Third Annual
Online Public Report
SALERO Deliverable 10.5.6
The work presented in this document was partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th framework programme for R&D.

This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

This document contains material, which is the copyright of certain SALERO consortium parties, and may not be reproduced or copied without permission. All SALERO consortium parties have agreed to full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the SALERO consortium as a whole, nor a certain party of the SALERO consortium warrant that the information contained in this document is capable of use, nor that use of the information is free from risk, and does not accept any liability for loss or damage suffered by any person using this information.
# Table of Contents

1 Project Overview ........................................................................................................................................ 1

2 Summary of Activities .................................................................................................................................. 1

3 Second Phase Experimental Productions ................................................................................................... 2
   3.1 MyTinyPlanets.com ............................................................................................................................. 2
   3.2 Turing Opera & Bot ................................................................................................................................ 3
   3.3 Triage Trainer ....................................................................................................................................... 4
   3.4 Interactive VJ ........................................................................................................................................ 5

4 Procedural Face Generation ......................................................................................................................... 7

5 Speech Synthesis and Audio Processing ..................................................................................................... 7

6 Intelligent Media Research .......................................................................................................................... 8
   6.1 Media Semantics ..................................................................................................................................... 8
   6.2 Context-Based Search & Retrieval .......................................................................................................... 8

7 Integrated Applications .................................................................................................................................. 10
   7.1 Program Editor and Renderer .................................................................................................................. 10
   7.2 Bones Dailies ......................................................................................................................................... 12
   7.3 Scalable Codec ....................................................................................................................................... 13
   7.4 Game Development Tools .................................................................................................................... 13
   7.5 Audio Transformation Tools ................................................................................................................ 14
   7.6 Vowel Builder & Animation Builder .................................................................................................... 16

8 User Involvement, Promotion and Awareness ............................................................................................ 17
   8.1 User Group ............................................................................................................................................ 17
   8.2 Scientific Publications .......................................................................................................................... 17
   8.3 Other Events ........................................................................................................................................ 17

9 Conclusions and Future Work .................................................................................................................... 17

10 Further Information ................................................................................................................................... 18
   10.1 Useful Web Links ................................................................................................................................. 18
   10.2 Contact ................................................................................................................................................. 18
   10.3 Project Partners ................................................................................................................................... 18
1 Project Overview

SALERO

http://www.salero.eu

SALERO aims to make cross media-production for games, movies and broadcast faster, better and more cost efficient by combining computer graphics, language technology, semantic web technologies as well as content based search and retrieval.

SALERO will define and develop 'intelligent content' for media production, consisting of multimedia objects with context-aware behaviours for self-adaptive use and delivery across different platforms. 'Intelligent Content' should enable the creation and re-use of complex, compelling media by artists who need to know little about the technical aspects of the tools that they use.

Based on research into methodologies for describing, creating and finding intelligent content, SALERO will develop toolsets to create, manage, edit, retrieve and deliver content objects, addressing characters, objects, sounds, language sets, and behaviours. The toolsets developed and the concept of intelligent content will be verified by experimental productions.

2 Summary of Activities

In the third year of the project major progress has been made in all areas. Examples are the development of advanced algorithms for facial animation, expressive speech synthesis, improved tools for game development, establishment of an ontology for modelling virtual characters and the implementation of the second phase experimental productions.

Research in facial animation yielded techniques for emotional animation based on the maskle approach and using the circumflex model to specify emotions. This is combined with work on highly realistic modelling and animation of the appearance of persons depending on medical state.

In respect to expressive speech synthesis a limited domain engine has been implemented. The work in respect to audio transformation has yielded different algorithms to alter the quality of a voice (e.g. converting a female voice to a male one, changing the perceived age of the speaker).

An ontology to describe virtual characters for 3D animation based and a first version of a multimedia annotation tool have been developed.

