

**Cultural, Artistic and Scientific knowledge
for Preservation, Access and Retrieval**

**On the horizon:
preserving digital information
with CASPAR**

Protect digital information for your users and
for future generations:
Join the **CASPAR** Preservation User Community now!

www.casparpreserves.eu
info@casparpreserves.eu

“ How can digitally encoded information still be understood and used in the future when the software, systems and everyday knowledge will have changed? ”

On the horizon: preserving digital information with CASPAR

Why digital preservation?..... 1
 Introducing CASPAR 2
 The CASPAR roadmap..... 4
 The CASPAR solution 6
 A user-oriented approach: The CASPAR Community 8
 Governance 9

Content Index

Work partially supported by European Community under the Information Society Technologies (IST) programme of the 6th FP for RTD - project CASPAR contract IST-033572. The author is solely responsible for the content of this document. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.

Community Members

Cultural Organisations:

- ASTRA Museum, Romania
- Edison Studio, Italy
- Getty Research Institute, U.S.A.
- IKA Hordaland IKS, Norway
- Italian Group of the International Institute for Conservation, Italy
- Jewish National and University Library, Israel
- Koninklijke Bibliotheek, National Library of the Netherlands
- Library of Congress, U.S.A.
- Netherlands Institute for Sound and Vision
- Phonos Fundación Privada, Spain

Educational Institutions:

- Arts and Humanities Data Service, United Kingdom
- Brigham Young University, U.S.A.
- Ecole Polytechnique Fédérale de Lausanne, Switzerland
- The Hebrew University of Jerusalem, Israel
- University of Oxford - Image Bioinformatics Research Group, United Kingdom
- Università Cattolica del Sacro Cuore, Italy

Public Administrations:

- Alpineon RTD, Slovenia
- Arquivo Municipal de Lisboa (Câmara Municipal de Lisboa), Portugal
- Bundesarchiv, Germany
- Camera di Commercio di Roma, Italy
- COST, Belgium
- Fonoteca Nazionale Svizzera, Switzerland
- National Archives Service of Finland
- National Library of Finland

Research Centres:

- Center for Computer Assisted Research in the Humanities, Stanford U., U.S.A.
- CN2SV, France
- Günter Grass Stiftung, Germany
- Kinneret Limnological Laboratory, Israel
- Ludwig Boltzmann Institute Media.Art. Research, Austria
- Medicenter Sarajevo, Bosnia and Herzegovina
- National Institute for Research and Development in Informatics, Romania
- UKOLN, United Kingdom

Technological Providers:

- Data Archiving and Networked Services, Netherlands
- Endeavor Information Systems, Inc., U.S.A.
- Epistemica, Italy
- HELETEL LTD, Greece
- Multimedia Flug, Italy
- StorageSwitch, U.S.A.
- The Virtual Experience Company, United Kingdom
- YDreams, Portugal

Other organisations:

- Fondazione Rinascimento Digitale, Italy
- Intercultura Consult, Bulgaria
- KLEIO heritage consultants, Belgium
- Level Consulting, Italy
- Oil IT Journal, France
- Scientific Frontline, U.S.A.

Individual members:

- Christoph Becker, Austria
- Joel Cordesses, France
- Kellie Snow, United Kingdom
- Neil Beagrie, United Kingdom
- Nicolae Constantinescu, Romania
- Noemi Scarpato, Italy
- Peter Van Garderen, Netherlands
- Ronny Rooman, Belgium
- Sidney Zafran, U.S.A.
- Victoria McCargar Consulting, U.S.A.
- Victoria Boyd, United Kingdom

Why digital preservation?

Digital preservation has quickly become a pressing issue. The challenge posing our generation is how to maintain and safely preserve the immense volume of information that we have accumulated as a global community. This global predicament will have huge impacts on both individuals and businesses if it is not addressed properly. Digital information preservation requires a commitment to advancing the state of the art, combined with an understanding of the social underpinnings of digital information, and an ability to foresee future developments in the computer industry.

Digital information is an extremely valuable yet vulnerable resource that is easily mishandled, corrupted, deleted, and lost. Exponential rates of technological evolution also render systems and files obsolete in just a matter of months. Our current global archives do not provide a stable database for the indefinite preservation of information in the future. As world markets change, the access to information becomes a fundamental societal asset. The systems that are currently employed provide at best a fragmented infrastructure unable to provide a secure and sustainable global archive. For some institutions, failure to comply with legal requirements on the preservation of records may have severe consequences.

