SIXTH FRAMEWORK PROGRAMME
PRIORITY 2
Information Society Technologies

Contract for:

SPECIFIC TARGETED RESEARCH PROJECT

Annex I - "Description of Work"

Project acronym:	**SEMANTIC-HIFI**
Project full title:	BROWSING, LISTENING, INTERACTING, PERFORMING, SHARING ON FUTURE HIFI SYSTEMS
Proposal/Contract no.:	507913
Related to other Contract no.:	*None*

Date of preparation of Annex I: ***28 November 2003***

Operative commencement date of contract: ***1 December 2003***

Table of Contents

Proposal summary page

Proposal full title :

BROWSING, LISTENING, INTERACTING, PERFORMING, SHARING ON FUTURE HIFI SYSTEMS

Proposal acronym : SEMANTIC HIFI

Primary objective : 2.3.1.8 Networked audio-visual systems and home platforms

Proposal abstract :

In the context of large scale digital music, the goal of the project is to develop a new generation of HIFI systems, offering new functionality for browsing, interacting, rendering, personalizing and editing musical material. This next generation of Hard-disk based HIFI systems will change drastically the relation of home users to music and multimedia content. Users will be able to interact with music up to the point of blurring the traditional limits between playing, performing and remixing. These HIFI systems will be listening stations as much as open instruments.

Technically, with IPv6, every HIFI system will have an IP address and systematic use of metadata extraction and exploitation techniques will allow semantic or thematic browsing in large content catalogues over Web and file sharing systems. Converging with the TV, DVD, Game Station and 5.1 set-ups, HIFI systems will also bring 3D audio real-time navigation in the sound scene.

The main innovations include a deeper access to the audio content and structure, through state-of-the-art semantic musical metadata extraction and exploitation (partially from MPEG-7): temporal segmentation, polyphonic, melodic, and high-level descriptions like genre, etc. Users will be offered innovative possibilities of manipulation, edition, re-composition, organization, sharing and interactivity with the audio material, and true 3D audio rendering and real time remixing of music.

Using music from CDs, consumers will be able to freely exchange metadata information enriching their music and allowing them to listen to remix sessions performed by other users. But with appropriate DRM systems sharing music files is also possible and may even support new profitable music business models for the music industry.

Bringing together leading European music research and industry partners, Semantic HIFI will deliver next generation tools for HIFI systems and also as modular PC applications, both validated by musicians and expert users.

B.1 Scientific and technological objectives of the project and state of the art

Within the last two years considerable technological mutations brought new social, cultural and economical schemes into the cultural domain. Music was the first content to benefit from, and stimulate, the development of network protocols (streaming) and compression schemes (mp3) and at the same time revealed new trends which the industry tried to exploit with anticipation with the failure we know. Although new music business models are still not in place, these trends are still valid and can be summarized as following :

- Content storage cost tends to decrease (down to zero in file sharing systems)
- value is brought by a tight and innovative combination of technology and content
- metadata management technologies are at the heart of future information and entertainment systems since they bring access to more (browsing over Internet and databases) and better (high level or semantic indexing) data
- DRM issues related to music content exchanged on the Web reinforce the interest of systems not using illegal music files but mostly based on metadata produced by the user himself, which can be freely exchanged and may even increase music sales
- easier access to content and to content description allow users not only to retrieve and listen but to process and perform
- as proved by the development of mp3 players on mobile, car and HIFI systems, valid business models come from a nice combination of freedom, instability, discovery, sharing (Internet) with comfort, quality and simplicity (dedicated equipments).

Much debate and competition address the issue of where this simplicity will end up, the issue of convergence, be it on a PC, on a TV set-top box, on a DVD player, on a gaming console or an audio HIFI system. More recently, leading PC consumer electronics and digital imaging companies announced their support for MPV(TM) (Music Photo Video), a new standard format to enhance the way consumers store and enjoy collections of personal music, photo and video content on storage media, such as data CDs and DVDs.

Some companies will release the first MPV creation applications and players in the second half of 2003. Companies announcing support for the MPV standard include Eastman Kodak Company, HP, LG Electronics Inc., Olympus Optical Co., Ltd., Royal Philips Electronics, Samsung Electronics Co. Ltd, and Sony Corporation. Also announcing implementation of the MPV format in future products are companies developing leading software and firmware applications for creating multimedia discs. These applications include ACD Systems ACD See, Ahead Software's Nero, ArcSoft PhotoBase, HP Memories Disc Creator, Planetweb Digital Photo and Audio Managers, Roxio Easy CD & DVD Creator, Sonic Solutions MyDVD and RecordNow Max.

Brought by leading European research organizations and industry in music, this project focuses on hard-disk based audio HIFI systems with visualization capabilities on TV. In addition, authoring and personal publication tools on a PC will be developed as compatible options for this system, and validated with the help of musicians and expert users. The goal of the project is to go beyond these standards and existing HIFI interfaces with limited control and access to audio material (track selection of a CD, Play, Stop, Record buttons, volume, treble/bass faders,...) and to design, develop and assess innovative and relevant new functionality for domestic HIFI systems, as a direct application of state-of-the art and specific research in audio and music technology.

The main principles on which these new features rely include :

- Deeper access to the audio content and structure, limited up to now to the digital audio signal, through adapted description structures (standardized in MPEG-7) in terms of temporal segmentation, polyphony, melodic structures, and high-level descriptions (music genre, etc.),
- As a consequence of the former point, increased possibilities of manipulation, edition, re-composition, organization and sharing, and interactivity, with the audio material,
- A new step forward into 3D audio rendering and mixing through high-level control.

The experience developed by partners in FP5 projects, including CUIDADO for content-based management of audio and music information, and CARROUSO and LISTEN for coding and rendering of 3D audio scenes, provides a solid basis in terms of content description models and client-server architectures. The goal of Semantic HIFI includes the capitalization of the results of these projects, by now focusing on the client side and consumer market.

Scientific and technological objectives can be mapped as follows :

System architecture and database management

The overall Semantic Hifi system will include several actors distributed over the Internet: a home HIFI system, a metadata authoring application running on a PC, a centralized online server for musical metadata. A peer-to-peer mechanism will also allow data exchanges between some of these elements. The system architecture will aim at providing Internet-based interoperability between these components, wherever users feedbacks will prove that this makes sense functionally, and within the limits of technical feasibility of such communication. The existence of multiple databases also raises serious technical challenges (data consistency over several databases distributed on the Internet, compatibility between these data...), for which solutions adapted to the means and objectives of the project will be studied in the very first conception phases. Specific data management and conversion must be provided for several formats (mp3, AAC, MPEG4, MPV, etc). Use of content description encoded in MPEG7 and monitoring of Digital Item Adaptations (Undo functions) are also targeted.

Indexing

The project will go beyond traditional classification, indexing and management systems of audio files in local hard disks, external storages (Audio CDs indexed or converted in files, CDROMs, DVDs) and Internet. The Indexing modules will help excerpts management :

- Automatic and manual segmentation, textual/iconic labeling for each segment,
- Navigation into a music piece (graphical representation, audio summary).

Browsing

Innovative browsing features will be the visible consequence of the above-mentioned indexing technologies with :

- Query by example (search by similarity with input excerpts)

- Query by humming
- Artists description, keywords, cultural attributes and play lists generation (using metadata management and constraints)
- Composition with segments using musical rules and constraints.

Listening and visualization

The impressive rendering capabilities of Personal Home cinema are useless without user control capabilities. This project will bring :

- Multi-format compatibility (stereo, DVD 5.1, MP4, etc.)
- Compatibility with multiple rendering set-up : multi-loudspeakers, stereo (with transaural option), binaural (with an online database of equalization data for a set of headphones).
- Automatic compensation of the acoustics of the listening room for several loudspeakers set-ups (context adaptation)
- Application of a given acoustic quality by modification of the recording acoustic quality of played files
- Use of perceptual parameters for acoustic quality modification
- Synchronization and manipulation of sound associated with images, videos, MPEG4 files, music scores, …

Beyond 3D audio rendering, using metadata associated with the signal allows interesting polyphony management capabilities such as :

- Use of standard DVD tracks (up to 5) for polyphonic decomposition, listening track by track and remixing
- Automatic source separation of certain polyphonic music
- Navigation into polyphony using constraints
- Interfaces for navigation into the music piece and synchronization with images and scores.

Interacting and performing

Even with limited control devices (TV command, microphone, Game joystick, DV camera,…) performing with the system is an important goal of the project bringing :

- Conductor instruments : tempo variations by time stretching in real time using home controllers, variable time-stretching producing segments to be used in the authoring process
- Percussion instruments : adding simple percussive tracks to the music
- Instrument Score follower (accompaniment system)
- Vocal instruments (karaoke like system).

Authoring

In addition to the HIFI system, an authoring environment is designed for use on a PC but in some cases can lead to simple commands integrated in the HIFI system for :

- Transposition of sources
- Replacing one instrument by another
- Filtering using FFT or higher level descriptors (moods, styles)
- Composition of segments : real-time mixing

- Recording of all filtering and processing parameters for non-real time authoring.
- Composition : sequencing, mixing, morphing using content information
- Publication of play lists
- Music analysis and editorial tools (text composition, synchronization of media)
- Education/ training with style rules using a simple music keyboard.

Sharing and DRM issues

Today content providers try to restrict the usage of their products by their customers in order to prevent unlicensed distribution. On the other hand, consumers ignore these restrictions and share multimedia files for free. Content providers and content users treat one another as enemies with conflicting interests. This project aims to bring customers and content providers back together to a common economic interest. In order to meet the demands of both at the same time a system is needed that offers a certain amount of security as well as easy handling of content for the user. Therefore two different use cases have to be handled:

1) A DRM system for popular multimedia content The basic idea is to allow copying of content if the user is willing to mark such content with his personal digital signature. This leads to a system that, depending on the country´s legislation, allows the user to transfer content to his favorite portable devices, his car stereo, or even his or her family and friends. Everything in the scope of fair use is acceptable just as it is today. However, if content leaks out to the general public, if it is found for example in open file sharing systems, it can easily be traced back to its origin: By analyzing the digital signature that was embedded during the act of signing the user who released this particular content can be tracked. In this way large-scale infringement can be prosecuted while the user experience remains as it is today. It is well known that many users are reluctant to register their identity with an unknown system. For this reason the system offers a step-by-step approach which allows the user to draw benefit from the system even when they are not yet registered.

2) Alternative approach for new or unknown multimedia content Providers of new or unknown content support their customers to re-distribute this content as they pay for any re-distributed file a percentage on commission. Customers have a choice to pay for a received copy or to use it freely. If they use it for free, they will get no commission on re-distribution. If they pay, they will become licensed re-distributors automatically. The goal is an alternative peer-to-peer file-sharing model of customers, which brings profit to both, content providers and their customers.

The central part of both IPR concepts is an accounting web-service. It uses the providers' accounts to register music and the customers' accounts to store the transactions. The server generates receipts for both, registration and payment. These complementary concepts will guaranteed IPR without restricting the consumer.

Build on top of IPv6 protocol, the HIFI system benefit from seamless connection to the Web for accessing other music files. The innovative point of this project is the ability of the user to share the way his collection is organized and all the metadata he has produced thanks to :

- Peer-to-peer file sharing
- Tight integration with Web based file sharing systems (networking protocol, metadata publication) and interoperability with mobile devices
- Share of metadata (play lists according to semantic paths, associated data, annotations, editorial, music analysis) to be applied on data locally
- Share of data (provided copyrights and DRM are managed) like files, segments, excerpts, sequences of titles, remix or processed sessions, 3D audio remix sessions, MPEG4 composed scenes, etc

B.2 Relevance to the objectives of the IST Priority

B2.1 Networked audio-visual systems and home platforms (2.3.1.8)

This project will take great advantage of previous research initiated in the CUIDADO IST project which was targeted towards content –based retrieval on music server applications. The goal of this project is to focus developments on browsing over personal music collections on a home platform in combination with Internet file sharing applications. This new context addresses the issue of personal indexing and classification tools and publication of metadata over file sharing systems. This system should bring to individual users, technologies which were previously accessible only to content providers. It relies on legal sharing of personal metadata applicable on musical content that can be purchased or accessed legally on appropriate sites. Content protection and tracing systems against illegal access and sharing can be added on top of the system, but will remain a separate issue not addressed in the project.

B2.2 Semantic Based knowledge systems (2.3.1.7)

Semantic HIFI is taking advantage of CUIDADO IST project ending in December 2003 directly related to Semantic Content management itself based on CUIDAD, a European initiative which contributed to audio standardization in MPEG 7. MPEG 7 is now the best candidate for a uniform standardized description of multimedia contents. CUIDAD has managed to input several ideas in MPEG 7 both in the Multimedia Description Scheme and in the Descriptors retained for Instrument Timbre Description.

B2.3 Multimodal interfaces (2.3.1.6)

Building upon experience from other projects related to new interfaces for music browsing and interaction (RIAM French project PHASE), past research projects in this domain, manufacturing of new controllers (AtoMIC Pro) and participation in international conferences on this topic (MIME, Dublin 2002, Montreal 2003) this project is the first attempt to include the latest kinetic and mapping strategies for existing low cost control devices. Beyond retrieval application using content extracted features. It has the ambition to use these content features for sound editing and processing thanks to the direct integration into a professional authoring tool brought by one of the few European professional audio company which is part of the project.

B2.4 Ipv6 EU priority

Among IPv6 related EU projects are :

- LONG

Laboratories over Next Generation Networks

- 6WINIT

IPv6 Wireless Internet Initiative

- 6NET

Large-scale International IPv6 Pilot Network

- Euro6IX

European IPv6 Internet Exchanges Backbone

- NGN-LAB

Next Generation Networks Laboratories
- 6LINK

IPv6 Projects Linkage Cluster - accompanying measure - cluster
- 6QM

IPv6 QoS Measurement
- IPv6 TF-SC

IPv6 Task Force Steering Committee
- Eurov6

The European IPv6 Showcase

And also : ANDROID, AQUILA, BRAIN, DriVE, GCAP, MIND, MobyDick, NETGATE, NGN-Inititiative, SUITED, WINE GLASS, WINE

But none of them are directly related to music and media access.

The rapid and continued growth of the Internet requires new measures to ensure that it can continue to meet emerging requirements. Europe's ambition to be the most competitive and dynamic knowledge-based economy by 2010 can only be realized if the EU is at the forefront of this expansion of the Internet's capabilities. Europe will only be able to maintain and build on its leadership in wireless and mobile communications if a rapid transition is made to the next generation Internet based on Internet Protocol version 6 (IPv6).

IPv6 is a new version of the Internet Protocol, designed as a successor to IP version 4 (IPv4), the predominant communication protocol in use today. The changes from IPv4 to IPv6 are primarily in the following areas:

- Expanded addressing capabilities
- Header format simplification
- Improved support for extensions and options
- Flow labeling capability
- Consolidated authentication/privacy capabilities.

