Int J Digit Libr (2004) 00: 1–14 / Digital Object Identifier (DOI) 10.1007/s00799-004-0102-z

InscriptiFact: A virtual archive of ancient inscriptions from the Near East

Leta Hunt[1], **Marilyn Lundberg**[2], **Bruce Zuckerman**[3]

[1] University of Southern California, Information Services Division, B25 DML, University Park Campus, Los Angeles, CA 90089-0189, USA
e-mail: lihunt@usc.edu
[2] West Semitic Research, 12 Empty Saddle Rd., Rolling Hills Estates, CA 90274, USA
e-mail: mlundber@usc.edu
[3] University of Southern California, Department of Religion, 328 Taper Hall, Los Angeles, CA 90089-0355, USA
e-mail: bzuckerm@usc.edu

Published online: [redacted] 2004 –

Abstract. The need for preservation and reclamation of ancient Near Eastern texts that provide the foundation and historical point of reference for Judaism, Christianity, Islam, and the cultures from which they emerged has become urgent – especially in light of recent cultural disasters. The West Semitic Research Project at the University of Southern California addresses this urgency by documenting ancient inscriptions in a highly meticulous fashion in order to reclaim readings from deteriorated documents using a wide range of technologies. The complementary InscriptiFact image database application employs data organization along with search and display strategies to facilitate easy, intuitive access by scholars.

Keywords: InscriptiFact – West Semitic Research Project – Near East – Inscriptions – Digital library

Introduction

There are two realities that recent world events have only served to reemphasize: First, that the fragile, written records from the ancient past are a precious legacy and a common heritage of all modern civilizations. Second, that this legacy and heritage are under growing threat, and that – unless decisive steps are taken – a significant part of this message from civilizations long past may be lost forever. Over the last two decades, two closely related projects have developed at the University of Southern California (USC) which endeavor to take these decisive steps. The West Semitic Research Project focuses on producing high-quality, high-resolution images of ancient inscriptions. The InscriptiFact project focuses on dissemination of those images with metadata to scholars, students, and the public. It is hoped that the results of these two projects may serve not only to preserve ancient texts from the Near East but as models for the use of advanced technologies to preserve and distribute image data in related fields of study.

Data acquisition: The West Semitic Research Project

Beginning in the early 1980s, faculty in the University of Southern California's School of Religion undertook a research initiative known as the West Semitic Research Project (WSRP) that addressed an essential, even urgent, need in the field of ancient Near Eastern studies: the long-term acquisition and preservation of high-quality image data, recording some of the earliest written records in the world. The aim of the WSRP is to document ancient Near Eastern and Mediterranean inscriptions in a highly meticulous fashion in order to reclaim readings from deteriorated documents using a wide range of technologies.

The target inscriptions come from an international array of museums and libraries and occasionally involve field projects to remote locations where inscriptions still remain in situ. They include the Dead Sea Scrolls; cuneiform tablets from Mesopotamia and Canaan; papyri and monumental inscriptions from Egypt; incised and *bas relief* stone inscriptions from Israel, Jordan, Lebanon, and Turkey; Hebrew, Aramaic, Ammonite, and Edomite inscriptions on a variety of hard media (e.g., clay sherds, plaster, silver, copper, semiprecious stones, jar handles – even an ostrich egg!); the earliest known protoalphabetic inscriptions, inscribed on cliff walls in the Sahara Desert; and the earliest complete Hebrew Bible.[1] These ancient

[1] For a selection of images produced by the WSRP, see the Educational Site at `http://www.usc.edu/dept/LAS/wsrp/` (9/2004). For an example of advanced technology, see further `http://www.usc.edu/dept/LAS/wsrp/information/IntroCBSpgW.html` (9/2004).

texts represent religious and historical documents that serve as a foundation and basic, historical point of reference for Judaism, Christianity, and Islam and the cultures out of which they emerged.

Virtually all missions initiated by the WSRP can be best described as rescue projects. While the reasons this is so have always been apparent to specialists involved in the study of the ancient world, they have now become far more readily apparent to the general public due to cultural disasters that have recently occurred in Iraq and Afghanistan (and earlier in Kuwait and Lebanon). The urgency of the initiative is in recognition that age, weather, substandard archival conditions, and human forces such as the frequent wars in the Middle East are resulting in the slow – and, all too often, not so slow – deterioration and/or destruction of these ancient artifacts. Despite the best efforts of archaeologists, librarians, and museum curators, many if not most artifacts preserving ancient texts are losing vital data at a steady, often alarming rate, and many artifacts have been lost altogether or have now deteriorated to the point that their data are no longer viable.

In many cases, photographic images from early research expeditions preserve the best, and sometimes the only, data of ancient inscriptions; yet such photographic images are also often fading and deteriorating or have been simply misfiled, lost, or destroyed.

The key philosophy of the WSRP has been to bring together under the most optimal conditions the essential data. Formidable obstacles make this a challenging task. Fragments of a given inscription or collection of inscriptions are often scattered among museums, libraries, and archaeological sites. Existing images of these texts can also be scattered among various institutions or reside in private hands. Hence, it is all too often the case that, in order to put together a single text or corpus of texts, one must track down the evidentiary remains in various institutions throughout the world, each with its own security protocols and national, political issues. Even where digital images are available and accessible over the Internet, they are usually not of high enough quality or resolution to be relied upon for serious analysis.

Obviously essential to the reading and interpretation of an inscription is a consensus on what, in fact, was actually written. This should be based on best visual evidence; nonetheless, scholars rarely have access to such quality data and, as a result, consensus regarding what is a reliable textual record is frequently difficult to achieve.

The publication of numerous differing readings based on the same poor photograph is not unusual.[2] In fact, new or revised editions of texts have often presented competing readings of incomplete or degraded areas – readings that a scholar or student has had no way of checking except by examination of the original text at one or more museums or archaeological sites far distant and not readily accessible to the scholar. Within this context, WSRP focuses on developing photographic techniques that maximize the potential for reclaiming data from ancient inscriptions.[3]

To aid in decipherment, WSRP includes in its objectives the acquisition of image data created at stages when the ancient manuscripts were in less deteriorated condition – copying negatives, including fragile glass plates dating from as early as the beginning of the 20th century. The quality of these early photographs is typically poorer than what can be achieved today. Nonetheless, early images frequently preserve data from texts when they were significantly less deteriorated than they are today. Thus, when early images are used in conjunction with the higher-quality, contemporary images produced by WSRP and its collaborators, gaps in the record are typically better clarified: the sum is greater than the individual parts.