Digital preservation refers to the management of digital information over time. Unlike the preservation of paper, the preservation of digital information demands ongoing attention. This constant input of effort, time, and money to handle rapid technological and organisational advance is considered the main stumbling block for preserving digital information beyond a couple of years. Indeed, while we are still able to read our written heritage from several thousand years ago, the digital information created merely a decade ago is in serious danger of being lost. Digital preservation can therefore be seen as the set of processes and activities that ensure the continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats.

Introducing CASPAR

CASPAR (Cultural, Artistic, and Scientific knowledge for Preservation, Access, and Retrieval) combines an advanced approach to storage and access of information with extensive scientific, cultural, and creative expertise. Designed as an implementation from the OAIS reference model (Open Archival Information System, ISO:14721:2002), CASPAR is an Integrated Project co-financed by the European Union within the Sixth Framework Programme. CASPAR also plans active collaboration and teaming with relevant digital preservation initiatives outside the EU, with National and International projects such as Chronopolis, INTERPARES and NARA (US National Archives and Records Administration) to bring a global dimension to its work.

An OAIS is an archive consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated user community. The information being maintained has been deemed to need indefinite preservation, even if the OAIS itself is not permanent. This preservation is concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. In this reference model there is a particular focus on digital information, both as the primary forms of information held and as supporting information for both digitally and physically archived materials.

The CASPAR mission is practical yet ambitious: Provide secure, reliable and cost-effective access for digitally encoded information for an indefinite time period. The motivation of the CASPAR project is to define the methodology and infrastructure for digital preservation in Europe. This objective will be pursued by following 4 guidelines.

- Building a preservation environment based on the OAIS reference model.
- Demonstrating its ability to handle the preservation of the digital resources of diverse user communities.
- Adapting and integrating current state of the art technology in digital preservation.
- Developing technological solutions aimed at sustaining expansion.

By implementing these guidelines with the appropriate scientific, cultural, and creative expertise coupled with commercial partners and world leaders, CASPAR will be instrumental in providing sustainable digital information preservation.

Today, many of CASPAR's users are not even born. The ability to grow and adapt in a constantly evolving technological environment is critical to its mission. This evolution comes in several forms: by maintaining current hardware and software systems, imitating existing systems, migrating to new systems, and "virtualising" systems by describing them - which will allow the archive system to operate with digital objects, not just files. By utilizing cutting edge technology and adapting to new trends in the market, CASPAR will set the standard of modern and personalized services.

One of CASPAR's most compelling features is the aspiration to be user orientated. CASPAR will not be a closed system; the aim will be to interoperate with as many different systems as possible and to be operated and re-implemented by others outside the CASPAR project. The ongoing and effective participation of the CASPAR Community will guarantee that the system fulfils the requirements for which it is designed to. The three areas to be evaluated by the overall CASPAR community will be: innovation, operation, and strategy. These divisions will be available for critical assessment from virtually anyone in the community, from system integrators, users, and even potential customers. By listening to the collaborative input from both clientele and system operators, CASPAR will provide active progress assessment.

CASPAR will focus on evolving with current trends in the market, as well as sophisticated representation from within the CASPAR community. Furthermore, CASPAR will comply with future Digital Rights management requirements, handle authenticity and provenance issues, and finally inform the brokering of services for information preservation. These anticipated areas of change will enable the existing framework to correspond with future trends and changes in the industry.

CASPAR Facts

- * CASPAR website: www.casparpreserves.eu
- * Project type: Integrated Project
- * Start date: 1 April 2006
- * Duration: 42 months
- * EU Funding: Euro 8 800 000
- * Projected Budget: Euro 16 000 000
- * Number of partners: 17
- * OAIS Reference Model: (ISO 14721)
- * Project co-ordinator: Council for the Central Laboratory of the Research Councils (CCLRC), UK

Methodology and Models: The CASPAR Roadmap

The core of CASPAR Methodology is the policy upon which the project was founded. The core of this policy is the commitment to the OAIS foundation, and the in-depth analysis of the framework to guide integrated and advanced digital preservation. The CASPAR methodology also strives to create the software platform that enables the building of services and applications that can be adapted to multiple areas.

Modelling of information objects will also comprise a fundamental role within the framework. By virtualising information to facilitate a description in order to process underlying digital data, CASPAR will be developing sequences for simple, complex, and high-demand objects – therefore the infrastructure will be highly efficient, stable and interoperable.

