players -- be they a small group focusing on an authored project or casual visitors to a waiting space -- submit and name their own images. The Textable Movie application thus came to serve as a platform for multiple video-oriented interactions along the lines of the play described in the scenario. By encouraging associative thinking in the verbal domain, the application leads participants into a new realm of story making.

Participant "projectionists" use Textable Movie to drive a movie experience made up of the story text and the association between particular words in the text and video clips stored in the database. Textable Movie (Vaucelle, C., et al., 2003) retrieves movie segments and sounds in a specified database, by analyzing textual input. The application loads and plays the clips in real time while the story is being typed; consequently, a novel movie gets created and generated in a very transparent, and easy manner everytime a new story is typed in. For example, a set of clips in the database might be variously labeled [forest, nature, tree, wood, leaves, Yosemite]; a user might then type in "when I walked through the forest at Yosemite kicking up the leaves that had fallen from the trees." This sentence would result in a short, personal 10-second video projection of my Yosemite experience which I might call "forest.mov" (see Figure 1.) By personally labeling their video clips, users are given maximum ability to control retrieval based on their own personal annotation scheme. The simplicity of use and immediate response of the system can help the user/projectionist focus on the story rather than on the technical concerns of editing.
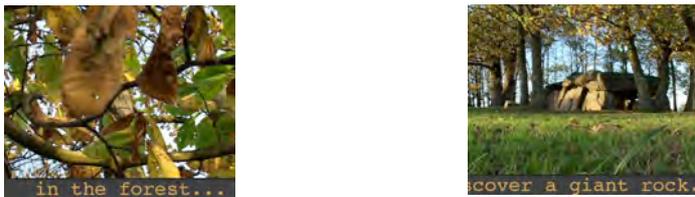

Fig. 1. Video segments play instantly following the story typed. For instance, the author types *forest*, and a video corresponding to the word forest starts playing instantly on the screen.

A series of simple commands add instant manipulations of the movie being played (see Figure 2.). These commands are typed directly in the text, and by deleting the text-command, the effect disappears, e.g. to zoom: [closeup], to change the speed rate [faster], to alter the overall coloration of the image [winter].

Fig. 2. A set of commands modify in real-time the video segments, e.g. [closeup], [winter]. The author types [closeup] and the rock instantly becomes bigger, then [winter] and the RGB values of the movie changed to a "winter" tone

Textable Movie version 1.0 was made with MAX/Msp/Jitter, a programming language that allows real time processing and analysis of the video and sound. Textable Movie, version 1.1, has been implemented in Java to allow easy public distribution. Pilot studies with teenagers at the Computer Clubhouse in Dublin informed us about areas of improvement. The system allows the user to play with a selected video database, to have access to a list of modifiable keywords available, and it offers the possibility to export a movie at the end made of the video segments in the order of the story typed.

IV. Contribution to the Digital Storytelling Community

Textable Movie offers a very natural method of combining "narrated" story and image collection. Because we can merge personal collections, Textable movie may offer a diverse community of users an activity that challenges their assumptions about their surrounding environment during storytelling. Textable Movie differs from previous systems have been built for dynamic video sequence generation by secondary image selection (Murtaugh, M., 1996) or textual specification (Lindley C. A., 2001) in that it provides sequencing by textual and spoken text input. By leading on natural language at the interface without becoming encumbered by speech understanding, we can provide a natural approach to story building and play for both the person who is trying to understand the world and for the person who engages in improvisational storytelling with someone else using the system. We hypothesize that collecting video contributions from many people world wide, Textable Movie will become more interesting to an international crowd of Video Jockeys. Through immersion in ones own personal life story, and by watching the stories of others, we hypothesize that the system will challenge young users to examine their life beliefs and expand their cultural awareness.

1. Communications Technology for Cultural Changes

We play, learn and exchange ideas about our identity using stories. For many Centuries, from family letters to Pen Pals, the letter writing has been a dominant form of information exchange and knowledge sharing. Today communications technology, with its spectrum of media potential, expands the resources we have at our disposal for exploration and sharing of our identity. Therefore it seems natural that our approach to real-time movie-projection can serve rich cross-cultural explorations, as it allows participants to investigate, express and exchange concepts of identity. We ask, can young adults, by sharing images and stories across

borders, gain a deeper sense of cultural identity? Can we transfer a sense of culture through a mix of text-based and visual play, collaboration, storytelling. Can we share cultural as well as personal realities, through the production of media? Will this sharing enable us to better understand relationships with others who live in alternative communities, with different rules, habits, and references to the world? With digital networked technologies, we are moving away from a dependence on letter writing as a principle mode of sharing our realities, and step toward a world in which we can share images as easily as we share local language, and in which in which we co-construct and share realities with others in near real time. Textable Movie provides a tool to engage people in multi-dimensional exploration, expression and exchange across diverse identities.

