

GSAS News & Views

A PUBLICATION OF THE GRADUATE SCHOOL OF ARTS AND SCIENCES

FALL 1997

New Technology for Old Writing

Roger Bagnall

Chairman, Department of Classics

"Five years ago I suggested at an international congress of papyrology both that digital imaging was going to offer us an opportunity to solve the problem of access to distant collections, and that software like Web browsers would make it possible to knit together textual data and images to form a single scholarly and educational system. I got mostly puzzled stares; many Europeans still thought that technology was an odd American obsession."

Well-intentioned friends among the Graduate School's alumni have occasionally wondered how I have kept busy since the end of my term as dean. Apart from the stock answers (teaching, writing, chairing the Classics Department), a big part of the answer is a project called the Advanced Papyrological Information System (APIS, which is also the name of an ancient Egyptian bull-god), in which Columbia is the lead partner of a six-institution consortium. This project aims to stretch the capabilities of the World Wide Web to link the six universities' papyrus collections and ultimately dozens more around the world into a seamless virtual collection.

Papyrus is a fibrous writing material made from a reed-like plant once common in most of the Nile valley. It was the most important writing surface of the ancient world and perhaps ancient Egypt's most important legacy; alongside it were used other (often cheaper) materials, like wood, clay and stone (broken pottery sherds or stone chips

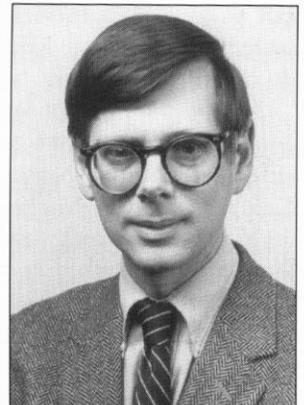
with writing are called ostraca). Columbia has a sizable collection of Greek papyri, mostly acquired in the 1920s and 1930s through an international buying collective, and some Arabic texts from later gifts. Today relatively few papyri make it into the market, and those come at enormous prices; the few thousand dollars a year that President Butler allowed William Linn Westermann (Professor of History, 1925-1944) to spend on papyri were a good investment. We also bought more than four thousand ostraca, mostly in Coptic Egyptian, from the Metropolitan Museum of Art in the early 1960s; these came from their excavations across the Nile from Luxor.

On these materials are recorded everything from high literature to the myriad of documents and other communications of daily life. Some are fragments of literature, either a far more ancient witness to a work known otherwise from medieval manuscripts or a text hitherto lost in antiquity; the other day I looked at a neatly lettered ostrakon in Coptic and realized that it was a few verses of St. Paul (2 Cor. 5:17-19). Mostly, however, our texts are private letters or documents of every conceivable sort—legal and business papers, government regulations, property records and transactions, petitions to high officials, tax and rent receipts, bank deposits and payments, and farm and crop reports. As such, these documentary papyri differ little from modern archival material; except for their usually fragmentary nature and extreme antiquity, they reflect the quotidian affairs of government, commerce, and personal life in much the same way that modern records do.

From such documentary papyri were born the fields of ancient social, economic, and administrative history, which have increasingly displaced the older histories of kings and battles. The papyri (using the term broadly) are thus the source of a large part of what we know about many aspects of antiquity, particularly those concerned with economic life, social relations, cultural interaction in a pluralistic society, and daily life.

