

Yale, AMEEL, and the Prospects of This Workshop

Building bridges of understanding to and from the Middle Eastern world is a fundamental task of our time. For a broad range of societies in the Middle Eastern and Arab worlds, the time is ripe for making important progress in building libraries of digital resources that embody the cultural heritage and social ambitions of those worlds. That is why we are here.

My participation in this workshop will be focused on selecting, organizing, and preparing *scholarly, text-based content for such libraries*; these are our targets in the projects described below. Others are here to speak of the multiple challenges of building technical and social infrastructures that will actually deliver the content to a broader public.

The vision that we see emerging in this workshop is one that we have been exploring for several years now at Yale, and we have already received on three occasions U.S. federal funding support for such projects. In this paper, I will describe a bit of what we are doing, say why we are doing it, and then extract from this report some suggestions for both vision and mission statements that could emerge from our conversations here. The work we are doing at Yale is deeply rooted in our institution's culture and strategies. In particular, Yale's President, Richard Levin, has been bold and outspoken in committing Yale to a new vision of international engagement for the benefit of global society.

Yale Library's Middle East Digital Initiatives

Yale Library's first Middle Eastern (ME) digital project was modest in scale but important for enabling us to explore the fundamental issues any such project would face. Project OACIS (Online Access to Consolidated Information about Serials) brought together metadata on a broad range of serial publications in Arabic and Islamic studies in an online representation that makes it possible for researchers around the world to discover what is available and to identify libraries where the material in these journals and other serials can be consulted (see <http://www.library.yale.edu/oacis/>). OACIS began in late 2002 with about \$460,000 of funding over three years from the U.S. Department of Education (Title VI TICFIA program). As of December 2005, OACIS contains 38,747 bibliographical records (of which about 14,000 are unique titles) and 51,278 holdings records. The database now includes serials published in or about the ME in 66 languages: the top three non-Western are Arabic, Turkish, and Persian. The group of project partners and participants are 19: 14 in the U.S., 4 in the ME, and one in Germany. We think of OACIS as a kind of pilot for the deeper involvement we are now continuing.

We have just been awarded about \$100,000 from the NEH to digitize some 100,000+ pages of Iraqi scholarly journals in Arabic, using OACIS as a metadata building block.

Additionally, our new major project, AMEEL (A Middle Eastern Electronic Library: <http://www.library.yale.edu/ameel/>) was funded by TICFIA in late 2005 and will run for four years for a total of around \$760,000. AMEEL seeks to lay the foundations for future progress in building an actual working library by a twofold strategy:

- 1. To integrate existing scholarly digital content** within a prototype Web-based portal for study of the ME, including its history, culture, development, and contemporary face so that such

material is easier to find and use efficiently and freely. Via this portal, scholars will be able to search, retrieve, and view a significant portion of our partners' currently available textual ME electronic resources, including bibliographic records, selected abstracts, full text, and other formats in Roman and particularly non-Roman, vernacular scripts

2. **To develop an environment for creating new digital resources**, as well as integrating existing resources, into a common structure. We will also tackle in particular the challenges of scaled up Arabic digitization and the opportunities for technologically assisted interlibrary borrowing and lending.

The first stage of AMEEL will establish an 'integration infrastructure,' both to create a useful resource rapidly and to put in place an architecture for future development. AMEEL will then work with project partners to articulate agreed-upon international standards for digitization of Arabic-language scholarly materials and to support the efforts of other would-be providers of digital information in Arabic. Training programs in Arabic digitization, along with digitization of selected scholarly content, are part of the AMEEL project deliverables. And, finally, AMEEL will conduct a pilot Interlibrary Loan project with a small number of libraries in the U.S., Europe, and particularly the ME, to prove the viability of the digital exchange of materials or technically facilitated representations of those materials (faxes, photocopies, photographs) where digitization has not yet occurred.

To create AMEEL, we will draw from international (US, European and Middle Eastern) expertise of academic and information technology specialists to produce a gateway to knowledge that is more than a collection of links. The search engine and its underlying data architecture will scale with the addition of new resources, while retaining its power to organize and assist the reader's investigations. In addition, we will bring deep expertise in handling technical challenges related to non-Western languages such as Arabic into the U.S. university arena. As a result of our partnered effort, (1) information creators in U.S. institutions will become proficient in Arabic text digitization and character recognition, qualified to lead digitization, preservation, and document delivery efforts, while (2) participating in the development of a digital library from which students and scholars may gain more comprehensive knowledge about the ME. In short, we plan to create mechanisms, skills, and information resources to be shared widely with others, in academe and beyond. Our ME digital library will go ambitiously beyond providing a proprietary subset of information with *ad hoc* interfaces that have, however broadly they "cover" the subject matter, an ultimately narrowing effect for the way they limit the reader to the horizons of the provider's resources.