As can be seen on the Cordis Web site (http://www.cordis.lu/ist/rn/ipv6.htm), within the IST Programme several IPv6-related projects were launched in 2001. Directly related are the projects 6NET, 6WINIT, Euro6IX and LONG, all of which are providing IPv6 platforms for experimentation. The LONG project addresses the design and deployment of IPv4/IPv6 transition scenarios.

6WINIT is setting up and validating an operational IPv6-3G Mobile Internet, with customers having native IPv6 access points and services in a 3G environment, using applications mainly in the clinical and m-commerce domains. Euro6IX and 6NET are both providing major real-scale IPv6 pan-European test-beds, using mostly native IPv6, with the objective of gaining a better understanding of deployment issues. Euro6IX involves a large number of European Telecom Operators. It focuses on the Internet Exchange aspects of the network implementation using 34 Mbps in the core. It also aims to validate the new technology in a realistic setting where the different factors which exist in the present Internet are extrapolated to the IPv6-based next generation Internet.

It provides a unique space where leading European ISPs can validate the technical and business cases for IPv6. 6NET is providing a high capacity core network, up to 2.5 Gbps, that will be used for experimental applications such as IBM France (e-business and GRID software) or Sony Europe (new IPv6-oriented applications). It involves also a large number

of European NRENs and will use resources from GEANT and NORDUnet networks. In early 2001, the IPv6 related projects expressed the need to have an IPv6 cluster which could provide a platform to the IST projects for the development of liaison mechanisms, ranging from the exchange of technical information to the organization of common trials. The IPv6 cluster (www.ist-ipv6.org) started in June 2001.

Its main objective is to set up a platform for exchange of information and experience on agreed methodology for the benefit of the IST projects. This is achieved through discussion groups, common reports suitable for publication, periodic common workshops, common trials and demonstrations. Standardization activities are addressed through concerted contribution and the distribution of reports.

Among the EU and international initiatives related to IPv6 are :

IPv6 FORUM: **www.ipv6forum.com**
A worldwide consortium of leading Internet vendors and NRENs are shaping the IPv6 FORUM. Its mission is to promote IPv6 by dramatically improving the market and user awareness, creating a quality and secure Next Generation Internet and allowing worldwide equitable access to knowledge and technology.

IPv6 Task Force: **www.ipv6-taskforce.org**
This is an initiative of the European Commission in April 2001. It includes representatives of European ISPs, telecom operators, mobile operators, equipment supply industries, Research Networking, and key “application” sectors. By the end of 2001 it had developed a comprehensive action plan, aimed at ensuring the timely availability of IPv6. The Task Force’s recommendations can be found on its web site.

It must be highlighted that 2 partners of the project are already participating in the IPv6 Task force : Sony Europe and Ircam.

B2.5 Creativity in RDT project

Semantic HIFI aims at including artists in many phases of the project and not only as beta-testers, thereby going along with one of the recommendations from the EU. It is also a pilot initiative in the sense that artistic content (e.g., music and sound effects) will be the core element of the project driving the technology development and not the contrary as in many past RDT projects.

B2.6 European Audio Standardization

The widespread acceptance and penetration throughout the industrial partners of the consortium and the already existing user base within various fields and markets (from home-user segment, pro-audio users like music studio professionals and radio stations either federal or private) are expected to cause a standardization effect supporting existing standardization efforts of such organizations as the EBU (European Broadcast Union) and the AES (Audio Engineering Society).

B.3 Potential impact

B.3.1 Contributions to standards

Using and contributing to the MPEG 7 Media standard

It is important to remember that MPEG 7 differs from precedent MPEG initiatives not only because it does not deal with the content itself (but rather with its description) but also because one key condition for proposing new contributions in this standard is to bring them into real applications. Version 3 of MPEG7 will start in 2003 and is likely to be followed by future versions during the course of the project. These are two reasons why the project will first select those schemes issued from MPEG 7 which where mainly proposed by Fraunhofer IIS and the CUIDADO project like :

- Low level audio descriptors for identification and retrieval
- Timbre description
- Multimedia DS
- Adapted profiles.

Other descriptors related to sound source localization and room acoustics, singing voice and other spectral and temporal features will be extracted, validated in perceptual experiments and encoded either within the MPEG 7 or with the SDIF format (the Sound Description Interchange Format developed by two of the CUIDADO partners and the University of Berkeley, see references). These new contributions developed in the project will in turn be submitted into MPEG 7 for instance in the form of an MPEG7 Music Profile (to be discussed in July 2003 MPEG Meeting in Trondheim).

Revealing the interactive potential of MPEG 4

MPEG4 has proved to be the most promising standard for optimized rendering (see CARROUSO project) because it brings mobility and flexibility with optimized and adaptable compression rates (AAC, …) but also because it brings interactivity inside data. This will be of particular importance in the rendering part of the project using 3D audio BIFS and other audio synthesis formats such as SAOL.

Exploring the MPEG 21 exchange platform

MPEG21 will be fully investigated in different aspects or the project although it is too early to say if DRM and data exchange systems will be managed by end users in the short term. However MPEG21 clearly brings the exchange format environment this project needs at the level of the Digital Item Adaptation (DIA) and Digital Item Identification (DII) if not for purely eCommerce purposes.

Connecting personal publication with the World Wide Web

Related to the need for synchronizing media on the TV set or for navigation purposed, W3C SMIL and SVG technologies, already supported in partner activities, will be of particular importance for the project.
Following this bi-directional scheme (using and contributing to standards) is an important measurable yardstick of this project.

B.3.2 Strategic Impact

Potential Impact on Industrial/research Sector

There are still few European projects on audio and music by comparison to the video industry. This may be explained by a traditional weakness of the music European industry mainly dominated by US and Japanese companies. But while the European music industry is weak, both DVD and online music services are increasing rapidly in Europe. Nevertheless it is important to mention that although our society provides more access to music material, there are still many important challenges, as addressed by a project like Semantic HIFI :

Information reserved to specialists and lack of interoperability

Most of the information and metadata is available within Content provider's back offices and is not shareable with other providers, very specialized or reserved to a music genre or not easy to find. In peer to peer and file sharing systems, large interoperability at the data level but lack of interoperability at the metadata level is obvious. The user has to know which specific music he is looking for in order to access to sound/music files through distinct not interoperable online catalogues.

Information overload

Information overload is currently prohibiting direct access to recordings by using Internet search engines. Even in his personal collection of CDs, an average user is listening to the same titles and has trouble browsing over all his recordings. There is also a need for interoperability between personal collection and large databases to benefit from metadata share and interoperability with other users.

Inadapted or unpersonalized interfaces

It is easy to figure out that, once an online database is identified, keyword interfaces are not always adapted for sound retrieval. If the user is looking for music with dijeridoo the best interface might be for him :

- A quick way to listen to a lot of music (audio summarization)
- A way to specify content keywords in his query such as timbre information
- A way to provide a given dijeridoo excerpts and listen to retrieved similar recordings.

Poor value-added to information

Most of the music resources available on line are from CDs, CD-ROM or audio tapes which gather several songs related to a given theme or editorial idea. This value is lost if no personal publishing tools allow for the re-introduction of this richness in the context of digital databases.

Proprietary access systems

Even if the Web protocol has brought a tremendous cross-platform and hypertext compatibility at the keyword level, there are very few distributed databases using a common front-end, and consequently there is a lack of interoperability at the content level. In the audio field, each database has its own interface, indexes and administration system. This

current drawback is probably the greatest challenge for distributed services and for our project too.

Facing this industrial and research context, the project has the following impact on social, trade and scientific issues :

Exchange free metadata instead of sharing copyrighted music

As an alternative to the current threat related to music exchange and DRM, the project proposes to build a system dedicated to creation, processing and exchange of metadata instead of infringing on copyrights through the use of music file sharing systems. On the contrary new business models for music will come up from this project since metadata should be used with purchased music content.

Enhance transparency and mobility in the HIFI domain thanks to IPv6

Mobility has been a tremendous gain in telecommunication. With the advent of IPv6 all listening devices can have their own IP address and connect anytime anywhere to Internet. With Bluetooth and similar wireless systems this mobility is even more effective for music listening anywhere at home. Transparency and mobility are key features for the applications targeted in the project.

Complement traditional keyword navigation content-based approach throughout the information process

The project aims to develop technologies for content-based products and related software using the MPEG 7 emerging standard with an innovative and comprehensive content-based approach. The project will cover content information extraction as well as exploitation : from the *analysis process* (extraction of descriptors), through the *navigation process* (retrieval methods and interfaces), the *creative process* (authoring tools) up to the *publishing process* (editorial and sharing). The scientific goal is to show how content-based technologies are not only for content providers back office but also for personal and home systems.

Design and validate new technologies for segmentation, source separation and connection between low level audio features and high level music features

Another key issue of the project is to design and improve segmentation and source separation in order to provide to personal users extraction tools which can be directly applied on their personal CD collections. In MPEG 7 and in content processing in general, several attempts to find the right relationships (or scheme in MPEG7 jargon) between low level spectral/temporal audio features and higher level musical representations (notes, chords, instruments,…) are promising. Combining these two levels ensure the reliability and grounding of the description necessary for the targeted application.

European dimension and Added value

Even if Europe has suffered at the first place from the Internet crash of the last two years (Vivendi/Universal pending case, BMG/Napster, etc), it must be recognized as the best place for the convergence of contents and technologies. With respect to this important macro factor, the Semantic HIFI partnership benefit from a good balance between advanced research and innovative product development usually brought in Europe by small companies. But European audio and music research has always suffered from not working closely with the large consumer electronics companies because their research centers are often located in

Japan and in USA. This project has managed to build a strong consortium of recognized partners including a world leader in consumer electronics.

Furthermore, this project is related to the Sony European Culture Policy in which the group has decided to commit to technological and content offers targeted to Europe cultural specificities.

B3.3 Innovation-related activities and competition

It can be assumed that challenging technologies for the Semantic HIFI project will be coming from all the converging devices identified in the introduction of this project :

- TV and set-top boxes (from companies such as Samsung, Thomson, Sony, NEC)
- Game Console (Microsoft, Nintendo, Sony)
- HIFI (Philips, Sony, Thomson)
- Creative PC and mobile music devices (Sony, Apple, Thomson, HP, Nokia).

At the scientific level, the recent ISMIR conference organized at IRCAM in October 2002 has shown most of the innovation-related activities in research centers, universities and industry groups such as :

Preprocessing, indexing, classification and analysis techniques

- Automatic melody Segmentation (U. of Padova)
- HMM for Music Retrieval (U. of Michigan)
- Indexing Music databases (U. of Helsinki)
- Voice Separation (Darmstadt U.)
- Combining Musical and Cultural Features (MIT Media Lab)
- Pattern Discovery Techniques (Carnegie Mellon U.)
- Music Summarization (3 systems presented : FX labs, NTT, IRCAM)
- Fingerprinting System (3 systems presented : Philips, Fraunhofer, IRCAM)
- Fast map for audio retrieval (UPF)

Similarity and recognition

- Similarity of Rhythmic Patterns (Tampere U.)
- Music Similarity (Sony CSL)
- Singer identification (MIT Media Lab)
- Music Artist Similarity (Columbia U./MIT/NEC)
- Rhythmic similarity (FXLabs/CCRMA)
- Music Pattern matching (IRCAM)
- Drums segments identification (Tampere U.)

Query systems

- Retrieval in Peer to Peer environments (Harbin Institute)
- Cuby Hum (Philips)
- Query by Humming (Fraunhofer)
- Mobile Melody Recognition system (Nokia)

Browsing Systems

- Playlist Generator (Philips)
- CUIDADO (Ircam/UPF/Sony CSL)
- Radio and TV Monitoring and mobile identification system(Yacast)
- Representation for playlists (Creative Labs)
- Content-based playlist generation (HP Labs).

B.3.4 Dissemination

In order to disseminate their results partners will develop a strategy in four points

Web dissemination

The project relies on the Web potential for sharing music and music metadata, this is why the Semantic HIFI Web site will be designed using the same interfaces, browsers and databases to be developed for the final application. This makes the Web site more attractive and extends its information role to a dynamic test-bed.

Scientific expertise in publications and conferences

The partners have a strong impact on the scientific scene through their publications in journals such as IEEE Signal Processing, JAES, Communications of the ACM, Journal of New Music Research but also by their regular participation in International conferences such as AES, ICMC, MIDEM, ISMIR (International Conference on Music Information Retrieval..

Some recent publications from the partners have been listed in chapter B.3.3 on innovation-related activities.

Standardization participation

Standardization is also a way to disseminate results of the project since proposition for standards are often based on real applications like issued from Semantic HIFI. This is of particular importance in the context of MPEG 7 where Ircam, Fraunhofer Gesellschaft and UPF are committed.

User’s validation in workshops and fairs

Along conferences and fairs, user’s tests and on site validation will be performed. Among the places targeted for these sessions are MIDEM, Siggraph, AES, NAMM, NAB, IBC.

B.3.5 Exploitation

This project is gathering different partners already involved in new projects for multimedia and network use on a national or an international basis. Their experience in cooperative work and shared resources has led us to concentrate on a Common Database and Plug-in architecture to be applied in the different partner contexts. The use of the scientific modules or plug-ins must be kept independent of, but compatible with, several applications. The Application Exploitation plan will mainly involve Sony and Native Instruments as the two main application developers. The Exploitation plan is summarized by the following table :

Partners	**Markets**	**Exploited results**	**Products/Services**
IRCAM	Musicians	Processing modules Performing modules Rendering modules	Diphone AudioSculpt Max/MSP Spat
	Archives Music labels	Indexing modules Database systems Browsing modules	Music Portals Studio Online Music Archives
	Music Software companies	Indexing modules Performing modules Authoring modules	Sound processing plug-ins NI's Traktor
	Listening Device manufacturers	Indexing modules Rendering Packages	Sony HIFI
FhG	Providers	Sharing system	Web services
	Listening Device Manufacturers	Encoding technology	Sony HIFI
Sony CSL	Listening Device Manufacturers	Performing & Browsing mod.	Sony HIFI
UPF	Music authoring	Performing modules	Sony HIFI NI products
BGU	Music authoring	Indexing modules	NI products
NI	Music authoring	Authoring soft	NI products
Sony NSCE	Consumer Electronics	HIFI system	Net-DAV system

This plan shall be specified and renewed according to the user feedback and market studies performed as part of the project, and to the evolution of the targeted markets all along its duration.

Sony market exploitation and strategy

Hifi Systems in Sony Europe, today

First, the domain of consumer and professional music listening has long been a core business of Sony Corporation. This is illustrated for instance by the history of Sony's top products in this domain (from the first consumer tape recorder to the first WalkMan and first CD player, to, more recently, the SACD format and the Atrack format for compression and protection of digital music). In Europe in particular, this activity has been of utmost importance for Sony, who is currently the leader of consumer Hifi products with about 20% market share in Europe (main competitors being JVC, Panasonic and Philips).

