Rethinking Collections in the Harvard College Library:
A Policy Framework for Straitened Times, and Beyond
Discussion Paper and Action Plan, June 2009

Summary

The Harvard College Library faces budgetary reductions without recent precedent. These cuts will compound a deepening structural mismatch between available resources and our community's expectations. Forces beyond the control of either Library or University mean, more generally, that our time-honored approaches to acquisitions and collections are no longer sufficient. The first section of this paper depicts our understanding of the emerging information environment. The paper then proposes redefined principles, goals, and priorities that will allow the College Library to respond to this information environment and also live within its means. It suggests an iterative and collaborative process with which to refine our approach, and summarizes HCL's short-term plans to cope with the budget cuts that have already been mandated for 2009-2010.

Introduction: our unviable status quo

The Harvard College Library is particularly known for its rich and varied collections. These holdings, carefully constructed over the centuries, are indispensable to both research and learning. Two sets of challenges now put this project at risk.

HCL's collections, while vast, are far from comprehensive. Moreover, our past acquisitions were never exhaustive. Even so, it is abundantly clear that the collections have in recent years lagged behind both current expectations and historic levels. Our persistently diminishing coverage contrasts with the increasing investments in libraries and collections at a number of peer institutions.

The HCL administration, the Director of the University Library, the FAS Library Committee, and individual faculty members and friends have all called attention to the enduring consequences of this decline. One-time cash infusions, administrative acknowledgements of our concerns, and commitments to a place in an eventual capital campaign provide some grounds for optimism. Nonetheless, our 2009-2010 collections budget will be more than 15% smaller than that for the current fiscal year. These reductions will compound the direct cuts and the erosion in purchasing power that have compromised our acquisitions for more than a decade.

We hope that additional support will eventually ease our predicament. In the meantime, however, our budget simply will not allow a "business as usual" approach. We instead need to revise our collections model to reflect the current limitations in funds and staffing. A process that is as open as possible should foster a shared understanding of what we can realistically achieve.

However, we need to do more than just prune and trim. Budgetary constraints, while a disturbing and very visible source of diminished collections coverage, are only part of the picture. Even had the budget held up, the Library could not now fulfill its mission merely by staying the course. The landscapes of information, scholarship, technology, and academic organizations are all in flux. New University programs and initiatives have broadened our mandate. Our collections strategy has to address a radically fluid context, brimming with new opportunities and demands.

This document briefly describes some key elements in today's collections landscape, and also offers a simple model for information types and their uses. This framework in turn suggests a set of principles to inform a redirected collections and content strategy for the College Library, whose initial implementation is then outlined in the second section.

I.   Universities, information, and library collections: an environmental scan

The information landscape: continuity and change

<u>Information:</u> The supply of information resources has mushroomed across all formats. Emerging countries and also traditional publishing centers are producing more than ever before. Recession and deflation may mitigate these trends, and some categories of publications—print newspapers are a likely example—may decline or even disappear. Nonetheless, predictions that hardcopy publications will soon be overwhelmed by an avalanche of electronic resources are emphatically premature, particularly in the developing world and other areas of emergent modernity. Analog materials remain both prevalent and indispensable, as the digital explosion continues apace.

Despite the persistence of print, large-scale digitization is transforming the library world. The scale of the resources available to Google and similar entities substantially exceeds that of any single research library, and also of the research library community as a whole. The immense range of materials available through Google has also shifted the value calculus by which we have traditionally judged our largest library collections. Libraries, along with other agencies devoted to our intellectual and cultural heritage, are experiencing a dual crisis of purpose and identity. Most students and scholars perceive less cause for concern.

The day's expanding array of information resources, in all formats, is complemented by intense price pressures that consistently outpace inflation. Publisher conglomerates, which already wield oligopolistic control over the STM (Scientific, Technical, and Medical) universe, are now expanding into other market segments. Greater outputs of increasingly expensive published materials characterize developing regions throughout the world. The weakened dollar, which lost about one-third of its value relative to both the euro and the pound between 2000 and 2008, particularly compounds the challenge for libraries with heavily international collections. Information may want to be free, but it's also a commodity—and scholarly resources are in thrall to the marketplace.

<u>Scholarship:</u> Research and teaching continue to evolve. Until recently, pedagogical models and research strategies privileged core or "canonical" writers and sources. In most fields, scholarship was considered

an orderly and necessarily cumulative enterprise in which new inquiries both relied and built upon the earlier studies that comprised the scholarly record. The bulk of research and learning was confined within rigid disciplinary boundaries, with each field claiming its own foundational literature and a unique suite of closely aligned methodologies.

Today's expectations are profoundly different. Cross-disciplinary inquiry, participatory learning, an obsession with primary resources and original documentation in all formats, and hybrid methodologies are increasingly the norm. The record of scholarship, while still important, has in many fields become less central. Multi-media research products, scholarship that relies upon massive and remotely hosted datasets, and team-based inquiry are other features of this emerging panorama.

Technology: In the past, libraries tended to acquire and warehouse hardcopy materials as passive objects for students and scholars to ferret out and then interpret on their own. Digital resources, by contrast, are energized from the start. The "if it isn't on Google it doesn't exist" catchphrase only begins to capture this dynamism. For a simple example, keyword searches across JSTOR can lead researchers to sources that would have remained invisible in a context limited by the traditional apparatus of field-specific bibliographies, indexes, and abstracts. On a more mechanical but likewise transformative level, linked footnotes among e-resources allow the seamless pursuit of citation threads that would be unrealistic to track through physically dispersed books and journals. Mash-ups and other digitally recombinant possibilities encourage projects that transcend an exclusively textual framework. Libraries are now called to help users by contextualizing all of these energized resources within a broadly activated system of information, tools, and expert staff. Integrating our deep stores of analog holdings into this high-energy e-network remains a central challenge.