2. Media Creation

In a world in which media is everywhere, media creation can provide a means of exploring self-identity both through individual sharing and group construction of media spaces. Constructionists have suggested that users could benefit from systems that support self-expression rather than create content for the user. With Textable Movie, our focus is to empower people to express what they think is meaningful in their culture, and what is optimally challenging to them. We base this research on the body of work regarding digital stories for documenting culture. A network of media could allow users to master the art of digital media for their own expression, developing powerful ideas through the presentation of interactive videos representing their own lives for the purposes of cross-cultural exchange. This would allow remote peers to reflect on their assumptions about other cultures through experiencing these videos, and instructive installations.

V. User Study

While our interface paradigm is specific, it was inspired by the need to make visual storytelling more accessible, engaging, and powerful for young people (Resnick, M., 2002). With that interest, we have designed an international workshop that focuses on visual storytelling using Textable Movie. Teenagers participating in these workshops go through the process of putting together a piece by first storyboarding, then shooting, editing, and projecting a video-story. We have started to plan a global strategy, which will allow us to compare how teenagers use this system across cultures. We have conducted video pilot studies in Computer Clubhouses, we observed teenagers create their own video database about Dublin, and we are running tests for a new design in Sweden.

Textable Movie acts as a projection device for a storyteller. It is specifically a tool for improvisational storytelling, and not a regular editing tool. It is based on the theory that people learn by constructing their knowledge (Papert, S., 1991) and that people are engaged in the process because they have fun, because they are challenged, and because they expect an audience for their creation. The extension of the constructionist field we explore comes from the Community Art notion of « challenging » the audience by the object created. In this workshop, we observe how teenagers can be empowered in the sense of involving them in creating pieces to challenge their recipient about their assumptions. These assumptions can be about their environment, their identities, and also their culture. The core issue of the research is to focus on storytelling in order to bring to light contradictions and resolve them. We are particularly working with teenagers because they are in an intermediate stage where they are questioning the contradictions in their environment.

1. Methodology

A one-week user study has been held at the Ark, Dublin, in Ireland with 11 teenagers. We divided the week into different activities: each day, the teenagers experimented with our video tools for 3-4 hours. Around ten adult mentors interacted with the teens, helping them to use the video equipment and Textable Movie, and making sure each teenager explored the equipment, and understood its features. We have first familiarized the teenagers with the traditional methods of movie making and documentary making. The teenagers had a hands-on introduction to the video equipment. We have introduced the notion of new ways of movie making by assembling video clips in a specific order, and new forms of stories by remixing them together. We have quickly shown Textable Movie in use for such purposes. They first created a storyboard telling a story about the building in which they were in. They filmed their own movie based on it, and used the Apple software iMovie to digitalize and segment their movies into video clips (see Figure 3.).

Fig. 3. Footage for the keyword *mobile phone* done by the participants from the workshop at the Ark, Dublin, August 2003

2. Results

We have found that the participants in our workshops exclusively used Textable Movie as a tool to show others short stories in real-time and make surprise effects by creatively associating keywords to visuals. They consistently gravitated towards iMovie and its timeline to create an edited movie for export. They never used the functionality of Textable Movie to create a whole movie out of the story segments. We can deduce that the specific utility of Textable Movie is to provide a response to the user input in real-time and then act as a projection device. A more quantitative analysis could support these results; however, we found very informative that, in a one week workshop in which the teenagers felt challenged by the activities, all of them have looked at Textable Movie as a real time projection device, and each has pushed the limits of its functionality, e.g. by making surprise effects. Thus, we have found it important in imagining a set of video editing tools that this set should include a regular editing tool, e.g. iMovie, as well as an interactive projection tool, e.g. Textable Movie. We finally observed that the teenagers were constantly into creating specific content for their audience, and were excited about sharing their database with other countries. From the beginning, one of our goals for this tool was to encourage teenagers from around the world to share views of their world with each other.