Photographing for decipherment

WSRP photographers and scholars have pioneered a number of procedures to document texts from the ancient world. They generally prefer to use a large format (4×5 in., or $\approx 10 \times 13$ cm) view camera on either a copy stand or highly flexible tripod. They use four or five different kinds of film: color reversal (i.e., transparency film), color negative, usually two types of black-and-white negative (a medium-resolution/medium-contrast film and a high-resolution/high-contrast film) and, occasionally, black-and-white infrared film. A Polaroid is used to check for the correct setup, lighting, and exposure, and then a series of exposures with the various other kinds of film are taken. Like commercial photographers, WSRP photographers vary the exposure settings of the film in order to ensure good coverage of the inscription. The use of strobe flash has been found to be the best way to provide the lighting. Not only can the light be controlled most effectively, but flash is far less harmful to most objects than any other form of lighting. The lighting itself can be manipulated (e.g., filtered to exclude all but light in the ultraviolet spectra, polarized or cross-polarized)¿.

For inscriptions of significant size at least three levels of documentation are preferable: reference shots (photographs of the whole inscription or column of text); sec-

[2] To cite but one example of this phenomenon, a crucial reading in an alphabetic cuneiform text from ancient Ugarit (c. 13th century BCE) was read in the literature in at least nine different ways by various commentators – primarily because the image data were so poor that they allowed for such latitude. After WSRP made detailed photographs of the area of the clay tablet in question, a definitive reading for this text was proposed (which turned out to be different from all nine previous proposals!) that has remained unchallenged since its publication in the mid 1980s; see [6]; for the previous nine readings, see esp. p. 20. For other examples, see, e.g., [7, 9, 10].

[3] For a description of WSRP photographic techniques, see [3, 11–14].

MS ID: JODL102
30 November 2004 11:18 CET

tional shots (for example, the top, middle, and bottom of a column); and selected detail shots (of particularly problematic areas of the inscription). For many different kinds of inscriptions on three-dimensional media it is necessary to photograph at several different light angles to capture all of the relevant data.

Each inscription, or type of inscription, presents a unique challenge to the photographer. The "tool box" of methods developed at West Semitic Research is an ever-expanding one as new inscriptions are encountered and recorded for research and preservation. Space does not permit a complete description of those methods, but below are examples of the kinds of methods used in the work of WSRP, involving both conventional and digital photography.

Manuscripts on soft media such as parchment (used here to refer to animal skin in general, regardless of preparation) or papyrus are best photographed using a system of three lights. Two of the lights (sometimes with diffusing "soft boxes" to spread the illumination) are set up at an equal distance on either side of a manuscript. A third light is used beneath the manuscript as a backlight to avoid the creation of shadows on edges or within holes in the manuscript (Fig. 1). This is particularly important if the manuscript is mounted in glass or similar media since otherwise this would result in the material casting a shadow onto the background. In all too many cases such shadows can be confused with ink traces. For many manuscripts this kind of setup is adequate for taking a high-quality image – that is, if the ink is visible

Fig. 1. Picture of a Dead Sea Scroll – 4Q109 Qohelet[a], a biblical text from the book of Ecclesiastes. Note that this picture is evenly lit from the top and is also backlit so that all the edges and holes are clearly illuminated. Photograph by Bruce and Kenneth Zuckerman, West Semitic Research. Courtesy Department of Antiquities, Jordan

and the manuscript in relatively good condition. Unfortunately, this is not always the case.

It has long been known that certain kinds of manuscripts that are unreadable in visible light respond well to the use of infrared (IR) photography. For example, as noted above, most Dead Sea Scrolls photographs taken in the 1950s and 1960s are IR images. Until recently, however, IR film was used without great concern for filtration, with the result that photographs were typically taken that sampled a broad spectrum of IR wavelengths. In the 1990s WSRP discovered that the use of a particular filter (Kodak Wratten 87C) used in conjunction with Kodak high-speed IR film cut off all but a narrow band of wavelengths in the IR spectrum – and that this bandwidth worked in a superior fashion to reclaim data from the Dead Sea Scrolls.[4] Dead Sea Scrolls are often unreadable because they were written with a carbon-based ink on animal skin or parchment, which over time has carbonized to the point that the skin is as dark as the ink – at least in visible light. At a particular band of the IR spectrum the dark background of the parchment "drops out" quite dramatically and becomes highly reflective while the ink remains black, providing the necessary contrast for distinguishing the data from the background (Fig. 2). IR works most effectively for these conditions and has also proven effective for reclaiming faded or erased ink traces on ancient papyri (e.g., the Jewish texts of Elephantine from the 5th century BCE). It is unhelpful for metal-based inks (in which case ultraviolet light can be effective) or for situations in which there is a light background and faded ink (in such cases colored filters for use with black-and-white film can provide some additional gain in data).

[4] This was also noted by Greg Bearman of the Jet Propulsion Laboratories, in this case utilizing digital imagery and a liquid crystal variable cutoff filter; see [1].

Incised inscriptions on hard media (such as stone or clay) require an entirely different approach. Rather than minimizing shadow, one needs to create it. The best results are obtained when one light is used as the "key," i.e., the primary light, placed at a raking angle, with a reflector or "fill" light or reflector placed approximately opposite the main light. The primary light is placed so that shadows are created along the strokes of letters or symbols, taking care not to place the light angle parallel to the strokes (Fig. 3). If an inscription has a rounded surface, two equal lights placed opposite one another might be used, each light thus acting both as a primary light and as fill for the light opposite, what might be called a "fill-fill" lighting setup. In many cases one particular light angle does not suffice to highlight all of the data, so images will be captured from different light angles and then studied side by side.

Cuneiform tablets have usually been photographed using a key-and-fill setup. Traditionally the light is placed at the top left corner of the tablet, since this is thought most closely to emulate the way ancient scribes would have seen the tablet as they were reading or writing (Fig. 4). Since many tablets are curved, they may need to be rotated so that all parts of the writing can be clearly seen. This method of photographing cuneiform, while effective, tends to be quite time consuming, especially con-

Fig. 2. Images of a fragment from a Dead Sea Scroll, 1Q20 Genesis Apocryphon ar, illustrate the gain in information through the use of infrared imaging. Photographs by Bruce Zuckerman and Marilyn Lundberg, West Semitic Research.
Courtesy Department of Antiquities, Jordan

