The Advanced Papyrological Information System: Past, Present, and Future

Roger S. Bagnall – Traianos Gagos

The Congress of Helsinki offered an opportunity to take stock of the work of the Advanced Papyrological Information System (APIS) project over the past nine years. We present this report here in three parts. The first is the response delivered by Roger Bagnall to a panel of three papers on the future of papyrological research tools at the Copenhagen congress of 1992. This was not published in the *Proceedings* of that congress, but colleagues' references to it during the Helsinki meeting led us to think that it would be useful to put it into the permanent record, because it was the first stage in the conceptualization of APIS. The second part contains Traianos Gagos' assessment of the work done by APIS up to the present and of the current state of the project. The third section is Bagnall's assessment of future directions for the project, particularly its integration with some of the other research tools in papyrology.

1. The Future of Papyrological Reference Tools [1992]

These three papers [by Dieter Hagedorn, Janet H. Johnson, and Theodore Brunner] offer a rich material for reflection. They give us accounts of the development and uses, both present and potential, of both textual and visual resources for our discipline, its neighbors and its context. In Dieter Hagedorn's remarks we have a valuable sense of the current thinking about the relationship of print and electronic resources by an experienced, balanced and trusted leader in the discipline. Particularly valuable for me is his frank admission that his own thinking about these matters has changed over time and continues to have a provisional character. Given the rapid change of technology and of our working habits, and given the unknowability of the future, a certain humility seems called for in confronting the questions we have posed today. That may be part of the wisdom Ted Brunner asks.

For my part, I will concentrate on two general problems raised by our panelists. First, I want to revisit some of the questions about the relationship between print and electronic tools that Hagedorn engages. I shall think aloud

a bit about the nature of different media of dissemination, and about how we actually work. These reflections lead me in part to conclusions similar to Hagedorn's, in part to more radical ones--a fact that will probably surprise no one in this room. Secondly, despite Hagedorn's warning and disclaimer, I shall offer some speculations about what electronic tools we might create in the next decade, building with the materials furnished by Ted Brunner, Jan Johnson, and others, along with materials yet to be brought into the electronic age. Some of you may regard this as science fiction, but I assure you in advance that it is nothing of the kind.

First, let us think together about the nature of printed books as reference tools. Dieter Hagedorn raises here and there throughout his remarks a series of points, which I think we may reasonably take as a group to be his criteria for whether or not a work should exist in print form. These are the following: (1) Fixedness or mutability of a text; that is, is a printed text needed as a permanent, unvarying point of reference? (2) Mechanical or thoughtful nature of the tool; that is, whether one has an assemblage of data or significant original intellectual and interpretive content; (3) Permanence or evanescence of material; that is, whether it will continue to be viable and available for a long time or deteriorate rapidly or unpredictably; (4) Comparative convenience of different media; and (5) the security of funding available for enterprises.

Of these, I would eliminate the last. The long-term availability of funding for projects from government or private grants seems to me an imponderable factor for all projects, electronic, print, or a combination. The funds available for such projects have in my country suffered a long-term decline in the past two decades, and yet major electronic projects have managed to flourish. We simply cannot tell what lies ahead. Some European academic economies have fared better, but I do not believe that any is invulnerable, and I see absolutely no reason to suppose that those who control these funds can be expected to support print projects in perpetuity. The currently disastrous situation for *L'Année philologique* seems to me regrettably exemplary of the threats that can emerge even for the most prestigious print-based reference tools.

The notion that books are a more secure repository of information than optical storage of data seems to me to rest on a confusion. It is certainly true that electronic storage does not of itself confer any immortality on data. Particular media go out of use or degrade. The life expectancy of CD-ROMs is completely unknown at this point, with industry estimates that I have

heard ranging from a few years to a half-century. But these points are significant only for two categories of material: those existing in unique copies, where physical degradation of a disk or tape would be disastrous; and those stored for long periods of time without being updated, used, or transferred to other media. It is exceptionally unlikely that either of these conditions is in any way relevant to any of the tools we are discussing, which are frequently updated and distributed in hundreds of copies.