Textable Movie as a Tool for Anyone to Become Video-Jockey (VJ)

The teenagers segmented their own movie, creating a palette of keywords to explore their movie. We have found that the storyboard segmentation of a movie into a sequence of shots was working very well with the segmentation of a movie with Textable Movie. The

immediate response from the system made it comparable to a video game, and not an editing tool. The teenagers were excited to "play" with Textable Movie, and wanted more videos about themselves. They were thinking about ways to connect their facial expressions to Textable Movie. We ended up with a palette of heads and the names of the teenagers as keywords. They wanted to redo more of their own footage for the system.

Textable Movie for Cross-cultural Exchange

We have explained the intent of the workshop as a way for teenagers from other countries to navigate through their life. To this end, we have asked the participants to shoot in the city as a reporter of their environment. They shoot elements presenting their city and what they liked or disliked in it, e.g. mobile phone conversations, couples kissing, fights. They watched their video, and decide how they will cut it for the interactive projection. One participant was designated to be the projectionist creating the city story using Textable Movie. However, all the others were 'shouting' the keywords they wanted him to type in order to have surprising footage.

Textable Movie as a Tangible Interface

From our observations of teenagers using video in workshops, we hypothesize that Textable Movie would be even more powerful if we could construct a physical a movie making device that would allow teenagers to easily understand and make videos using traditional cinematic language. For this purpose we are prototyping a mobile video table that contains cameras at the four corners, making it easy to shoot panoramas as well as action-reaction sequences and create set-ups that are optimized for video effects. It is our hope that this platform will help teenagers discover how to incorporate point of view into a scene. In parallel we are exploring how we might create a tangible puzzle like construction set that enables a more rapid and sociable approach to image selection and sequencing.

In collaboration with the UMEA Design School, we are currently developing a concept: *Moving Pictures* that extends Textable Movie to a tangible interface (see Figure 4.). Moving Pictures aim to be an accessible, robust multi-user unit and a set of physical tools that enables young users to explore, manipulate and share video content with others.
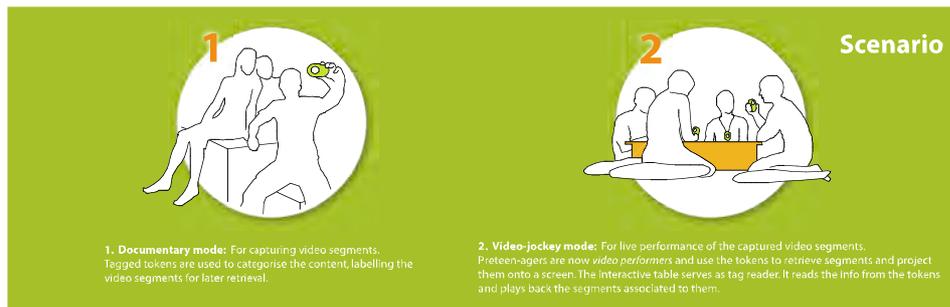
Fig. 4. Example of scenario with the tangible video platform

Moving Pictures involves a collaborative approach that supports social interaction in group work. It proposes exploration, methods and a tangible tool set for creation of co-present video play. Moving Pictures will be a tangible interface for pre-teenagers, which allows users to establish a cultural exchange through playful storytelling using motion picture capture. Within this future step on the Textable Movie interface, we will explore different types of input tools appropriate for young users. Based on our finding from our video workshops, this new design will allow a group to explore concepts such as simultaneous perspective taking, sequential perspective taking, as well as the relationship between space and time (see Figure 5.).
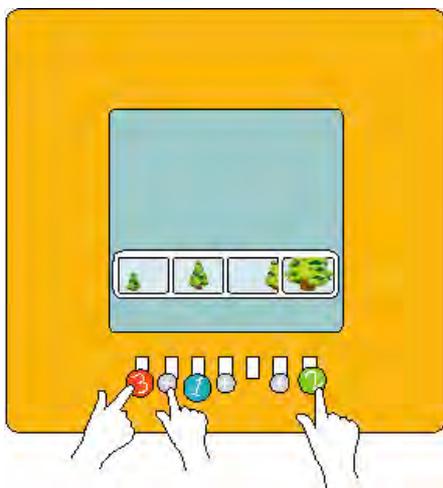

Fig. 5. Example of collaborative interaction exploring different perspectives on the same scene.