Our plan begins with bibliographical records, initially those developed in our earlier OACIS project. To these we will add full text content, followed by services for users. Early exploratory discussions with producers of fully digitized scholarly resources in and about the ME have been conducted with publishers such as MultiData Online (Beirut), Brill Academic (The Netherlands), Oxford University Press (UK), JSTOR, and others who have agreed to work with and support AMEEL. The time is ripe to create an appropriate test bed to support a robust, rich, and well-structured portal to ME digital information, arranged in ways that make it a genuine digital library for students, scholars, and information seekers at every level.

In the AMEEL partnership, a core of subject experts will debate, define, and refine the conception of the project's range of coverage. While an ongoing, international advisory body, comprising the core partners, will have ultimate "board of directors" authority over the presentation and organization of the project, a larger group of librarians and scholars will be called upon for guidance on specific or special areas of endeavor.

The component of Project AMEEL that integrates distributed resources will be adaptable and extensible as (1) new ME e-projects emerge and are chosen for the portal; and (2) OAI harvested materials grow in number. After working with our partners and advisors to make fundamental choices about data architecture and protocols, we will proceed to identify 2-4 initial, significant full text projects, to be linked and made accessible and interoperable. This segment of the enterprise will require the most sophisticated technological work and will form a basis and proof of concept for later resource integration.

Likewise, our plans for digitizing selected resources will require the input of other experts such as ME scholars or leaders of other projects that have developed methodologies and priority lists for digitizing periodicals and/or other sources. See, for example, the work of the CAORC¹ in its IMLS-funded Middle East Research Journal project, MERJ.² CAORC is a collaborator with AMEEL. We plan to work with them in the creation of new digital works under this project.

Once we can measure the size, shape, and feasibility of AMEEL, we will address policy questions related to access and equity of access, to assure that the broadest range of resources is available to the broadest range of users, from western scholars to students and teachers in countries with fewer traditional paths of access to western library resources. We cannot resolve all questions of policy and access, but discussions of this nature will play an important role in our work with our international and publishing partners.

Summary: AMEEL Will Fill Two Key Needs

1. **Digitization:** Scholars value information about resources available in traditional libraries but what they increasingly prefer, for reasons of utility and productivity, is access to online full-text content. Catalog information about materials no longer satisfies today's scholarly users. But publishers and vendors will never digitize many older full text items such as books and journals from the ME. The business case is not strong enough: there is not a large enough paying market for such materials and the costs of digitization are significant. These materials, however, particularly many journals, are vitally important for education and research. AMEEL will digitize a significant sampling of such materials.

2. **Interlibrary Loan (ILL):** Scholarly information in libraries in the ME has to date been either extremely difficult or altogether impossible to identify and obtain from ME libraries. ME libraries have important and unique collections not available in the U.S., and they do not have a tradition of sharing resources through ILL, which means that not only is sharing among their libraries limited, but also that U.S. libraries cannot borrow materials from those libraries except in rare circumstances using personal connections. ILL involves policy at an institutional level, development of procedures that fit institutional workflows, an understanding of copyright law both nationally and internationally, and the potential to expand a circle of participating libraries both within a country and beyond its borders. AMEEL will lay the further technology foundations for such ILL and will promote a culture of ILL-based resource sharing among libraries in the region.

Obstacles

¹ CAORC is the Council of American Overseas Research Centers <<http://www.caorc.org/>>.

² The Center for Research Libraries in collaboration with CAORC has initiated the Middle East Research Journals Project or MERJ <http://www.crl.edu/FocusArticles/MERJ_Article.htm>.

During the course of the three-year OACIS Project, we encountered certain difficulties in executing our work; surmounting these has provided invaluable experience as we move forward with our newly-funded projects, and a brief listing may have some utility. We have no magic to bring to these issues but our expectations are now more realistic and we have developed at least some connections and contacts that will advance our work. Note that most of the problems do not require from us any advanced technological solutions.