The matter of fixity of printed text vs. mutability of electronic text, about which Brunner has also spoken, is much more interesting and important. There has been an extensive discussion in recent years – for American classicists associated above all with the name of Jay Bolter[1] – about the implications of the user's ability to change electronic files, whether that means altering numbers in a data file or rearranging and editing text. Electronic text is certainly neither fixed nor linear in the way that a book is. Some see this as liberation, some as threat, and still others as a great unknown. I belong to the last group. Not all electronic media offer changeability. Very importantly, CD-ROMs are not alterable by users, although there are other optical storage systems that are. If I copy a papyrological text from the CD-ROM of the Duke Data Bank to my hard disk, I can, as Ted Brunner points out, change it on the hard disk, and for myself I see nothing wrong with this; but it remains unaltered on the CD-ROM. The CD-ROM thus behaves exactly in the way that Hagedorn describes a book as acting, preserving a fixed text to which one can refer. On the other hand, it has been the policy of the Duke Data Bank to eliminate texts replaced by reeditions, thus removing them from the permanent record. It must be remembered, in addition, that the Data Bank does not record much of the information about a papyrus that a standard edition provides, such as Greek numerals, the form of abbreviations, and other palaeographic and diplomatic details. It is thus not a replacement for books.

The issue of comparative convenience of use is the issue most likely to appeal to the average user, who may find more abstract considerations uninteresting. I would describe this question generally in terms of the nature of effective use. The combination of printing with the book codex was one of the great technological achievements of humanity, and the book remains

[1] Among other works, Jay David Bolter, Turing's man: western culture in the computer age (Chapel Hill 1984) and Writing space: the computer, hypertext, and the history of writing (Hillsdale, N.J. 1991).

unmatched above all for reading. It also offers superior portability (without the aid of additional technology), high quality resolution (which is important for how well we learn, according to specialists in vision), and (at least historically) cheapness of reproduction. It is losing its advantage in this last respect because of many changes in the economics of technology, typesetting, printing, and institutional budgets, but it remains the case that we learn better from books and work better with them for some purposes. (I think that this will remain true even for a generation raised on television.) That we do not yet fully understand why this is so does not reduce the importance of such ergonomic factors.

Perhaps it is only because I come from a once-rich country, now fallen on harder times, that I think we must add to Hagedorn's criteria that of total cost--not just the cost to one user, but the cost to the entire system of scholarly communication. For all that the American economy has had some rough times and probably faces more, there are still many more countries poorer than the United States than there are richer, and the dramatic improvement of openness and freedom of travel and communication in much of the world just since our last congress will matter little if most people cannot afford to buy what is available to them. In the long run I am absolutely confident that the basic workings of a worldwide market economy will force all of us to consider the cost-competitiveness of the choices we make about how to distribute information and disseminate our scholarly work.

Where would these various criteria take us in considering the various tools that Hagedorn discusses? With respect to the *Wörterbuch* and *Namenbuch/Onomasticon*, I would agree fully with his conclusions: we do not need printed lists of citations, but we need as much as ever thoughtful lexicographic tools with selective citation and good analysis. These would be handy to have online, but my instincts tell me that most people will still, probably for good reason, want to use them in print form at least part of the time.

For the *Sammelbuch*, the case seems to me less clear. It seems to me obvious that it should at least be created in electronic form for direct loading into the Duke Data Bank (typing it all twice is senseless), as Hagedorn suggests. That these are texts does argue for reading them in printed form, but the increasingly insupportable cost of the printed volumes makes it hard to be enthusiastic about them. Users can, of course, print out individual texts or groups of texts from the electronic form in order to read them. It would

perhaps be a good idea to survey users to see what *they* think. For the moment, *non liquet*.

In the case of the *Berichtigungsliste*, I can see absolutely no argument for a printed version. An electronic version, linked to the Data Bank and stored on a CD-ROM, would offer immediate cumulative access, while the archaically produced volumes are now virtually unaffordable. Since this is purely a tool of reference, not one of reading, the considerations that weigh significantly in favor of a printed *Sammelbuch* seem to me irrelevant here.

The *Bibliographie Papyrologique*, now available in electronic form, has two uses: informing scholars promptly of what has recently been published, something it does better now than at any time in my memory, thanks to the energy of Georges Nachtergael and Alain Martin; and providing a long-term repository of bibliographic information. It is to me unquestionable that the latter use is much better supplied by electronic means than by filing cards. But for current browsing through a set of references, received six times a year, print strikes me as superior. I am inclined to think, however, that the index card has outlived its usefulness. The new citations, printed on standard paper, could be produced and mailed out much more cheaply, then thrown away when a new electronic version was received (perhaps annually or biennially).