Digital information predominates within some fields and is increasingly prevalent across the board. Digital technologies are also affecting scholarly inquiry and output, as well as teaching and learning. Large datasets—numerical data, text corpora, image banks—invite structured inquiries across masses of information at a scale that can easily exceed the capabilities of any single institution. Tools for analysis, manipulation, and visualization may best be developed as community efforts. The "cloud" is becoming the locus for more and more data and applications, in contrast to past models with their readily identified developers and sites. Community engagement is likewise the byword for social networking initiatives, whose shorthand is the panoply of "Web 2.0" products and services. Many new approaches to teaching and learning similarly rely upon open-source collaborative and participatory tools.

We do not yet understand the scholarly significance of large swaths of the digital universe. Blogs are often compared to diaries; e-mails are likened to letters and memos. The analogies are not only imperfect, but they may also complicate our decisions about what we need to capture and preserve. More familiar products like learning objects and computer software can be difficult to assess. Websites are typically dynamic and multi-layered, requiring thoughtful protocols to determine what to retain. Social networking spaces are again unfamiliar. Instant Messages or cellphone videos pose challenges of their own. Scholars, users and creators, technologists, librarians, and digital objects themselves, all have roles to play in clarifying our possibilities and needs.

One critical, perturbing, and unresolved element within the electronic universe concerns the requisites for preservation, which are most effectively addressed at the moment of digital conception. Today's technologies for access control—DRM, streaming systems, legal and contractual limitations—often work at cross purposes to permanence. Consensus-based regimes to ensure digital persistence are far from certain.

Organizations and institutions: The academic world is moving beyond structures defined primarily by discipline. Newly minted centers, institutes, programs, and initiatives today provide homes for interdisciplinary scholarship, even as traditional departments remain strong.

The scholarly community is also affected by other kinds of structures and constraints. Intellectual property regimes channel access to and uses of many information resources. Google is an archetype for the commercial players that now occupy an expanding and disquieting space within the realms of information and academia. Traditional higher education is itself struggling against intense financial pressures, with for-profit institutions promoting an essentially distinct vocational model.

Cooperative arrangements and consortia are further reshaping the institutional environment. Economies of scale, aggregated expertise, new synergies and unexpected opportunities, and strengthened political coalitions and operational capacities are among the potential benefits. Institutional autonomy is ever less possible or desirable.

Harvard: While Harvard participates fully in all of these trends, its responses are shaped by three additional considerations. The University is a flagship institution for research and teaching. Even as it takes the lead in some areas, its size and sense of self can render it cautious and risk-averse. Harvard also has a long and successful history of pronounced decentralization, which sometimes seems out of step in a world of seamless connections and crumbling boundaries. Finally, the University's dependence on endowed funds, ordinarily a major strength, has exposed significant fiscal vulnerabilities. Today's economic uncertainties, paradoxically, may overtake our typically measured processes in unexpected ways.

Modeling information, collections, and content

Academic institutions create and also consume information. Libraries play a critical role within this ecology as they ensure the community's continuing access to the information resources that sustain research and learning. Conceptual frameworks, as well as practical tools, enable libraries to understand and then manage the torrents of information that now overflow the landscape. One such heuristic model has proved helpful in clarifying our local options and thought. This model asserts that information resources in all forms and formats, whether viewed individually or in broader groupings, can be clumped into four "ideal" categories that reflect their academic uses as well as their origins. Libraries, along with scholarly disciplines, departments and programs, and individual students and scholars, play critical roles in enacting this classification.

Core resources and curricular support. All academic libraries provide the basic bibliographies and reference works, reading list materials, foundational literatures, and other core sources that are required for teaching and learning. Curricular support is a fundamental activity for every college and university library. Local definitions of each field's core resources also tend to carry across from one institution to the next.

The record of scholarship. Academic libraries in institutions that support original research and advanced study further aspire to capture some or all of the record of scholarship. New studies in many fields are framed within a broader context of ongoing inquiry, as manifest in the scholarly record. Holdings that recapitulate this record thus remain critical in sustaining the cumulative process of creating new knowledge. This category includes the published outputs of colleges and universities, think tanks and scholarly societies, commercial laboratories and trade organizations, academies and associations, specialized agencies and ad hoc research groups. A particular library's collecting appetite may vary within this large realm—only American university press publications, perhaps; or a multinational, multilingual sampler; or (at least in theory) exhaustive coverage. Levels of coverage can also vary among fields. Electronic publishing and new access technologies may mitigate the need for comprehensive local collections of the scholarly record by providing alternate ways to locate and make use of these resources.

The record of scholarship manifests itself above all in books and journals. Scholarly journals, which are important in all fields, comprise the primary vehicle to validate new findings in science and technology. Market dynamics are pushing serials toward digital formats, through which they can also be disseminated at multiple levels of aggregation (bundled journal packages, individual serial titles, specific articles)—always within a context of escalating costs. Scholarly monographs, then, are particularly central to the humanities. Despite experiments in electronic publishing and forecasts of ubiquitous print on demand, these materials are currently at risk.

Primary resources. An immense third category consists of all organized human expression, or the full range of primary sources. These raw materials for scholarly work have become ever more eclectic. Many libraries have always pursued a broad range of "non-canonical" creative writing—novels, drama, poetry, and on. Local and international newspapers, as well as government documents, are enduring mainstays as well. The scholarly record and synthetic works themselves serve as primary sources for researchers studying intellectual history and broader shifts in "mentalités" and ways of thought. Rare book holdings and many special collections fall within this category as well.