However, current audio products are mostly targeted at so-called "passive" listening. Much of the technology has concentrated so far on audio quality (from 44.1Khz to the one-bit stream format), on portability and miniaturization (the portable mp3 players), with a clear separation storing (on dedicated supports such as MemoryStick) and playing (on dedicated hardware). PCs are now changing the relationship between storing and playing as they can hold large amounts of tracks and play them directly from the storage location (its file system, local or remote). PCs also allow users to alter, edit, mix, visualize music, although this usage is not yet a mass market habit. It is however likely that users will progressively modify their relationship to music and ask for a more interactive relationship than in the past, as illustrated with file sharing. The theme of *active listening*, was precisely introduced at CSL laboratory in 1998, with the aim of designing such new ways of accessing music (playing) and considering music collections (storing). The theme led to the development of prototypes such as MusicSpace (control of spatialization), Continuator (accessing content from performance characteristics) or PersonalRadio (accessing music collections through metadata) and many others. In some sense, the SemanticHifi project is a direct continuation of these precursor research activities.

Second, it is important to note that Sony's product have always been ahead of the technology, for various reasons, including the fact that Sony's strategy is based on the development of cutting edge products, rather than in mass marketing of common technologies. Also, Sony has constantly been investing in research and development to release products with advanced audio/music features.

On the industrial front, we can quote the *DJ Mix* Hi-Fi system (trademark of Sony) released by Sony in Europe in 1996. This product allowed users to add various effects ("Loop", "Groove", "Flash", "Non Stop Play) and drum machine loops, to the tracks listened to. This product had a good impact on the market, and showed that such active listening features are enjoyed by customers on a relatively large scale. These features can be considered as basic seeds on which SemanticHifi will elaborate. So end products built on the results of SemanticHifi are considered a priori as having great potential for Sony European products.

Prospects for Semantic Hifi

More Precisely the results of the Semantic HiFi project are expected to create new ways to store, index, browse, listen to, manipulate and share audio content. These new functionalities are potentially applicable to future audio products and services that Sony is currently envisioning.

The functionalities and architecture explored by the Semantic HiFi project are of very high interest to Sony to try to meet the new market demands for new types of interactions with

music. The most obvious ways to introduce these new functionalities to the public are the following:

Extending the functionality of existing products:

This can be done by adding new features into current portable mp3 or audio CD players, or into home HiFi systems. Networking features is an obvious choice in this category, whether mobile or fixed. Networking will enable sharing from fixed to mobile devices belonging to the same owner or sharing between different people. Sharing can also be thought of as sharing the repositories without duplicating them before playing them.

Local storage will also increase as well as the need for intelligent indexing and browsing. For home systems, technologies allowing to modify, edit and combine audio content are candidate extensions to traditional HiFi systems. Intelligent indexing either by meta-data or by music signal analysis will clearly create a new way for users to interact with their content base.

Creating new products:

Home servers will be created to store content, including music for the family in the home. Servers will also required network connectivity to receive and exchange (music) content with the outside world and within the home. Intelligent indexing and browsing capabilities will become differentiating factors between products, more so that raw storage capabilities.

Music editing machines can also be envisaged either as part of a home theatre setup or as separated, possibly portable, machines. Some of the technologies envisaged in the SemanticHiFi project will allow users to act as musicians and not only as passive listeners. This is also true for indexing and browsing where the users will be able to use tools mimicking the skills of professional musicians.

Creating new services:

Storage management services, digital rights management services, intelligent or aided indexing and browsing services, distribution, editing and (on-line or on media) publishing services will also meet market demands for these new types of interaction with music. Intelligent indexing may require intense pre-processing of large amounts of audio content that can be done on-line as a service to users. Updates and sharing of metadata content may also become a service, possibly in a peer-to-peer environment where contributions could come from the users themselves.

Combine new services with new features on existing or new products:

The new features and services provide many possible combinations on existing or on new devices. Clearly there is a marketing effort to be done here to define which combinations would meet the needs and demands of the customers (and of which customers). The “packaging” and the design of the user interfaces for these new combinations is also a challenge given their innovation and the fact that users simply cannot imagine that these features are even possible.

Native Instruments market exploitation and strategy for SemanticHIFI

Music software at Native Instruments, today

Since its foundation in 1996, Native Instruments implements sophisticated technologies into easy to use audio products. *Reaktor*, the first application of Native Instruments, was the first modular development environment for digital musical instruments which was accepted and used by a broad audience. Native Instruments also developed the first successful emulations of vintage keyboards and introduced FFT based audio effects to a broader usership. *Kontakt*, the first modular sampler on the market allows extremely creative work with audio material. Native Instruments' products are targeted to beginners as well as to high end professionals, concentrating more and more the entry level.

The development and marketing of the *Traktor* line of DJ products is strongly related to the line of musical instruments. Firstly, because sophisticated audio technologies are both being used in products for DJs and for musicians. Secondly, the difference between musicians and DJs gets more and more blurred. Native Instruments is considered by many observers as the leading brand both in the markets of software instruments and professional DJ software.

Native Instruments presented the first version of *Traktor* in early 2001. The software rapidly gained interest by professional and amateur DJ's. A strategic alliance with the DJ equipment manufacturer Stanton Magnetics resulted in the recent release of *Traktor Final Scratch*, a DJ software which uses vinyl records to control the playback of audio files on personal computers. The system is regularly used on stage by some of the most known contemporary DJ's, such as Paul Van Dyk (Trance), Richie Hawtin (Techno), John Acquaviva and Francois K (House), Coldcut (Breakbeat), and DJ Craze (Hip Hop).

Recording a DJ's performance as a control file is a core feature of *Traktor*. It allows the playback of a mix without duplicating the audio material. This technology is being used in Germany's most successful dance compilation *Dream Dance* by Sony Music. Eight *Dream Dance* editions with a total of over a million copies have been released since 2001, each containing a bonus DJ mix to be played back by a *Traktor* Player software included in the data track of the CD.

Prospects for Semantic Hifi

The Semantic Hifi project includes technologies which are demanded by many customers and which have already been envisaged for the future of *Traktor*. The project will add even more possibilities for advanced users which are interested in the additional functionality. At the same time, sophisticated automation features will allow semi-automated mixing for beginners which will strongly profit from the software assisted analysis and performance modes. Having a strong impact on the development of a major software of Native Instruments, the Semantic Hifi project is being considered as one of the key projects for the next years.

Extending the functionality of existing products:
The remixing features of *Traktor* will be extended by technologies for segmentation and source separation of musical material. In addition, automated mixing will be extremely easy to use. *Traktor*'s current ID3 based database will include advanced indexing and browsing features. A seamlessly integrated server access for providing music, metadata and control information can deliver an easy way of buying and legally sharing different types of data. Special editions of *Traktor* might also be ported to hardware devices such as PDAs or mobile phones.
Also Native Instruments' products for musicians which use sample playback, such as *Kontakt* and *Reaktor*, will profit strongly from enhanced indexing and browsing of sample content as well as from methods for the segmentation and manipulation of audio.

Creating new products:
The preparation of audio files and mixes for playback and interaction on home Hifi systems will become increasingly important for music playback applications. The software will tightly work together with engines on dedicated hardware platforms with clearly defined access to the mixes and metadata.
The newly developed software versions will offer an easy to use interface for algorithms which will be developed within the project. Two versions are envisaged:

1. A version of the authoring software running on standard desktop PC's for indexing and preparing musical content,
2. A version of the software running on the Hifi System as plug in, that extends the inbuilt features to a highly interactive music centre or even a DJ console. The semi-automated remixing and recording functions will allow even inexperienced customers to produce professionally sounding results.
3.

Developing a new market:
DJs are at the forefront of funful and creative usage of music. They are in the focus of public interest and among the few contemporary idols, especially for young customers. The position of Native Instruments in the professional market allows the authentic involvement of a large number of prominent DJs. These are most suited to work not only as content producers but also as efficient promoters of the new technologies.

B.3.6 Risks in the Proposal and steps to minimize

Long term/short term balance

This project combines long term research on metadata which requires frequent exchange with the applications workpackage at important milestones. This is the first challenge of this project which has to ensure the availability of adequate resources for the coordination and users feedback in this parallel process.

Competition between partners

A beneficial competition / complement already exist between the research partners (Ircam/UPF/FHG/Sony CSL). The same situation may occur between the industrial partners. Over the 3 years of the projects industrial strategies will evolve, Sony may get more and more involved in software authoring tools and Native Instruments may get more involved in listening / performing devices. However, the current positioning of the partners is quite clear and transparent: research partners know and respect themselves because they already exchange and work together. Industrial partners have a lot of opportunities to work together.

Balance between proprietary and open standards

One threat is particularly sensitive in the MPEG context : the adoption of proprietary standards (such as Microsoft de facto standards) instead of using open standards which brings more interoperability and more competition. MPEG4 is currently struggling for life at this level, MPEG7 is not even at this stage but will be the first to be challenged when large industry groups will start thinking about interoperability and scalability.

Competition with large groups

Contacts initiated during the project preparation revealed hard competition between the major Consumer Electronic Multinationals. It can be predicted that these groups will get involved in the field rapidly and may even propose similar projects in the FP 6. The partners must be able to move quickly following industry moves and monitoring technology.

Poor availability of IPv6 addressing

Ircam is currently part of the IPv6 task force in France and is active in the group for Educational and Cultural markets and also in the Home devices domain. A lot of political issues are speeding up and slowing down the adoption process. Beyond W3C decisions, the integration of IPv6 addressing in consumer products such as the Semantic HIFI is an important issue for the spreading of such systems even more when downsized to mobile systems and portable devices.

B.4 The consortium and project resources

B.4.1 Tabular Description of the Consortium

Participant	Country	Role	Function	Note
IRCAM	F	Coordinator Research Lab. Music Production Music Archive	Coordination Research Partner Application developer	User feed back issue is supported by the running interaction between scientists and musicians at Ircam
Fraunhofer-Gesellschaft	D	Participant Research Lab.	Query by Humming Audio management and conversion Peer-to-peer file sharing and messaging Audio rendering	Peer-to-peer file sharing will be developed together with 4FriendsOnly.com Internet Technologies AG
Sony CSL	F	Research Lab.	Browsing Modules Rendering Modules Performing modules	Natural links with Sony Music and Sony NSCE
UPF	ES	Research Lab	Performing Modules	Previous experience in working with the music industry
BGU	IS	Research Lab	Research	Joint research with Ircam
NI	D	Integrator	Development & validation	Leader in music software
Sony NSCE	B	Integrator	Development & validation	Leader in Consumer Electronics

B.4.2 Sub-contracting

Fraunhofer AEMT is associated with 4FriendsOnly.com Internet Technologies AG (4FO) acting as spin-off and subcontractor.

Fraunhofer AEMT conducts research in the fields of security of virtual goods including cryptography and legitimation-testing according to legal sharing of multimedia files. Peer-to-peer sharing, messaging and payment technology is developed and provided by 4FriendsOnly.com Internet Technologies AG that holds a patent on a peer-to-peer sharing system called "Potato System". This technology will be the basis for the Fraunhofer AEMT developments of legal file sharing and messaging capabilities in context of Semantic HIFI. Together, Fraunhofer AEMT and 4FriendsOnly.com will develop DRM and accounting technology for the peer-to-peer file sharing system. By including messaging capabilities into file sharing, they aim to raise the attractiveness of new business models for consumers.
During the project and afterwards 4FriendsOnly.com will run a sharing and accounting web-service that cannot be provided by Fraunhofer AEMT as Fraunhofer-Gesellschaft is not allowed to do consumer business.

4FriendsOnly.com Internet Technologies AG (4FO) (www.4fo.de/en) was founded in June 2000 as a spin-off from Technical University of Ilmenau. 4FO provides technical solutions and business models according to game software distribution, micro-payment and legal file sharing systems. With the help of 4 employees and many students 4FO developed a micro-payment system, that lets the customer choose the way he wants to pay (www.paybest.de). The game feature platform allows for the dissemination of basic parts of computer games in order to sell further parts that raise attractiveness of the game play. 4FO holds the patent for the game feature platform technology. To encourage legal sharing of music and other multimedia files a sharing and accounting network called Potato System has been developed. (www.potatosystem.com , http://www.4friendsonly.org/eng/).

In the Potato System, which is applied for patent, the user plays an active distribution part. Our approach motivates the users to re-distribute content they have paid for and earn money with it. This allows a fast distribution of new content. The Potato System provides its own P2P clients which contact a central web-service. This will open new opportunities for music exploitation. In 2002, the 4FO AG established a close co-operation with the Fraunhofer Arbeitsgruppe für Elektronische Medientechnologie in Ilmenau, Germany. One topic of this wider co-operation is the realization of a joined patent in the field digital music business. After Jürgen Nützel received his diploma in electrical engineering at the University of Applied Science in Schweinfurt, he joined in 1991 the company Siemens AG in Fürth/Bavaria. He developed several i/o-boards for automation systems. In 1990 he started in parallel his studies at the Technical University of Ilmenau. In 1994 he received his diploma and joined the University for his PhD thesis in computer science.

After acquiring his PhD in 1999 with object-oriented design of embedded real-time systems he shifted his research focus to virtual goods. He started his "Habilitation" in this field. In 2000 he founded the company 4FrindsOnly.com Internet Technologies AG (4FO). 4FriendsOnly.com Internet Technologies AG (http://www.4fo.de) develops and runs payment services and content management systems for virtual goods. This company develops and runs payment services and content management systems for virtual goods. Jürgen Nützel currently works as an "assistant professor" at the University in Ilmenau. He is also CEO of the 4FO AG. In 2003 he organizes together with Prof. Grimm the international Workshop "Virtual Goods".

B.4.3 Other countries

BGU participation is important for succeeding in achieving one major Workpackage related to source separation and polyphony management which of particular importance both for the rendering system (ability to perform spatial mixing in real-time) and the performing system (real-time mixing). BGU has been an important participant in the CUIDADO project. This fundamental research lab in mathematics and statistics has a strong experience in collaboration in the music and audio field and has been publishing several papers in Journals and conferences such as :

- [DubTish97] S.Dubnov and N.Tishby, “Analysis of Sound Textures in Musical and Machine Sounds by means of Higher Order Statistical Features”, Proceedings of the International Conference on Acoustics Speech and Signal Processing, Munich, 1997.
- [DubRod98a] S.Dubnov and X.Rodet “Timbre Recognition with Combined Stationary and Temporal Features”, Proceedings of the International Conference on Acoustics Speech and Signal Processing, Ann Arbor, 1998.
- [DubRod98b] S.Dubnov and X.Rodet “Study of spectro-temporal parameters in musical performance for expressive instrument synthesis”, Proceedings of the International Conference on System, Man and Cybernetics, San Diego, California, 1998.

B.4.4 Description of the participants

P1 - IRCAM (F)

1. Company profile.

IRCAM (Institut de Recherche et Coordination Acoustique/Musique) is a research center, a music production resource (concerts, CDs, books) and an education place gathering musicians and scientists to work on music and new technologies. Basic and applied research covers 4 main areas :

? acoustics (psychoacoustic, room acoustics, instrumental acoustics)
? sound analysis/synthesis and sound design
? musical data representation/manipulation and composition
? real-time systems and online services

Ircam basic research activities are performed in the framework of an Ircam-CNRS joint laboratory, named Science and Technology of Music and Sound (STMS, UMR Ircam-CNRS 9912).