The fundamental outline for CASPAR is the OAIS reference model (ISO:14721:2002). OAIS is a very powerful framework, because it is both organisational and technical. OAIS describes all the functions of a digital repository: how digital objects can be prepared, submitted to an archive, stored for long periods, maintained, and retrieved as needed. This is the reason it can be used as a common thread to combine diverse levels of knowledge and information. OAIS does not prescribe specific technologies, archiving techniques, or types of content. However, in order to change OAIS recommendations and structures into a functioning preservation framework, the operational system should be refined by distinguishing what are the relevant knowledge and information components by designing them in terms of structures and tools, then integrating them into the OAIS organisation. This framework provides CASPAR with the discipline within which the components work together to ensure preservation. The overall CASPAR system will depend on a software framework that will be the technological basis for the implementation and testing of different testbeds.

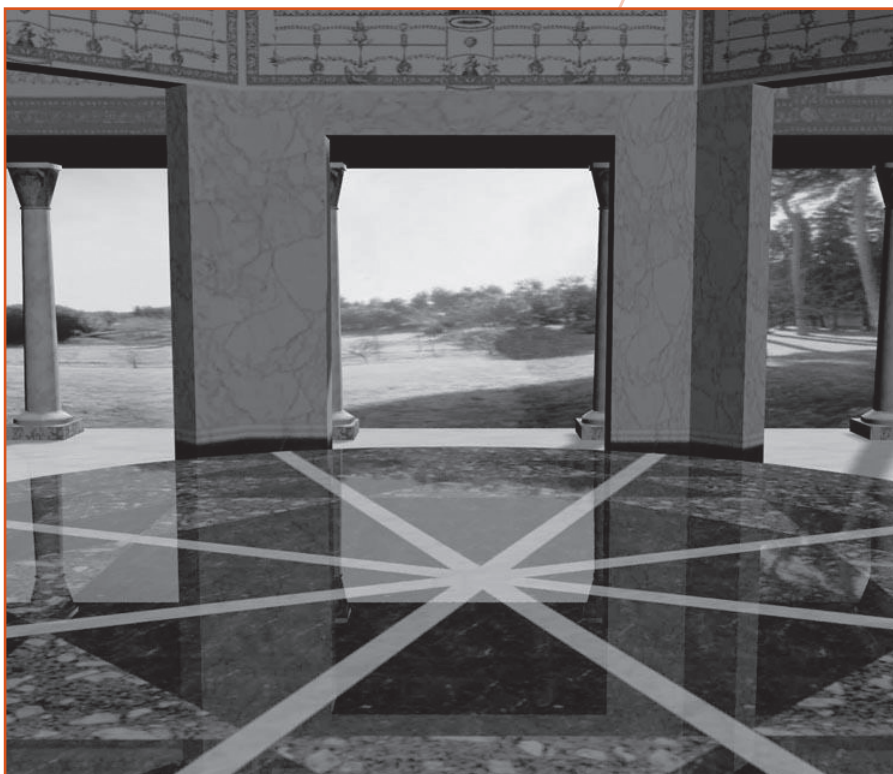
Standards, policies, strategies, guidance, and tools for a broad genre of users and industries will govern the CASPAR project. CASPAR will produce a coherent preservation methodology, an implementation of a support infrastructure, and a validation methodology with case studies in science, cultural heritage and performing arts, with replicable testbeds. CASPAR's mission com-

plements the 'Lisbon Strategy', which includes the aim of making Europe 'the most competitive and dynamic knowledge based economy in the world.' CASPAR will be a significant contributor to the enlargement of digital preservation and digital services in Europe and will help support the European body of shared digital scientific, cultural heritage and performing arts data.

The OAIS reference model

In 1990, the Consultative Committee for Space Data Systems (CCSDS) launched a cooperative arrangement with the International Organization for Standardization (ISO) to recommend solutions for data management issues. The CCSDS decided to create a unifying framework that established shared concepts and terminology associated with digital preservation. In 1995, the CCSDS convened to develop a reference model for an open archival information system (OAIS). The OAIS reference model was approved and published in January 2002 as international ISO standard 14721.

The reference model provides a framework, conceptual outline, and terminology for digital information preservation. The model provides a starting point for the implementation, because it characterizes the high level responsibilities, services, and information requirements that must be incorporated into the system. The capacity both to create and consume digital information has advanced exponentially, but the ability to manage this irreplaceable record has not; OAIS strives to produce a tangible environment to preserve encoded digital information and provide a coherent accessible archive for users in the long term future.