Other primary sources have only more recently gained a place within the library (and scholarly) pantheon. Ephemera and grey literature, pamphlets, popular magazines, comic books, visual imagery, films and video, manuscript and archival collections, and sound recordings are all by now accepted as legitimate collections categories. Websites, blogs, and other digital outlets, sometimes created as social endeavors and sometimes to represent a single perspective, are more recent additions. The relative centrality of digital versus print primary sources correlates roughly with gross domestic product: emerging areas continue to reveal themselves largely through analog materials.

Data. Unorganized or minimally structured raw data, finally, represent a category of information that we are only beginning to understand. Scholars' unprocessed laboratory notes and research transcripts—

unruly file cabinets, boxes of scribbles and scrawls—provide a quaintly venerable and readily managed example. The realm of raw data has assumed greater importance as the research tools associated with "big science" drive more and more scholarship. Digital satellite imagery, DNA and genome sequences, remote sensing data, raw survey responses, meteorological measurements, and text and image corpora are among the datasets and data streams that now pose daunting challenges of capture, interpretation, and curation.

Each field's scholarship and teaching draw upon different blends of information from these four categories. Research in Medieval Studies, for example, relies upon an array of original sources and texts that is by now pretty much fixed, at least when compared to the endless tidal wave of new materials that inform scholarship in fields like film studies, chemistry, and political science. For medievalists, exhaustive access to contemporary scholarship is therefore essential. Even this ground, of course, is not entirely solid—for instance as the field's research has broadened beyond a confined textual canon to include the evidence of archeology and material culture. The High Energy Physics community, by contrast, relies heavily on the almost comprehensive availability of research findings in the "arXiv" preprint server. Peer-reviewed journals then invest specific reports with validation and prestige. Vast streams of raw data, for example those generated by CERN's Large Hadron Collider, are crucial as well. Benchmark monographs serve to recapitulate the field's overall state of the art at particular points in time.

The academy's lore depicts the library as the humanist's laboratory, implying that historical materials and primary sources are less central in other scholarly realms. We need a more nuanced understanding. For example, historic field surveys are indispensable for botanical and zoological research. So, for astronomers, are star maps and celestial observations from both past and present. Scholars' uses of non-current literature in disciplines like chemistry or physics may follow in the path of medical researchers, who have fruitfully engaged in text-mining across large sets of historic data and reports.

The academic uses of information resources are shifting, sometimes in unexpected ways, across all four categories. While discipline-specific research paradigms often remain important, more agile models for scholarship and inquiry also suggest more fluid approaches. Today's pedagogical models routinely require students to grapple with primary sources and special collections, as well as secondary works and synthetic texts. Just as individual resources have become energized in the current environment, so has the entire structure of information. Our traditional collecting expectations, which were far more static and staid, no longer serve us well.

The changing contexts and expanding scale of collection development

Models for information resources provide one useful window into library collections and collecting. Another perspective focuses on the changing context within which our collections are now being built.

Collections of record, collections for use: With a few exceptions such as consciously duplicated core materials, reserve readings, and high-use recreational works, research libraries have sought to build enduring collections of record. Carefully selected individual items, in their aggregate, comprise definitive collections for the associated topics and fields. Libraries then care for these assemblages so that they will

be permanently available. Creating and stewarding this patrimony constitutes a vocation of broad cultural consequence.

Looking to the future, research libraries will in some areas continue to build enduring collections of record. In others, they will settle for use-driven holdings while seeking neither comprehensive coverage nor long-term retention. The availability of digital surrogates or of remotely maintained "archival" copies may also affect local choices. Ideally, libraries will seek to ensure that some institution is providing ongoing preservation and care for everything they hold—but there may be instances in which current-use materials are acquired and discarded regardless of provisions for persistence. The continuum of curation will become more diverse.

From "collections" to "collections and content": Most academic libraries will continue to acquire both analog and digital materials for their on-site collections. However, their focus will expand ever more emphatically beyond acquisitions as they also provide access to intellectual content that is leased rather than acquired, or to which they only point. Some libraries will likewise continue to create new, primarily digital resources on their own. The increasing ubiquity and utility of highly diverse digital resources will require adjustments in all library operations. "Content"—a category that encompasses everything to which a library enjoys ready physical or digital access, regardless of ownership status—is central to all that we do.

The diffuse knowledge that is embedded within and suffused throughout every university is a form of local content that most institutions have barely begun to tap. Energizing and leveraging this largely latent capacity is critical to the academy's future. The process will most fruitfully engage faculty, staff members, and also the students whose research pilgrimages—mental and physical—foreshadow tomorrow's scholarly agendas. At Harvard, the Medical and Business Schools have already made substantial progress in knowledge management. This will be a necessary element in our emerging content strategy.

Enlarging the field: partners and players: All academic libraries are under intense financial pressure. The possibilities and the shared challenges associated with digital resources, the scale of today's information needs, and examples of consortial achievements together make cooperation more appealing than ever before. The production of information resources, as well as conjoined consumption and processing, can become shared functions within a virtual environment.

Collective action may allow libraries to more fully shape both the landscape and the marketplace for electronic resources. Collections cooperation has traditionally emphasized the obscure, low-demand, sometimes expensive resources that can be shared among partners with minimal inconvenience to occasional local users. The compelling logic holds that shared physical resources made available through interlibrary loan can effectively reduce the need for redundant acquisitions at many different sites. Collecting scale, geographic and programmatic proximities, and resonances with other cultural institutions further shape the potential results. Structured acquisitions programs and streamlined processes for resource sharing have particularly allowed progress in some relatively specific collections niches.

This model might now also be turned on its head as members of consortia together identify and arrange for digital access to core materials. Particularly in the electronic age, cooperative activities can cut across all four categories of collection resources. For large datasets, collaboration will be essential in building both infrastructure and tools due to the sheer scale of the task.

