

The Livingstone Online Enrichment and Access Project (LEAP)

A. Significance of LEAP

Importance of Livingstone's Works to the Humanities, Science, Medicine, and Education

Many critics and historians regard David Livingstone (1813-73) as the single most important British imperial travel writer of the nineteenth century. His writings illuminate European imperialism and colonialism, African history, and the nineteenth-century slave trade, all of which continue to affect contemporary geopolitics. In his manuscripts, he documents the cultures and geography of diverse African populations -- their social organization, languages, religions, politics, and technologies.¹ His letters offer striking meditations on numerous topics, including his first-hand impressions of East African slavery: "I once saw a party in the slave yoke singing merrily & thought my these fellows have taken to it kindly [...] - I asked the cause of their mirth & was told that they laughed at the idea 'of coming back after death and haunting & killing those who had sold them'" (Letter to Horace Waller, 5 Feb. 1871). Through such descriptions, Livingstone championed the rights of Africans and helped bring an end to Zanzibar's notorious East African slave trade. His letters and other manuscripts also provide an exemplary case of how scientific and political networks across the Victorian globe were set up. Livingstone, both when abroad and at home, mobilized domestic and foreign resources to create a career mostly spent 5,000 miles from Britain. Livingstone's correspondence reveals the practicalities of how empires are created and managed -- elements of history crucial to our understanding of present-day globalization

Because of the breadth of Livingstone's writings and experience in precolonial Africa, his manuscripts contain a range of otherwise unavailable primary data. Thanks to his skills as a medical observer, geographer, natural historian, and cartographer, Livingstone's field diaries, for instance, describe African diseases, medical practices, and technologies; record zoological, botanical, and geological information; and catalogue geographical, magnetic, and astronomical data. Livingstone's three long-term visits to Africa (1841-56, 1858-63, 1866-73) took him through the countries that today constitute South Africa, Botswana, Zambia, Angola, Zimbabwe, Malawi, Mozambique, Tanzania, and the Democratic Republic of the Congo. He became the first European to make contact with a range of previously "unknown" African tribes, to cross Africa from coast to coast, and to describe and name Victoria Falls and many other interior geographical features. Together, his manuscripts, which are scattered across archives in the UK and elsewhere, form an under-utilized resource for literary scholars, historians, scientists, geographers, and anthropologists.

Fortunately, Livingstone's original manuscripts survive in abundance, unlike those of many other Victorian travelers and explorers. The *Livingstone Catalogue of Documents* (Clendennen and Cunningham 1979; Cunningham 1985)² records over 2,000 extant Livingstone letters; nearly 70 journals, diaries and notebooks; and a variety of other manuscripts and fragments. Additional items of significance continue to come to light and, in several instances, have in fact been recovered and digitized through *Livingstone Online*,³ the subject of this grant application. Livingstone's manuscripts as a whole enable scholars to track the representation of Livingstone's observations through a variety of revisions and to study the many ways in which the Victorian publishing industry reshaped Livingstone's words for public consumption. Access to these manuscripts -- only a small percentage of which had been edited to modern standards and published prior to the launch of *Livingstone Online* in 2006 -- promises to enhance our understanding of the production of both nineteenth-century scientific knowledge and nineteenth-century African and British imperial history. Study of the manuscripts will also illuminate broader nineteenth-century Anglophone strategies for representing foreign peoples, cultures, and practices. Far from being

¹ Livingstone "is the principal and sometimes the only source of information on certain areas of Africa in certain periods" (Bridges 1987).

² See Appendices for a bibliography of works cited in this narrative.

³ <http://www.livingstoneonline.ucl.ac.uk/>

relics of the past, these strategies continue to influence how our own culture represents people around the world.