Before turning to describe my dream machine, I will close the first part of my remarks by asking if others feel, as I do, that an encyclopedic dictionary of papyrology would be of extraordinary utility, whether to practiced scholars, students, or scholars in other fields – helping to open up our field to those interested in its riches, in the way that Ted Brunner describes as happening already as a result of the Data Bank. It is possible to imagine a reference work combining the virtues of the *Grundzüge* and the *Fachwörter*, in which one could find the most useful and recent discussions of particular terms or institutions. I can easily imagine its utility in print form, but of course it could be kept up to date on a regular basis through an electronic version. Perhaps some courageous person has already launched such a project, in which case I look forward to hearing about it.

Much of what you have heard this afternoon concerns individual research tools, whether those be the splendid historic print tools of papyrology or the dramatic new ones produced by the TLG, the Duke Data Bank, the Chicago Demotic Dictionary and others. I turn now to ask how all of these fit together. How, in other words, might we integrate the electronic resources now existing and yet to come into existence into a comprehensive system of

scholarly information for our discipline? Let me say at the outset that absolutely nothing of what I will be wishing for requires technology still to be invented or any other development not readily foreseeable. My system of the future would, it is true, require more memory and storage, and better displays and printers, than most of us now have available. But if anything at all is sure, it is that all of these will be available much better and more cheaply in a few years than they are now. I am confident that what I propose would work very well on a system that today would cost about $10,000 in the U.S.; by our next congress it will probably cost a little over half that. Nor is the diversity of programs and operating systems understandably deplored by Hagedorn a significant barrier. By 1995 there will be commonly available systems capable of dealing with DOS, UNIX, and the Macintosh operating system impartially. My proposal will, however, require a considerable investment in integration and in creating new electronic resources--an investment in human time and energy, mainly.

Imagine, if you will, a few simple and even trivial experiences. You sit down to read a papyrus, and you find a name unfamiliar to you. It happens to most of us regularly. You ask the computer to search for it, and it finds two examples. As an experienced editor, you are naturally suspicious. You move the cursor to the name and hit a function key, or perhaps move your mouse to it, pull down a menu, and start a search. After a few seconds you learn that there is nothing in the *Berichtigungsliste* about this line. Another search produces a blank for the *Bibliographie Papyrologique*, too. You check the other reference and have the same result. Now you are really doubtful. You click again, and this time up comes in a window a photograph of the papyrus; you move to the name in question, then enlarge the place until it is a convenient size to read. You do the same with the other passage. This time you find it is damaged; a few quick actions enable you to enhance the image or draw in some missing strokes to see how plausible your idea for a reading really is.

Good fun, and a lot more efficient than the way we work now. It will of course not remove books from our lives. Texts lead to commentaries, bibliographies to articles, electronic imagery to much higher-resolution photographs where needed. It requires integrating the Duke Data Bank and the *BP*, adding an electronic *Berichtigungsliste*, and adding as well a large corpus of photographic imagery scanned at a good resolution--the International Photographic Archive of Papyri would make a good start. The software tools involved in such a system all exist now, though they do not to

my knowledge exist in a single package, and I do not wish to minimize the work necessary to make all of these pieces work together.

But that is not all. A well-designed modular system would allow the addition of other material as well. Another function key might take one to an electronic *Wörterbuch* for a definition (or even parallels--sometimes searching the Data Bank turns up too much); another one to the corresponding entry in a Demotic dictionary or onomastic repertory; another to the encyclopedic dictionary I would like to have. A flick of the wrist and Willy Clarysse's date-computation program will tell you what the Julian equivalent of Choiak 29 in year 13 of Antoninus Pius is. A modified version of Dieter Hagedorn's database might tell you what other papyri of the same decade are available in photographic imagery or plates. The *Prosopographia Ptolemaica* or Dominic Rathbone's prosopography of Roman Egypt will tell you if an individual in a given text is attested elsewhere. Such a system, properly maintained, would accommodate new tools as they are created, growing constantly in utility, always evolving. The extraordinary burst of electronic creativity on the part of members of our discipline who a decade ago knew absolutely nothing about computers makes me absolutely certain that more tools will come our way.

When I have been describing this conceptual system for a while, I have to stop and remind myself that it is only a concept; this is not a new product announcement, and this is not even vaporware! The problems that stand between us and its realization, however, are not those of technology. We require time, money, leadership, and collegial cooperation. These are not insurmountable obstacles.