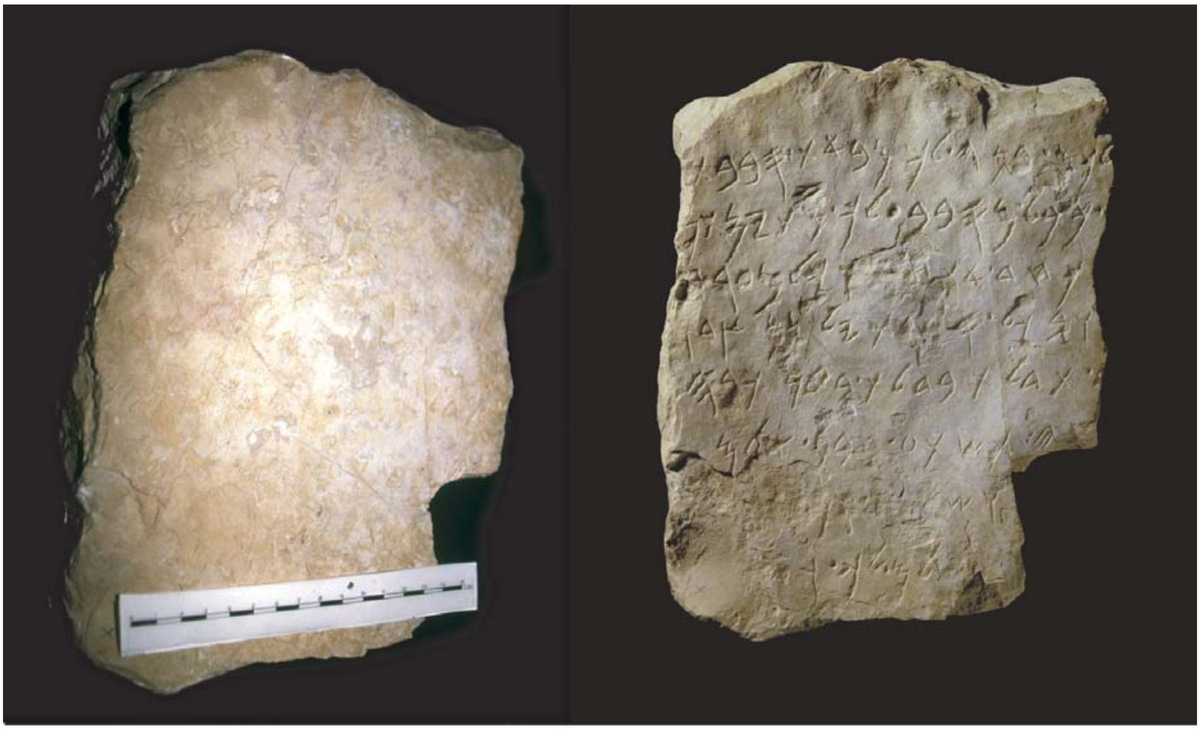


Fig. 3. These two images of the Amman Citadel Inscription (c. 800 BCE) illustrate the importance of proper illumination. The photograph at *left* shows the inscription as it appears in diffuse light while the photograph at *right* shows the same inscription when illuminated by a raking diagonal light from the left-hand corner. Photograph by Bruce and Kenneth Zuckerman, West Semitic Research. Courtesy Department of Antiquities, Jordan

Fig. 4. Clay tablet recording an economic transaction from ancient Mesopotamia (c. 2000 BCE). This tablet has been illuminated to bring out the details of the seal impression (in bas relief) rather than the details of the transaction (impressed wedges). Photograph by Kai Quinlan, West Semitic Research. Courtesy University of Southern California Archaeological Research Collection

MS ID: JODL102
30 November 2004 11:18 CET

sidering the hundreds of thousands of tablets in museums and collections all over the world. Lately we have been experimenting with digital technologies that might be both faster and more effective.

One technique especially useful for cuneiform tablets and other three-dimensional objects is a technology called "image-based relighting" developed by Hewlett Packard scientists Thomas Malzbender and Daniel Gelb. A digital camera is mounted on a device with as many as 50 strobe flashes and attached to a computer. The device fires off a succession of digital pictures, each one illuminated from a different lighting angle – some 40 to 50 lighting angles in all. Then software developed by Malzbender and Gelb amalgamates these images into a single, integrated view that allows one to see the target at any light angle – including angles that were not actually shot by the camera. As one moves a control device around, e.g., a computer mouse, the light also moves and plays at different angles across the surface – like having the equivalent of a very precise beam of light that one can place at any angle at any given moment. As the virtual light moves, the play of shadows changes, making the subtle indentations on a clay surface jump out for the observer in a way never before possible. Moreover, with a couple of clicks, one can change the reflectivity of the surface to startling effect. For example, one can make the dull matte finish of a clay surface "super shiny," as though it had been dipped in molten metal. Once again, with this lighting effect nuanced details suddenly become clear that were not visible before (Fig. 5).

Among incised inscriptions, stamp seals provide their own challenges. Seals that are relatively flat and nonreflective can often be photographed using a key-and-fill technique. If a seal is rounded, the fill-fill approach may be needed, with at least two lights, but as many as four. In such situations fiber-optic light sources and a time exposure are useful – in effect one "paints" with the light. If a seal is highly reflective, it can almost act like a mirror when light is shone on the surface at a particular angle. That reflection cannot be captured if the seal is laid directly under the camera lens and both surface and lens plane are flat. However, if the object is tilted at the optimal angle, the light bounces off the surface into the lens like light from a mirror to show every aspect of the surface, down to the smallest detail. Some distortion of the object will occur, but this can be minimized if the lens plane and film plane are also tilted in parallel to the object's tilt (as can be done when employing a view camera). We call this approach the "tilt technique" (Fig. 6).

One of the most recent digital techniques to be developed by WSRP uses a panoramic digital camera to capture 360° images of cylinder seals. Cylinder seals were used in ancient Mesopotamia and elsewhere as a form of identification for validating contracts, economic transactions, or any document that needed the equivalent of an official signature. A piece of stone, fashioned into a small cylinder (approximately the size of a AAA battery) and incised on its rounded surface typically with a combination of intricate designs and language-signs (normally giving the name of the owner), would be rolled across a piece of wet clay on which the transaction in question was written, making a continuous image. This impression, then, is the equivalent of the signature/identification of the seal's owner. Typically when scholars want to see the inscription, they roll the cylinder seal in wax, clay, or other appropriate medium, then photograph the seal impression. While this can be useful, the surface of the seal *itself* contains a level of data that often cannot be seen in an impression. For example, sometimes tiny details such as tool marks or errors that were scratched out can be seen that do not show up in the impression.