2. Number of employees: 120

3. Budget: 10 MEuros

4. Experience in other EU project

Ircam has been participating in the 4th framework program in two working groups :

? Harmonica (Music Information in Libraries) as a partner in the Working group 2 on User needs and interfaces.
? CUIDAD (Content Processing of Music) as a coordinator for gathering research, industry and users input on new applications of content description, navigation and processing with inputs in the current MPEG7 standardization process.

In the 5th framework program Ircam has been involved in 5 projects :

- WedelMusic (Access and navigation into music scores in a protected environment)
- Carrouso (3D audio authoring , broadcasting and rendering)
- Listen (Virtual audio system for museums)
- Cuidado (Content-based interfaces for audio/music databases available online)
- Agnula (A Gnu Linux Audio distribution)

Ircam is currently leading the Multimedia Standards Working group of the MusicNetwork NoE.

6. Role in the project

Ircam will be the coordinator of the project in charge of Specification and Architecture aspects of the project. Indexing, rendering and Dissemination along the project are also to be managed by Ircam.

7. Curriculum Vitae of key personnel

Hugues Vinet (Scientific Director, IRCAM)

Hugues Vinet has been managing all the scientific activities at Ircam since 1994, he is currently the Coordinator of the CUIDADO FP5 project and will be project Coordinator.

Pascal Mullon (head of Online Services team)

He is the software architect of the CUIDADO project and will coordinate the architecture design and development, as well as the development of specific applications.

Xavier Rodet (head of the Analysis/synthesis team)

He will be supervising all the developments on WP related to extraction of audio parameters.

Geoffroy Peeters (analysis synthesis team)

He is senior researcher in DSP and will be coordinating the WP related to audio extraction

Olivier Warusfel (head of Room Acoustics)

In charge of Ircam contribution in Listen and Carrouso projects he will supervise the rendering WP.

Olivier Delerue (Room Acoustics team)

Conducts basic research in the field of constraints and music, and has developed a variety of constraint solvers for music interaction, as well as musical feature extraction techniques.

Norbert Schnell (Head of Real time Applications team)

He will be involved in the WP related to Performing

Diemo Schwarz (Real time Applications team)

He will be in charge of research and development related to the Performing WP.

Vincent Puig (Marketing Director, EU projects coordinator, IRCAM)

He is managing Ircam Hypermedia Studio and will be coordinating Partner relations and consortium agreement, Intellectual Property Management and results exploitation.

P2 – Fraunhofer-Gesellschaft, Institute for Integrated Circuits (D)

1. Company profile.

The **Fraunhofer Gesellschaft** (FhG) is the leading organization of applied research in Germany. A staff of around 13,000 is employed at 57 research establishments throughout Germany, most of them scientists and engineers.

The Applied Electronics Centre of the **Fraunhofer Institute for Integrated Circuits, IIS,** develops microelectronics circuits and systems, from microchips to digital broadcasting systems.

Fraunhofer IIS is the leading research laboratory in the area of audio coding. Scientists and Engineers of IIS have participated in the development of several audio coding algorithms including MPEG Layer-3, MPEG-2 AAC and MPEG-4 Audio. Fraunhofer helped the European companies Micronas Intermetal and Thomson Semiconductors to develop the first MP3 decoder chips. A key activity is digital audio broadcasting and data systems also based on other wireless transport systems (DECT, GSM) as well as handheld data terminals.

The **Fraunhofer Arbeitsgruppe für Elektronische Medientechnologie, AEMT,** situated in Ilmenau, is a subsidiary of Fraunhofer IIS founded at the end of 1999. Its focus is on extraction and handling of metadata, virtual acoustics, audio coding for special applications and DRM business solutions.

Information about Fraunhofer IIS and Fraunhofer AEMT is available at http://www.iis.fraunhofer.de and http://www.emt.iis.fraunhofer.de.

2. Number of employees:

13,000 related to Fraunhofer Gesellschaft in all.Fraunhofer IIS has a permanent staff of about 270. At Fraunhofer AEMT about 50 permanent employees are engaged.

3. Budget: One Billion Euros for Fraunhofer Gesellschaft as a whole.

4. Experience in other projects

Fraunhofer IIS receives contracts from companies and participates, as a member of consortia of several R&D institutes and industrial suppliers, in national and European programs such as ESPRIT, EUREKA, FUSE, LEONARDO, MEDEA, RACE. Fraunhofer AEMT participates in European programs such as CARROUSO.

5. Company (Group) skill

At Fraunhofer AEMT two teams work on metadata extraction and classification. Two major research topics are identification of any audio material (AudioID) and melody recognition also called Query by Humming (QbH).

6. Role in the project

Fraunhofer AEMT will conduct research on integration of melody recognition, audio indexing, management and conversion of several audio formats. Further it will develop technologies to support legal file sharing and messaging in peer-to-peer networks under control of accounting services that ensure DRM.

7. Curriculum Vitae of key personnel

Rüdiger Grimm is University professor for multimedia applications in the Technical University of Ilmenau, Germany. He is Deputy Director of the University Institute for Media and Communication Studies, and head of the research group "Security for Virtual Goods" in the Fraunhofer Institute for Integrated Circuits IIS.

1985-2000 Rüdiger Grimm was scientist in the Research Institute of Secure Telecooperation in GMD, Darmstadt, and concerned with communication technology and IT security. 1995-2000 he was head of the research area "Marketplace Internet". Since 1993, he gave lectures at the University of Frankfurt on IT security. One of his specialities is electronic payment. His professional interest is in the commercial usage of the Internet and related questions on security and legal binding.

Markus Cremer

Markus Cremer joined the Fraunhofer Institute for Integrated Circuits (IIS). He there was involved in the development of real-time audio codec platforms and the design of prototype and test systems for satellite radio broadcast systems. In 2000 Cremer joined the newly founded Fraunhofer Arbeitsgruppe fuer Elektronische Medientechnologie in Ilmenau, Germany, to work in the field of automated metadata extraction from audio signals. He currently heads the team "Automated Audio Analysis". Markus Cremer is member of the Audio Engineering Society (AES). He is co-author of 4 papers in the field of audio metadata.

Frank Klefenz

In 1992 he became leader of the division "Massively Parallel Processors" at the University of Mannheim with several projects granted by GSI – Gesellschaft für Schwerionenphysik, Darmstadt.

In 1997 he joined the french company Tekelec. In 2000 he joined the newly founded Fraunhofer Arbeitsgruppe fuer Elektronische Medientechnologie in Ilmenau, Germany, to work in the field of automated metadata extraction from audio signals. He currently heads a team that develops melody recognition technologies.

Frank Klefenz published 28 papers on massively parallel FPGA computers for high-speed pattern matching.

P3 - Sony France (F)

1. Company Profile

Sony is an established consumer electronics manufacturer, supplying products and systems related to audio, video and broadcasting. It devotes a substantial part of its turnover to research and development activities, carried out by dedicated corporate research laboratories, active in the fields of components and audio visual technologies, telecommunications, multimedia information and data management systems. Sony France more particularly houses a laboratory specialized in computer science, CSL, which conducts basic research in the field of music, particularly content-based access to music catalogues.

2. **Number of employees:** 2484 persons

3. **Budget:** 1 672 845 585 Euros

4. Experience in other EU projects

As a leading manufacturer of consumer and professional electronic equipment, Sony has participated to a number of other EU research projects. These include:

- SPEECON (recognition technologies)
- BRAIN (Radio access related technologies)
- VIRTUE (creation of virtual environments)
- NEXTV (development of interactive TV)
- SCARE

5. Company (Group) skill

Since its creation in 1946, Sony has always been a leading figure in the world of applied hardware and software. In addition to its expertise in electronics and information technologies, Sony now benefits from its ever increasing presence in the audio and musical domains through its subsidiary Sony Music which creates and distributes musical recordings, amongst which, music indexing, throughout the world. The Sony CSL laboratory has developed and accumulated expertise in combinatorial optimization in the multimedia domain, through extensive collaborations with Sony Music experts. Applications investigated range from constraint-based mixing to automatic play-list generation from music catalogues.

6. Role in the project

Sony CSL will lead research on the combinatorial optimization aspects of the project, and develop research on constraint satisfaction techniques (sequence generation system for sounds and music titles) and will involve Sony Music France in the validation phase.

7. Curriculum Vitae of key personnel

François Pachet (Ph.D., Eng., Team leader). François Pachet, got a Ph.D. in Computer Science/Artificial Intelligence from University of Paris 6, where he has been assistant Professor since 1994. He has been doing research in computer music for 10 years. He now heads the music research team at the Sony Computer Science Laboratory in Paris, where he develops techniques, including constraint-based, for designing and building interactive multimedia systems.

Atau Tanaka is a permanent researcher at Sony CSL. He is specialized in music network systems and music interaction.

Jean-Julien Aucouturier is a PhD student, working on music timbre modeling and representation. He is also involved in human centered approaches to musical listening system.

Aymeric Zils is a PhD student working on the extraction of high level music descriptors from acoustic signals. He is the main designer and developer of the EDS system.

Amaury Laburthe is a research associate working on architectures for music browsers.

P4 - UPF / Pompeu Fabra University - MTG (E)

1. Company profile.

The Music Technology Group, MTG (http://www.iua.upf.es/mtg) is part of the Audiovisual Institute of the Universitat Pompeu Fabra of Barcelona and specializes in audio processing technologies and their musical and multimedia applications. MTG carries out research and development projects in areas such as audio processing and synthesis; audio identification; audio content analysis, description and transformation; singing voice synthesis and transformation; interactive systems; and software tools. The MTG is currently working in many public funded projects, both at the EU and national levels, and private funded projects with companies like Yamaha, SGAE, Telefónica I+D, Roland, Tape Gallery and DUY. The main research areas of the MTG are Audio Processing and Synthesis, Audio Identification, Audio Content Analysis, Audio and Internet, Singing Voice Processing, Interactive Systems, Software Tools.

2. Number of employees: 814 professors, 8703 students, 349 administrative (UPF). MTG: 40

3. Budget: 52 MEuros

4. Experience in other projects

? RAA - Recognition and Analysis of Audio (IST-1999-12585),
? CUIDADO - Content-based Unified Interfaces and Descriptors for Audio/music Databases Online (IST-1999-20194),
? OPENDRAMA (IST-2000-28197),
? AGNULA - A GNU Linux Audio Distribution (IST-2001-34879),
? AUDIOCLAS (EUREKA E! 2668),
? MOSART - Music Orchestration Systems in Algorithmic Research and Technology (HPRN-CT-2000-00115),
? CUIDAD - Content-based Unified Interfaces and Descriptors for Audio/music Databases (ESPRIT-28793)
? TABASCO - Content-based Audio transformation (TIC-2000-1094-C02-01),
? Digital Audio Effects (COST G6 – DAFX)

5. Role in the project

MTG-UPF will contribute to the automatic analysis of music and will lead the research on the Performing workpackage, mainly contributing to the Voice Analysis and Processing in a real-time environment and Variable Audio Time-Scale Modification for rhythm and tempo variations.

6. Company (Group) skill

The Music Technology Group has a strong background in research in the area of audio signal processing, carrying out basic and applied research in music applications. In the last few years the MTG has been very active in the area of automatic audio description and has contributed to the MPEG-7 standard and on real-time analysis and synthesis of musical signals.

7. Curriculum Vitae of key personnel

Xavier Serra (Director of the MTG)

Xavier Serra is the director of the Audiovisual Institute (IUA) of Pompeu Fabra University (UPF) where he is also founder and director of the Music Technology Group and director of the doctoral program in Computer Science and Digital Communication. He holds a Master degree in Music from Florida State University (1983), a Ph.D. in Computer Music from Stanford University (1989) and he has worked for two years as Chief Engineer at Yamaha Music Technologies USA, Inc. His research interests are in sound analysis and synthesis for music and other multimedia applications.

Jordi Bonada (MTG team leader), Jordi Bonada joined the Music Technology Group of the Audiovisual Institute of the UPF as a researcher and developer in digital audio analysis and synthesis. He is currently involved in research in the fields of spectral signal processing, especially in audio time-scaling and voice synthesis and modelling.

Alex Loscos (MTG team leader)

Alex Loscos joined in 1997the Music Technology Group of the Audiovisual Institute of the UPF as a researcher and developer. In 1999 he became a member of the Technology Department of the UPF as lecturer and his main research interests are voice processing / recognition, digital audio analysis / synthesis and transformations, and statistical / digital signal processing and modelling.

P5 – Ben Gurion University (Israel)

1. Company profile

Multimedia Communications (MMC) Laboratory at the Communication Systems Engineering Department (CSE) in Ben-Gurion University is dedicated to research and advanced project development in the areas of Multimedia Signal Processing and Multimedia networking. Research areas cover a broad range of information technologies dealing with multimedia signals, including digital video, computer music, audio compression and retrieval, computer generated music, and learning in Human-machine interaction. Current Research Projects include:

I. Multimedia information technologies:

Advanced sound analysis using spectral, polyspectral and time-frequency methods. Statistical learning of temporal structure in sound and MIDI using variable length Markov models and information theoretic methods. Multi-disciplinary issues including cognitive and expressive aspects of man-machine communication.

II. Networked multimedia systems:

Multimedia Systems Technology: Real time sound synthesis and protocols for transmission of real-time data, Multimedia system architecture, Multimedia system design: Buffer design, Traffic shaping, Scheduling, Congestion Control, Distributed Multimedia Server and Data Management, Distributed Multimedia Applications.

2. Employees: 13500 students in BGU and 500 in the R&D Faculty

3. Budget (1998/99) : USD$ 164 million

4. Experience in EU projects:

Ben-Gurion University of the Negev was founded by the government of Israel in 1969. The University is committed to the advancement of education, teaching, and research in all the fields of Human knowledge. The University has endorsed a special commitment to broad based research activities and especially seeks joint research projects which can benefit mankind at large. Towards this end, Ben-Gurion University has successfully participated in both the 4th and 5th Framework Programs in a wide variety of funded projects.

5. Group skill

Mr Dubnov lab is specialized in Information management and statistics applied to media.

6. Role in the project:

The BGU expertise in advanced sound analysis methods and statistical learning for content-based audio retrieval are expected to bring an important contribution related to source separation. During the years 1996-1998, Shlomo Dubnov was a Co-Investigator at IRCAM and at the Ben Gurion University in a Franco-Israeli Research Cooperation Project on ``High level structuring and intelligent content based retrieval in audio databases", supported by Israeli Ministry of Sciences and AFIRST.

7. Curriculum Vitae of key personnel

Shlomo Dubnov: (researcher, in charge of the MMC Laboratory).