With i-VJ, Triage Trainer, Tinyplanets.co.uk and the Turing Opera and Bot the second phase experimental productions were implemented and subsequently evaluated. Based on the experiences during the implementation of the experimental productions and the results from the evaluation, the planning for the third round of experimental productions has started.

The intermediate results of the research activities and the results from the experimental productions are a perfect starting point for the next experimental productions and the integration steps planned for the final project year.
3 Second Phase Experimental Productions

3.1 MyTinyPlanets.com

In the second experimental production phase of SALERO, Peppers Ghost Productions produced animation tests for the My Tiny Planets browser based virtual world using both methodologies devised by Dublin Institute of Technology and traditional toolsets, and evaluated the integration of these into a variety of on-line delivery mechanisms.

Having established an initial test group of several thousand users, PGP was better able to understand behaviours and requirements of players in virtual worlds, and revise development plans accordingly. It became apparent that animated sequences as a support for game-play, narrative and immersion represent a significant commercial advantage, and will be pursued aggressively during the next phase of development and experimental production.

DIT’s approaches for both lip-synch (driven primarily by audio amplitude) and activation of animation proved to fit well as a layer over the top of more usual approaches such as layering these over the top of ‘ambient’ motion capture files to prevent stilted movements.

![Figure 1: Screenshot from mytinyplanets.co.uk](image-url)
3.2 Turing Opera & Bot

Crucible Studio produced the Turing Machine cross-media theatre and online production in collaboration with Helsinki Skaala Opera. Its opening night was on 28th March 2008 at the cultural event venue Kulttuuritehdas Korjaamo in Helsinki, Finland. The opera premiere was followed by a further four performances between 30th March and 2nd April at the same venue.

Crucible Studio produced 3D-scenography for the opera performance and a related web-based module, which further explores the life of Alan Turing. The web site that features Alan Turing as an enigmatic artificial chatterbot character, Turing Enigma, works by itself or as part of the Turing Machine Opera. A DVD compilation, containing the recording of the Turing Machine Opera performance, trailer and a “Making of” was subsequently produced.

The production investigated repurposing media objects between the two genres, between physical theatre space and an online site. The same objects that appear in on-stage projections also pop out from the Turing Enigma machine.

The Crucible Studio experimental production was an audience and review success, though the linkage between the opera and online production could have been improved upon according to user studies. Six newspapers, four magazines and one television programme provided media coverage and critique for the production. According to Hufvudstadsbladet 30.3.2008 “in Turing Machine, the visuals fall perfectly into place and are a virtuous handwork on their own account.” The biggest Finnish newspaper Helsingin Sanomat, in turn, summarized 2.4.2008: “The end result is a fascinating one-hour experience, which should appeal to a wide audience and fit for international export.”

Tools research found the SALERO emotional framework useful for script-writing and reviewed and improved upon the use of facial animation tools (Maskle). Creating a linear performance and a non-linear artificial chatterbot character from the Alan Turing character was also Crucible Studio’s artistic and design research project on storytelling, in which the chatterbot as an intelligent media production application repurposed automated dialogue technologies to narrative media.
3.3 Triage Trainer

Triage Trainer supports the development of critical decision making in the life saving skill of Triage. This is the process of prioritising casualties for treatment after a major incident. The training game is set at the scene of an explosion in a high street. The multiple casualties look and behave realistically to make the experience as near real life as possible, which supports training. Trainees’ decisions are tracked through the game, when they have assessed all casualties they are given feedback on their
performance, compared to the protocol laid down by Advanced Life Support Group, which advised NATO on triage.

The SALERO tools and methodologies used in the production were:

- Improved facial animation processes, animation rigging, procedural generation of eye movement, respiration, blood flow, sweating, pallor and flushing
- First release Shader Editor
- First release Distributed Asset processing tool
- Re-usable asset processes

Evaluation was carried out with 3 audiences:

- Trainees using the game as part of a learning package
- Professional users of the tools within Blitz Games Studios
- Professional peers within Blitz Games Studios and the wider industry

Surveys among game industry professionals using a Likert Scale asking for the level of agreement to statements about the use and impact of the tools confirmed that the state of the art in animation and procedural generation had advanced and that production workflows had significantly improved.