The photographers and scholars at WSRP have built a miniature turntable on which the cylinder seals rotate, while the panoramic digital camera records the data through a slit in front of the lens as a continuous image. Hence, images in 360° can be recorded as though one had peeled off the rolling surface and flattened it out for opti-

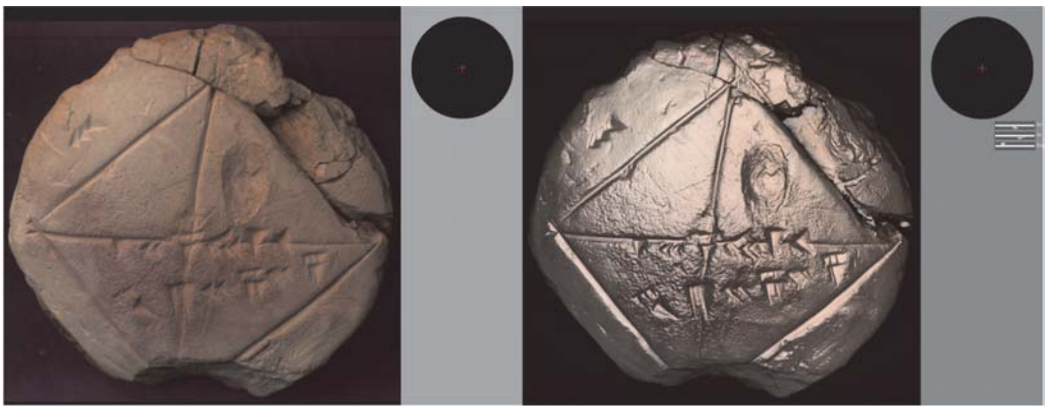


Fig. 5. Images of a cuneiform tablet captured with "image-based relighting." The image at *left* represents conventional lighting; the image at *right* made "super shiny" by specialized algorithms. Images by Thomas Malzbender and Daniel Gelb, Hewlett-Packard Imaging Labs. Courtesy Yale Babylonian Collection

MS ID: JODL102
30 November 2004 11:18 CET

Fig. 6. An ancient Moabite seal in its original bronze setting (c. 700 BCE). This seal was photographed using the "tilt technique" in which not only the seal but also the lens and the film plane are all tilted at the same angle so as to optimize the reflective qualities of the polished jasper stone surface without distortion. Photograph by Bruce and Kenneth Zuckerman, West Semitic Research. Courtesy Shlomo Moussaieff

mal viewing. The use of fiber-optic light for illumination allows the seal to be photographed in a variety of ways to capture both details of the artisan's work as well as colors and textures of the stone (Fig. 7).

Data distribution: the InscriptiFact project

While the data acquisition of the WSRP is the essential first step for long-term preservation of ancient texts, a second step is also of crucial importance: data distribution. InscriptiFact,[5] a database application, is therefore the natural extension of the WSRP, since its goal is to make the WSRP image archive broadly available to specialists, teachers, and students throughout the scholarly world via the Internet. InscriptiFact extends the value of the WSRP by using appropriate, well-tested technologies to facilitate access to its images in high-resolution formats in an intuitive, "scholar-friendly," presentation.

InscriptiFact has been available as a platform-independent database application since May 2003 with a test data set of 840 images. The current version (v.4.0.1 as of this writing) includes 5200 images. InscriptiFact has funding to make 20 000 images widely available by 2006 as the deliverable for its stage one database application.[6] As part of stage two development, InscriptiFact will make available over 100 000 images from the WSRP and other archives. The long-term vision is to make InscriptiFact a distributed database of images of ancient artifacts found all over the world.

A number of features make InscriptiFact especially serviceable to the community of scholars who study, analyze, and decipher ancient inscriptions. These include:

- Organization of metadata and access points based on user needs
- Transparent Boolean query construction based on multiple access points
- Drop-down lists to communicate possible selections and to preclude typing errors
- Clear, constant, graphic-based communication designed to show the user always where he or she is and

[5] The name was coined to represent a compound of "Inscription" and "Artifact" with the "F" capitalized to connote that the database application conveys factual data about ancient artifacts as inscriptions.

[6] We wish especially to acknowledge the support of the Andrew W. Mellon Foundation, the Underwood Family Trust, USC's Annenberg Center for Communication, The Ahmanson Foundation, USC's Zumberge Research and Innovation Fund, Oracle Corporation, LizardTech, BEA, Science Systems Applications, and 3Caos for their support of the West Semitic Research and InscriptiFact projects.

Fig. 7. A cylinder seal shown as a 360°+ image. Image by Kenneth Zuckerman, West Semitic Research. Courtesy Vorderasiatisches Museum, Berlin

what further options are available, given current selections
- Various retrieval mechanisms, depending on what is most appropriate for a particular user-selected text
- Variety of image display options:
  - Capacity to preview before selecting images for detailed view
  - Capability to view and compare up to five images at a time at high resolution
  - Flexibility in resizing and moving digital objects
  - Full-screen viewing capacity
  - Various conventional image display tools (pan, zoom, select zoom, etc.)

In the discussion that follows, we will focus on some of the more important features that make this database application especially useful to its target audiences.

InscriptiFact data organization, metadata, and access points[7]

Data in InscriptiFact are organized around the concept of a text, rather than a digital object or a collection containing texts. A "text" in this context is a virtual object in that a given text may not physically exist at any one place in its entirety. That is, since text fragments are often found in scattered locations in various collections, InscriptiFact brings together images of a given text regardless of the location of individual parts of that text in institutions around the world (Fig. 8).

7 The InscriptiFact team gratefully acknowledges Wayne Shoaf, the digital archivist at the University of Southern California, for his significant aid in developing the cataloging approach of InscriptiFact and for his advice and editing of this paper.

Numerous images of a given text, captured at various dates and locations, using a variety of film types and/or digital methodologies, taken in different light spectra (ultraviolet, visible, infrared), at several scales of magnification, and focused on specific fragments, typically coexist within the WSRP archive. Thus a single search query could yield hundreds of images. The organization of the metadata must address this issue in such a way as to facilitate efficient retrieval.

Explanation of the means by which InscriptiFact addresses the organization of its data requires a brief digression into two approaches to cataloging: The Dublin Core Metadata Initiative (DCMI)[8] and the Functional Requirements for Bibliographic Records (FRBR) [4]. The Dublin Core and FRBR approaches were formulated to address the complexities of cataloging increasing amounts of information (particularly digital information) and greater variety in information types. The Dublin Core approach involves a set of 15 core metadata elements as follows:[9] Title, Creator, Subject and Keywords, Description, Publisher, Contributor, Date, Resource Type, Format, Resource Identifier, Source, Language, Relation, Coverage, and Rights Management. Each element can be repeated as many times as necessary in a given catalog record, and any element could be omitted if not relevant. Additionally, there are two sets of qualifiers for each element:[10] first, to facilitate the refinement of an element in order to meet the specific needs of particular users and

8 See "History of the Dublin Core Metadata Initiative" (`http://www.dublincore.org/about/history/`).
9 See `http://www.dublincore.org/documents/dcmi-terms/` "Section 2: The Dublin Core Metadata Element Set." (8/2004)
10 See `http://dublincore.org/documents/usageguide/qualifiers.shtml` (8/2004).