Consequently, Livingstone's manuscripts are an exceptional primary resource for a range of disciplines. The materials that we propose to digitize through our new project will be of value to specialists from literary studies (Victorian literature, travel writing, colonial and postcolonial studies), history (Empire, colonialism, Africa, slavery & abolition, science, technology, medicine, globalization), cultural studies, political science, the medical humanities, geography, environmental studies, cartography, and anthropology. In recent decades, scholars have drawn on Livingstone's primary texts and manuscripts to study a broad range of topics, including:

- the impact of missionary activity in Malawi (McCracken 1977)
- the origins of South African apartheid ideology (Du Toit 1983)
- fungi in tropical Africa (Pearce 1985)
- the role of slavery in Zanzibar's East African commercial empire (Sheriff 1987)
- collaborative editing practices in the Victorian publishing industry (Helly 1987)
- local trade along the Botletli River in Botswana (Cashdan 1987)
- the prevalence of malnutrition in precolonial Africa (Rijpma 1996)
- the contribution of non-Western populations to Victorian geographical discourse (Bridges 1998)
- climactic variability in southern Africa during the nineteenth century (Endfield and Nash 2002)
- state building in south central Africa (Flint 2003)
- the ophthalmological consequences of exposure to tropical environments (Larner 2004)
- the role of travel in the construction of nineteenth-century scientific knowledge (Dritsas 2005)
- the development of Victorian travel book illustrations (Koivunen 2009)
- East African urbanization and Central African village market economy (Wisnicki 2011-12)

The scope of these studies testifies to the range of unique information available in Livingstone's texts. However, in the majority of cases, manuscript use has been restricted to a handful of edited collections, mainly focusing on letters.⁴

Expanded and enhanced access to a variety of Livingstone's manuscripts will launch new, much more comprehensive research in the many disciplines noted above -- especially as the documents from the David Livingstone Centre (DLC) and the National Library of Scotland (NLS) that we propose to digitize and make available through our project together cover all of Livingstone's expeditions to Africa. Major collection highlights include:

- NLS: The manuscript of *Missionary Travels and Researches in South Africa* (1857). This is the book that established Livingstone's reputation as a writer. There are significant discrepancies between the manuscript and print versions of the book. The NLS also holds the publishing records, such as marked page proofs from the book (including illustrations), and other publishing reports detailing the costs of production, sales and the reception of the work. Justin Livingstone (no relation) has noted that the manuscript contains radical critiques -- later excised -- of the British upper classes and the Cape Colony and British governments. These changes along with Livingstone correspondence from the period already digitized by *Livingstone Online* cast insight into the process of negotiation by which Victorian travelogues reached their published form. Access to the manuscript and other associated documents promises to reveal unique information relating to the reader networks involved in the Victorian publishing industry. Additionally, historians of the British empire and missionary activity in Africa will take particular interest in the manuscript of *Missionary Travels* because it foregrounds the process by which Livingstone both represented his own life story -- which would become an inspiration for the many

⁴ E.g., Schapera 1959, 1960, 1961, 1963, 1974; Foscett 1964; Clendennen 1976, 1992; Boucher 1985; Holmes 1990.

missionaries that traveled to Africa in his wake -- and crafted his highly influential imperialist vision of developing Africa through the combination of Christianity, commerce, and civilization.

- NLS: The partial manuscript of *Narrative of an Expedition to the Zambesi and its Tributaries* (1865). The published narrative details Livingstone's second expedition to Africa (1858-64). The correspondence (some of which is already digitized on our site), manuscripts, and printed materials relating to this expedition are significant for the light they shed on four areas. First, they show how a national hero (Livingstone) and an ambitious scientific community could mobilize private patrons and a reluctant British government to sponsor an overtly imperial enterprise for its supposed scientific worth and economic benefits (identification of raw materials and promotion of commerce). Second, they exemplify how Victorian exploration science was carried out and how its concrete rewards -- notably geographical and geological exploration, magnetic measurement, and botany -- were estimated and incorporated into metropolitan knowledge networks.⁵ Third, Livingstone's observations are a unique record of the Zambezi area at mid-century: its populations, the slave trade, and Portuguese imperial policy. Fourth, Livingstone's expressions of his ambitions reveal aspects of his character and, more broadly, a Victorian vision of an African future. Livingstone envisaged turning the Shire Highlands area (in what is now Malawi) into a network of Christian villages populated by natives and emigrant British artisans producing cotton goods. The conflict of that vision with slave trading, local tribal warfare, and Portuguese social organization is a crucial part of the history of south east Africa.
- DLC: 12 of the extant 17 Field Diaries from Livingstone's last African expedition (1866-73), the "Unyanyembe Journal" (1866-72), five notebooks of miscellaneous observations from the same period, a book of astronomical observations (1869), and seven maps from the 1870s. This is possibly the most comprehensive and diverse surviving collection of manuscript documents related to any single nineteenth-century British expedition to Africa. Collectively these manuscripts encompass a broad range of subjects, including the complexities of Livingstone's relationships with Arab and African slave traders; the famous meeting with Henry Morton Stanley ("Dr. Livingstone, I presume?") and the subsequent travels of the two men together (1871-72); African fables recorded by Livingstone in the field; geological and astronomical notes; observations on the African climate; Livingstone's theories of the Nile River system and the central African watershed; vocabularies from African populations such as the Swahili, Batusi, Baganda, Manyema, Masai, Lunda, and Kavirondo; reflective essays on diverse topics, including the East African slave trade, Darwin and natural selection, the American Civil War, and African history; and day-to-day diary entries and sketches from Livingstone's travels over large portions of eastern, south central, and central Africa.