(b. 1962) Received both scientific and musical education: B.Sc in physics and mathematics from the Ben Gurion University, B.Mus from Rubin Academy of Music, Jerusalem, M.Sc in Electrical Engineering from Technion, Haifa, and a Ph.D. in Computer Science at Ben Gurion University, Jerusalem. During the years 1996-98 he worked as an invited researcher and co-PI at IRCAM.

Dubnov is in charge of Multimedia Program at Communication Engineering Department in Ben-Gurion University, Israel, where he teaches information theory and multimedia signal processing.

He is a founder of the Israel Computers and Music Forum and serves as its chairman ever since. He also regularly teaches courses in computer music and computer applications for music composition and research.

As part of its AC cost model, BGU will dedicate a variety of its own resources to the project, including permanent staff (Academic faculty, Technical support staff, Administrative personnel), as well as other resources, including computers and network, University support services and support units, physical plant and depreciation on buildings and other capital items. As an approximate, but realistic estimation, this investment of BGU own resources represents **49% of the total cost of BGU activities in the project.**

P6 – Native Instruments (D)

1. Company profile.
Founded in 1996, Native Instruments was one of the first companies to use the possibilities of real-time sound synthesis on Macintosh and Windows platforms. The current products include a modular system for the development of audio and event processing structures, emulations of classic keyboards, innovative synthesizers, samplers, sample libraries, audio effects, and computer DJ products. The products which have earned a reputation for exceptional quality and innovation are being developed in close cooperation with musicians and DJs.
2. Number of employees: 70
3. Budget: 6.5 MEuros
4. Experience in other projects
Native Instruments has participated in PAIDFAIR (Protecting Accumulated Intellectual Data For Accounting In Real Time), a project in the 5th framework program.
5. Company (Group) skill
Native Instruments has extensive knowledge on the development of software instruments and DJ applications. Traktor is a complete, software-based DJing solution, including decks, beat matching, mixer, filters, recording of realtime edits, and music database functions.
6. Role in the project
Native Instruments will be responsible for the authoring part of the project and will contribute instruments to the work packet related to performing.
7. Curriculum Vitae of key personnel
Egbert Jürgens (Business Development Director)
He is responsible for business development and manages the product line of vintage emulations. Before joining Native Instruments in 1999 he studied physics and received a PhD on signal analysis and modeling in brain science at the University of Marburg. He will be the Project and Exploitation Coordinator.
Friedemann Becker (Traktor Product Manager)
He is product designer and product manager for DJs applications. Before joining Native Instruments in 2000 he studied physics in Tübingen, Marseille and Milano and made his thesis in computer aided pattern recognition. He will be the Technical Manager of the project.
Toine Diepstraten (Traktor Developer)
He is the main developer of Traktor and will be responsible for software development. Before joining Native Instruments in 2000 he acquired a Diploma in Computer Science at the FH Karlsruhe.
Florian Haver (Marketing Director)
He is the head of the marketing department and will be involved in the exploitation of the project results.
Daniel Haver (CEO)
He is the CEO of Native Instruments and will be involved in the specification of the authoring software and the exploitation of the project results.

P7 – Sony – Network & Software Technology center Europe NSCE (B)

1. Company Profile
Sony is an established consumer electronics manufacturer, supplying products and systems related to audio, video and broadcasting. It devotes a substantial part of its turnover to research and development activities, carried out by dedicated corporate research laboratories, active in the fields of components and audio visual technologies, telecommunications, multimedia information and data management systems. Sony NSCE is responsible for supporting the development and maintenance of Software platforms for various products developed in Europe such a digital TV and set top box products, mobile phones, and Aibo.

2. **Number of employees:** 80 persons

3. **Budget:** 11.5 M Euros

4. Experience in other EU projects

5. Company (Group) skill
Since its creation in 1946, Sony has always been a leading figure in the world of applied hardware and software. In addition to its expertise in electronics and information technologies, Sony now benefits from its ever increasing presence in the audio and musical domains through its subsidiary Sony Music which creates and distributes musical recordings, amongst which, music indexing, throughout the world. Sony NSCE has extensive experience in the development of real-time software for embedded systems, ranging for operating systems, to HW drivers, to middleware and to applications.

6. Role in the project
Sony NSCE's main role in the project is to participate in the design and development efforts of the project by bringing an industrial view on the developed technologies and to integrate these technologies into actual Sony products or prototypes in order to demonstrate their applicability to actual products.

7. Curriculum Vitae of key personnel
Michel Feret (Ph.D.). Michel Feret, received a Maitrise d'Informatique from the University of Paris 6, and a Ph.D. in Computer Science from Queen's University, Kingston, Ontario, Canada. His main research focus has been Case-Based Reasoning (CBR), a research domain where he was active until 1993. Since then, he has worked for Northern Telecom on real-time operating systems (94-95) and at Sony in Japan in the areas 3D virtual words and again on real-time operating systems (95-96). He currently leads NSCE whose mandate is to develop software for various Sony products, such video-on-demand systems, digital TVs and set top boxes, mobile phones, and networked products.

B.4.5 Quality of partnership, involvement of users and SMEs

The Consortium has managed to involve an innovative professional music software company in Europe despite the lack of European companies in this field. Sony CSL part of Sony France, with close collaboration with Sony Music as the well known music major, brings visibility and marketability of the developed technologies associated with its music contents and its own research in the field. Sony, as the leading consumer electronics manufacturer investing research at the European level shows the strategic importance of content based technologies. This combination of industrial interests at very different levels is very promising when associated with internationally recognized research centers in Music with cutting-edge content-based technologies (IRCAM, UPF, Ben Gurion University, Fraunhofer Institute).

Geographic balance
Partners are from important countries where consumer electronics can compete with USA and Japan with strong research and cultural background (France, Germany, Spain). A partnership with Israel at the research level is also bringing high-end technology in the content processing domain.

Research / Industry balance
There is 5 research entities (IRCAM, UPF, BGU, FHG, Sony CSL), 1 industrial partner world leader in consumer electronics (Sony represented by its R&D center on Network and Software Components NSC Europe) and 1 industrial partner involved in the development of music software (Native Instruments).

Standardization experience
IRCAM, UPF and FHG are actively participating into standardization processes and at the first place MPEG and W3C.

Multi-level Marketing Exploitation
There are at least one industrial partner for each Workpackage of the project. One industrial partner is involved in the Consumer Electronics industry (Sony) and the other one is involved in the music software/hardware business (Native Instruments). Another important factor of success is the fact that the Research centers involved in the project have their own software development policy and means which guarantees a smooth transfer from their labs to the industrial partners.

Involvement of users
Involvement of music fans, music creators and sound designers in all phases of the projects is of key importance. Similarly to value being added to music by composers, value and creativity will be brought in this project by music fans which is not often possible in European RDT projects but here fully justified by the very nature (content) of the project scope. For instance intuitive and perceptual entries used by artists must be superimposed to the low-level control reserved to machines or experts but keeping a scientific approach. The number of workshops, conferences and publications presenting the project not only in the scientific and industrial domain but also in musical and artistic networks is also one key strategic impact.

Involvement of SMEs
Apart from the 5 research entities and the large company (Sony) involved, the project has managed one of the most innovative SME in the field of music software tools (Native

Instruments). This shows the challenging aspects of the project in developing not only access and listening tools but creative authoring tools both based on advanced metadata research.

B.4.6 Resources to be deployed

The project duration is 36 months since it is mainly a research project where new methods, algorithms and software cannot be developed in a short period and furthermore must be integrated into a platform with high constraints. However the research results will regularly feed into the application development and in addition will contribute as new features to industrial partners existing partners. The research & development oriented criteria is also visible in the MM split between the 3 main tasks:

Management, Specifications and Technical coordination (WP1) : 18 %
Research : 50 %

WP2 Signal Indexing:	12 %
WP3 Browsing :	10 %
WP4 Rendering :	12 %
WP5 Performing :	16 %

Applications Development, integration, dissemination and validation: 32 %

WP6 Authoring:	13 %
WP7 Sharing :	7 %
WP8 HIFI System :	9 %
WP9 Dissemination and user groups:	3 %

B.4.7 Overall Financial Plan

	Title	IRCAM coord.	IRCAM	FHG	Sony CSL	UPF
Hour rate		45	45	51	53	28,
Hours/month		140	140	140	151	14
Mmcost		**6 300¤**	**6 300¤**	**7 209¤**	**8 003¤**	**4 000**
Personnel		108 234¤	1 008 000¤	338 804¤	448 168¤	195 990¤
% of pers costs		**4%**	**37%**	**12%**	**16%**	**7**
Equipments			10000	0	0	80(
Consumables			3500	3000	1700	20(
Travel & Subsistance			20000	14000	8500	100
Computing						
Protection of Knowledge						
Other costs			3078			50(
Overheads rate		80%	80%	80,13%	80%	108,9
Overheads		86 587¤	835 662 ¤	285 106 ¤	366 694 ¤	240 81
Subcontracting				64000		
TOTAL		194 821¤	1 880 240¤	704 910¤	825 062¤	461 803
Cost basis		FC	FC	FC	FC	FC
Requested funding		100%	50%	50%	50%	50
Funding		194 821¤	940 120¤	352 455¤	412 531¤	230 902

B4.8 STREP Project Effort Form

Full duration of project

Project acronym – SEMANTIC HIFI

	IRCAM	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE	TOTAL PARTNERS

Research/innovation activities	IRCAM	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE	TOTAL PARTNERS
WP Spec Architecture	24	3	5	1	1	5	5	**44**
WP Indexing	26	8	3	0	10	3	0	**50**
WP Browsing	9	18	12	0	0	3	0	**42**
WP Rendering	36	2	12	0	1	0	0	**51**
WP Performing	17	0	12	39	0	0	0	**68**
WP Authoring	17	0	0	6	0	33	0	**56**
WP Sharing	12	13	0	0	0	1	3	**29**
WP HIFI System	4	1	8	1	0	1	24	**39**
WP Dissemination	3	1	3	1	1	4	2	**15**
Total research/innovation	**148**	**46**	**55**	**48**	**13**	**50**	**34**	**394**

Demonstration activities	IRCAM	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE	TOTAL PARTNERS
Total demonstration								**0**

Management activities	IRCAM	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE	TOTAL PARTNERS
WP Management	29,18	1	1	1	1	1	1	**35,18**
Total management	**29,18**	**1**	**1**	**1**	**1**	**1**	**1**	**35,18**

	IRCAM	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE	TOTAL PARTNERS
TOTAL ACTIVITIES	**177,18**	**47**	**56**	**49**	**14**	**51**	**35**	**429,18**

B.5 Project management

The Men-months allocated to various partners in WP1.1 correspond to participation in meetings and administrative management. The effective technical management tasks appear in cross-participation of partners in various WPs, with a limited number of MM (1 in the majority of cases).

B5.1 - Project Management Structure

IRCAM as prime contractor is responsible for the overall project management but will ensure that users and artists are involved in each phase of the project. IRCAM will nominate a Project Manager who will act both as the focal point for all activities in the project, and as the primary interface between the project and the Commission. The Project Manager's responsibilities are as follows:

? controlling but facilitating the information flow among partners through regular contact with the partners to ensure that the project program, milestones and time scales are maintained and deviations resolved and recorded with change control
? to interface with the Commission for all matters associated with the project and maintaining relationships between the EC and the partners
? to lead the project and hold regular Technical Progress Meetings with the Technical Managers
? to hold Aministrative Committee Meetings (ACM) with Administrative Managers, where deviations from the project cannot be resolved in regular Project Management Committee meetings.
? to co-ordinate the preparation and distribution of all major reports
? to maintain accurate records of costs, resources and time scales for the project and to produce and circulate agreed minutes of each meeting.

In order to monitor the project, each Partner will produce a brief report (in a standard Excel form) and deliver it to the project manager by the end of each month. Partners will meet every six months for formal reviews. Each WP and sub WP leader will be responsible for the following:

? plan the detailed technical work of the WP
? ensure that project timetables are maintained, and flag any discrepancies immediately to the Project Manager
? initiate corrective action for project deviations, if any
? consolidate partner information and prepare short monthly reports for submission to the Project Manager
? ensure the objectives and milestones of the whole WP, and of the detailed activities within the WP, are achieved
? ensure deliverables are available on time.

Project Management Committee

A Project Management Committee (**PMC**) set up by the partners will come under the overall control of the Project Manager and will handle all questions related to the project execution : coordination and information exchange between partners, technical and scientific coordination between WP. The PMC may also meet at the coordinator's initiative in order to find appropriate solutions to problems and conflicts which may occur during the project execution. Each partner will nominate a Technical Manager for each WP, who will be represented in the PMC and whose role will be to supervise all technical activities at the partner's level, as well as coordinating the work of all partners related to the WP. Technical Managers will report to their Administrative Managers.

Administration Committee

Each partner will nominate an Administrative Manager whose role will be to supervise the administrative work within his company and who is duly empowered to make all necessary decisions on behalf of his organization. During the project the tasks of the Administrative Managers will comprise :

- ensuring that all necessary contractual and cost information is supplied in a timely manner (once every six months) to the Project Manager,
- making all resources and facilities required for the project available,
- attending meetings of the Administrative Board,
- participating in contractual negotiations on behalf of the organization he represents,
- serving as a contact for any project administrative matters that relate to her organization.

Exploitation Committee

Furthermore, each partner will choose an Exploitation Manager. The Project Manager, the Exploitation Managers, or the Administrative Managers in those organizations that have not appointed a specific Exploitation Manager, form the *Exploitation Committee***.** The Exploitation Board supervises product development, dissemination and marketing efforts to keep them in line with the market requirements. It observes the market situation and recommends actions to be taken by the consortium to ensure maximum market impact. Exploitation Managers agree on terms for project collaboration (consortium agreement) and exploitation (exploitation agreement), specifying the IPR and licensing rights involved and how these will be protected/rewarded.

B5.2 - Legal obligations and IPR issues

Before the signing of the contract the project partners will sign a Consortium Agreement in which the rights of each partner will be specified. The agreement will take into account the typical rights that each partner has according to the specification given by the EC in the 6 th Programme.

The Contractors in the Consortium have a joint responsibility for achieving project objectives within time and budget. The contractors recognize their special responsibility towards the other participants in project regarding planning, managing shared activities, obtaining software or licenses and music rights to use deliverables and products for clearly agreed duration. The contractors plan to agree that availability, usage and IPR rights issues relating to Deliverables will be governed according to the ownership and availability categories set out in the technical annex to the contract. The categories listed against each deliverable may be adjusted and new deliverables added during the project only by formal agreement of the Project Committee. Any such change will be recorded by the Project Manager in the minutes of the meetings of the Committee meeting.

The consortium agreement will start from the main guidelines provided from the EC. Moreover, for the definition of the agreement the following aspects will be taken into account :

? The rights to Affiliates for the partners with mother companies out of Europe.
? The availability and the protection of already available material provided by the partners (background technology) with their corresponding value on the market.
? The percentage of the investment performed by the single partners in each specific WP and subWP that will provide the basis for producing the results and the demonstrators. The investment may include also the additional internal investments that each partner could perform for adding additional unplanned results to the project.