Fig. 8a–c. These three images show fragments of 1Q20 Genesis Apocryphon ar, located in Amman, Jordan (**a**), Jerusalem (**b**) and Norway (**c**). **a** Photograph by Bruce and Kenneth Zuckerman, West Semitic Research. Courtesy Department of Antiquities, Jordan. **b** Photograph by Bruce and Kenneth Zuckerman, West Semitic Research. Courtesy Israel Antiquities Authority. **c** Photograph by Bruce Zuckerman and Marilyn Lundberg, West Semitic Research. Courtesy M. Schøyen

particular materials, and second, to notate the "encoding scheme" that governed entry of the element content. (For example, a "Title" can be refined as an "Alternate Title." A "Subject" can be encoded from the list of "Library of Congress Subject Headings.") Most importantly for InscriptiFact, catalogers are allowed to develop other local qualifiers as well.[11]

The FRBR approach is that of an entity-analysis technique normally used in the development of conceptual data models for relational database systems. It involves three groups of entities. Group 1 (primary entities) is especially relevant to InscriptiFact.[12] It consists of the entities:

– *Work*: A distinct intellectual or artistic creation (e.g., Shakespeare's *Hamlet*)
– *Expression*: The intellectual or artistic realization of a work (e.g., the intellectual content of Folger's edition of Hamlet or the intellectual content of the Yale edition of the play).
– *Manifestation*: The physical embodiments of a particular version of a work (e.g., the published edition of Folger's *Hamlet* or the published edition of the play done by Yale).
– *Item*: A single exemplar of a manifestation (a specific copy of the Folger's edition. e.g., the one that is "dog-eared," that has a handwritten note from "Aunt Emma" on the inside front cover, a gift book plate, and in which p. 62 is partially ripped out).

According to FRBR, a *work* may be realized through one or more than one *expression*. An *expression* is the realization of one and only one *work*. An *expression* of a *work* may be embodied in one or more than one *manifestation.* A *manifestation* may embody one or more than one *expression.* A *manifestation* may be exemplified by one or more than one *item*. An *item* may exemplify one and only one *manifestation.*[13] The FRBR concept is to provide the capability to group *expressions*, *manifestations*, and *items* under a single parent record for the *work*; whereas using the previous standard, the bulk of cataloging was accomplished at the level of *manifestations* with additional minor work in order to track *items*.

The general approach of the InscriptiFact team is to combine the Dublin Core and FRBR approaches (with notable exceptions; see below) within the context of meeting the needs of InscriptiFact users for access and for descriptive information.[14]

The Dublin Core standard is well recognized among libraries, and, in fact, Qualified Dublin Core[15] serves as the standard for the home university of InscriptiFact, the University of Southern California (USC). However, it is not easily representative of the metadata desired by scholars of ancient texts. Usually, in the Dublin Core approach, cataloging is assigned to a given physical or digital object, and information is confined to this manifestation alone.[16] In contrast, for specialists who study inscriptions, a text is an intellectual concept, and they need appropriate scholarly cataloging that must include information about several manifestations of a given text. For example, information must be included about the physical object(s) containing the inscription, the intellectual work of the inscription itself, the photographic images, and the digital version of the images. This complexity dictates a variation in approach if meaningful access and descriptive cataloging appropriate both to the scholarly clientele of InscriptiFact and to the library world is to be provided.

At first glance, the FRBR approach seems closer to matching InscriptiFact data than Dublin Core. The FRBR concept of *work* seems, initially, generally consistent with the concept of a text that may exist as a whole only in virtual form. Similarly, the FRBR concept of *manifestation* seems to fit generally with various manifestations (physical object, photographic image, digital version), and the FRBR concept *item* is generally consistent with the InscriptiFact concept of a given digital representation of an inscription.

More detailed analysis, however, raises questions. The foremost problem is that in the case of ancient inscriptions, the actual intellectual *work*, as defined by FRBR, is unknown, uncertain, or questionable. That is, scholars of ancient inscriptions are presented with a physical object(s) (*manifestation*) that is (are) inscribed. Over the course of time in which a given text is studied and deciphered, there may be ambiguity regarding the extent and specific organization of the text. Fragments thought to be a part of one text in a given organizational scheme may be considered part of another text in a different organizational scheme, or a given text may be subdivided into several texts as it is studied. It is the job of archaeologists, linguists, epigraphers, philologists, and other specialists to try to reconstruct the original text, that is, figure out what pieces fit together, how the text is organized, when it was inscribed, and what, in fact, the intellectual content of the inscription might be.

These uncertainties in the accurate reconstruction and reading of an ancient text must be reflected in the data model. FRBR assumes the fundamental capability to determine what is, in fact, the intellectual content (the *work*). But this assumption is inappropriate when applied

[11] Ibid.

[12] Group 2 included "responsibility entities" (e.g., person; corporate body), and group 3 included "subject entities" (e.g., concept; object; event; place); see [2].

[13] Cf. n. 18, IFLA Sect. 3.1.1 Group 1 Entities: Work, Expression, Manifestation, Item.

[14] The design of InscriptiFact, including data organization, graphical user interface design, and engineering, reflects requirements specifically developed in accordance with the specific needs of its scholarly audience.

[15] See `http://www.dublincore.org/` (8/2004).

[16] The one exception to this general rule is that additional information about other manifestations can be included under the "Source" element.

MS ID: JODL102
30 November 2004 11:18 CET

to ancient inscriptions. Therefore, it is inappropriate to separate the intellectual *work* of an ancient text from the physical object or objects upon which it is inscribed. The intellectual *work* in InscriptiFact is, for this reason, defined as the Inscription in the context of the physical object(s) upon which it is inscribed.

Moreover, there is interpretative latitude inherent in the process of determining the difference between a *work* and an *expression* in the context of ancient inscriptions – far too little is known about either the extant *expressions* or the possible *works* behind them to make this kind of determination.[17] For this reason, each text [in context with physical object(s)] must be considered separately from all other instances with the same textual content. Therefore, the FRBR concept *expression* is inadequate in the context of InscriptiFact. Rather, relationships between an instance of a given textual content of an inscription and other instances of the same content are better expressed using the Dublin Core element Relation.

A separate problem is that there is a many-to-many relationship between *work* and *manifestation* in the context of ancient inscriptions. Consider, for example, the case of a palimpsest – a document written on top of an earlier, erased, text. A good example of this is Syrus Sinaititicus, a 5th-century codex found at St. Catherine's Monastery in the Sinai Desert. In this case a wholly separate work, "The Lives of the Female Saints," is written in Syriac (an eastern dialect of Aramaic) over one of the earliest extant Syriac translations of the gospels. Of course, while both texts are intrinsically important, the historical significance of the early gospel has overshadowed the text that has overwritten it. In FRBR terms this illustrates *expressions* of two wholly separate *works* written on the same *manifestation*.