Domus Aurea, Rome – Italy
Copyright: © Advanced Computer Systems

The CASPAR solution

As the evolution of modern technology creates more effective services and components, CASPAR will promote the integration of both new hardware and software. In order to maximize efficiency, the existing information management shall be integrated and customised as far as possible into the CASPAR system. The means of testing and validation throughout CASPAR is generated by three interdisciplinary user communities intended to cross-evaluate the developed solutions. These testbeds allow CASPAR to cross reference diverse applications, and increase the wide range of diverse needs.


Cultural Data Preservation Testbed

This testbed focuses on the preservation of all data necessary to document, visualise and model heritage sites and will provide a valuable resource to assist conservation experts in restoring the associated site while keeping its original integrity. The documentation, visualisation, and modelling of natural and cultural heritage sites is a complex task that requires large amounts of data and information. The objective of this testbed will be to test the solutions developed throughout CASPAR, including virtual digital objects, spanning between processing and display.

6

Contemporary Art Testbed

This testbed is focused on the implementation for contemporary arts: contemporary music, performing arts, and other forms of technology-enhanced arts like video games. These objects also include dynamic interactive digital objects oriented towards presentation and replay. Contemporary music, as well as performing arts and video games produce very complex objects including specific hardware, instructions, and specific equipment. These electronic models often contain highly complex extensions such as interaction devices and instructions. Interaction devices include specific sensors, instruments, and equipment while instructions often include score, software components, and audio files. The objective of this testbed will be to study these specific extensions, their relationship with the generic models, procedures, and services, and to start the analysis and set-up of both a generic and specific infrastructure.