SCATTERED FRAGMENTS

Papyri are often fragmentary, not only because water and insects have done their



Professor Roger Bagnall

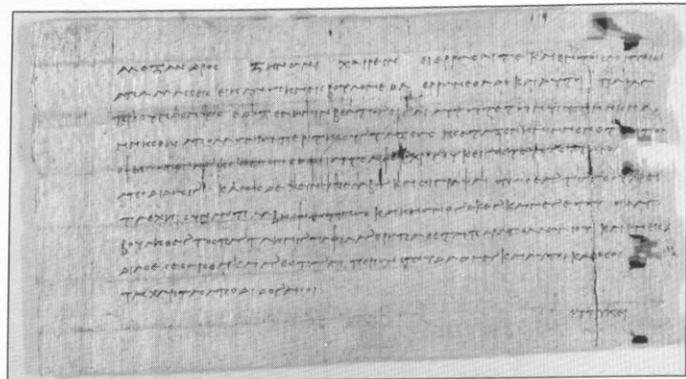
worst, but because their discoverers and sellers have cut them up to sell them in several pieces for more money. For example, one roll in our collection with receipts for money paid to donkey- and camel-drivers for delivering wheat taxes has long been known to join up with pieces in Berlin's Staatliche Museen. More recently a fragment in Groningen (Netherlands) turned out to belong to the same roll, and just last month I received a new volume in which a papyrus in Paris is shown to be another sizable part of the same document. Some parts are apparently still at large! Even when individual pieces are not divided, archives almost always are, and papyri from individual localities are spread over many collections. As a result, papyrologists spend an inordinate amount of time and money getting photographs from distant collections, or travelling to them in order to work with the originals. As more and more papyri are published (the total is now about 50,000), the body of material becomes ever more unmanageable.

Papyrologists' traditional working tools, as is true in many text-based humanities fields, are dictionaries and bibliographies. These have, over the years, entered the electronic age with the digitization of the basic textual material—all papyrus documents can now be searched in the Duke Data Bank of Documentary Papyri—and the bibliography long produced in Brussels on 3x5 index cards, which has in large part been converted to electronic form at Columbia and is now wholly electronic. But the actual

“The concept of a distributed virtual library is easy to imagine but much more complicated to bring to reality.”

physical images of the papyri have long been the missing element. Five years ago I suggested at an international congress of papyrology both that digital imaging was going to offer us an opportunity to solve the problem of access to distant collections, and that software like Web

browsers would make it possible to knit together textual data and images to form a single scholarly and educational system. I got mostly puzzled stares; many Europeans still thought that technology was an odd American obsession. But at the classics meetings after Christmas the next year a younger colleague at Michigan asked when I was going to do something about it. Eighteen months later, an application was at the National Endowment for the Humanities, representing the six most important papyrus collections in the US: Berkeley, Columbia, Duke, Michigan, Princeton, and Yale. And in 1996 we began work, with a grant from NEH and a lot of local support, including private donations. We have now submitted our application for the second phase, to end in 2000, which should leave us with a fully working system and completed cataloguing and imaging at all except the two biggest collections (Berkeley and Michigan). In this phase 2, moreover, the most important collection in



A Papyrus in the Columbia University Collection

Britain, that of the Ashmolean Museum at Oxford (including the Oxyrhynchus papyri, the single most important and respected publication series of papyri) will join APIS with its own funding. They will be the first of what we expect to be a series of other collections wishing to join, whom we will supply with detailed technical specifications to allow this to happen without extensive work on the part of the existing partners. The system will thus be indefinitely expandable with only modest central effort.

CENTRAL METADATA, DISTRIBUTED DATA

What exactly will we have produced? From a technical point of view, we will have a central meta-database, with a catalogue record for each object in all of the collections; in that record will be a pointer to locations—usually on other machines—where the text (Greek, Coptic, Arabic, or whatever), a translation, an image (or multiple images, for larger objects), commentary, and bibliography can be found. To the user, all this technical wizardry, spanning many computer systems, will be invisible except for the line in the browser which shows the “address” (URL) one is currently visiting. The user’s view, instead, will be of an informative interface, offering many different routes into the material. The specialist, for example, might wish to search a Greek term or ask for a specific reference (P.Oxy. IV 815, for example, in papyrologist’s lingo). An economic historian of antiquity not trained in papyrology might want to find all the land leases, and searching by document type will be available. Someone interested in diseases of the eyes could search the translations for “eye” and related vocabulary. A student writing a paper on taxation could search the subject field of the catalogue records for appropriate terms. At any point one will be able to move from text to translation to image to bibliography, whether to follow up a discovery with similar texts or to check the image to see if the papyrus was read correctly by the editor.