Essentially, the objects that are delivered to the user of InscriptiFact are images. The FRBR concept *item* is generally consistent with the InscriptiFact concept of a given digital representation of the inscription. Note, however, that an important difference exists in that an image, for example a plate of fragments (Fig. 9), may include representations of numerous texts. Therefore, a many-to-many relationship exists between *works* and *items*.

In addition to the above considerations, the InscriptiFact combined approach includes two variations: separating the cataloging for the text (that is, the inscription; the *work*) from the images (that is, the digital objects; the *items*) and extending the Qualified Dublin Core presentation to include an additional qualifier to denote *manifestation* [e.g., physical object, intellectual work (the inscription), photographic object, or digital object] described by a given attribute. The *items* (that is, the digital images) are retrieved hierarchically under the *work* (the inscription).

The InscriptiFact data team developed a basic set of attributes for these two sets of cataloging (for the text and for the image) that addresses the needs of the researchers and scholars of ancient texts. The resulting cataloging for the text includes information about the physical object and the intellectual work, while cataloging for the image includes information about the digital version, the photographic image, the inscription as represented by each specific image, and, to a much lesser extent, the physical object itself (e.g., the museum accession numbers of fragments that are actually represented in a given image). The data team then mapped these attributes to Dublin Core elements[18] and presented them as qualifiers.

Access to InscriptiFact data is provided through nine of the Dublin Core elements: Title (Main Text or Publica-

[17] Good examples are the two Ketef Hinnom Amulets, dating from the 7th–6th centuries BCE. These inscriptions quote, each in a slightly different version, the prayer, known in Judaeo-Christian tradition as "the Priestly Benediction" and found in Numbers 6:24–26 ("The LORD bless you and keep you . . . "). These artifacts are the earliest known texts to cite text also found in the Bible. On both these amulets, however, are separate preambles that invoke the prayer as protection against evil. Scholars therefore have to consider: Are the texts on the amulets the earliest examples of an *expression* of a biblical verse, or are the texts in both the Bible and on the amulets actually *expressions* of some prayer that preceded both?

[18] http://dublincore.org/documents/dces/ (8/2004)

Fig. 9. The fragments on this plate of papyrus fragments (Berlin P13461A) represent, or are part of, many texts. Photograph by Bruce and Kenneth Zuckerman, West Semitic Research. Courtesy Aegyptisches Museum, Berlin

tion Number), Description (Medium, Script),[19] Relation (Is Part of Corpus; Corpus Category; Corpus Subcategory; Is Part of; Is Version of), Language, Geographical Coverage (Find Site), Temporal Coverage (Timeline; Named Time Period), Source (Alternate Text or Publication Number, Collection Owner), Rights (Collection Owner), and Subject (Keywords, Text Division). Additional descriptive cataloging includes a plethora of information deemed useful for users of InscriptiFact.

Access by Text or Publication Number is somewhat complicated in that the identification number for one text in a given identification system might refer to more than one text in other identification systems.[20] Scholars may wish to retrieve a text based on any of these identification systems. InscriptiFact addresses this problem by defining each text according to one "main" identification system and mapping all other organizational systems to the "main" system. Users may therefore request a text or group of texts by any of the conventional identification systems.

Access to the images is accomplished by first selecting the text. Query refinement is then accomplished as a function of the type of text selected. For example, if the text is long with several columns, the metadata may include text divisions as a means of refining the query. In some cases, numerous images may be associated with a given surface area of a text (for example, the front, back, sides, etc. of an inscribed clay tablet). In such cases, it may be more useful to facilitate retrieval of images of a geometrically defined area of a selected surface of the text.

Metadata not associated with typical cataloging[21] are necessary for access to images using InscriptiFact's spatial search (see "InscriptiFact graphical user interface" below). In spatial search each detailed image to be retrieved by InscriptiFact must be "spatially referenced" to one or more reference images. The reference image presents a small-scale overview of a given surface on which a text is written. A detailed image is a larger-scale image of a section or portion of that surface. The reference image is used as a "map" on which scholars can define an area (e.g., by using his/her mouse to draw a box around that area). Spatial referencing is accomplished by the catalogers as part of data input using a specialized software prepared for InscriptiFact by System Science Applications (EASy).[22] The cataloger selects a relevant reference image and a group of detailed images to be referenced to the selected reference image. The EASy software applies a transparent grid system to the selected reference image. InscriptiFact catalogers are then able to overlay a selected detailed image on top of the reference image. The edges of the detailed image are then adjusted until it "fits" the appropriate section of the reference image. The coordinates of the edges are captured into a database.

InscriptiFact graphical user interface

InscriptiFact's graphical user interface has been designed to aid scholars in searching for texts using means that are appropriate to their work and perspective. However, the breadth and variety of ancient texts and the diversity of scholars who study them complicate the creation of a process by which a given text and its images are selected. InscriptiFact addresses this inherent complexity by providing a number of features to facilitate appropriate retrieval and preclude confusion.

Texts may be retrieved in response to queries defined by any one or more of the following: **corpus** (i.e., any generally recognized collection of texts, e.g., Dead Sea Scrolls, Elephantine texts, Ammonite inscriptions, etc.), **medium** of the physical object (i.e., what it was written on, e.g., clay tablet, stone, copper, pottery shard, etc.), **find site** (i.e., provenance, e.g., Qumran, Ugarit, Amman Citadel, unknown, etc.), **language** (e.g., Hebrew, Aramaic, Phoenican), **script** (e.g., alphabetic cuneiform, demotic, paleo-Hebrew, Jewish), **collection owner** (e.g., Los Angeles County Museum of Art, USC Archaeological Research Collection, Staatliche Museen zu Berlin), **time period**, and **keyword**. In the case where a scholar knows the publication number in any conventional identification scheme, the **Text or Publication Number** may be entered directly.

Query options are displayed in drop-down lists. The Boolean query construction is transparent to the user in order to facilitate ease of operation. That is, Boolean "OR" relationships are assumed between choices within the same category, while "AND" relationships are assumed between choices across different categories and between categories and keywords.[23] For example, a user may define a query for texts written on stone in Phoenician script. Another example would entail selecting the corpus "Aramaic Texts from Egypt" with a corpus category of "Contracts" and entering the keyword "marriage." This will retrieve Aramaic marriage contracts from ancient Egypt. In defining a search, scholars may select from three chronologies, in accordance with the customs of var-

---

[19] Note that the attributes Medium and Script could also be encoded using the Dublin Core element Format.

[20] See note 26 above in reference to the hypothetical aspect of reconstructing an ancient text.