Collaborative action might encompass other dimensions as well. Libraries, archives, and museums are often co-located. They also share similar aspirations and missions. New opportunities for service and for deeper complementarity may be at hand. Research library cooperation has been most successful in focused efforts among groups of limited size: intensive partnerships among two or three peer institutions, and relatively compact consortia such as the Committee on Institutional Cooperation or the California Digital Library. Cooperative initiatives that achieve enduring operational success seem to be bound by intractable limitations of organizational structure and scale, even in today's technological age.

Libraries as storehouses, libraries as tool sheds: The mass of information resources now available on the Web, many of them free, is fundamentally changing the library community's collections calculus. High-quality and openly accessible scholarly resources—digitized maps and medieval manuscripts, books and journals, images from archives and art museums, music scores and sound recordings, and on—can be found in staggering profusion, without even considering the medium's less scholarly emanations. Links to freely available digital content, meta-search capabilities that cut across products and platforms, and local aggregations of e-resources, will all play a growing role in libraries' collections and content strategies. This in turn will also reduce the physicality of library holdings and alter the functionalities of their spaces. But we need to go further.

Three aspects of Web-based content require close attention. In the first place, the search engines that today allow users to find materials on the Web are neither transparent nor fully revealing of useful content in predictable ways. Google Scholar, for example, relies upon opaque search algorithms and relevance rankings that appear not to fully exploit the wealth of standards-based metadata that libraries routinely provide. But most libraries do little better, investing their cataloged resources with robust metadata that our discovery tools rarely handle well. Secondly, sources on the Web—whether websites themselves or the data, images, objects, and documents embedded within them—are notoriously unstable. Content is added, changed, and removed; links shift around and disappear. Scholarship relies on enduring access to constant content, a goal which remains elusive in the digital domain. Capture, curation, and digital preservation are all implicated in this conundrum. Finally, dispersed and disparate Web content requires tools that can work across amalgamated sets of sources in predictable and repeatable ways. Some of the uses are well understood, while others reflect a new realm of inquiry that includes text mining, pattern recognition, visualization, and simulation. The needs are perhaps most pressing around massive accumulations of raw data.

Libraries, working together and also with academics and information technologists, have an evolving role in creating and supporting the tools that will enable students and scholars to take full advantage of the digital world. It is not yet clear whether lead roles can or should be pre-ordained: arrangements that embody flexibility and contingency seem most likely to succeed.

Scarcity: measure of prestige or consequence of manipulation: Research libraries have traditionally built their reputations on the basis of their overall size, and also the depth and breadth of their rare book holdings and their special collections. Scarce or unique artifacts, as well as uniquely comprehensive collections, remain primary measures of quality. Prestige based on both size and scarcity may diminish as large-scale digitization weakens the once obvious benefits of local ownership. The "structural" scarcity associated with rare artifacts is ever less compelling in a rich digital environment.

Paradoxically, our most coveted resources now include those digital materials whose uses are limited by contractual restrictions. Electronic gatekeepers can create scarcity (and also compromise long-term persistence) by manipulating license agreements and relying upon restrictive delivery technologies, even as the underlying resources could in theory be available without limit. "Scarcity" in a traditional sense reflects materials that are physically rare or unique. Today's environment adds in the artificial scarcity that results from restrictive manipulations of the digital marketplace.

Authorship and authority: Academic libraries have historically served as custodians for carefully selected, authoritative information. Library holdings were then taken to embody the highest standards of analytical and methodological rigor. The weighty bound tomes associated with research libraries and traditional scholarship carried their own aura of permanence and security. Norms for careful reading and for measured scholarly discourse further suggested prudence, stability, confidence, and authority. Deeply embedded synergies between artifact and text played an essential role in research and teaching.

Our excursions into the exuberantly expressive realm of primary resources have pretty much destroyed these presumptions. All manner of deliberately ephemeral products circulate at high velocity, undermining anyone's attempts to delimit agency, define a canon, or codify quality. Ours is instead a prolific universe of spontaneous, unmediated, unvalidated information. "Web 2.0" both reflects and engenders "Authority 2.0" as users, singly or in cohorts, participate in an electronic free-for-all. Platforms and formats are likewise provisional: experimental and ephemeral expressions may evolve into dominant manifestations and forms, though extinction (think "WordStar" in the pedestrian realm of word processing programs) is a real possibility as well. Libraries are on uncertain ground as they engage with this fractious, seductive, alien, and essential universe.

Guiding principles for collections and content

The trends here described suggest several general principles to guide academic libraries as they move toward the future. All of these statements apply to the Harvard College Library.

1. Most information—core materials, the record of scholarship, trade publications, an increasing proportion of recorded human expression, and data—is becoming available in digital formats. The emergent electronic realm will, in time, pretty much relegate new analog materials to a diminishing subset of primary sources. Digital resources will increasingly define both the information and the scholarly landscapes. **Our future is digital: libraries must prepare for and promote this shift.**

2. Digital resources are produced, become available, and then behave differently than hardcopy objects. Among many other features, few of them can be "owned" in the same way as books or journals. Libraries must therefore frame their information goals in terms of providing access to content that they do not possess, as well as on-site holdings. **Libraries must broaden their focus to encompass both collections and content.**

3. As budgets decline and priorities shift, many academic libraries will steer their acquisitions toward the basic texts and sources required for curricular support. These holdings will be heavily redundant across different institutions. Conversely, more and more non-core materials may be entirely missed. Cooperative efforts—international, national, regional, and local—can at once increase efficiencies around everyone's need for duplicative materials, and also maximize the collections reach of those libraries that are capable of pursuing scarce or unique resources. **Cooperative activities will become increasingly central to library programs and strategies.**

