



Keeping knowledge alive:

## Tessella archiving and preservation research

**Digital preservation is an emerging discipline that allows information stored today to be read and understood in the distant future. The data may be documents, media, raw data from instruments, engineering or financial information, regulatory records or operational emails. As well as providing archiving solutions, Tessella is at the forefront of researching methods to allow this information to be available when required many years from now.**

### Challenges

The information saved today is usually saved in files with some metadata to identify and describe the content. Often files can be read on their own, for example, documents. Some information is saved as a number of related files, such as a website. Some is in complex binary formats for instance, a database. All may need to be accessed in the future.

To make sense of the files you have retained in your archive, you need a combination of application software running on a supported operation system sitting on running hardware. If any of these are missing, the objects will not be viewable. The trouble is that it is almost certain that in the distant future, one or probably all will no longer be available. To cope with this, many different strategies must be employed. Tessella has been a key part of some of these major innovations that will lead to standard

ways of preserving digital information as long as it is needed.

### Planets

Planets (Preservation and Long-term Access through Networked Services) is a four-year project co-funded by the European Union under the Sixth Framework Programme to address core digital preservation challenges. The primary goal for Planets is to build practical services and tools to help ensure long-term access to our digital cultural and scientific assets.

The strong Planets consortium, including three technology companies (Tessella, Microsoft and IBM), brings together expertise across Europe from national libraries and archives, and leading research universities.



Tessella has conducted research into the characterization process which is critical to archiving – how can you preserve digital information unless you know what you have? Tessella has also developed an extended registry of information of all the knowledge required to preserve data including file formats, software packages, hardware platforms and storage media. This is key to driving preservation policy.

### KEEP

KEEP (Keeping Emulation Environments Portable) is a 3-year, EU-funded project aimed at developing an Emulation Access Platform to enable accurate rendering of both static and dynamic digital objects: text, sound, and image files; multimedia documents, websites, databases, videogames, etc.

KEEP is striving to facilitate universal access to our cultural heritage by developing flexible tools for accessing and storing a wide range of digital objects.

Tessella is working on the KEEP framework and implementing an example emulator.

### Testbed

Tessella was a key technology provider to the Dutch Government's Digital Preservation Testbed project. This evaluated possible strategies for long-term preservation of born digital government records, leading to a set of recommendations to the Dutch Government on the creation, management and long-term preservation of key electronic record types.

### DIOSCURI Emulator

Tessella has worked with the Koninklijke Bibliotheek (The National Library of The Netherlands) and the Nationaal Archief (The Dutch National Archives) to develop Dioscuri, an x86 computer hardware emulator written in Java. It is designed by the digital preservation community to ensure documents and programs from the past can still be accessed in the future.

The Dioscuri emulator has two key features: it is durable and flexible. Because it is implemented

in Java, it can be ported to any computer platform which supports the Java Virtual Machine, without any extra effort. This reduces the risk that emulation will fail to work on a single architecture in the future, as it will continue to work on another architecture.

Dioscuri is flexible because it is completely component-based. Each hardware component is emulated by a software surrogate called a module. Combining several modules allows the user to configure any computer system, as long as these modules are compatible. New or upgraded modules can be added to the software library, giving the emulator the capability to run these.

### Preservation Tools

Tessella has worked with several leading memory institutions to develop the key tools now in use to deliver digital preservation technology. Although in widespread use now, these were radical developments at the time that moved both the technology and philosophy behind preservation forward.

Tessella has worked on the UK National Archives PRONOM system since its inception in 2001. This web-based repository provides information on file formats and the technical components (especially application software) needed to create or access files in such formats.

The latest release of PRONOM also includes the freely distributable Digital Record Object Identification (DROID) tool, to allow files in hundreds of file formats to be appropriately identified by detecting format-specific byte sequences.

Tessella has also developed a range of migration tools for various format transformations. This includes the handling of complex components such as websites where a number of files are interdependent and must be migrated in unison.

**Tessella plc** 26 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS, UK  
T: +44 (0)1235 555511 | F: +44 (0)1235 553301 | E: info@tessella.com

**Tessella Inc** 233 Needham Street, Suite 300, Newton, MA 02464, USA  
T: 1 617 454 1220 | E: info@tessella.com

**Tessella – successfully delivering IT and consulting services to world leaders in R&D, science and engineering.**

For decades, Tessella has been successfully delivering IT and consulting services to world leaders in R&D, science, and engineering. Through the application of scientific methods and rigorous quality procedures, we enable clients in life sciences, energy, the public sector, and consumer industries to achieve a wide range of objectives, including, forecasting floods, developing fusion power, enhancing military sensor capability, improving drug discovery and development efficiency, and reducing risk to health and the environment in the extraction and production of oil and gas. With offices in Europe and North America, global companies rely on Tessella for business critical assignments.

Copyright © Tessella plc 2009, all trademarks acknowledged. Issue: V1.R0.M0 | Jul-09



[www.tessella.com](http://www.tessella.com)