Trainees using the game in controlled trials performed significantly better in a common assessment exercise than those who had used a tabletop exercise during training. The improved performance may reasonably be ascribed to the immersion achieved with realistic human figures.

![Figure 5: The Triage Trainer application](image)

### 3.4 Interactive VJ

Activa Multimedia has developed an experimental prototype called an interactive Video Jockey (i-VJ). The production is a digital television broadcast over IPTV. This is in line with two growing trends in digital content, making the TV experience more interactive and allowing for customization.

The main objectives of this experimental production are to explore:

- Customization: users can customize distribution lists and their virtual presenter.
• Interactivity: users can see what they want, when they want, and how they want.
• Automation: control systems that automatically generate content and minimize human intervention.
• Re-use: maximizing the re-use of assets, particularly 3D models and animations.
• Compatibility with different ways of generating audio: a low-cost, highly-automated audiovisual production must be compatible with various video sources, ranging from pre-recording to voice transformation and text-to-speech. The efficient use of audio sources crucial is for cost reduction, the re-use of assets and automation of production.
• Automatic and semi-random script modification: automatic productions that meet users’ tastes but are different every time.

This experimental production is designed as a 24-hour music IPTV channel, a virtual Video Jockey as a 24/7 automatic broadcast and VoD (Video-on-Demand) service. The content of some sections can be customized by users.

The following are the main services:
• My i-VJ: allows users to decide the virtual VJ’s appearance.
• My Playlists: users can search for music videos and create their own playlists.
• My News: users can activate music news headlines.
• My Weather: choose a city for a TV weather forecast.
• My Account: users can set up and manage their account.

Figure 6: Music video channel broadcast with the button to enter into the application
4 **Procedural Face Generation**

Blitz Games Studios has built upon previous work in improved animation techniques and procedural generation of head and eye movement to develop an infinitely flexible tool for the rapid creation of realistic people. Initial work has focused on the face and head and the construction of a deformable mesh to enable the automatic conformation of multiple ethnicities, ages and weights as well as the ability to display emotion.

The system enables additive morph blend offsets on top of existing meshes, which supports blending from one base mesh to up to three ethnic variations of this mesh while retaining features from all of the morphs. The animation bones in the rig are set to fixed points in the base mesh and then procedurally altered with numerical offsets generated by the blends on the base mesh. Animations are additively blended at the root of the animation tree so that any further animations can be blended afterwards while retaining the desired offsets. This way, animations created for the original head are just as valid for any other generated heads.

5 **Speech Synthesis and Audio Processing**

Speech synthesis research by Ramon Llull University was oriented to improve the expressiveness and the quality when the text input does not belong to a limited domain (e.g. weather forecast). In this sense, a new speech database in English was recorded (by Pepper’s Ghost Productions), segmented and tagged. The database was initially designed from texts related to the MyTinyPlanets experimental production, but some general texts were added in order to improve the phonetic coverage. The final goal is to use this voice with the voice conversion tools (see section 7.5) in different experimental productions with more open domains than MyTinyPlanets.

More precisely, the current version of the text-to-speech tool is corpus-based and it incorporates some improvements related to the unit selection function. Ramon Llull University also worked in other research topics:

- Analysis and synthesis of Voice Quality parameters
- Prosody estimation of expressive speech
- Hidden Markov Models (HMM) based speech synthesis

The expected advances in these research lines will be integrated in the text-to-speech tool (Figure 8).
6 Intelligent Media Research

6.1 Media Semantics

The SALERO project researched methods and tools to formally capture and exploit the semantics of media objects in media productions. This is enabled through a semantic workbench which provides lifecycle support for multimedia ontologies. This includes tools to lift existing multimedia metadata to ontologies, to engineer ontologies, tools to annotate media objects using these ontologies and facilities to exploit the annotations in semantic search.