A lot of hard work lies between us and this vision, although we hope to have an experimental first version of the system up this fall. Most of Columbia’s collection—and it is not alone in this—has never even been catalogued. Some has hardly been looked at. Associate Curator Raffaella Cribiore ('93Ph.D.) is creating the catalogue papyrus by papyrus, ostrakon by ostrakon, with the help of consultants

in Arabic and Coptic. Graduate student assistants Maria Wenglinsky (Classics) and Anna Trumbore (History) have been removing papyri from thick glass sandwiches with oozing tape (vintage '60s) and reframing them in lighter glass, then making digital images with the Phase One digital camera acquired with support from the Provost's Strategic Initiative Program (directed by Vice Provost Michael Crow). This wonderful device—a scanning back attached to a Hasselblad camera body and lens—will take an image with 600 dots per inch in each direction (360,000 pixels per square inch!) of an object up to a foot square. Even a CD-ROM will hold only four images at that rate. The ostraca, long wrapped up in ratty cotton wool and packed away in rarely-opened boxes, are being rehoused in drawers where they can be found instantly; in the process, long-separated fragments are being rejoined.

On the technical side too there is much to be done. The concept of a distributed virtual library is easy to imagine but much more complicated to bring to reality. David Millman of Academic Information Systems, working with other colleagues in AcIS and the University Libraries, is leading this effort in collaboration with his counterparts at the other institutions. Getting six institutions to agree on standards for representing and storing data is a big challenge. Today's NEH can't afford us a budget big enough to pay for travel to talk face-to-face, so we work by e-mail.

The result, however, should be well worth the labor. The opening up of these collections to anyone connected to the Internet around the world will be a dramatic

change. It is not only papyrologists who will experience the transformation of how they work; they have at least been able to look at the papyri freely if they visited a collection. For the larger public, even within academe, these objects have been hidden in two ways: the originals have been locked up and invisible except in the occasional display; and the publications have been almost as impenetrable, rarely designed to convey information to anyone but a papyrologist. APIS will provide tools to open up this immense body of information to those who do not know the languages in which the papyri are written, who have not been trained in the technical arcana of papyrology. We probably cannot—and should not—turn back the tide of increasing scholarly specialization, but we can do a lot more to make the results of that specialized study available to anyone interested. Electronic tools like APIS will help us to do that. ■

Roger S. Bagnall is Professor of Classics and History, Chair of the Department of Classics, and curator of the papyrus collection; he was Dean of GSAS from 1989 to 1993. He has led APIS as president of the American Society of Papyrologists.

“Getting six institutions to agree on standards for representing and storing data is a big challenge.”

A NEW SPECIES OF CHIMPANZEE? PERHAPS.

Professor of Anthropology and Biological Sciences Don Melnick and a team of anthropologists have analyzed DNA from chimpanzees in southern Nigeria and western Cameroon. Their findings show that the population is genetically distinct from the three recognized subspecies of common chimpanzee.

This could mean that a new subspecies of chimpanzees has been discovered or that the existing division into three subspecies is erroneous and a new family tree must be devised. Melnick and Katy Gonder, a graduate student at CUNY, conducted the field research. As Melnick put it: “The fact that the Nigerian samples all sorted out together in our tree and that they were well separated from the other named clusters, led us to the conclusion that we might have a new subspecies.”

Three subspecies of common chimpanzee have been



Professor Don Melnick

identified: *Pan troglodytes verus* is found in West Africa; *Pan troglodytes troglodytes* in Central Africa; and *Pan troglodytes schweinfurthii* in East Africa. Though there are scattered groups of chimpanzees in Africa west of Ghana and east to Nigeria, all genetic samples of the *verus* subspecies prior to this study had come from west of Ghana, which may account for their considerable genetic divergence from the others.

The study found as much difference between the Nigerian chimps and the *verus* subspecies as between the two other recognized subspecies. That, in itself, is not enough to declare a new subspecies, but it does throw the existing schema into question. If the Nigerian chimpanzees turn out to be not different enough to be regarded as a distinct subspecies, then the other groups of central and eastern chimpanzees now recognized as subspecies should only be considered populations of *Pan troglodytes troglodytes*.

Lab work for the study was supported primarily by the Center for Environmental Research and Conservation (CERC). The researchers, who are members of the New York Consortium in Evolutionary Primatology, point out that more time is needed before an informed decision can be made.