2. The First 8 Years of APIS (2004)

What in the early 1990s papyrologists was seen as science-fiction vision, in the past 8 years has become a Janus-like project that brings together the past and the present of our field through the technology of the future and unites the papyrological universe at the speed of light. Though still relatively young, APIS has already taken us to places and collections that few of us had seen before. APIS and its forerunners exemplify in electronic form the long-established *amicitia papyrologorum*, a commitment to friendly scholarly symbiosis and a loyal effort to share information continuously and globally.

In-house electronic cataloging, digital imaging, and preservation of the University of Michigan papyrus collection began as early as 1991; the original goal, believe it or not, was to share information on disks! Between that time and 1995, developments in digital technology and the growth of the Internet attracted the attention of a growing number of Papyrologists. In 1993-4 Duke University began cataloging its collection in US-MARC format. In 1995 the Michigan Papyrus Collection tested on the Internet several images captured with different media hoping that other universities would follow suit in making their collections available in digital form. It is in this period that Roger Bagnall, President then of the American Society of Papyrologists, presented the proposal for the creation of an integrated electronic system at our congress in Copenhagen. APIS was formally established in 1996.

Conceived originally as a cooperative project among the six larger papyrus collections in the U.S., APIS, now well in its Phase 4, has evolved into a global consortium effort that at present encompasses virtually all North American institutions with papyrus collections. American participants are Columbia University, Duke University, New York University, Princeton University, Stanford University, The University of California at Berkeley and its regional partners, The University of Chicago, The University of Michigan, The University of Pennsylvania, Yale University, Washington University, and the University of Wisconsin.[2] But APIS is no longer simply an American project. It has several important European partners that have committed to the idea of data and image integration: The Halle-Jena-Leipzig consortium project, the Herculaneum Papyri Project, the University of Oslo (images and data from which will appear on APIS before the end of 2004), Oxford University, Heidelberg, the University of Toronto and the University of Crete in Greece. As the project grows, new partners join our efforts, some from collections almost forgotten or neglected. I would like to mention in particular two collections here: (1) Interest in the Tebtunis papyri at Berkeley revived as a result of APIS; (2) the collection at the Hermitage in St. Petersburg, Russia, which has been completely inaccessible to the scholarly community for the past 80 years. Thanks to the efforts of the curator, Dr. Mariam Dandanaeva, APIS will soon make available

[2] This last collection will be transferred to Michigan where it will be conserved, scanned, and cataloged. This is one of the services that APIS partners with experience and appropriate personnel can offer to smaller collections that have no papyrologists.

approximately 70 images of the best preserved papyri. And we will have the opportunity to hear more about this collection during our visit to St. Petersburg.[3]

Funded primarily by the National Endowment for the Humanities (NEH) and the participating institutions themselves, APIS seeks to integrate into a global "virtual" library the holdings of all papyrological collections worldwide. The project encompasses both preservation of a large and important body of ancient manuscript material, through conservation and imaging, and improvement of intellectual access to this material, through cataloging and an innovative electronic system linking catalog records with images, Greek text, bibliography and published literature.

Underlying APIS are two main goals: first, to transform instruction and research in papyrology; and, second, to make papyrological material readily accessible to non-specialists. The vast resources of the papyri have, as we have all frequently complained, been used relatively little either by scholars of most fields concerned with antiquity (literature, history, philosophy, religion, archaeology) or by a broader educated public. Changing that situation is central to APIS's existence.

APIS and the homepages of the individual institutions have a unique opportunity to reach out to wider audiences. In particular, APIS can provide resources and tools for students in secondary and tertiary education to attract the next generation of papyrologists, Egyptologists, ancient historians and philologists. A gradually increasing number of college students in the US has in fact been using APIS, while at my institution, the University of Michigan, we have began creating special resources for secondary education students (http://www.lib.umich.edu/pap/k12/k12.html).

What is distinctive about APIS is that the institutions involved have adopted and implemented collectively a set of standards for imaging, for the format of cataloging and for the linking of the several types of electronic data generated. From its conception the entire project was carried out with a view to the creation of an integrated information system, available on the Internet and also in other forms, as the technology evolves. Thus, the cooperative aspect of this project is central to its existence, for it is gradually replacing a world of incompatible, separate systems, each with its own standards, with that of a single, seamless system that is readily usable not

[3] Postscript: A grant from the Dorot Foundation will make it possible to extend work on the Hermitage collection during 2005.

only by a handful of papyrologists but also by thousands of scholars and students at all levels of education and in a wide range of other fields.