[21] Note that spatial referencing information could be included as a variation of the Dublin Core element Coverage. Currently, Coverage is divided into Geographical Coverage and Temporal Coverage. But other variations could easily be defined. In this case the deciding factor was that the coordinates on the grid applied to each reference image are not meaningful, for example, in the sense that coordinates on the earth's surface are meaningful. Therefore, the spatial referencing information is not meaningful to InscriptiFact users outside its ability to facilitate spatial search and thus is not necessary in the display of cataloging.

[22] More information on System Science Applications (EASy) can be found at `http://netmarkscom.com/ssa/` (8/2004)

[23] Note, however, that within the category "Corpus," Corpus Categories and Subcategories are defined using Boolean AND relationships.

MS ID: JODL102

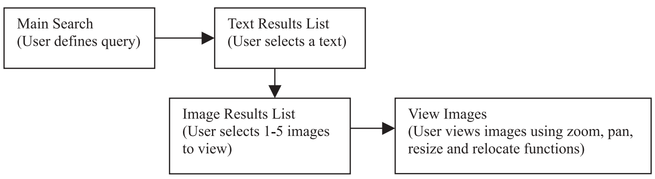
30 November 2004 11:18 CET

ious subdisciplines within ancient Near Eastern studies: Egyptian, Mesopotamian, or Syro-Palestinian.

The many images of a given text could potentially yield hundreds of images in a single search query. Scholars therefore need to be able to retrieve images in a logical, time-efficient, and intuitive fashion depending on what particular "needles" they need to study at a given moment in an expansive "haystack" of images. InscriptiFact addresses this variety and breadth by providing three retrieval options that are specified when the texts and images are input into the system. A "Page Navigation Menu" at the top of each InscriptiFact screen communicates to the user where she or he has been and where she or he may go, given previous choices, at any particular time. The options are signaled to the user by being either highlighted (indicating current choice options) or "grayed-out" (indicating that given choices in this context are excluded).

Simple presentation

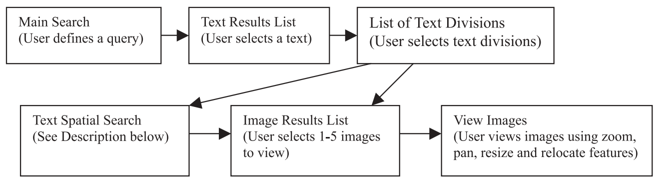
Some texts, such as sealing stones with short inscriptions designed to be stamped on clay surfaces, only entail a simple presentation of the associated images. The screen navigation for these simple presentations is as follows:



The **Main Search** screen facilitates query definition. The **Text Results** screen lists the retrieved texts. The user can view the cataloging for any of the listed texts and select a text for which images are desired. The **Image Results** screen presents the retrieved images. In the simple presentation, no intermediate stage occurs between the **Text Results** list and the **Image Results** list. The user may select 1–5 images from the **Image Results** list for detailed viewing. Note that a 512×512 dpi enlargement function aids the user in making selections from the **Image Results** list (Fig. 10a). In the current stage one version of InscriptiFact,[24] these images can be sorted by date of photographic image, film type, or degree of magnification. In subsequent generations of InscriptiFact, it is expected that these (and perhaps other) sorting criteria will be usable in combination. The cataloging for each image is available from both the **Image Results** list and the **View Images** screen. The **View Images** screen (Fig. 10b) facilitates high-resolution viewing of multiple images in MrSID[25] format. Zoom, pan, relocate, resize, and full-screen viewing functions are available to aid the user in selecting images for download.

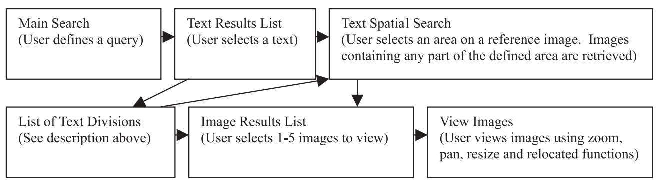
Presentation for texts with many columns or parts

Sometimes there are many images for a text, and an intermediate selection may be necessary to aid the scholar in making the best, most efficient choices. For example, for texts with many columns or parts, e.g., scrolls or codices, a list of text divisions may be appropriate. The screen navigation for the presentation of a text with divisions or parts is illustrated below.



Text spatial search presentation

Other texts that are documented in numerous, detailed images might be better accessed by the option of a two-dimensional geometrical spatial search. This is an innovation unique to InscriptiFact. A reference image (e.g., of the entire front, various sides or back of an inscription) is used as a "map" on which scholars can define a particular area to focus upon. Once a given "view" of the text is selected and displayed, the user employs his/her mouse to draw a box around an area of interest. The search, once initiated, will then retrieve all the images that contain the defined area encompassed by the box. The screen navigation for this presentation is illustrated below. Note that a spatial search may also follow a list of text divisions when appropriate.



Intellectual rights

The issue of intellectual rights is a complex one for InscriptiFact given the distributed nature of the physical objects and the multiple manifestations of a text. The InscriptiFact team has chosen to abide by cooperative agreements. Access to InscriptiFact is therefore password protected.[26] In order to gain access to InscriptiFact, a user must agree to restrict use to educational and research purposes only. InscriptiFact users must sign an agreement to this effect. Subject to this restriction, images are accessible from InscriptiFact free of charge.

[24] Ibid.
[25] See www.lizardtech.com (8/2004)

[26] To get a guest userID and password, send an email to lihunt@usc.edu.