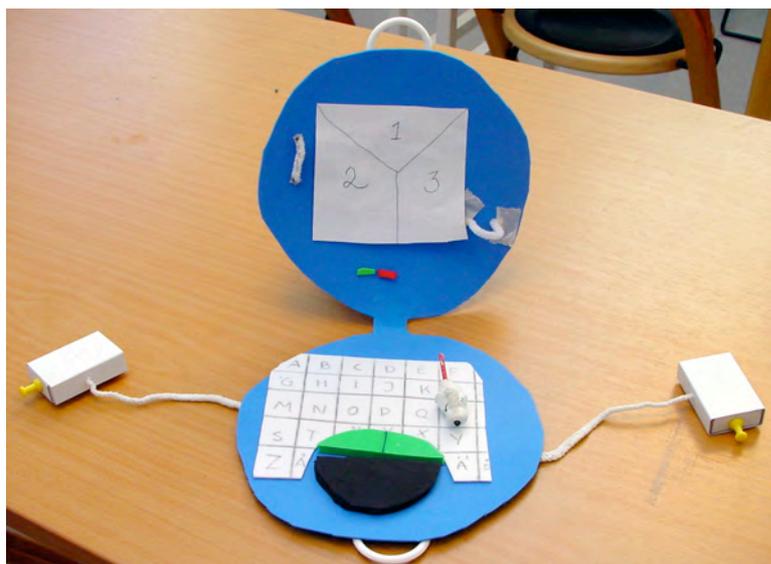We are also running participatory design on the tangible platform (see Figure 6.).

Fig. 6. Example interface created by the group of participants from School of Ostermalms, Sweden.

VI. Future Work

We found that Textable Movie is a natural tool for improvisation and projecting audio-visuals while telling a story, and we have also begun a new multi-cultural database of teenagers in their environment. We have prepared the setup for use in different countries where equivalent workshops will take place, and the cooperative results will be a permanent kiosk to navigate among other's culture. In its future versions, our system will be networked, and used as a multimedia tool to reflect into someone else's culture. Our plans include a further analysis of the new possibilities created for participants of different countries who share a database of clips from different places, installed in the form of a kiosk in children's museums internationally. Textable Movie will also be part of the worldwide activities of Pangaea (PANGAEA, 2003) that aim to let children around the world feel "bonds" personally.

VII. Conclusion

In this paper we present an alternative framework for video editing and storytelling motivated by the desire to reproduce the playful improvisational environment of child storytelling. We have discussed that Textable Movie can be a tool to engage young users in multi-dimensional exploration, expression and exchange across diverse identities. We have outlined the specifications of Textable Movie during workshops that engage teenagers to become video-jockey, projecting media during storytelling in front of an audience. The power and flexibility of Textable Movie can be evaluated in part based on applications that we have created which

we did not initially envision; for instance, the teenagers have used Textable Movie to build video games. We have finally introduced the idea of a tangible interface for teenagers to interact with Textable Movie.

Acknowledgments

References

Book:
Lindley C. A. (2001). *A Video Annotation Methodology for Interactive Video Sequence Generation*, Digital Content Creation, Earnshaw R. and Vince J. (eds).

Chapter in a Book:
Davis, M. (1995). *Media Streams: An Iconic Visual Language for Video Representation.* In: Readings in Human-Computer Interaction: Toward the Year 2000, Morgan Kaufmann Publishers.

Papert, S. (1991). *Situating constructionism.* In Papert & Harel, Eds., Constructionism. Cambridge, MA: MIT Press.

Resnick, M. (2002). *Rethinking Learning in the Digital Age.* In The Global Information Technology Report: Readiness for the Networked World, Oxford University Press.

Proceedings:
Lieberman, H. & Liu H. (2002). *Adaptive Linking between Text and Photos Using Common Sense Reasoning*, Conference on Adaptive Hypermedia and Adaptive Web Systems.

Vaucelle, C., Davenport, G., and Jehan, T. (2003). *Textable Movie: improvising with a personal movie database*, Siggraph, Conference Abstracts and Applications.

Theses:
Murtaugh, M. (1996). *The Automatist Storytelling System*, M.S. MIT Media Lab.

Web Site:
PANGAEA. (2003). (http://www.pangaean.org/)
The ARK. (2004). A Cultural Centre For Children (http://www.ark.ie)
UMEA Design School. (http://www.dh.umu.se/)