Digitizing and describing the DLC and NLS materials will ultimately facilitate the unified study of diverse Livingstone manuscripts by enabling unique cross-referencing of materials from these two collections as well as between these materials and manuscripts previously digitized by *Livingstone Online*. For instance, scholars will have opportunity to study an early draft of *Analysis of the Language of the Bechuanas* (1852) from the DLC alongside a more mature version of the manuscript (1852-57) from the NLS. Or, to give a more comprehensive example, there will be opportunity to consult the major NLS book manuscripts described above in conjunction with diverse items from the DLC and *Livingstone Online*, including maps, hundreds of letters (including some 80 NLS letters already digitized, from Livingstone to his publisher John Murray), and other miscellaneous documents such as an 1838 London Missionary Society questionnaire detailing Livingstone's religious beliefs, 37 folia of penciled notes made during the Zambezi Expedition in 1860, and astronomical observations recorded on the Shire River and Lake Nyassa in 1861. As a result, our proposed project will allow for comprehensive comparative study

⁵ Only one scholarly study of this aspect of the expedition exists (Dritsas 2005).

of Livingstone manuscripts drawn from Britain and around the world in a way that has never been possible before. Such study promises to enrich our understanding of Livingstone, history, and the diverse international contexts in which he traveled, wrote, and published.

LEAP and its Long-term Value to Research, Education, and Public Programming

Since its inception in 2006, *Livingstone Online* has established itself as the leading, open-access Internet resource for Livingstone's primary writings. The *Livingstone Online* project team has partnered with some of the leading archives and libraries in Britain and elsewhere to produce digitized, searchable, and rigorously encoded transcriptions and/or images of nearly 500 Livingstone letters, an updated digital version of the *Livingstone Catalogue of Documents*, and several essays detailing Livingstone's historical, scientific, medical, and cultural contexts. Over 150 additional letters from libraries in Britain, Belgium, South Africa, and New Zealand are at various stages of production.

In collaboration with the UCLA Digital Library Program, *Livingstone Online* has also released the results of the NEH-funded *Livingstone Spectral Imaging Project*⁶ as a separate website. This internationally recognized, multidisciplinary project applied advanced imaging technology to restore a series of illegible diaries and letters produced when Livingstone was stranded without ink or writing paper in central Africa. For the first time since the nineteenth century, a selection of Livingstone's original texts was united and made legible. The project collected some 770 GB of "raw" spectral data of Livingstone's 1870 and 1871 Field Diaries as well as select letters from the same period, and produced processed spectral images and transcriptions that restored about 99% of Livingstone's once mostly illegible original texts. In addition, this project released critical electronic editions of Livingstone's primary manuscripts (the *1871 Field Diary*⁷ and the *Letter from Bambarre*⁸). Finally, the core primary data from this project has also been made available separately as the *Livingstone Spectral Image Archive*,⁹ a database created to international library standards. This archive ensures the long-term viability of the data and provides direct, unmediated access to the data from which the critical editions were developed -- thereby readily enabling our data to be incorporated into other digital projects and online scholarly works.