B5.3 - Communication flow Mechanisms and Procedures

For distributing information among partners Internet technologies will be mainly used. A specific private WWW site for the project partners will be set up. Every partner will be capable of downloading from the WWW site the most updated version of the deliverables. This WWW site will be also used for posting and exchanging reports and cumulating common documents useful for the project. Some of these will be public. Specific working meetings will be organized according to the work planned. Every meeting will be documented. All the communications will be collected and made available on the WWW site of the project.

All deliverables will be subject to peer review both by E-mail and at work package meetings before formal quality assurance activities are undertaken.

Potential conflicts will be brought to the immediate attention of the project manager by the appropriate technical or administrative manager. The project manager will then attempt to resolve this by discussion or by the calling of an ad hoc technical or management meeting. If the management team(s) within the relevant partner(s) cannot produce a written resolution of the potential conflict within four weeks, the issue(s) will be brought to the attention of the PMC. If despite all efforts no decision can be reached on the conflicting issues unanimously, the prime contractor has the right to reach a decision by majority vote of the Administration Committee.

The reference people for each partner are reported in the following table.

Partner	***Project Management Committee Members (incl. Technical Managers***	***Administrative Committee Members***	***Exploitation Committee Members***
IRCAM	Hugues Vinet, Pascal Mullon, Xavier Rodet, Olivier Warusfel, Norbert Schnell	Hugues Vinet	Vincent Puig
BGU	Shlomo Dubnov	Sharona Riffberg	Shlomo Dubnov
NI	Egbert Jürgens, Friedemann Becker, Toine Diepstraten, Daniel Haver	Egbert Jürgens	Egbert Jürgens, Daniel Haver, Florian Haver
FhG	Markus Cremer, Rüdiger Grimm, Matthias Kaufmann, Christian Sailer Jürgen Nützel	Matthias Kaufmann, Frank Klefenz	Frank Klefenz, Matthias Kaufmann
SONY CSL	François Pachet, Jean-Julien Aucouturier	François Pachet	Philippe Boulanger
UPF	Xavier Serra, Perfecto Herrera	Xavier Serra	Joaquim Solana
Sony NSCE	Michel Feret	Michel Feret	Michel Feret

B.6 Workplan

B6.1 Workplan structure and objectives

The workplan is organized across the objective of producing three complementary and interoperable applications:

- the **Smart HIFI system**, which provides a set of basic features, mainly targeted to various access modes ands interfaces to the audio data and metadata materials, and with functionality limited by hardware constraints according trade-offs to be studied and adjusted along the project :
    - title database management and browsing, with basic metadata generation and play-list generation
    - rendering : virtual acoustics and interactive high-level mixing
    - performing : basic performance algorithms and content-based browsing inside the music piece,
    - sharing : access to sharing server

  The system also includes an architecture compatible with the two other applications, in terms of interoperability and possible integration of additional features.
- the **Authoring tool software**, which may run on a separate PC and/or on the HIFI hardware and will complement the HIFI System by providing authoring features on the audio data/metadata materials, as well as more advanced interactive and editing functions :
    - full-featured indexing and metadata generation from audio files
    - full metadata editing for access in HIFI and sharing systems
    - audio file segmentation and source separation
    - real time and non real time audio file processing and personal authoring tools
    - rendering authoring features
    - full-featured performing features
- the **Metadata sharing system**, available online as a file sharing system connected to a centralized server which enables multiple users to share :
    - their metadata on audio files (play-lists, segments, high-level indexing editing, interactive mixings, performance recordings, ...)
    - their authoring work on audio files through Authoring tool scripts, which enable re-producing all the processes performed on audio files provided that these files are referenced on the user HIFI System.

An important aspect of the project will be to specify and implement a set of functionalities for the three applications which will take into account various factors such as :

- the target market and users which may be different for each of the applications,
- the consistency, from the user's point of view, of the system made of the three interoperable applications, in terms of functionalities and interfaces, through the study of various use cases which may include part or all of the applications.
- the technical architecture, including implementation constraints, and data exchange formats and protocols between applications.

In order to achieve its goals, Semantic HIFI workplan has been organized into 9 main workpackages. The first one is dedicated to management, specifications involving users throughout the project and technical coordination. The second aims at performing basic research on the core modules both at the content extraction level, classification and indexing methods adapted to personal use on top of which 3 Creative research & development workpackages are based: WP3 is related to browsing features, WP4 is dedicated to audio and multimedia rendering, WP5 addresses performance and real-time processing. The next

workpackages are dedicated to target applications: WP6 for the authoring and publishing a PC connected to Internet, WP7 for the metadata online server, and WP8 for the HIFI System. WP9 is dedicated to dissemination. Research & Development, demonstration and management phases are running in parallel thanks to a user-feedback loop methodology.

The SEMANTIC HIFI project has specific transversal objectives :

Creativity in the working methods (objective 1)

As a first objective an innovative approach will be implemented with feedback loops involving music creators and consumers through the entire project and ensuring real-time online co-ordination through a collaborative web site with quality control. Technical and functional specifications defined alongside with users will remain open to new emerging standards coming from the MPEG 7 process. The user feed-back loop process has been one key element in the success of software development thanks to Internet discussion lists not only used for technical support or beta-testing but also for directly send reactions and suggestions to the software developers. For testing the HIFI system, software emulations will be delivered to users and real testing on PC emulations will be preformed during workshops and panels.

Rather than driving the development by pure research orientations, it is experienced now thanks to Internet communication putting developers directly in contact with users that both in industry and research centers this kind of short time feed back loops are beneficial even in advanced research fields. This model is primarily applied in IT currently where changes are faster and justifies the shortening of the innovation process. Partners in this project are already applying these methods in their products and services using tools such as wish lists, open documentation servers, updated frequently asked questions, discussion lists, forums, chats, etc. Furthermore partners of this project experienced this model in MPEG 7 where the community of participants agreed to validate and experiment only real applications involving users at the early stage.

The workplan of the project takes this objective into account in the early phase by first committing the management to design this methodology with the necessary tools. Then by creating user groups participating in the first specifications and by involving these users in a re-specification phase in the middle of the project and for the final validation of the 3 applications.

Content oriented rationale of the technical workpackages (objective 2)

There are three key assumptions in this project which build the rationale of the technological part of the workplan :

Reusable modules
Quite often technologies are adjusted for a specific application and cannot be extended beyond. In Semantic HIFI, several critical technologies are addressed thanks to the partner diversity. Their first challenge will be to provide open modules (mainly in the Basic research WP 2 to 5) for easy collaboration and for reuse by several integrating companies in different applications:

? Extraction technologies as the heavier most critical starting point especially regarding the production of segmentation and source separation modules,
? Statistical indexing and encoding as a promising way to come up to classification modules and high level description of contents for intelligent browsing and personalization,
? Real-time and sound control technologies to be packaged in open modules or plug-ins for industry standards,
? 3D audio and rendering technologies which could be integrated in several equipments for a variety of audio formats and set-ups,

? Database and Networking technologies packaged in middleware modules for building sharing systems and servers since the largest content repository is formed by the Web to which the user can connect, browse and share,
? Audio processing and mixing technologies reusable in several music composition contexts.

Music descriptors based on objective features
One key assumption of the project is that high level descriptors of audio such as music descriptors are the key of future successful content-based applications. For providing these descriptors with an objective and replicable method and thus being able to apply them to several applications, the high level descriptors must be based on low level audio features extracted automatically. Otherwise only manually entered meta-data or subjective descriptors will be used introducing more overload and inoperability when sharing information and not allowing smart processing of this content.

Knowledge consistency
In order to make the extracted audio features attractive (even if the latest DSP analysis methods have been implemented) for content based applications, they must be organized to come up to valuable retrievable entries close to human practices. For instance, latest information filtering methods based on constraint management introduce a certain degree of artificial intelligence with automatic music play list generation or user's profiling.

Creative Applications (objective 3)

Content-based applications are still to be developed. Users are waiting for simple and reliable automatic indexing tools and before this will happen they will continue to index manually their collections or not to index them at all. A third objective of Semantic HIFI is to develop pilot applications convincing for industrialization. This is done in 2 steps:

- development of Indexing, Browsing, Rendering and Performing modules as a first result of corresponding R&D tasks,
- integration of the obtained modules into a unified middleware architecture (WP1.2), including the following final applications :
    - the HIFI System
    - the PC Authoring application
    - the Sharing online server

A typical user's scenario could be:
? input a given title and get proposed matches (adjusting simple parameters such as "timbre" or "energy") : Sound similarity and Query by Example
? find music files which share some high level features like music genre, melody or instrument played : Music similarity
? input a melody in the microphone and ask for the closest match : Query by humming
? choose two music pieces and ask for a choice of play lists between both (or music compilations): Constraint-based music selection
? link personal keywords (in any language) to a given sound/music. This personal indexing tool or learning system will then allow the user to input its personal keyword for the retrieval.
? user's profiling may also be automated by recording user's choices and mapping them with the extracted meta data. This allows to propose another way of searching for similarity taking the user's tastes as the guideline

Validation with artists and music consumers (objective 4)

Gathering artists and scientists to work together on music technologies and applications is the central aim of most of the partners involved in this project. This is why one of the particular objectives of this project workplan is to involve music fans but also artists and creators in the music domain. Sound designers, multimedia authors, sound engineers, composers, musicians and music students will be involved through user groups open to discussions on the Semantic HIFI issues. Artists won't be treated just as beta testers but as co-developers in close collaboration with the teams through creativity meetings, workshops and continuous Internet exchange in all Workpackages.

B6.2 - Implementation Plan

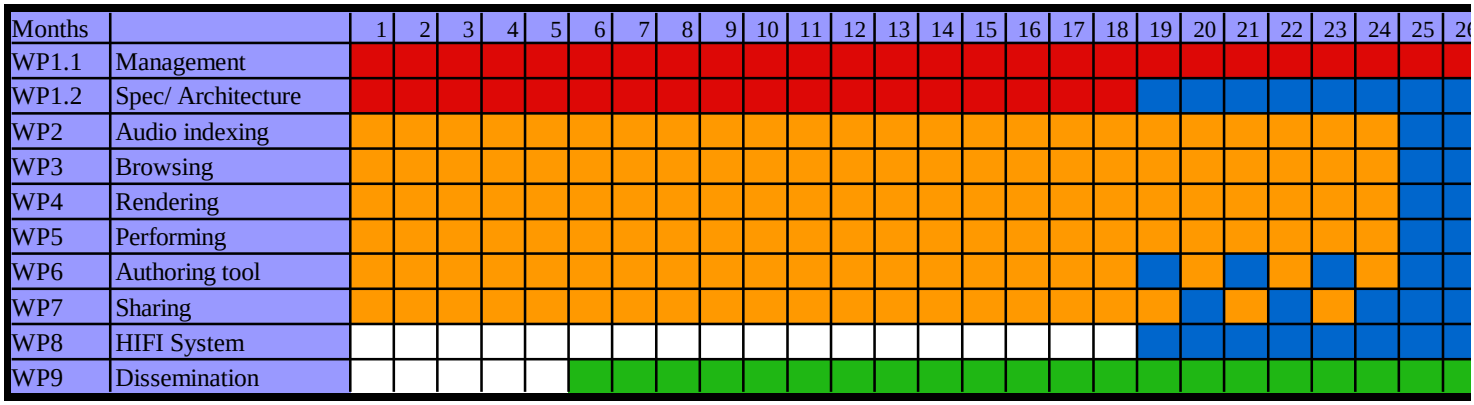
The project will start with a 5 month long specification phase, which will specify and coordinate the work to be done in all workpackages in terms of sub-workpackages (SubWPs), refined planning and partner roles, and issue a first version of functional specifications (Milestone M1). This phase will include a market and user requirement study coordinated by application WP leaders (WP 6&8). In parallel, R&D activity in all functional WP (2 to 6) will have been started. They will provide intermediate results as functional demos (M2), extraction modules (M3), first functional prototypes (M4) and final functional prototypes (M6). All these prototypes and demos will be developed as independent applications and will aim at implementing target applications features (including data representation and management, analysis and extraction algorithms, and user interfaces) in specific software environments adapted for prototyping. User feedback will be performed from M4 to M5.
On the basis of M2 to M4 prototype delivery and user feedback, a re-specification phase will start after M3, and end up in M5 with the delivery of the final technical and functional specification concerning target applications (WP 6 to 8).
After that specification (M5), the development of final applications will start by implementing the technical architecture required for the 3 target applications (WP6 to 8) which will be available at M6, as well as user interface mockups.
The next phase between M6 and M7 will be dedicated to the integration of functions issued from WP2 to WP6 into target application. It will include in particular the technical adaptation of functional modules in terms of external interfaces (APIs, data structures), implementation issues (tradeoffs between functionality and hardware limitations) and specific user interfaces. This phase will end up at M7 with the final delivery of functional modules from WP2 to 6.
The next phase will concern the final application developments, which will end up at M8.
M9 marks the project end with the delivery of the final dissemination report and implementation plan.

In addition the implementation plan uses quality control mechanisms agreed during specifications:

? Number of projects and initiatives that will use the solution proposed or part of it in the close future, or even during the last phase of the project
? Number of attendees that will participate to the public meetings and user groups in which the project and corresponding final and partial results will be presented
? Number of articles published on national and international conferences by project partners on the results produced by the project including proposals submitted in the standardization committees processes.
? Number of articles published on national and international journals and magazines in the field of both music and computer science by project partners on the results produced by the project.
? and international journals and magazines in the field of audio, broadcast, music and computer science by project partners on the results produced by the project.