The U.S. partner institutions have altogether about 50,000 items, of which about one-eighth have been published to date. The APIS interface gives access at present to more than 20,000 of these items (including several thousand unpublished objects), plus documentary texts (through links to the DDBDP), several thousand translations, and images for varying smaller number of items, in an integrated system available over the WWW.[4] By the end of 2004, this number will increase substantially, since Alan Bowman and the Centre for the Study of Ancient Documents in Oxford will provide for APIS several thousand images scanned from photographs taken as part of the missions of the Association Internationale de Papyrologues to the Egyptian Museum in Cairo.[5] By the end of APIS 4 (June 30, 2005) we hope to have several thousand more images from collections such as Berkeley, Chicago, Michigan, New York University, Pennsylvania, Wisconsin, the Hermitage, and Oslo. The system also provides a manual, specifying standards for supplying metadata, including remote image links, dates and personal names cataloging methods, and database design. This is freely available in the central APIS webpage for all those who wish to become partners.

We shall take now a cursory, but more specific look at the elements of APIS.

(a) Catalog of papyri and ostraca, in a format acceptable nationally and internationally to bibliographic utilities. This is the heart of the system. The base format used in the central system is the US-MARC record type in its specific version for manuscript collections (AMC), using standard Library of Congress Subject Headings and *Art and Architecture Thesaurus* index terms. For the most part, these headings do not correspond to traditional papyrological classifications, but substantial additions to these standard repertories of headings have been made and an on-line thesaurus has been created. Librarians at Michigan are currently working at standardizing further the use of subject headings in the local database and creating a more homogeneous thesaurus that can be used by the old and the incoming partners.

---

[4] For current statistics, see http://www.columbia.edu/cu/libraries/inside/projects/apis/statistics/.

[5] Funding for this project comes from the Andrew W. Mellon Foundation.

(b) Images of the majority of the papyri that have been cataloged, at present nearly 14,000 unique images. As digital production continues, several thousand more images will gradually appear both in the central interface site and in the homepages of the individual institutions. With some exceptions, most of the images reside in the corresponding local institutions; hence a request for an image listed on the record that is housed in the central system takes users to the server at the local institution. Users are allowed to view images at a variety of levels of resolution, ranging from simple thumbnails to jumbo 600 dpi. For views of the entirety of large documents, 300 dpi images have been captured, but 600 dpi or higher have been used for images of portions of large originals. Papyrologists using these images in conjunction with image enhancement software, have consistently found them much more valuable than traditional photographs, either in black and white or in color, and 600 dpi provides sufficient resolution for almost all purposes.

(c) Textual links, drawn from the Duke Data Bank of Documentary Papyri. All published documentary texts from all the collections (including the new partners), up to a few months ago as for instance *SB* XXIV, *CPR* XXII and *P.Dryton*, are now included in the DDBDP. All the published texts on APIS have now established, permanent links to the proper text of the DDBDP. The texts in the DDBDP that were encoded after June 30, 1996 are not searchable at present, but this temporary problem will be solved by cooperative efforts between Duke University and the Perseus project.

Adding languages other than Greek and Latin can be handled by entry in coding, with tagging that allows user software to display the proper font if the user's computer has the font in question. Relatively few of these texts (Arabic, Coptic, and Demotic, in the main) from participating institutions have so far been published. The limited funding available for cataloging such material has also contributed to making this task a less pressing issue that it had originally been expected to be. However, the issue has been addressed and the University of Chicago has begun work on its Egyptian texts and the Arabic papyrologists, as we note below, are embarking on their own digital projects, with which APIS will collaborate.

(d) Translations: The central database accommodates also translations when provided by partner institutions. So far three of the major contributors to APIS, Berkeley, Columbia, and Michigan, do provide translations, and plans are under way for inclusion of more translations of published texts in the next couple of years.