4. The commercialization of scholarly information, on top of longstanding trends toward monetization and privatization in the realms of mass expression and entertainment, threaten the free flow of information that the academy requires. Prohibitive costs and artificial scarcity are among the consequences. Many experiments and initiatives, with those broadly clustered under the "open access" rubric among the most promising, are now in play. **Academic libraries must actively engage in reformulating information flows and scholarly communications in order to protect future research and learning.**

5. Libraries have always sought to make information both accessible and usable. Catalog records link users with the sources relevant to their interests; reference (or "research and learning") services then help those users extract the fullest possible benefit from what they have found. Digital resources, and particularly large-scale, cloud-based data, require new, standards-based tools and services. **Libraries, acting among themselves and through external partnerships, must participate in developing all of these tools and services.**

6. Academic libraries must be aligned with and accountable to their parent institutions. Yet information is becoming more diffuse and library activities, across the board, are ever more cooperative in nature and expansive in scope. Closely consultative processes within each campus will remain essential, but may no longer be sufficient. **Universities and libraries must devise models for governance that both ensure local accountability and encourage cooperative activities.**

The world of library collections is one in which once solid certainties no longer obtain. The range of relevant materials has shifted and grown, though the relative centrality of tangible resources under the library's direct control is in decline. Libraries will increasingly work to identify information that they will never own, and to provide the tools that enable their students and scholars to use these resources effectively. The sources themselves will take on new dimensions whose continued usability will demand different kinds of support. More and more, libraries will have to engage in partnerships and collaborative

efforts to achieve their goals. While the mandate to ensure ready access to a comprehensive array of information resources will remain, the what and how will seem quite different.

II.　　　Strategies and goals for the Harvard College Library

Harvard's collections: mission and goals, past and present

The Harvard College Library organizes its activities in accordance with general statements that embody the priorities and goals of the Faculty of Arts and Sciences as a whole. At the most fundamental level, HCL defines its mission in the following terms[1]:

> The Harvard College Library supports the teaching and research activities of the Faculty of Arts and Sciences and the University. Beyond this primary responsibility, the Library serves, to the extent feasible, the larger scholarly community.

> The Library acquires, organizes, preserves, and makes readily available collections of scholarly materials in all media and formats.

> The Library fulfills its mission by providing intellectual access to materials and information available at the University and elsewhere, by providing assistance and training in the location and use of these materials, and by providing facilities and services for research and study.

The Library's "collections overview" statement further delineates three broad goals:

1. The collections provide exhaustive, comprehensive support for courses and programs of study within the Faculty of Arts and Sciences. Items from reading lists, reserve readings, and the associated core literatures and reference works are acquired as a matter of routine.

2. The collections aspire to comprehensive coverage of the record of scholarship, from all countries and in all relevant fields. … These holdings provide a knowledge base for much of the research conducted within the University.

3. The collections seek to selectively capture the overall record of human expression, across time and also across formats and media. Primary materials comprise the resource base for future scholarship. …

While the Library's acquisitions have never been comprehensive, past accounts would have it that we once approached this happy ideal. Three elements contributed to this illusion. Traditional acquisitions models privileged canonical works and the record of scholarship as the collections categories appropriate for academic libraries. These domains were finite and clearly bounded, particularly when research institutions were few and core literatures thin. The range of appropriate formats was likewise limited,

---

[1] This Mission Statement, and the "Collections Overview" that follows, are taken from the HCL website (http://hcl.harvard.edu), under the "General information" tab and the "Library Business" category.

centering on books and journals. Within this circumscribed realm, Harvard's budgets indeed did permit us to acquire pretty much everything that we could identify.

This paper has rehearsed numerous contextual changes that militate against any hopes that our acquisitions might now be comprehensive. Our starkly finite capabilities have been impressed upon us with relentless force since the Library's most recent mission statement was put into place. Even more, our universe of possibility has expanded as the model of an information firmament centered on our own perspectives, collections, and needs has shifted to a Copernican universe of infinite resources and no fixed center. A current statement of our collections goals will therefore carry qualifications and limits that were unnecessary and perhaps unthinkable in the past. These goals might be reasonably framed as follows:

> The Harvard College Library supports local research and learning through sustained efforts to capture, organize, preserve, and project the human record. The Library will meet the needs of the Faculty of Arts and Sciences for core and course-related materials. As deemed appropriate through consultations with local constituencies, it will pursue the record of scholarship in all languages and from all parts of the world. It will aggressively but selectively document the broad realm of human expression, in all its modalities. The Library, working with academic units and drawing upon new sources of support, will also expand its programs to capture and curate digital data in fields spanning the humanities, social sciences, and sciences. Explicit interdependence with other institutions will inform all of these activities.

Process and implementation

Our collections ambitions remain both generous and open-ended. However, the preliminary acquisitions budgets in place for 2009-2010 have dropped in every case, sometimes by 25% or more. Other libraries within the Harvard University system face cuts of their own, some of which will affect users within the Faculty of Arts and Sciences. The College Library's collections are also heavily dependent upon endowment income, which will decline over and above the general cutbacks already mandated within FAS. The reduced support for acquisitions requires adjustments as of July 1, before even an initial round of community consultations. Our immediate plans will focus on coping with the drop, as we also prepare a more deliberate process.

Over time, HCL's revised approach to collections will coincide with shifts in the Library's broader portfolio of activities and priorities, and also our evolving institutional context. Our collections goals will thus find synergies with new expectations (and metrics) for library programs that put particular weight on outreach and service. Within FAS, the Dean's four priorities announced in the spring of 2008 emphasize undergraduate learning and the student experience. Even as user studies at other institutions document an enduring preoccupation with collections as the library's quintessential service, these are no longer the be-all and end-all of our endeavor.