Parts of the semantic workbench are provided as Web Services with the intention to include functionalities for annotation and semantic search into other applications. The workbench is accompanied with a set of ontologies built in the course of the last year, namely the SALERO Virtual Character Ontology and the SALERO Annotations Ontology. The first one supports formal description of virtual characters including geometry, animations, and behaviour, the latter is an abstraction of this ontology to be used for end user annotation.

6.2 Context-Based Search & Retrieval

Two themes were evident:

- to continue improving the underlying search and retrieval algorithms (e.g. feature selection)
- to develop novel interfaces aiming at aiding the user in expressing complex search needs.

In the first case, work has continued on underlying content-based image and video retrieval techniques, including participating in the TRECVID 2008 effort. To the second case, a “FacetBrowser” interface was created (Figure 9) which allows users to create and organise their search process by creating facets. Each facet represents a single narrow search need, but within the context of a larger information need as represented by a sequence of facets. This interface was designed for the image and video domains.
Figure 9: The Facet Browser interface, showing three search facets

A second interface called “AspectBrowser” was developed to search the web using Yahoo’s BOSS service (Figure 10). This is a development of the FacetBrowser interface, integrating the lessons learned from user evaluations, and providing a user with a greater range of visualisation options than that originally provided.

Investigations of definition and use of contextual factors in search and retrieval has concentrated on taking advantage of the possibilities of FacetBrowser style interfaces. One technique uses the global query information from all facets to focus the query executed by the system for an individual facet.
7 Integrated Applications

7.1 Program Editor and Renderer

The Program Editor is a tool for automatic generation of audiovisual pieces. Given a set of assets as 3D meshes, virtual characters, animations, virtual sets, etc. it is possible to generate an animated sequence.

The sequences can be edited by hand, through the editing timeline or by preparing templates for automatic generation. Given the production requirements, it is possible to create templates that will be used to automate the process of creating several videos with dynamic contents.
Latest enhancements to the Program Editor include improvements to the interface and interaction flow. Moreover the results of the research on emotional animation can be accessed thanks to an interface that gives direct access to emotional states.

Once the sequence has been edited or the template for automatic generation is ready, it is possible to render the final output video, thanks to an OpenGL based render engine developed using the GTI-Framework from FBM-UPF. This render engine uses state-of-the-art algorithms for dynamic lighting and shadowing, complex shaders and post-processing.
7.2 Bones Dailies

The Bones framework is a next-generation environment for digital intermediate and post production that lets users create workflows tailored to their needs. It serves as the heart of an end-to-end post-production workflow and manages the process from data ingest to finished digital master.

BONES Dailies is a fast and efficient workflow tool for producing sound synchronised and colour corrected dailies (rushes). The workflow provides a benefit for the entire production and post production chain.

BONES Dailies allocates advanced features for audio and image ingest, metadata capture, colour grading, Audio/Video synchronization and play-out of dailies deliverables. The framework interfaces directly with postproduction products such as film scanners and other high end capture devices. The open environment supports third-party applications as well as render-clusters to deliver optimum rendering times tailored to specific project requirements. Using Bones applications, users can edit and conform transitions, rescale images and convert clips to a variety of output formats such as 4k film data and HDTV and SDTV. Where third-party applications are desired, the environment manages the data transfer of images to and from workstations in a facility and collects them for conformance into a completed sequence.

![Figure 14: User interface for image manipulation and timeline processing](image)

The Bones framework runs on a general-purpose Linux computer that can be connected to a storage area network (SAN) and its applications operate under a common user interface. Several Bones workstations can operate in a cluster enabling collaborative work of several operators.