B6.3 Work Planning (Gantt chart)

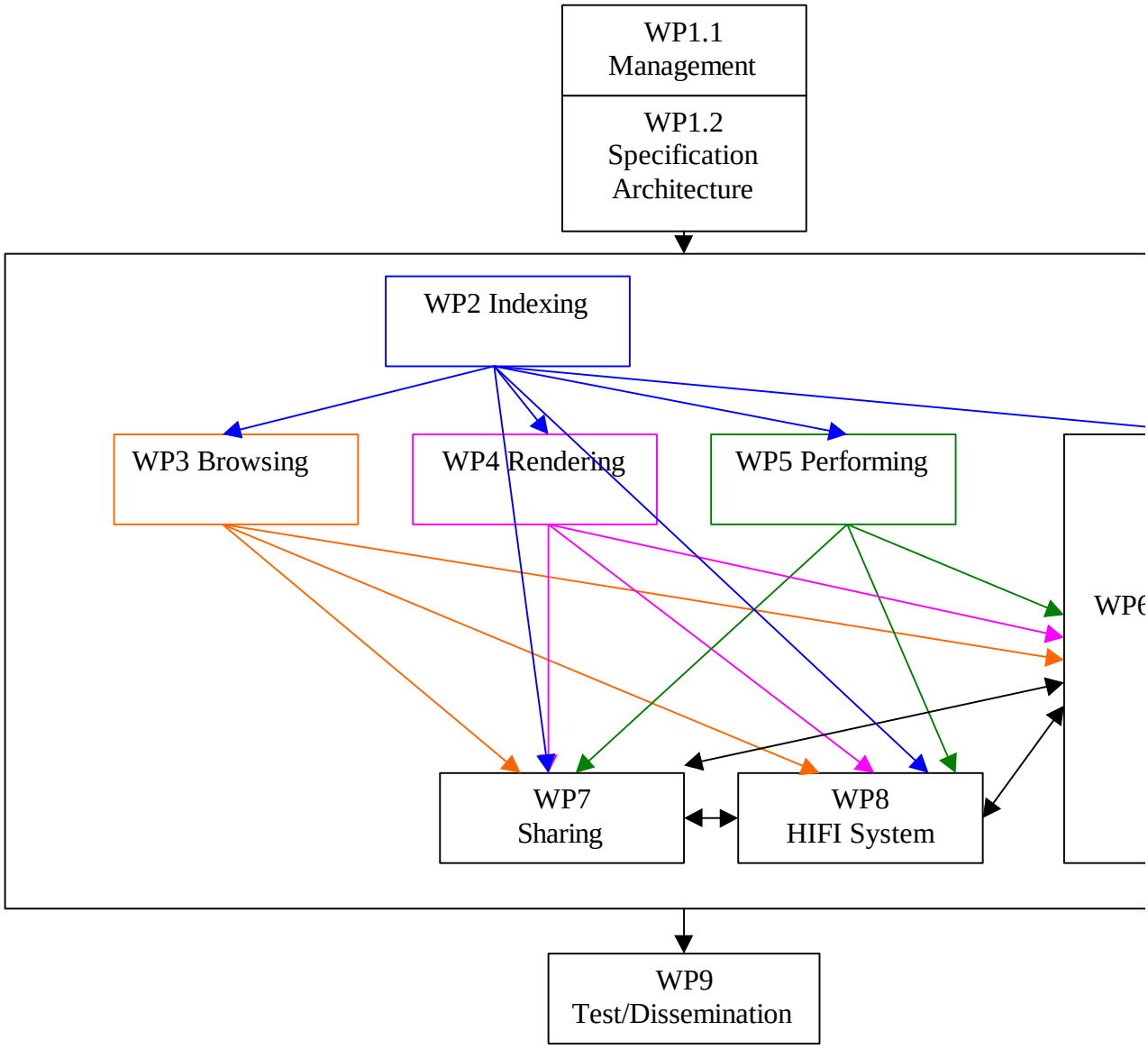
Planning of the WP M1 M2 M3 M4 M5 M6



Milestones

M1	SPECIFICATION
M2	FIRST FUNCTIONAL DEMOS
M3	FIRST EXTRACTION MODULES
M4	FIRST FUNCTIONAL PROTOTYPES
M5	FINAL SPECIFICATION
M6	FINAL FUNCTIONAL PROTOTYPES - ARCHITECTURE
M7	FINAL FUNCTIONAL MODULES
M8	FINAL APPLICATIONS
M9	VALIDATION/DISSEMINATION RESULTS

B6.4 Pert Diagram



B6.5 - Detailed Workplan

Workpackage list (full duration of project)

Work-package No	Workpackage title	Lead contractor No	Person - months	Start month	End month	Deliv-erable No
WP1.1	Management	Ircam	35,18	1	36	D1.1.1, D1.1.2, D1.1.3,
WP1.2	Specifications/architecture	Ircam	44	1 19	18 30	D1.2.1 D1.2.2 D1.2.3 D1.2.4
WP2	Audio indexing	Ircam	50	1 25	24 30	D2.1, D2.2 D2.3
WP3	Browsing	Sony CSL	42	1 25	24 30	D3.1, 2, 3, 4 D3.2, D3.3 D3.4, D3.5
WP4	Rendering	Ircam	51	1 25	24 30	D4.1, D4.2 D4.3, D4.4
WP5	Performing	UPF	68	1 25	24 30	D5.1, D5.2 D5.3, D5.4
WP6	Authoring	NI	56	1 19 25	18 24 35	D6.1, D6.2 D6.3, D6.4 D6.5, D6.6 D6.7
WP7	Sharing	FhG	29	1 19 25	18 24 35	D7.1, D7.2 D7.3, D7.4 D7.5, D7.6 D7.7
WP8	HIFI System	Sony NSCE	39	18	35	D8.1, D8.2 D8.3
WP9	Dissemination	Ircam	15	6	36	D9.1, D9.2 D9.3, D9.4
	TOTAL		430,18			

Deliverables list (full duration of project)

Deliverable No	Deliverable title	Delivery date	Responsible Partner	Nature	Dissemination level
D1.1.1	Management and quality control methodology	05	IRCAM	R	CO
D1.1.2	Internal Web site	05	IRCAM	R	CO
D1.2.1	Functional specification and refined workplan	05	IRCAM	R	CO
D6.1	Market and user requirement study for the Authoring application	05	NI	R	CO
D8.1	Market and user requirement study for the HIFI System	05	SONY NSCE	R	CO
D9.1	Public Web site	05	IRCAM	O	PU
D9.2	User feed back reports	05	IRCAM	R	CO
D3.1.1	Concept and specification for melody association application	06	SONY CSL	R	CO
D7.1	Concept and architecture for messaging and peer-to-peer-system	06	FHG	R	CO
D9.3	Progress reports on Dissemination	06	IRCAM	R	PU
D3.2	First Browsing functional demos	09	SONY CSL	P	RE
D4.1	First Rendering functional demos	09	IRCAM	P	RE
D5.1	First Performing functional demos	09	UPF	P	RE
D6.2	First Authoring functional demos	09	NI	P	RE
D2.1	First indexing modules and associated data models	12	IRCAM	P	CO
D7.2	Software documentation for Internet accessible web-service and right management	12	FHG	R	PU
D9.3	Progress reports on Dissemination	12	IRCAM	R	PU
D3.3	First Browsing functional prototypes	15	SONY CSL	P	RE
D4.2	First Rendering functional prototypes	15	IRCAM	P	RE
D5.2	First Performing functional prototypes	15	UPF	P	RE
D6.3	First Authoring functional prototypes	15	NI	P	RE
D1.2.2	Software and documentation for Internet accessible metadata look-up database used in WP 6, 7 and 8	18	FHG	R	CO
D1.2.3	Final specification	18	IRCAM	R	CO
D3.1.2	Software and documentation for internet accessible melody look-up database	18	FHG	R	CO
D9.2	User feed back reports	18	IRCAM	R	CO
D9.3	Progress reports on Dissemination	18	IRCAM	R	PU
D2.2	Indexing modules and associated data models (descriptor schemes)	24	IRCAM	P	CO
D3.4	Final Browsing functional prototypes with documentation	24	SONY CSL	P	RE
D4.3	Final Rendering functional prototypes with documentation	24	IRCAM	P	RE
D5.3	Final Performing functional prototypes with documentation	24	UPF	P	RE
D6.4	Final Authoring functional prototypes with documentation	24	NI	P	RE

D6.5	Implementation of the Authoring application architecture	24	NI	P	CO
D7.3	Software and documentation for embeddable P2P client	24	FHG	P	CO
D8.2	HIFI System architecture implementation	24	SONY NSCE	P	CO
D9.3	Progress reports on Dissemination	24	IRCAM	R	PU
D1.2.4	Software and documentation for embedded identification and metadata database management used in WP 6, 7 and 8	30	FHG	P	CO
D1.2.5	Software integration support	30	IRCAM	O	CO
D2.3	Final indexing modules optimized and adapted for target applications	30	IRCAM	P	CO
D3.1.3	Software and documentation for embedded melody recognition	30	FHG	P	CO
D3.5	Final Browsing functional modules adapted to target applications	30	SONY CSL	P	CO
D4.4	Final Rendering functional modules adapted to target applications	30	IRCAM	P	CO
D5.4	Final Performing functional modules adapted to target applications	30	UPF	P	CO
D6.6	Final Authoring functional modules adapted to target applications	30	NI	P	CO
D7.4	Integration of metadata look-up database	30	FHG	R	CO
D9.3	Progress reports on Dissemination	30	IRCAM	R	PU
D3.1.4	Software integration support and test support	35	SONY CSL	O	CO
D7.5	Software integration support and test support	35	FHG	O	CO
D7.6	Distribution of Mix files from NI Traktor DJ environment	35	NI	O	PU
D7.7	Online Database of sharable data and metadata	35	FHG,IRCAM	P	PU
D6.7	Final Authoring application with documentation	35	NI	R	RE
D8.3	Final HIFI System with documentation	35	SONY NSCE	R	RE
D9.2	User feed back reports	30	IRCAM	R	CO
D1.1.3	Final Public Report (FPR)	36	IRCAM	R	PU
D9.4	Final report on dissemination	36	IRCAM	R	PU

Workpackage description (full duration of project)

Workpackage number	1.1 Management		**Start date or starting event:**		Month 1		
Participant id	**IRCAM**	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE
Person-months per participant:	**29,18**	1	1	1	1	1	1

Objectives

Coordinate partners' effort in order to reach the project objectives with the minimum time loss, good quality and partners satisfaction involving user feed-back loops methodology and Quality Insurance with online workplan monitoring

Description of work

Overall management is ensured by the coordinator and by WP leaders in proportion with their involvement in WPs. IRCAM coordinator responsibilities include:
- liaison between the EC and the partners
- administration, e.g., progress reports and management reports for EC
- collection of Cost Statements every 6 months
- preparation of the reviews
- information flow among partners, management of related documentation
- management of project Web site (public and private)
- quality control over all activities and deliverables via specific tools.
- user feedback loops methodology used for achieving the best usability of all the project results. This last process will be organized with each WP leader and by the coordinator. It Is based on regular assessment by user groups set up by each integration partners. This feedback loop methodology is linked with a quality assurance method relying on continuous monitoring and upgrade of the workplan accessible online. Technical meetings will be planned according to the needs of the detailed workplan. A set of yardsticks will be used for controlling the project evolution. To this end, a detailed list of points (sub phases, functionalities) to be reached during each WP will be defined. The reaching rate of these functionalities and points will be measured.

Partner roles

Ircam-Coordination : project management, methodology design and implementation
All partners: participation in management activities and meetings.

Deliverables

On a periodicity defined with the EC (every 6 or every 12 months) : a progress report (PR) and a cost statement (CS).
- D1.1.1 Management and quality control methodology (End of month 5)
- D1.1.2 Internal Web site (End of month 5)
- D1.1.3 Final Public Report (FPR) (End of month 36).

Milestones and expected result

- Milestone 1 : specifications, efficiency of coordination procedures.
- Milestone 9 : completion of the project according to the workplan and the budget.

Workpackage number	1.2 Specifications, Standards & Architecture			**Start date or starting event:**	Month 1		
Participant id	**IRCAM**	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE
Person-months per participant:	**24**	3	5	1	1	5	5

Objectives

Ensure specification process along the project with particular emphasis on architecture choices, database management systems and networking and User feed back implementation (continuous feed-back).
Coordinate all technical aspects between research, development and application WPs.
Specify the global architecture and the data formats for exchanges between functional modules and applications.

Description of work

Implement the various specification phases, taking into account simutaneously:

- the project organization among partners
- user feedback
- research objectives and results
- targeted application functionalities

Develop specific middleware and data conversion modules (such as for mp3, AAC, MPEG4 files). Use of content description encoded in MPEG7. Monitoring of MPEG 21 Digital Item Adaptations (for instance for Undo functions), Automated association of metadata to format independent legacy audio and radio content using MPEG-7 technology. This will be used to identify audio items on e.g. mp3 CD's or to help the user find and purchase music that has been aired over radio.
- Set up of a small local metadata databases that will reside in the HIFI system itself, and a large database that is accessible through the internet connection and will be used to look up more rare audio material. Latter database will also be suited to validate P2P traffic as described in workpackage 7.

Partner roles

Ircam-Coordination : coordination of specifications
Ircam-SEL : technical architecture specification, design and development of specific data management and middleware modules
FHG : development of data management and conversion modules (mp3, AAC, MPEG4, MPEG 7).
Sony CSL, UPF, BGU : participation in specification and architecture design
NI, Sony NSCE : participation in specification and architecture design.

Deliverables

D1.2.1 Functional specification and refined workplan (End of month 5)
D1.2.2 Software and documentation for Internet accessible metadata look-up database used in WP 6, 7 and 8 (End of month 18)
D1.2.3 Final specification (End of month 18)
D1.2.4 Software and documentation for embedded identification and metadata database management used in WP 6, 7 and 8(End of month 30)
D1.2.5 Software integration support (End of month 30)

Milestones and expected result

M1 (first specification with broad standards selection), M5 (re-specification with users), M7 (final middleware modules with selected standards)

Workpackage number	2. Indexing		**Start date or starting event:**		Month 1		
Participant id	**IRCAM**	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE
Person-months per participant:	**26**	8	3	0	10	3	0

Objectives

Conduct basic research on algorithms for signal extraction, automatic indexing, segmentation, audio summary and source separation.

Description of work

Classification, indexing and management for audio files stored over the local hard disk or on external storages (Audio CDs indexed or converted in files, CDRoms, DVDs) and Internet. Development of excerpts management :

- automatic and manual segmentation, textual/iconic labels for each segment
- Extraction of high-level music descriptors from audio signal :
  - navigation into a music piece (music representation, audio summary)
  - tempo/phase recognition.

Partner roles :

Ircam-AS : WP leader, high-level music metadata extraction modules and audio summary for the Browsing WP.

Sony-CSL : Systematic extraction of high level music descriptors (global timbre energy descriptors, live/studio discrimination, etc.) as well as design of an optimized architecture for computing these descriptors for the corresponding browsing functions in WP 3.

BGU : segmentation , source separation and remixing

NI : tempo/phase recognition

Deliverables

D2.1 : First indexing modules and associated data models (descriptor schemes) (M3 – month 12)

D2.2 : Indexing modules and associated data models (descriptor schemes) (M6 – month 24)

D2.3 : Final indexing modules optimized and adapted for target applications (M7 – month 30)

Milestones and expected result

M3 : First version of indexing modules (month 12)

M6 : Indexing modules and associated data models (month 24)

M7 : Final indexing modules optimized and adapted for target applications (month 30)

Workpackage number	3. Browsing		**Start date or starting event:**	Month 1			
Participant id	IRCAM	FhG	**Sony CSL**	UPF	BGU	NI	Sony NSCE
Person-months per participant:	9	18	**12**	0	0	3	0

Objectives

Design and implement an integrated browsing tool allowing all browsing scenarios in a home-based context where navigation time must be reduced and graphical interfaces must be simple.

Description of work

Design of technological modules for query by example (search by similarity with inputted excerpts, Sony-CSL), Query by Humming (FhG), Artists description, keywords, cultural attributes and play lists generation using metadata management and constraints (Sony CSL) and relevant properties for music mixing and manipulation (matched tempo, keys, mood,) (NI)
Design and implement an integrated system reusing all the browsing modules.