Using Digital Technology to Preserve Livingstone's Writings. Our proposed project -- The *Livingstone Online* Enrichment and Access Project (LEAP) -- is a 36 month project (May 2013 - April 2016) that focuses on both preservation of and access to David Livingstone's writings. The project builds on and extends our previous work with Livingstone's manuscripts. Our *preservation activities* will ensure the long-term sustainability and use of our entire Livingstone digital collection. Our *access activities* will enable users to engage in detailed and dynamic analysis of our entire digital data collection. The *Livingstone Online* and the *Livingstone Spectral Imaging Project* share a core set of team members and a common focus on Livingstone, his life, and his manuscripts. Our preservation and access activities will integrate these two Livingstone sites and core digital data sets by 1) updating all our transcriptions to the P5 version of the Text Encoding Initiative (TEI) guidelines, 2) developing a customized TEI ODD ("One Document Does it all") schema to guide future transcription efforts, 3) creating a new, unified, and enhanced site to be hosted by the UCLA Library Program, and 4) building a *Livingstone Online* data archive that will provide unmediated access to all core images transcriptions, and metadata. Our emphasis on standardized encoding will facilitate global access and collaboration. We will also develop linked data and metadata to bolster the interoperability of our data with other digitization projects, and will submit the site to cooperative networks and digital resource aggregators. To ensure the data's long-term sustainability, we will migrate the *Livingstone Online* files to the Islandora digital library framework, and will develop a set of functionalities to streamline our production processes. We will make all data freely

⁶ <http://livingstone.library.ucla.edu/>

⁷ <http://livingstone.library.ucla.edu/1871diary/>

⁸ <http://livingstone.library.ucla.edu/bambarre/>

⁹ http://livingstone.library.ucla.edu/livingstone_archive/

available under Creative Commons License, allowing data hosting by multiple sites not only for ease of access, but also to enhance data preservation through replication and use.

We will apply the same program management practices used for the successful *Livingstone Spectral Imaging Project* and other manuscript imaging projects, including those at the Walters Art Museum funded by the NEH. This work will include various site enhancements, such as integrating and exploiting metadata and TEI markup, the creation of the capability to harmonize and update catalog and transcription metadata automatically, and an automated system to index and archive incoming TIFF images. These functionalities will simplify the the addition of new manuscript images and transcriptions to the site, enabling consistent improvement and expansion of our core collection in the future.

Most importantly, the manuscripts we propose to digitize from the David Livingstone Centre -- the archive and museum now based in Livingstone's childhood home in Blantyre, Scotland -- and the National Library of Scotland span Livingstone's entire career, represent a range of manuscript types, and cover a broad variety of subjects related to all aspects of Livingstone activities. For the proposed project, we have chosen to partner with these two archival institutions due to the extent of their Livingstone holdings. In addition, the *Livingstone Spectral Imaging Project* (2010-12) has helped us establish excellent working relationships with relevant institutional staff since these were our two primary archival partners on that project. The proposed NEH grant will enable us to digitize and archive all relevant DLC materials. The NLS is already in the process of digitizing the Livingstone collection from its John Murray Archive and will make these images freely available to our site for archiving.¹⁰ *Livingstone Online* will begin releasing new materials to coincide with and capitalize on the wide-spread interest generated by the David Livingstone 2013 bicentenary.

The digital incorporation of these new materials into *Livingstone Online* will result in free online access to a significantly expanded portion of the Livingstone corpus (see Appendices for a list of DLC and NLS items to be digitized and archived). The items currently archived by or in pre-production from *Livingstone Online* can be divided roughly into two categories: 1) some 650 letters with notable medical and scientific content from, mainly, the early and middle portions of Livingstone's career (1841-66), and 2) two diaries and a few letters from the 1870-71 period, which falls during Livingstone's final travels (1866-73). By contrast, the items to be digitized from the David Livingstone Centre (DLC) include letters, maps, miscellaneous documents, over 20 field diaries and notebooks, and the massive 763-page "Unyanyembe Journal" (1866-72), which covers most of Livingstone's last expedition including the famous meeting with Stanley. Together, the DLC collection totals some 2,500 manuscript pages. The items from the National Library of Scotland include significant book manuscripts, and unique papers and illustrations related to the publication of Livingstone's books. These NLS materials total over 1,000 manuscript pages. The DLC items give particular emphasis to Livingstone's last expedition to eastern and central Africa (1866-73), while the NLS items focus on the first and second expeditions to southern Africa (1841-56 and 1858-64). Overall these new items document Livingstone's life, family relations, and private thoughts; his cultural, scientific, and medical observations; his theories of Empire and Christianity; his extensive travels of southern, eastern, and central Africa; and his observations on the social, political, and economic repercussions of the East African slave trade during a thirty-year period.