1. **Political and sociopolitical.**

Year One of OACIS coincided with the Iraq incursion, which complicated national and international relationships. For example, we were bringing ME interns to Yale, for partnership, metadata loading, and development work. Some visas were relatively easy to obtain; others were very time-consuming and labor intensive. The visa permissions process appeared almost totally arbitrary; for example, one of our interns from the BA received his visa in one day and the other took months. This had a domino effect at our end as we attempted to schedule such basics as housing, health coverage, and orientation for our interns, let alone to move forward on schedule.

The OACIS project does not include Israeli content; the decision to proceed in this way was very difficult; developing lists of countries and languages to be included has been fraught in various ways, and explaining this, however rationally (we think), can leave inquirers at best unsatisfied and at worst angry.

2. **Urgency and Priority.**

Cooperative projects are often not the highest priority for participants. OACIS depends on regular (quarterly) dataloads, but many of our partners – in the U.S. as much as anywhere – have had difficulty providing the needed records (Yale is one of these!) or responding to queries in a timely way. Yet our schedule depends, to a large extent, upon the efforts of others. In our attempts to survey key ME libraries about their ILL and technology situation, out of dozens of surveys sent (by mail, fax, and e-mail – all three; following personal contact by phone), only six were returned.

3. **Technological Issues.**

Disparities. Online catalogs and technological advances did not have the same meaning for our partners in the ME as they did for us. OACIS depended on loading metadata records into a database. But online catalogs for library partners such as the University of Jordan and Tishreen University in Syria bore no resemblance whatsoever to populated MARC or Dublin Core records. Promised bibliographical records from two major national libraries have yet to materialize.

Also, equally seriously, the libraries in the ME whose users would most benefit from access to OACIS and then AMEEL do not have the network connectivity or bandwidth (cannot afford it) to access the database.

Standards. Agreeing in advance to standards and then conforming to them, is absolutely essential to the success of such a project. For example, we have found that Unicode does not miraculously support digital representation of Arabic materials, and that the technological solution alone is incomplete and imperfect without previous agreement on best practices. Digital scanning of Arabic, moreover, faces the preliminary challenge (better addressed at the BA than anywhere else in the world we have found) of accurately capturing a wide variety of fonts, scripts, and diacritical marks (often, even in printed works, hard to distinguish from blots, paper

imperfections, and the like). The accuracy of OCR scanning for Arabic remains in the best shops well below what we take for granted with European alphabets.

Desired Outcomes for This Workshop

In preparing to return to the Bibliotheca Alexandrina and its commanding view of the ancient harbor, I have made my own list of criteria for success of this workshop, a list I hope will differ as little as possible from that of others. It is to:

1. Leave with a good understanding of what are the major ME DL initiatives happening today and where. Who are the players?
2. Leave with a sense of what we see as our greatest challenges and how to deal with them.
3. Agree that there are no monolithic or single sources or solutions to an ME DL, but that facilitated and distributed collaboration will bring us forward as rapidly as possible.
4. Recommend a non-directive vehicle for facilitating that collaboration. Can we imagine seeking funding for a series of these workshops, for a continuing working group, for a clearing house? Perhaps the ideal means is a coffee house of the Egyptian or Viennese variety, where all who work in these fields will *want* to visit regularly, to share the intellectual excitement, and in doing so, share in creating an ever-expanding community of participants in the common project with a minimum of cost, effort, and superstructure imposed on the good work and eager collaboration of participants.

Vision and Mission

As we advance in our ME projects, we have been guided by a larger sense of vision and mission that I would summarize as follows at the end of this document. We present these texts here as drafts and suggestions of what we might discuss in Alexandria, for we very much see the larger project and set of projects that will be discussed as achieving on a global scale what AMEEL has begun to undertake from a single lead institution. Yale would warmly welcome working with an extended set of partners on a fuller and more ambitious set of projects designed to achieve our original goals and others besides.

A Proposed Vision statement:

We propose a distributed, open, and highly interoperable digital library of materials from and about the cultures and societies of the Middle Eastern, Arabic, and Muslim worlds that will make it possible for people all over the world to have easy and meaningful access to high quality information of every kind. We propose further that we collaborate with appropriate partners to build the technological infrastructure that can assure access to such resources and to the resources of the international research community generally from all these lands.

A Proposed Mission statement:

This workshop proposes creation of a working group based in multiple countries that will set as its goals the creation of content appropriate to the Middle Eastern digital library, integration of that content, and the amelioration of infrastructure problems in countries and regions where there is interest in accessing this material. We envision a distributed community of projects and sites, but at the same time we propose creation of an integrating function, with BA and partners, to coordinate projects and assure that the desired interoperability is achieved.

Ann Okerson
Yale Library
28 December 2005