The process and the specific measures through which we will ratify our new collections model and adjust to a significantly smaller collection capacity need to be as thoughtful and as smooth as possible. We here

turn to these medium-term shifts, which will be enacted over the next three to five years. Three distinct time frames are particularly salient:

1. *July 2009.* All HCL units will face substantially lower budgets for collections, staff, and services. Staff reductions and the concomitant organizational adjustments will not yet have taken hold. A great deal of energy will be required simply to sustain our basic services. In the short term, our acquisitions will reflect ongoing patterns and priorities, albeit at a diminished level.

2. *Ongoing policy planning and implementation.* Concrete measures to ratify our collection and content priorities will begin in earnest as soon as the Library has regained its balance. Intensive, data-informed consultations with specific local constituencies will be matched by careful planning for cooperative collecting arrangements and evaluation of consortial programs such as BorrowDirect. The result will be an evolving matrix of benchmarks and collections goals, informed by a rigorous consultative process.

3. *Three- and five-year benchmarks.* The Library will significantly shift its collection and content profile in broad accordance with our expectations for overall change, but also following the process outlined below. Other operations including cataloging and conservation will likewise adjust to these new priorities.

Collections goals and our information model

Our model for information and its uses provides one organizing structure with which to anticipate the future of HCL's collections and content. The following reflections and expectations respond to today's general information landscape; they can also be matched to changing budget realities. While the specific features of our emerging collections efforts will vary across constituencies and fields, we expect that the general scenario will include the following components:

<u>Core resources and curricular support:</u> The Library will continue to support the curriculum and to provide core resources in a comprehensive fashion. Reserves, reference sources, textbooks, and other basic materials will whenever possible be sought in electronic versions. Hardcopy materials will only be acquired when essential for user convenience (for instance very long texts), when necessary to document a particular phenomenon or trend (for example by acquiring annual print "snapshots" of databases that are constantly updated in real time), or when preservation considerations so require. We expect that user enthusiasm for digital reserves and reference sources will vary among disciplines, and also that acceptance will by and large continue to grow. The shift will reflect a process rather than a single transformative moment.

<u>The record of scholarship:</u> The record of scholarship is expanding, assuming new forms, and in some fields becoming less axiomatically central to research and teaching. We will document circulation patterns, interlibrary loan transactions, and requests for new purchases, and then consult Harvard's faculty constituencies on a department-by-department and field-by-field basis in order to more closely match the Library's future acquisitions with actual need. The (evolving) extent to which we can rely upon

interlibrary loan and digital access will further inform our choices. Our working hypothesis is that some constituencies will continue to require exhaustive in-house coverage of scholarly monographs and serials. (Classics and Medieval Studies are among the bibliocentric fields that immediately come to mind.) Other groups, including some that particularly rely upon other kinds of data, may be less insistent in seeking inclusive local holdings of the scholarly record in all languages and from all parts of the world.

Secondary materials and trade publications will become ever more readily available as other academic libraries focus on the record of scholarship and core literature, and as consortial arrangements and print on demand then reduce the need for comprehensive on-site holdings. The deals being struck by entities like Google are also serving to monetize and bring to market a substantial share of in-copyright materials from the United States that had gone out of print and were largely dismissed as carrying very little value. We expect that deferred purchases, as well as interlibrary loan, will emerge as increasingly viable options for access to mainstream materials from North America and Europe that we did not acquire upon publication. Trade imprints from other parts of the world will follow similar patterns, albeit at their own rhythms.

HCL will thus extract ever less benefit from sustaining a uniquely large assemblage of easily obtained and heavily duplicative materials. The current panorama nonetheless is sufficiently uneven that any shift away from ostensibly duplicative coverage must be cautious. Harvard is also one of few or occasionally the sole owner of some categories of scholarly monographs and journals. (Reports from WorldCat, our closest approximation to a universal union catalog, can help to clarify this universe.) In these cases, new collecting arrangements will reflect carefully constructed divisions of responsibility with trusted peer institutions. Planned reductions in local holdings will be enacted in tandem with broader processes to ensure that our students and scholars are not inadvertently disadvantaged.

In sum: HCL will continue to purchase and provide access to the scholarly record in journals and in monographs alike, at levels that we will establish in consultation with our users. Digital versions will be preferred whenever price, licensing terms, user needs, and preservation considerations so permit. Harvard's traditional collecting strength in scholarly monographs also suggests the importance of interim measures, pending full consultations, that balance a respect for continuity with recognition of new opportunities and constraints.

Primary sources: Pending systematic consultations, we anticipate that Harvard users will ultimately be best served if the Library focuses as much as possible on specialized materials and primary sources. In the realm of primary sources, of course, we can never hope to be more than selective. This inherent limitation in turn mandates close attention to cooperative efforts and to shared collections. Our efforts concerning primary sources will privilege original materials over the prepackaged aggregations assembled by third parties. They will also encompass resources captured or created by our own students and scholars. (This is already a mainstream activity embodied, for example, in the field recordings that our ethnomusicologists routinely deposit with the Library.)

Data: The Harvard College Library has engaged only minimally with data and datasets, and even then primarily through allied units and organizations. Its capacity to house, serve, and curate data is barely beginning to take shape. Some other units within FAS, and also the University, have been more active.

While duplicative efforts are in no one's interest, HCL will become increasingly involved in this arena through direct activities and, even more, through participation in community projects both within and beyond the University. Unfortunately, even a modest pool of funds for strategic exploration and investment is beyond our immediate reach. This remains a deferred priority.

The elements of action

Implementing shifts in HCL's collections and content will necessarily center on matters of policy and strategy. Four major areas demand particular attention: our evolving collection development policy; guiding principles for cooperative activity; resource strategies; and the College Library's role vis-à-vis new approaches to scholarly communication. Addressing each of these broad topics will require adjustments in the operational plane as well as the strategic realm.