During the proposed grant period, we will also transcribe, encode, and provide critical introductions to the "Unyanyembe Journal" and the 12 Field Diaries from the Livingstone's last expedition held by the DLC (2,200 manuscript pages). These materials complement those already published by the *Livingstone Spectral Imaging Project* and, moreover, represents the first- and second-stage records of Livingstone's last expedition to Africa (1866-73). These records were posthumously be edited, radically revised, and

¹⁰These materials are not described in the NLS manuscript digitization list (Appendices) because the materials have not yet been identified. However, the NLS is committed to the *Livingstone Online* project as a whole and has pledged to make any additional Livingstone manuscripts digitized for the bicentenary available to us free-of-charge.

published as *Livingstone's Last Journals* (1874) -- a work that has been the primary source for this period of Livingstone's life and his travels for the last 140 years. Unified, digital access to the original manuscripts from this period promises to transform our understanding of Livingstone's last journey and the many facets of African life that he recorded during the journey. As a result, through images and transcriptions of materials from the DLC and the NLS, *Livingstone Online* will provide free access to a notably expanded portion of the Livingstone corpus by the end of the NEH grant period. Our preservation activities will furthermore enable our project to be largely self-sustaining beyond the grant period.

Using Digital Tools to Provide Access to Livingstone's Writings. Currently, there exists no web-based, interoperable study environment that allows users to manipulate and digitally process both images -- particularly spectral images -- and richly encoded textual materials. We will fill this gap by incorporating a handful of complementary, open-source, analytical tools into *Livingstone Online*. These tools will allow users to take full advantage of our digitized source materials. For enhanced image study, we will explore the use of IIPImage,¹¹ which facilitates online streamed viewing and zooming of ultra high-resolution image, in conjunction with the ImageJ¹² processing program and Paleo,¹³ an ImageJ toolbox. For textual study, we will incorporate the Meandre¹⁴ Workbench into our site. Meandre will allow users to process our transcription data through various textual analytics and visualizations, all within the same web environment. The complexity of the Livingstone corpus -- which covers the explorer's entire adult life and spans a variety of manuscript types -- makes his texts ideal for such data processing. Meandre, for instance, will allow users to build interactive simile timelines and perform entity extraction and topic modeling. In addition, Meandre's GIS tools will enable diverse mapping capabilities, such as using our data coding to visualize the geographic distribution of place names. Ultimately, we will integrate online use of these tools (IIP Image, ImageJ/Paleo, and Meandre) through the SharedCanvas data ontology.¹⁵ SharedCanvas provides a digital model for linking, sequencing, rendering, and annotating complex bodies of images, texts, and other documents in a standardized, interoperable manner. Our web development work will, therefore, provide analytical resources useful to a variety of users in their study of our primary data collection.¹⁶

Innovation and Knowledge Transfer. The proposed project will also have broader institutional and extra-institutional resonance. Our documentation and dissemination practices will enable our project to serve as a model for other teams and archives interested in digitizing and providing access to humanities collections and reference resources. Our development activities will meet project needs by building on the experiences of our previous NEH-funded Livingstone project as well as other digitization programs, including the NEH-funded manuscript digitization at the Walters Art Museum and the privately-funded Archimedes and Syriac Galen Palimpsest Spectral Imaging Programs. LEAP, in turn, will also help establish program management processes, work flows, metadata collection protocols, data delivery practices, and digitization infrastructures that can be implemented across the digital projects and holdings of our project partners (UCLA, NLS, DLC). Consequently, our proposed project will contribute to internal standardization, encourage the international exchange of resources, and lay the groundwork for future digitization and collaboration.

¹¹ <http://iipimage.sourceforge.net/>

¹² <http://rsbweb.nih.gov/ij/>

¹³ This toolbox was developed through our NEH-funded *Livingstone Spectral Imaging Project*. See http://livingstone.library.ucla.edu/1871diary/documents/04_Equipoise_Toolbox_Christens-Barry.zip

¹⁴ <http://seasr.org/meandre/documentation/>

¹⁵ <http://www.shared-canvas.org/>

¹⁶ For more on all these tools, see the "Methodology and Standards section, below.