In the course of the past eight years APIS has made big strides. The original partners, together with those joining the project subsequently, have tried hard to make life easier for all papyrologists, but also to create tools useful for additional projects interested in digitization. APIS has created standards that are followed in projects beyond APIS and has developed data entry and retrieval tools for cataloging papyrus collections. For data entry the new partners can choose three different templates running on different platforms: the Filemaker for Macintosh, used at Michigan; Access for Windows, used at Berkeley; and the APIS central interface format. They are all now fully compatible and freely available for all those who wish to join the project. APIS has also helped smaller collections in cataloging, scanning and preservation: here we should mention Berkeley with its regional partners on the west coast and Michigan in the mid-west. This effort will continue. This outreach is international: APIS has already assisted several partners outside the U.S. which have no local experts or support available. I already mentioned the Hermitage collection. I should add the Oslo collection and other collections that have profited indirectly from the standards and templates created by APIS such as the Jena, Halle, Leipzig project.

We invite all potentially interested partners to join APIS and to use our programs, templates and experience gained so far. With APIS the field of papyrology is setting new standards in preservation, access, research, and instruction by creating the first global repository of ancient manuscript material. No other field had achieved so much in so little time and no other field has been such a keen technological leader.

3. The Future of APIS

APIS, as just described in section 2, grew out of Traianos Gagos' challenge to me to do something about the vision I rashly projected in Copenhagen 12 years ago (part 1, above) of how we might move from the disjointed situation as I saw it then to one in which our working environment was more integrated. APIS has grown up as a project based on institutional collections, including much unpublished material. But the original vision of bringing together the different parts of the working tools of the discipline, which mainly concern published texts, still is far from being realized. Here I shall reflect on how the possibilities for realizing that vision look in the environment a dozen years after the original proposal.

The goals have not changed significantly. They include (1) the integration of papyrology across linguistic and disciplinary boundaries; (2) the integration of various types of research tools into a working environment of maximum convenience that uses the technology of today, not that of a bygone era; (3) long-term stability of financial support for maintenance of tools projects in the face of many recent threats, particularly in Europe – making my remarks of 1992 look unfortunately prescient – and the inherent instability of grant support; and (4) the widest possible access for everyone, everywhere, to the richest possible array of text, data, and images. These last two goals, it must be added, may be in competition with one another, for long-term stability is best assured by earned income, and earned income means charging for information rather than giving it away – that is, instead of moving part or all of the burden onto the user rather than the producer.

The research resources available in digital form to papyrologists today are much richer than in 1992, and more are on the horizon.[6] The older members of our discipline, who grew up without such things, know how much difference these tools have made to our daily lives. I know that I personally have undertaken many research projects that I would never have attempted without these tools.

In the realm of what today is usually called metadata—the cataloguing of external information about written materials and the physical objects on which they are inscribed – we have the Heidelberger Gesamtverzeichnis for Greek documents, the Leuven Database of Ancient Books and part of an electronic Mertens-Pack third edition for literary texts (LDAB being about to add Coptic books to the Greek); the Ptolemaic documentation collected in the *Prosopographia Ptolemaica*; and a database of Egyptian texts, both literary and documentary, in hieroglyphic, hieratic, and demotic now planned by Mark Depauw. The Leuven Homepage of Papyrus Collections also deserves mention in this category of metadata.

On the textual front, we have most notably the Duke Data Bank of Documentary Papyri, a demotic project underway under the direction of H.-J. Thissen, and active discussions about a Coptic databank. At the same time, reports at the Helsinki congress showed how much progress is being made in the domain of the paraliterary papyri. Bibliography is now abundantly

[6] We have not given online addresses for the projects mentioned in what follows. They can be found easily in the "Links" page of the AIP web site, http://www.ulb.ac.be/assoc/aip/liens.htm.

available in the digital Bibliographie Papyrologique, taken back to its origins, and the list of editions of papyri, ostraka and tablets, and their abbreviations, is online in the electronic version of the *Checklist of Editions*. There are active plans to create counterparts to these various tools for the Arabic papyri, led by Andreas Kaplony.

Amid all these riches, however, there are aspects of the present situation that seem to me scarcely an improvement on the situation in 1992. The new projects have largely been conceived with incompatible characteristics and data structures. The era of Filemaker and the consequent easy creation of databases has allowed a high level of amateurism. The task of integration has not been made simpler. At the same time, in many respects the existing projects still seem to me to betray an emphasis on managing information of a traditional type, i.e., the tools are still seen to a large extent as a way of finding printed information. Moreover, the problems of sustainability of the separate projects, with constant threats from turnover of personnel, institutional funding cuts, and government and foundation unwillingness to provide permanent support, have by no means eased since 1992, even if they seem just at this moment worse in Europe than in the United States, the reverse of the situation a dozen years ago.