MS ID: JODL102
30 November 2004 11:18 CET

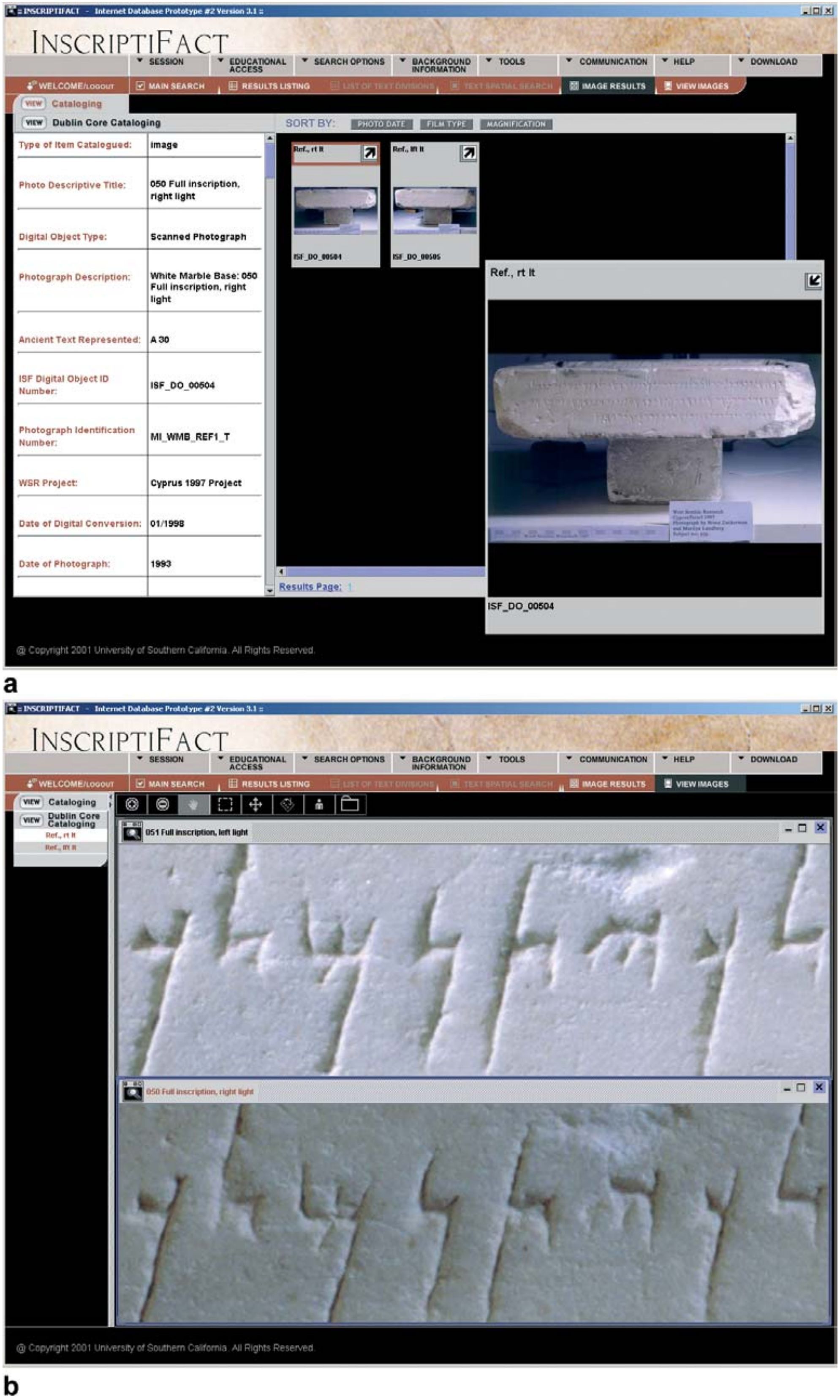


Fig. 10. **a** The Image Results screen presents the retrieved images for the text "A30. Inscription on White Marble Base." Note that a 512×512 enlargement serves to aid the user in making selections. The cataloging for a selected image displays to the *left*. **b** The View Images screen facilitates high-resolution viewing of multiple images in MrSID format. In this example, two images taken with opposite lighting are compared at high resolution. The viewer is in partial screen mode leaving the Page Navigation Menu displaying at the top.

Permission from the institution holding the physical object(s) must be obtained if an image is to be used for publication, reproduction on the Web, or any other similar public purpose. Images of a text are presented only after the institutions holding the physical objects agree to allow the images to be viewed and distributed for educational or research purposes.

Summary

InscriptiFact is a database application designed to facilitate optimal use and distribution of images of ancient Near Eastern inscriptions and to supply a permanent resource for the reclamation of inscriptions using current and future technologies. It is expected to facilitate many new techniques for analyzing these documents that serve as the foundation of Western culture. Most importantly, InscriptiFact is helping to establish a standard methodology for the analysis of ancient inscriptions that will facilitate scientific inquiry and full interchange of information. It is hoped that the InscriptiFact project will serve as a model for other interdisciplinary efforts that can benefit from the integration of new and continually advancing technology – especially in fields that have not relied significantly on technology in the past.

References

1. Bearman G, Spiro S (1996) Archaeological applications of advanced imaging techniques. Biblical Archaeol 59:56–66
2. Hickey TB, O'Neill ET, Toves J (2002) Experiments with the IFLA functional requirements for bibliographic records (FRBR). D-Lib Mag 8(2)
3. Hunt L, Lundberg M, Zuckerman B (2001) Eyewitness to the past: reclaiming ancient inscriptions with modern technologies through USC's West Semitic Research and InscriptiFact Projects. Biblos 50:79–100
4. IFLA Study Group on the Functional Requirements for Bibliographic Records (1998) Functional requirements for bibliographic records: final report. Saur, Munich. http://www.ifla.org/VII/s132/frbr/frbr.pdf
5. Lim TH, MacQueen HL, Carmichael CM (eds) (2001) On scrolls, artifacts and intellectual property. J Study Pseudepigrapha Supplement Series 38. Sheffield Academic Press, Sheffield, UK
6. Ratner R, Zuckerman B (1986) 'A kid in milk'?: New photographs of KTU 1.23, line 14. Hebrew Union College Annu 57:15–60
7. Schniedewind W, Zuckerman B (2001) A possible reconstruction of the name of Haza'el's father in the Tel Dan inscription. Israel Explorat J 51:88–91
8. Lubetzky S, Svenonius E, McGarry D (2001) Seymour Lubetzky: writings on the classical art of cataloging. Libraries Unlimited, Englewood, CO
9. Zuckerman B (1991) The Nora puzzle. Let your colleagues praise you: studies in memory of Stanley Gevirtz I. Maarav J Study Northwest Semit Lang Lit 7:269–301, plates 1–2
10. Zuckerman B (1995) On being 'damned certain': the story of a curse in the Sefire inscription and its interpretation. In: Beck A, Bartelt A, Raabe P, Franke C (eds) Fortunate the eyes that see; essays in honor of David Noel Freedman in celebration of his 70th birthday. Eerdmans, Grand Rapids, MI, pp 422–435
11. Zuckerman B (1996) Bringing the Dead Sea Scrolls back to life; a new evaluation of the photographic and electronic imaging of the Dead Sea Scrolls. Dead Sea Discov 3:178–207
12. Meyers E (1997) The Oxford encyclopedia of archaeology in the Near East (5 vols). Oxford University Press, New York, vol 4, pp 336–347
13. Zuckerman B, Zuckerman K (2000) Photography and computer imaging. In: Schiffman L, Vanderkam JC (eds) Encyclopedia of the Dead Sea Scrolls. Oxford University Press, New York, pp 669–675
14. Zuckerman B (2002) Working with a little more data; new finds in the 20th century: the semitic languages of the ancient world. In: Izre'el S (ed) Israel oriental studies XX; semitic linguistics: the state of the art at the turn of the twenty-first century. Eisenbraums, Winona Lake, IN, pp 481–497

MS ID: JODL102

30 November 2004 11:18 CET