Bones Dailies will be used as a platform for the third series of the Experimental Productions.
7.3 Scalable Codec

Based on the validated approach and framework, the scalable audio codec was further developed to accommodate both multi-channel (up to five primary channels) and extended bit-rate range (24-96 kbps/channel). This was delivered as the second version deliverable to the project during the year. Current work continues to develop GUI-based tools and to integrate the compression technology to partners’ production tools and work flow. The work at this stage aims at the applications of developed scalable codec in the third stage experimental productions. Meanwhile, parallel work is carried out to develop and enhance a multi-channel to two-channel audio spatial virtualization to support the concept of media intelligence and reusability from the play-back device perspective. This module is to be delivered to project for integration and evaluations in the third experimental productions, either as stand-alone tool or in connection with the scalable codec.

7.4 Game Development Tools

In the last year work on the component based editing interface for shaders has continued and the interface has been leveraged to create new tools for visually editing both animation blending trees and finite state machines. The state machine system has been targeted at core game logic control, player cause-effect scripting, artificial intelligence decision making and cut-scene animation sequencing when used in conjunction with the animation blending tree editor.

The bi-directional real time network link between game engine and editing interface has been developed to provide rapid and accurate change feedback on the target platform for edits and diagnostic feedback in the editing environment for runtime data state changes.

Usability and performance tuning work has continued on the work distribution system employed by the distributed game data deployment pipeline. Management of remote agents has been significantly improved along with robustness, automatic failure recovery and 32bit/64bit platform independence.
7.5 Audio Transformation Tools

The Audio Tools are a set of cross-platform VST plug-ins and offline tools including Voice Transformation, Tempo Transformation and Advanced Audio Equalizer. VST (Virtual Studio Technology) is an audio plug-in standard created by Steinberg to allow any third party developers to create plug-ins for use within VST host applications.

Voice Transformation allows several singing and speech voice transformations using spectral techniques for modifying the character of the voice. Transformations include: transposition, quantization, vibrato, roughness, breathiness, whisper, timbre mapping and other spectral transformations. This tool is targeted to recording or post-production studios that require voice processing.

Figure 16: Screenshot of the Visual Editor for Post Production Shader Processing
Tempo Transformation combines automatic rhythm analysis and time-scaling in order to transform the rhythm of an audio mix. An input audio is first pre-analyzed in order to extract a musical rhythm description. After this process, the user can vary the tempo or add/subtract swing by using the GUI controls during the playback.

Advanced Audio Equalizer is aimed for the advanced equalization of commercial music productions. This is a powerful tool able to equalize the audio using energy histogram-based on different criteria. It is not limited to the typical energy distribution versus frequency, but allows working with other useful energy distribution criteria such as panning, inter-channel phase difference, or energy variance. Some of its applications include: Remixing a piece by changing the volume or panning of each instrument independently or removing instruments from the mix and then re-adding them processed by some effects or isolate one or several instruments to perform a musical analysis of a piece.
7.6 Vowel Builder & Animation Builder

The Vowel Builder and Vowel Player tools are an implementation of the vowel stress tagging speech analysis framework for semi-automated character animation. Dublin Institute of Technology has developed an online prototype animation authoring tool in Adobe Flex called vowel builder, and an output application for the authored assets in Adobe Flash. The authoring method is both character and language independent, allowing media productions to leverage an asset repeatedly in different situations.

The authoring tool (Figure 20) allows an animator to author an asset once, and then repurpose the output with any character loaded into the Vowel Player application. In this way, a built character can be re-used as often as required, in any production context desired.

Figure 19: GUI for Advanced Audio Equalizer VST plug-in

Figure 20: Example Vowel Player Output
8 User Involvement, Promotion and Awareness

8.1 User Group

The user group has been revised and an active participation schedule has been set up for the group members’ involvement with the project. Three user groups have been held in 2008: In Barcelona (during 3GSM) in Soho, London and in Amsterdam (during IBC).

A questionnaire including details about each of the SALERO Tools and Experimental Productions shown and demonstrated during the events were sent after each event to gather feedback from attendants. Detailed reports on the events have been created including an analysis of the questionnaire results, in order to help in the improvement of further development of the SALERO Tools and Experimental Productions.