But there are also benefits to this kind of diversity. For one thing, people work differently; imposing a single approach to information would probably harm more than it would help, in the sense that no one person can be wise enough to legislate for all. That is why even seemingly divisive forces have benefits. An integrated working environment will have to offer options if it is to achieve its maximum benefit to the papyrological community and the wider range of potential users.

In June of 2004, a group of representatives of projects and institutions responsible for digital information in papyrology was convened at the Papyrological Institute of Leiden University.[7] It sought to consider the problems and issues set out above, in order to make sure that planning for integration could go forward with a full understanding of the situation as seen by a full range of participants. In the intense and cordial discussions there was by no means complete agreement on everything, but there was an unmistakable universal desire to work together towards a more integrated future.

---

[7] We record with gratitude that this meeting was made possible by financial support from the Andrew W. Mellon Foundation and hosted by Klaas Worp.

We did, however, agree on a great deal. The main lines of these agreements are the following:

(1) We need to manage the coexistence of the very real political, institutional, and intellectual imperatives for the continuation of the separate existence of both traditional and newer research tools with the desire for a cooperative creation of an integrative database.

(2) This balancing act probably implies the necessity of producing a capacious central database drawing from numerous projects and offering multiple forms of display, ranging from a wholly integrated one drawing data from many projects to purely project-specific views.

(3) The scholarly-political dimension of what we're doing is primary; technological and financial aspects are secondary. If we can agree on what is needed and how to govern it, the technological challenges can be solved and funding is far more likely.

(4) We need to accommodate at least the three languages dominant in the digital tools we already have (English, German, French). The most likely approach to this question is through some sort of invisible thesaurus; work on such a thesaurus is urgently needed.

(5) At least for the foreseeable future, an integrated system probably will not have everything on one computer but draw images (and perhaps text) from servers wherever they are. But an integrated system itself needs multiple locations for convenience of access and reliability.

(6) An integrated system, drawing on resources in many countries, needs international governance with representation of the full range of agents involved in the creation and maintenance of our research tools.

(7) The financing of long-term stability remains very difficult; relying on earned revenue from charging for access is problematic for many reasons (including the diversity of ownership of intellectual property), even though most traditional print versions have always charged, most of those on distributed digital media come only at a price, and some online ones currently require subscription fees (e.g., Tito Orlandi's Coptic manuscripts site). Participants generally did not favor charging for access to a universal system. But further consideration is needed of the other options. These include mainly institutional budgets. At any given moment these are in diverse conditions, some vulnerable, but there is some strength in the fact that they do not vary in the same direction over time. We have also considered sponsorships and advertising, although the potential for these is untested.

(8) We need to move in the direction of digitally conceived and initiated types of information and away from mopping up information from print sources. Ultimately, it would be highly desirable for information (especially texts in ancient languages, along with translations and commentary) to be typed only once in a form that would allow it to be captured for all time and for multiple uses. Duke University is exploring the possibility of creating a data-entry tool that would lead in such a direction.

The Leiden meeting also agreed on several concrete steps to take next:

(1) As a test project, the central APIS technical staff at Columbia will work with Heidelberg and Duke on creating a site integrating the HGV and APIS, along with a connection to the DDBDP, so as to have all Greek documentary texts in a form allowing the ready integration of text, metadata of several types, and images.

(2) In the context of that demonstration project, an English-German thesaurus will be created to allow users to work in either language.

(3) Over the next few years, more translations of texts will be added to this integrated site (following the APIS model; but possibly in some cases in both German and English).

(4) We will work on resolving differences among the projects over the bibliographic and cataloging unit of analysis, which varies considerably, as anyone who has tried to define what should get a separate publication number knows.

(5) The next priority once the question of the unit of analysis is resolved will probably be an integration with LDAB.

(6) A restrospective conversion of the *Berichtigungsliste* into digital form, with provision for future digital availability linked to APIS, is under active consideration.

Generally, the Leiden meeting found itself working toward a sense that an integrative tool was likely to be a flexible information gatherer capable of working with different data structures and presenting its results to the user in a perhaps deceptively coherent format, rather than a single, rigid database structure into which everything would have to fit. It has to be admitted that we do not yet know exactly how such systems will work, but it seems clear that the future of the Internet does not lie in increased uniformity and central management. Instead, it lies in devising clever tools for collecting and presenting data from multiple sources. Papyrology will undoubtedly do best to go with that technological and sociological trend.