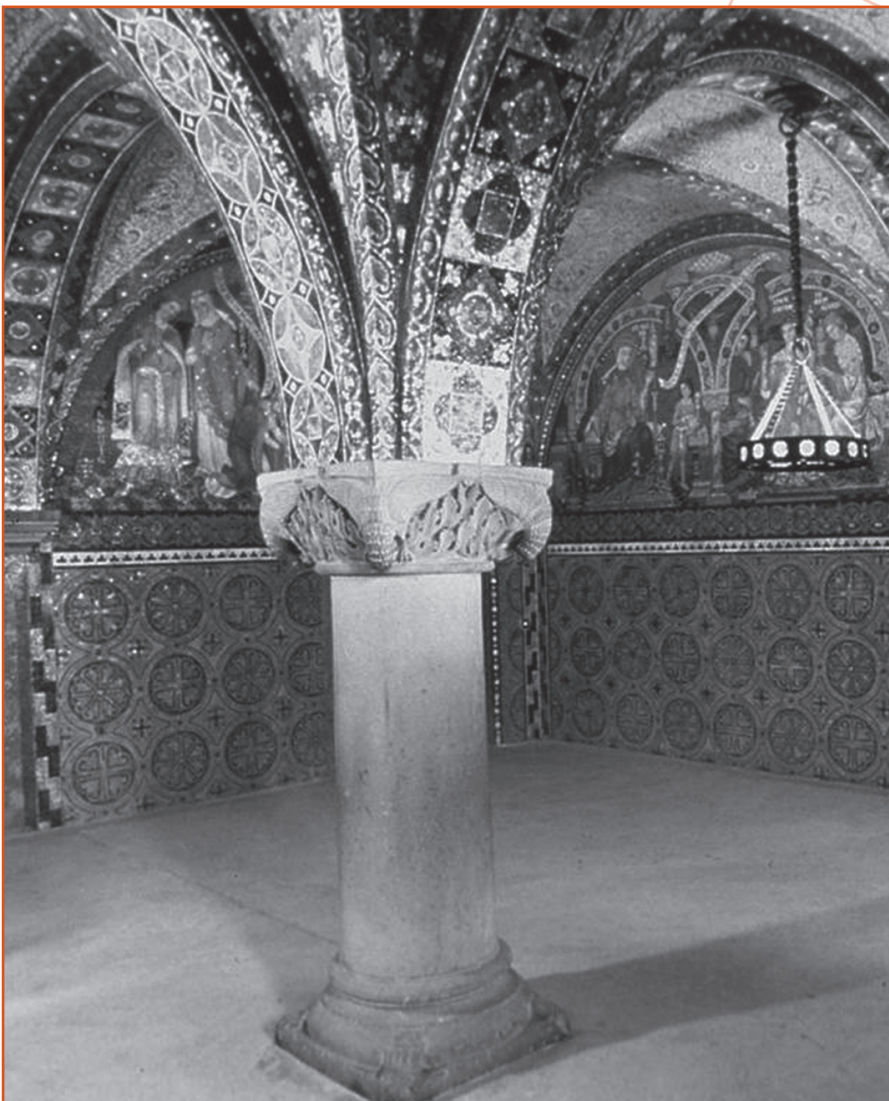


Science Data Preservation Testbed

This testbed will assess the Earth Science community requirements, develop the necessary specific services and prototype an Earth Observation science data preservation. Also, the preservation life cycle of scientific data and the preparation of specific tools that will allow visualisation and navigation of the complex metadata associated with science data complex objects. In terms of preservation, these are often very high volume, complex digital objects, oriented towards processing. The solutions being evaluated by these testbeds consist of 3 separate domains.

- Methodology, standards, and existing resources of the project
- Component level research and technology development.
- Integrated Framework of the project.

All services provided by these testbeds will be capable of working in a completely distributed environment and with a reliable solution even after migration to new systems or the reorganisation of resources used in the testbeds. This framework is intended to be used throughout the project and should constitute the foundation for an infrastructure even beyond the project timeframe.



Wartburg Castle – Germany

Copyright: © UNESCO / no photographer registered

A user-oriented approach: the CASPAR Community

Although CASPAR employs an intricate infrastructure designed to cross evaluate basic functionalities, it also is designed to incorporate exterior input from the steadily growing CASPAR Community. Membership to the User Community is free. The Community consists of a diverse pool of international organizations, individuals and projects who want to have a voice in protecting digital resources. Users range from universities and hospitals, to international institutions and research centres. Joining the CASPAR Community is possible by registering on the website www.caspar-preserves.eu. By registering as a member, you immediately gain access to a wealth of information and current events, as well as an opportunity to put your own personal touch by posting articles, events, journals, links, and presentations. Members also enjoy an opportunity to connect to this broad international community and share ideas on current trends and new technology.

Public Awareness: Reaching out to the Global Community

Focused events are being initiated internationally to highlight the critical importance of digital preservation to individuals. For example, personal and family data such as digital photographs and recordings have significantly diverse preservation demands, opposed to a picture album on a bookshelf. Obviously, digital information preservation is a social phenomenon which if not already, will affect us all in the short future.

The CASPAR Community is busy promoting social awareness among several different public sectors. At the National level, the CASPAR consortium has organisations from 7 countries, and several members are significant memory institutions. Therefore, the diverse Consortium panel is in a position to service more local concerns and interests by participating in local events.

Project governance

The CASPAR Consortium consists of a selection of industrial, academic, and cultural partners with proven track records. The Consortium's partners have a long experience in at least one of the following areas: preservation, knowledge management, access techniques, advanced storage technologies, software, exploitation, or content provision technologies. The Consortium's founders have long experience with digital preservation research projects, providing a solid base concerning the features and drawbacks of existing solutions. Furthermore, the Consortium aims at

- Maintaining a strong correlation with European and national research programmes by each EU Member State
- Guaranteeing an adequate level of manageability
- Including the main value chain actors in the digital preservation domain
- Providing an adequate level of dissemination and exploitation of results

The project is structured as follows:

IP Coordinator (IPC):

Dr. David Giaretta

Council for the Central Laboratory of the Research Councils (CCLRC)

www.cclrc.ac.uk

The management structure of CASPAR was specifically designed to distribute responsibility, while maintaining a constant flow of information. Project management is divided in several different sections, intended to maintain a balance between the Commission and Consortium, while simultaneously making sure that the project remains on course. The mediator between this balance is the IP Coordinator (IPC). The IPC is responsible for a variety of responsibilities including the handling of payments and maintenance of accounts, day-to-day operation, keeping partners informed of progress, and serving as the administrative liaison with

the Commission. The current IPC, Dr. David Giaretta has extensive credentials in digital preservation, including Chair of the CCSDS Panel, which originated the OAIS Reference Model.

The IPC is directly assisted by the IP Secretariat, who provides secretarial, administrative, financial, and legal support to the IPC. The IP Secretariat will be a permanent contact person who also supports projects participants. The IP Secretariat will be based in the CASPAR Project Office at CCLRC, and serves as project Secretary and archive.

The IPC is complemented by the Stream Directors, who will have a supervisory brief over all the Work packages in his/her stream in order to ensure coherence and consistency between work packages and interdisciplinary ideas between stream, work packages, and tasks. The membership will be reviewed by the General Assembly annually.