HCL's policy for collections and content. The Library could be purposefully vague in describing its collecting goals when budgets were sufficient to meet most needs. It was both economical and roughly accurate—albeit a bit pretentious—to declare that we would acquire whatever we should. Straitened times and increased demand now require us to articulate priorities that are more transparent, and also more attentive to local research and teaching. An iterative process that draws from both empirical data and inclusive consultations with faculty, students, and other collections constituencies will thus become central to our planning. In the near term, we will convene open discussions to consider the general causes and contours of our current situation. The Library's corps of departmental liaisons and bibliographers will then hold more intimate meetings to elicit targeted collection assessments from specific constituencies including departments, disciplines, affinity groups, and concerned individuals.

The Library's new collections and content priorities will be most directly shaped by the expressed needs of our faculty and students, as augmented by reports from our research and learning librarians. Data on actual collections use will further inform these exchanges. The automated systems upon which the Library and the University rely can make it difficult to represent even basic data in useful ways. HCL has also lagged in devising reliable structures for user feedback and advice. These deficiencies may have been (barely) tolerable when funds were ample. We now need a more rigorous approach.

Circulation statistics will allow us to identify holdings segments—variously defined in terms of subject, format, language, and year of publication—that are used heavily, a little bit, or not at all. Harvard's interlibrary borrowing requests may suggest patterns of collections insufficiency and strength. User requests for book purchases can likewise indicate areas of high demand or lagging performance. WorldCat data on areas of particular collection strength, as measured by size and by unique holdings, will be brought to bear as well. Together, discussions and data will lead to a detailed public statement of aspirations and goals that will lend itself to ongoing adjustment and review. The document will also delineate clear centers of accountability as to its accuracy and currency.

HCL's emergent collections and content policy will necessarily strike a new balance among the various perspectives that have traditionally underlain these documents. Some such policies privilege ongoing collections. "Building from strength" acknowledges the enduring value of focused effort over time, albeit with the risk of creating holdings that are peripheral to current research and learning. A different starting

point instead emphasizes the real-life information needs of flesh-and-blood students and scholars: concrete demands here assume pride of place. A third variant looks toward the academic programs and scholarly disciplines supported by the library's parent institution. Collections priorities are then driven by what each disciplinary community deems most necessary. (Such supposedly objective measures as "impact factors" can inform specific decisions within some fields.) Finally, some collections and content policies reflect analyses of the changing nature of information resources and scholarly communication, and therefore emphasize strategic measures to anticipate and shape longer-term trends.

Different constituencies within the Faculty of Arts and Sciences are likely to espouse collections priorities that draw differently from each of these starting points. Acknowledging and where possible reconciling the divergences will require both political skills and technical knowledge. It falls upon the Library to make these decisions.

Guiding principles for cooperation. We will continue and whenever possible expand our participation in effective cooperative programs as a means to broaden our collections reach. Harvard has traditionally held back from many such initiatives, not least from fears that our collections would be overwhelmed by external demand. Experiences at other large research libraries suggest more benign results. Carefully codified collections interdependencies can ensure enduring support for scholarship. Other academic libraries, of course, are cutting back as well: ill-considered interdependence could simply degrade the community's overall capacity. These are pitfalls that we can and will avoid.

Some decisions appear straightforward. For example, we will routinely avoid duplication of large or expensive microform sets that are already held by the Center for Research Libraries, or that are available through similar membership organizations in which we participate. Physical proximity, analogous missions and aspirations, and complementary library programs and systems then suggest that Harvard and MIT might benefit from closer collaboration. Enhanced access to collections, shared purchases, and focused efforts concerning some user services and internal operations may all warrant consideration.

The BorrowDirect consortium consists of an Ivies-centric group of libraries that together provide expedited interlibrary loan. The traffic tends toward uncommon imprints on one hand, and heavily used materials on the other. The potential benefits in user service and increased collections depth appear substantial. As BorrowDirect further evolves toward consciously coordinated collection development and processing operations, the advantages may become even greater.

Closer to home, the College Library accounts for about sixty percent of the University's traditional library activities and expenditures. Students and scholars from the Faculty of Arts and Sciences nonetheless rely on materials and services from all parts of the University Library, just as those in other schools and tubs rely on HCL. Tensions among units, each with its local priorities and commitments, are inevitable. Financial plenty has heretofore tended to mitigate the effects, though new initiatives in some units and reduced resources throughout may now threaten this enviable convergence. The University Library Council and its fledgling committee structure need to bolster the infrastructure of policy and process that at once protects local autonomy and upholds community expectations.

Resources and their deployment. Strengthened financial support is central in repositioning the Library to achieve its evolving goals for collections, content, and services. Aggressive fundraising and a prominent place in the next capital campaign are essential. Developing a message that addresses our emerging priorities and needs is crucial, now and to move ahead. On a more modest scale, HCL requires ongoing structural funds in order to meet the requirements of new academic programs. Current arrangements within FAS simply presume that the Library can stretch to meet increased programmatic needs, with obvious consequences in light of our finite resource pool.

Internal allocations. The College Library is unusual for its substantial reliance upon endowed funds. Some endowments are broad in their terms, though others can only be applied to narrow fields. Individual collections budgets may then include highly restricted endowments, and also funds with more flexible terms. Whatever the restrictions on use, most funds have typically carried forward within the original host unit from one year to the next. Our annual allocations for collections will henceforth more actively reflect the new policy directions that we articulate in concert with our users. New allocation procedures, including an internal advisory committee, will oversee this process.