Partner roles :

Ircam : Modules for browsing using audio fingerprints, Graphical interfaces for innovative browsing
FhG : Query by humming functionality with robustness, recognition rate, search performance, metadata integration, database scalability on :

- pitch estimation
- segmentation by determination of tone onset, tone offset, tone duration
- database search tolerant to inaccurate sung input

Sony CSL : WP leader, high-level music descriptions, similarity computation, play-list generation, design and development of prototypes, composition with segments under musical rules.
NI : Browsing by contextual and mixing relevant properties

Deliverables

D3.1.1 Concept and specification for melody association application (End of month 6)
D3.1.2 Software and documentation for internet accessible melody look-up database (End of month 18)
D3.1.3 Software and documentation for embedded melody recognition (End of month 30)
D3.1.4 Software integration support and test support (End of month 36)
D3.2 : First Browsing functional demos (M2 – month 9)
D3.3 : First Browsing functional prototypes (M4 – month 15)
D3.4 : Final Browsing functional prototypes with documentation (M6 – month 24)
D3.5 : Final Browsing functional modules adapted to target applications (M7 – month 30)

Milestones and expected result

M2 : First functional demos (month 9)
M4 : First functional prototype (month 15)
M6 : Final prototype (month 24)
M7 : Final functional modules adapted to target applications (month 30)

Workpackage number	4. Rendering	**Start date or starting event:**		Month 1			
Participant id	**IRCAM**	FhG	Sony CSL	UPF	BGU	NI	Sony NSCE
Person-months per participant:	**36**	2	12	0	1	0	0

Objectives

Bring simple and intuitive interface for an optimum audio rendering quality with real time spatial mixing features through high-level control in home context.

Description of work

3D audio rendering

- Multi input format compatibility (stereo, DVD 5.1, MP4, SACD, etc.)
- Output compatibility with multiple rendering set-ups : multi-loudspeakers, stereo (with transaural option), Binaural (with an online database of equalization data for a set of headphones).
- Automatic compensation of context for several loudspeakers set-ups
- Application of a given acoustic quality by modification of the recording acoustic quality according to the detected segments
- Use of perceptual parameters for acoustic quality modification

Polyphony Management

- Use of standard DVD tracks (up to 5) for polyphonic decomposition, listening track by track and remixing
- Re-mixing for certain polyphonic music from automatic source separation (BGU Virtual Mixer)
- Navigation into polyphony using constraints (MusicSpace, Sony CSL)

Browsing within music pieces / Interfaces and representation

- Graphical interfaces for navigation into the music piece (Ircam Hypermedia Studio)
- Synchronization and manipulation of sound associated with images, videos, MPEG4 files, …

Partner roles

Ircam Room Acoustics: WP leader, research on 3D audio simulation, context compensation and rendering modules, and design and development of the prototypes.
Ircam Hypermedia Studio : Navigation interfaces within music pieces for Computer Assisted Listening
FhG : MPEG4 coders
Sony-CSL : Tool for performing human-centered real time, on-the-fly music mixing
BGU : Virtual Mixer features

Deliverables

D4.1 : First Rendering functional demos (M2 – month 9)
D4.2 : First Rendering functional prototypes (M4 – month 15)
D4.3 : Final Rendering functional prototypes with documentation (M6 – month 24)
D4.4 : Final Rendering functional modules adapted to target applications (M7 – month 30)

Milestones and expected result

M2 : first demos (month 9)
M4 : first prototype (month 15)
M6 : final prototype with documentation (month 24)
M7 : final functional modules adapted to target applications (month 30)

Workpackage number	5. Performing		**Start date or starting event:**		Month 1		
Participant id	IRCAM	FhG	Sony CSL	**UPF**	BGU	NI	Sony NSCE
Person-months per participant:	17	0	12	**39**	0	0	0

Objectives

Performing music while listening with simple "instruments", home devices and voice as an addition to the played music, in an edutainment context (instrumental performance learning) or in a creative context. Real-time transformations of the music by using different input controls. Real-time interaction and performance with the music collection using Voice or "conducting instruments" : tempo variations by time stretching in real time using existing control devices (TV command, microphone, Game joystick).

Description of work

5.1 Real-time analysis of control signals
Analysis of different controls such as: hand tapping, voice, conductor gesture, …
- Hand tapping: Rhythm pattern and tempo detection (UPF)
- Voice: Singing Voice Specific Analysis. Systematic Extraction of vocal descriptors (vocal tract and excitation related) (UPF)
- Conductor gesture: Gesture recognition. (UPF) Tracing of objects. (UPF)
- Instrument Score follower for automatic performer accompaniment (IRCAM)

5.2 Real-time music transformations
Music transformations based on the real-time controls. The songs on the personal music collection are processed in real-time.
- Variable time stretching (UPF)
- Transposition with timbre preservation (UPF)
- Groove transformation: Swing modification (UPF)

5.3 Real-time voice transformations
Karaoke-like singing in which the user's voice is transformed in real-time
- Voice Excitation based Transformations (UPF)
- Vocal Tract based Transformations (UPF)
- Voice as a Real-Time midi controller: Voice2MIDI (UPF)

5.4 Performing on a simple keyboard
In this part Sony CSL will implement style and performing rules in a simple keyboard able to follow and continue the user play according to simple style constraints.

Partner roles :
UPF (WP leader) : design and development of audio analysis and transformation algorithms and prototype applications
Ircam-RTA : design and development of score following and audio analysis and processing algorithms.
Sony CSL : Interactive systems for exploiting music database for education and performance

Deliverables

D5.1 : First Performing functional demos (M2 – month 9)
D5.2 : First Performing functional prototypes (M4 – month 15)
D5.3 : Final Performing functional prototypes with documentation (M6 – month 24)
D5.4 : Final Performing functional modules adapted to target applications (M7 – month 30)

Milestones and expected result

M2 : first demos (month 9)
M4 : first prototype (month 15)
M6 : final prototype with documentation (month 24)
M7 : final functional modules adapted to target applications (month 30)

Workpackage number	6. Authoring		**Start date or starting event:**		Month 1		
Participant id	IRCAM	FhG	Sony CSL	UPF	BGU	**NI**	Sony NSCE
Person-months per participant:	17	0	0	6	0	**33**	0

Objectives

Development of a PC based Authoring tool for preparation of music materials to be played, browsed, performed and shared on the HIFI system.

Description of work

Development of the whole Authoring tool application, including

- development of a dedicated software architecture
- adaptation and integration of functional modules from WP 2, 3, 4 & 5
- development of specific functional modules, including:
    - real-time processing
        - transposition of sources (Ircam)
        - interpolation from one instrument to another (Ircam)
        - composition of segments: real-time mixing (NI)
    - recording of all filtering and processing parameters for non-real time authoring (NI)
    - morphing using content information (Ircam)
    - advanced remixing functionality (loop processing) (NI)
    - composition tools: sequencing, mixing (NI)
    - graphical mix-editing (NI)
    - music analysis and editorial tools (text composition, synchronization of media) (Ircam)
    - publication of play lists (NI)
- development of dedicated user interfaces and menus, user tests, documentation (NI)

Partner roles:

NI : WP leader, technical development and integration, development of specific functional modules (real time mixing, non-real time editing, loop processing), user tests, documentation

IRCAM : development of specific functional modules (real-time and non real time audio processing functions) and adaptations of Indexing (WP2) and Rendering (WP4) developments according to implementation constraints.

UPF, SONY-CSL : adaptations of Browsing and Performing developments according to implementation constraints.

FhG : adapted implementation of audio data and metadata management modules – coordination with Sharing server development.

Deliverables

D6.1 : Market and user requirement study for the Authoring application (M1 – month 5)
D6.2 : First Authoring functional demos (M2 – month 9)
D6.3 : First Authoring functional prototypes (M4 – month 15)
D6.4 : Final Authoring functional prototypes with documentation (M6 – month 24)
D6.5 : Implementation of the Authoring application architecture (M6 – month 24)
D6.6 : Final Authoring functional modules adapted to target applications (M7 – month 30)
D6.7 : Final Authoring application with documentation (M8 – month 35)

Milestones and expected result

M1 : market and user requirement study (month 5)
M2 : first demos (month 9)
M4 : first prototype (month 15)

M6 : final prototype with documentation and target architecture implementation (month 24)
M7 : final functional modules adapted to target applications (month 30)
M8 : Final application with documentation (month 35)

Workpackage number	7. Sharing System		**Start date or starting event:**			Month 1	
Participant id	IRCAM	**FhG**	Sony CSL	UPF	BGU	NI	Sony NSCE
Person-months per participant:	12	**13**	0	0	0	1	3

Objectives

Develop functionality to share (peer-to-peer) personal data over the Internet. Add messaging capability to the system. Build up necessary infrastructure for legal business.

Description of work

Peer-to-peer, file sharing and messaging

- tight integration with file sharing and messaging systems (net protocol, metadata publication)
- interoperability with mobile devices (mobile phones)
- share of metadata (play lists according to semantic paths, associated data, annotations, editorial, music analysis) to be applied on data locally
- share of data (provided copyrights and DRM are managed) like files, segments, excerpts, sequences of titles, remix or processed sessions, 3D audio remix sessions, MPEG4 composed scenes, etc.
- needed central infrastructure (web-services) for legal file sharing (accounting server with audio-ID)

The system will be developed as an online server, accessible through standard IP technologies. Two compatible and complementary versions of the server will be developed:

- one by FhG as a reference metadata repository for the HIFI systems
- one by IRCAM, as a reference of an online music database (Ircam Archives and Studio Online) with a metadata database associated to audio and music material produced at IRCAM.

Partner roles:

FhG : WP leader

Ircam-SEL: participation in design (data structures, DBMS, middleware) and development of a sharing server online at Ircam.

NI : Coordination with Authoring tool development

Sony CSL/Sony NSCE : Adaptation of sharing system to HIFI and existing Sony home network architecture.

Deliverables

D7.1 Concept and architecture for messaging and peer-to-peer-system (End of month 6)
D7.2 Software documentation for Internet accessible web-service and right management (End of month 12)
D7.3 Software and documentation for embeddable P2P client (End of month 24)
D7.4 Integration of metadata look-up database (End of month 30)
D7.5 Software integration support and test support (End of month 35)
D7.6 Distribution of Mix files from NI Traktor DJ environment (End of month 35)
D7.7 Online Database of sharable data and metadata (Ircam Studio Online and Ircam Music Archives) (End of month 35)

Milestones and expected result

M6 : Implementation of the target architecture for the Sharing server (month 24)
M8 : Final Sharing server implementation with documentation (month 30)

Workpackage number	8. HIFI System		**Start date or starting event:**		Month 19		
Participant id	IRCAM	FhG	Sony CSL	UPF	BGU	NI	**Sony NSCE**
Person-months per participant:	4	1	8	1	0	1	**24**

Objectives

Development of the HIFI System (personal Intelligent listening system), user tests and documentation.

Description of work

Development of the whole HIFI system application, including :

- identification of transferable modules and features,
- development of a dedicated software architecture and runtime,
- optimization and integration of functional modules from WP 1.2, 2,3, 4& 5,
- development of dedicated user interfaces and menus,
- user tests and documentation.

Partner roles:

Sony NSCE : WP leader, technical development and integration, user tests, documentation

Ircam : Direct compatibility with online databases

IRCAM, UPF, SONY CSL : adaptation of Browsing, Rendering and Performing developments according to implementation constraints.

FhG ; adapted implementation of audio data management modules, coordination with Sharing server development

NI : coordination with Authoring application development

Deliverables

D8.1 : Market and user requirement study for the HIFI System (M1 – month 5)
D8.2 : HIFI System architecture implementation (M6 – month 24)
D8.3 : Final HIFI System with documentation (M8 – month 36)

Milestones and expected result

M6 : First version on PC emulation, M7 : Integration of research modules, M8 : Final version

Workpackage number	9. Validation & Dissemination		**Start date or starting event:**		Month 6		
Participant id	**IRCAM**	FhG	Sony CSL	UPF	BGU	NI	Sony NCSE
Person-months per participant:	**3**	1	3	1	1	4	2

Objectives

Performed focused validation both for technical and usage aspects.
Provide expert users feed back and final user testing sessions.
Ensure the dissemination of the project results to all targets : scientists, professionals, consumers.
Participation and costs related to Standardization are covered in WP 1.2

Description of work

Following user's feed back loop methodology implemented with the management in WP1.1, working groups will be organized with users (Internal beta testers for the Scientific modules and selected external users for the applications (Forum Ircam, NI and Sony users).

Work plan on scientific conferences, training workshops, fairs and user group meetings including at MIDEM, IBC, NAB, AES, NAMM, Frankfurt Messe).

Partner roles

Ircam Hypermedia Studio and Forum Ircam : WP leader (Public Web site test repository, Internet lists, validation tests, coverage of invited experts and beta testers)
Ircam, FhG, SonyCSL, BGU : scientific and public dissemination
FHG, Ircam, NI, Sony NSCE : professional dissemination, public advertising
Sony NSCE: world wide dissemination in professional fairs, elaboration of business and exploitation plans, agreements on licensing with other partners.
Organization of user's feed back input (Web site, lists and validation tests).

Deliverables

D9.1 : Public Web site (M1 – month 5)
D9.2 User feed back reports (End of month 5, end of month 18, end of month 35)
D9.3 : Progress reports on Dissemination (Every 6 months)
D9.4 : Final report on dissemination (M9 – month 36)

Milestones and expected result

M1 : Public Web site (month 5)
M9 : final dissemination report (month 36)

B.7 Ethical issues

Since this project addresses communication and use of technologies widely used for illegal content access, it will be made clear in all documents that partners do not encourage illegal use of their technologies.

However the current context of illegal exchange of music material has an influence on the development of the project mainly in the Sharing WP since this system developed at the first place for metadata exchange can also be used for data. As stated in the introduction of this project profitable music service will probably have to balance freedom and comfort as planned in the project objectives.

Recent studies on Music sales show that illegal exchange of music does not affect directly music sales on traditional media. On the contrary one strong argument of this project is that the exchange of metadata will support music consumption through legal purchase, be the music copy protected or not in the future.

A. Proposers are requested to fill in the following table

Does your proposed research raise sensitive ethical questions related to:	**YES**	**NO**
Human beings		X
Human biological samples		X
Personal data (whether identified by name or not)	X	
Genetic information		X
Animals		X

B. Proposers are requested to confirm that the proposed research does not involve:

- ? Research activity aimed at Human cloning for reproductive purposes,
- ? Research activity intended to modify the genetic heritage of Human beings which could make such changes heritable
- ? Research activity intended to create Human embryos solely for the purpose of research or for the purpose of stem cell procurement, including by means of somatic cell nuclear transfer ;
- ? Research involving the use of Human embryos or embryonic stem cells with the exception of banked or isolated Human embryonic stem cells in culture

	YES	**NO**
Confirmation : the proposed research involves none of the issues listed in section B	X	