The three divisions combine to form the Executive IP Management (EPM) for the day-to-day project management for the entire Inte-

CASPAR Stream Directors

Seamus Ross

(University of Glasgow - Humanities Advanced Technology and Information Institute - www.hatii.arts.gla.ac.uk)

Stream 1: Methodology, Standards and Existing Resources

Carlo Meghini

(Consiglio Nazionale delle Ricerche – ISTI - isti.cnr.it/ResearchUnits/Labs/nmis-lab)

Stream 2: Component Level Research and Technology Development

Ugo Di Giammatteo

(Advanced Computer Systems - www.acsys.it/it/index.php)

Stream 3: CASPAR Framework)

Luigi Fusco

(European Space Agency – ESRIIN - www.esa.int/esaCP/index.html)

Stream 4: Testbeds, testing and validation

Silvia Boi

(Metaware - www.metaware.it)

Stream 5: Innovation-related activities, Training & Demonstration

David Giaretta

(Council for the Central Laboratory of the Research Councils, CCLRC - www.cclrc.ac.uk)

Stream 6: Administration, Scientific and Technical Project coordination and Management

grated Project. The EPM will define and periodically update the technical project management plan, and will clearly maintain the WP interfaces and dependencies. The EPM will also be involved in resolving conflicts and attending all meetings and events when they are needed. The EPM should consult the General Assembly on important issues and aim at achieving consensus. All project reports, financial transactions, and day-to-day operations will be overlooked by the EPM.

The External Review Committee (ERC)

The External Review Committee (ERC) advises the EPM on its scientific direction. It reviews the progress made on an annual basis and gives advice on the scientific aspects of the IP (e.g. new academic or technological achievements CASPAR should consider, new important trends, new societal developments the project should take into account, etc.). The ERC is also invited to propose ideas on generating new business and exploiting project results further. The IP Coordinator, with the agreement of the Commission, selects and nominates the members of the ERC among high-profile industry and/or academic representatives. The IPC chairs the ERC. The IP Stream Directors shall be represented at the meetings, and the Commission IP Officer may also be represented. Suggestions from the ERC may be forwarded to the Commission.

In addition to the annual meetings, members of the ERC will perform the independent external audit of the effectiveness CASPAR in terms of digital preservation, including

- evaluation of the methods of validation of the CASPAR methodology
- evaluation of the effectiveness of the Testbeds.

The reports from the ERC will be made available to the Commission.

The Exploitation Advisory Committee (EAC)

CASPAR will explore viable means to sustain itself and further its mission beyond the project timeframe by assessing possible industrial and commercial exploitation alternatives.

An Exploitation Advisory Committee (EAC) will be composed of high-profile business managers recruited from the Consortium partners. The EAC shall develop concrete proposals on how new business will be pursued and how exploitation should be organised from the project results. Any General Assembly (GA) member can propose EAC members, while the EPM decides on the composition of the EAC.

General Assembly (GA)

All partners of the Consortium are represented in the GA. The GA takes final decisions on the Consortium Agreement. The IP Coordinator will keep the GA informed about progress and achievements. The GA shall solve conflicts within the Consortium that cannot be resolved by the EPM or the IPC. The GA meets bi-annually, or if more than 30% of the GA members require it. It can take decisions by correspondence, and is regulated by the IPC.

The CASPAR Consortium

Council for the Central Laboratory of the Research Councils – United Kingdom

European Space Agency, ESRIIN - Italy

University of Glasgow, Humanities Advanced Technology and Information Institute - United Kingdom

UNESCO (United Nations Educational, Scientific, and Cultural Organization)

Advanced Computer Systems S.p.A. - Italy

@semantics S.r.l. - Italy

IBM Haifa Research Laboratory - Israel

Consiglio Nazionale delle Ricerche – Institute of Information Science and Technologies - Italy

Metaware S.p.A. - Italy

Institut National de l'Audiovisuel - France

University of Leeds, Interdisciplinary Centre for Scientific Research in Music - United Kingdom

Engineering – Ingegneria Informatica S.p.A. - Italy

Foundation for Research and Technology - Hellas - Greece

Centre National de la Recherche Scientifique - France

Institut de Recherche et Coordination Acoustique/Musique - France

International Centre for Art and New Technologies – Czech Republic

Work partially supported by European Community under the Information Society Technologies (IST) programme of the 6th FP for RTD - project CASPAR, contract IST-033572. The author is solely responsible for the content of this paper. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.