We will also create a central pool of funds to support large-scale purchases of special collections materials and primary sources. Purchase proposals will be evaluated by an advisory body of bibliographers and research librarians, though routine acquisitions of primary materials will remain within the purview of each bibliographer. The proportion of collections funds allotted to primary resources and special collections materials will build slowly from the levels typical of the recent past. The implementation of new collections policies, the Library's adjustment to significantly lower budgets, and field-specific benchmarks for trustworthy cooperative programs will all inform the pace of change.

Core collections infrastructure. Most HCL acquisitions will continue to be retained for the long term. Appropriate arrangements for bibliographic description and inventory control, conservation, and storage and access, will all follow apace. However, different protocols, capabilities, and services may apply to materials that are not destined for such permanence.

The administrative and funding structures underlying some components of the University Library's core collections infrastructure do not serve us well. Clumsy charge-backs are utilized to fund the Harvard Depository and the Digital Repository Service, as well as HOLLIS and other basic tools. The logic of these fees can encourage narrowly self-interested behaviors at the local level, sometimes at the community's expense. Central support for core activities, in conjunction with collaborative policy oversight and review, could both streamline operations and enhance services.

Information technology. HCL's capacity to capture, curate, and deliver electronic content is at present mediated primarily through the University Library's Office for Information Systems, the Digital Repository Service, and "DASH," the Institutional Repository associated with the Office for Scholarly Communication. Cross-unit efforts to ensure that digital information resources are fully utilized in teaching and learning will continue to intensify, along with complementary measures to make analog holdings more visible. Relationships within FAS, across all of the University, and with external agencies will deepen as well, also more directly accommodating a full range of faculty-generated research materials.

17

Selectors and collections. Materials selection within HCL splits between librarians who focus on specific subject areas such as music or anthropology, and those—mostly in Widener and the Harvard-Yenching Library—who work with publishing areas defined by language or geography. Other research libraries often rely on subject-based selection, charging staff members (who may carry additional responsibilities for public service or cataloging) with field-specific selection assignments that cover all parts of the world. Harvard's approach cultivates bibliographers who are deeply familiar with publishing areas and information markets. It also reflects our expectation that specialized staffing structures will permit us to acquire whatever we identify as carrying potential utility for our collections and users.

Most of the Library's bread-and-butter acquisitions are today enacted through approval plans, through which specialist vendors supply materials in accordance with detailed subject profiles. Our selectors both oversee these plans and supplement them with individual orders for additional items. Our model continues to presume that linguistic and marketplace expertise trumps detailed subject knowledge in today's complex and atomized information environment. Looking ahead, our reduced purchasing power and the burgeoning output of resources will require greater selectivity in an increasingly labor-intensive endeavor. More complex and diverse selection assignments may result.

Scholarly communications. Harvard as a whole is moving toward open access as an organizing principle for its scholarly output and also as it expands the content made available through the Library. The University Library's Office for Scholarly Communication is sponsoring open-access alternatives for Harvard-based scholarship, constructing an institutional repository, and espousing open access language in our e-resource licenses. The College Library will continue to support all of these efforts.

Greater efficiency and reduced redundancy. The College Library will continue to pursue operational efficiencies wherever they can be found. Close cooperation with our vendors and commercial partners can foster both cost-effective relationships and "downstream" improvements within the Library, for instance as dealers provide catalog records, spine labels, and barcodes for new acquisitions. Reduced collections redundancy is another area to explore, for example with new tools that may allow us to minimize inadvertent duplication across units. We are also reviewing our approach to books and journals that are available in both print and electronic formats in order to minimize duplication across versions whenever our basic collections conditions, typically including preservation criteria, can be satisfied. We will steer our users toward digital versions of holdings that are also available in analog formats, especially when expensive retrievals from the Harvard Depository would otherwise come into play.

In a different vein, Harvard is known both for its administrative complexity and its many libraries. Diffuse structures for administration and funding, dispersed collections, and distributed service points can reflect compelling needs. Tightly focused libraries, for example, help to delineate discrete mental spaces for particular constituencies—though interdisciplinary perspectives may thereby receive short shrift. Simpler arrangements may in some cases warrant our attention.

The Library engages in myriad negotiations with external agents. Blanket orders and approval plans are by and large arranged via technical conversations involving bibliographers and acquisitions staffs.

Partnerships with external reformatting agencies do not necessarily reflect consistent principles or a coordinated approach. Vendors offering pre-packaged sets of primary sources often approach the Library through piecemeal conversations, at times obfuscating our eligibility for the special discounts associated with large net outlays.

Electronic packages and serial subscriptions sometimes bring annual price increases that substantially exceed inflation. E-journals in the Scientific, Technical, and Medical (STM) realm, while often central to scholarship, tend to be particularly expensive. The correct balance between outrage over pricing and terms, the centrality of some of these materials, our very real budget limitations, and the feasibility of other approaches (for instance "pay by the drink" document delivery), is often difficult to discern. Each of these examples, and others as well, requires specialized expertise to establish our priorities and needs. Each also involves transactional relationships with one or several agents, both internal and external. A dedicated staff negotiations specialist might improve our results.

The Harvard College Library will necessarily curtail its acquisitions in the wake of the University's financial crisis. Our efforts have been flagging already, to widespread consternation and alarm. More essentially, the Library's collections program has not yet adjusted to fundamental and continuing shifts in the broader information landscape. This discussion paper has characterized this new context in order to address the interrelated challenges to our shrinking capacity, our enduring mandate to meet Harvard's requirements for information and library services, and HCL's larger role within a changing environment. The measures here described will position the College Library to construct a vibrant structure of collections and content that continues to support local research and learning, and also sustains our longstanding leadership role within the research library community.

Dan Hazen
Associate Librarian of Harvard College for Collection Development
June 2009