8.2 Scientific Publications

SALERO’s research results were presented at the following conferences and in the following journals:

- 5th European Semantic Web Conference (June 2008, Spain)
- Joint Conference on Digital Libraries (June 2008, USA)
- 31st Annual International ACM SIGIR Conference (July 2008, Singapore)
- 5th Sound and Music Computing Conference (August 2008, Germany)
- 5th International Conference on Visual Information Engineering (August 2008, China)
- ACM International Conference on Multimedia (October 2008, Canada)
- 3rd International Conference on Semantic and Digital Media Technologies – SAMT 2008 (December 2008, Germany)
- TRECVID 2008 (November 2008, USA)

8.3 Other Events

Technologies developed within SALERO project have been demonstrated at the following events:

- NAB 2008, Las Vegas, United States (April 2008)
- IBC 2008, Amsterdam, The Netherlands (September 2008)
- Broadcast India 2008, Mumbai, India (October 2008)

9 Conclusions and Future Work

Based on the promising results to date the focus of work for the final project year will be to further integrate the technologies, e.g. a search engine combining textual, semantic and content based search technology. Sets of interoperable modules will build the thorough base for the implementation and evaluation of the 3rd phase experimental productions.
A major focus for the final year will be the presentation of project results: Beside presentations and
demonstrations at various events, the most prominent events will be IBC 2009 towards the industry and
in SAMT 2009 (hosted by JOANNEUM RESEARCH) for the scientific community.

10 Further Information

10.1 Useful Web Links

- [http://www.salero.eu](http://www.salero.eu): SALERO project website including a showcase of SALERO’s results.
- [http://www.youtube.com/intelligentcontent](http://www.youtube.com/intelligentcontent): SALERO’s YouTube channel “Intelligent Content” publishing results from the projects research and the experimental productions.
- [http://ninoscompe.wordpress.com](http://ninoscompe.wordpress.com): Introduction to the Program Editor.
- [http://195.12.19.232/mytinyplanets/index.aspx](http://195.12.19.232/mytinyplanets/index.aspx): SALERO users are invited to become alpha testers for future My Tiny Planets developments. Test accounts may be created at the given address. Please note that this is a development environment, and that accounts may be changed or deleted without warning, but it is the point at which Peppers Ghost Productions will be initially trailing SALERO technologies.
- [http://mir.dcs.gla.ac.uk:8080/AspectBrowser/AspectBrowser.html](http://mir.dcs.gla.ac.uk:8080/AspectBrowser/AspectBrowser.html): The “Aspect Browser” interface.

10.2 Contact

Georg Thallinger
JOANNEUM RESEARCH Forschungsgesellschaft mbH
Institute of Information Systems & Information Management
Steyrergasse 17
A-8010 Graz
AUSTRIA

Email: georg.thallinger@joanneum.at
phone: +43-316-876-1240
fax: +43-316-876-1191

10.3 Project Partners

- Blitz Games Studios Ltd, [http://www.blitzgamesstudios.com](http://www.blitzgamesstudios.com)
- Pepper’s Ghost Productions Ltd, [http://www.peppersghost.com](http://www.peppersghost.com)
- Fundació Barcelona Media Universitat Pompeu Fabra, [http://www.iua.upf.es](http://www.iua.upf.es)
- Universitat Ramon Llull, [http://www.salleurl.edu](http://www.salleurl.edu)
- Dublin Institute of Technology, [http://www.dit.ie](http://www.dit.ie)
- Taideteollinen Korkeakoulu, [http://www.uiah.fi](http://www.uiah.fi)
- The University of Glasgow, http://www.dcs.gla.ac.uk
- Leopold-Franzens Universität Innsbruck, http://www.sti-innsbruck.at
- Universitat Pompeu Fabra, http://mtg.upf.edu/pages/home/link