B. History, Scope, and Duration: *Livingstone Online*

Initial Development (2005-10). Two grants from the Wellcome Trust enabled initial *Livingstone Online* development. Under the leadership of Christopher Lawrence, the project team of six initially designed *Livingstone Online* to explore the circulation of scientific knowledge, especially medical knowledge, between the metropolis and the colonial periphery in the nineteenth century. As a result, the project first focused on Livingstone's letters, particularly those with notable scientific and medical content.

Work began with a pilot phase (2005-06) that involved preliminary research, outlining costs, and constructing a website. *Livingstone Online* went live in May 2006 featuring all the Livingstone letters in the Wellcome Library. The operational project phase (2006-10) continued these efforts. The Livingstone team digitized (photographed and/or transcribed and digitally encoded) large portions of letters from several major Livingstone collections in the UK, including those of the Royal Geographical Society, the School of Oriental and African Studies (SOAS), and the National Library of Scotland.¹⁷ The team also digitized Livingstone letters in partnership with fourteen institutions with smaller holdings such as the Royal Botanic Gardens, Kew; John Rylands Library, Manchester; and the Fitzwilliam Museum, Cambridge. In addition, the team identified and digitally archived over 30 previously uncatalogued Livingstone items from archives and private collections.

The *Livingstone Online* site currently offers images and/or text of nearly 500 Livingstone letters. An additional 80 letters digitized during this phase remain in various stages of pre-production. Collectively, the site's digitized letters focus on the early and middle portions of Livingstone's career. However, the entire contents of each letter have been digitized, thereby providing access to materials with a much broader scope than just scientific and medical content. The team has also transformed the printed *Livingstone Catalogue of Documents* (Clendennen and Cunningham 1979; Cunningham 1985) into an online, searchable database¹⁸ and updated the database to include any material discovered or moved since the publication of the *Catalogue* supplement in 1985. Finally, the team has produced an array of short essays on Livingstone's life and historical, scientific, and medical contexts.¹⁹

NEH-funded Spectral Imaging Phase (2010-12). Adrian S. Wisnicki, the principal investigator of LEAP, joined *Livingstone Online* in 2009 and became site co-director in 2010. Under Wisnicki's impetus, the site launched the *Livingstone Spectral Imaging Project* and expanded its focus to include literary, historical, and geographical content. Grants from the National Endowment for the Humanities and the British Academy funded work during this phase (2010-12). The *Livingstone Spectral Imaging Project* applied advanced imaging technology to restore a series of faded, illegible Livingstone letters and diaries written in 1870-71. The team initially intended to embed this new project within *Livingstone Online*, but a failed UK grant bid prevented realization of this goal. As a result, *Livingstone Online* partnered with the UCLA Digital Library Program to publish the results of the spectral imaging project in a separate website.²⁰ This website includes electronic critical editions of the *Letter from Bambarre* (2010-11)²¹ and the *1871 Field Diary* (2011-12)²² as well as the *Livingstone Spectral Image Archive* (2011-12).²³

As noted in our original NEH Digital Humanities Start-Up Grant application (2009), our spectral imaging work on Livingstone's manuscripts had two primary objectives:

¹⁷ Items archived during this phase included Livingstone's letters to his publisher John Murray.

¹⁸ <http://www.livingstoneonline.ucl.ac.uk/catalogue/view.php>

¹⁹ <http://www.livingstoneonline.ucl.ac.uk/companion.php?category=historical>

²⁰ <http://livingstone.library.ucla.edu/>. Our long-term goal, however, remained integrating the *Livingstone Spectral Imaging Project* site with the main *Livingstone Online* site.

²¹ <http://livingstone.library.ucla.edu/bambarre/index.htm>

²² <http://livingstone.library.ucla.edu/1871diary/index.htm>

²³ http://livingstone.library.ucla.edu/livingstone_archive/

1. Build on previous spectral imaging of medieval palimpsests to devise and implement new processes and technologies to spectrally image Livingstone's *1871 Field Diary*, and process the resulting image data with open source software in order to recover the faint and often illegible text of the diary.
2. Produce an online critical edition and image database so that Livingstone's unredacted diary is accessible to scholars for the first time, with both marked up transcriptions and processed images.

We realized both these objectives in full and met or exceeded the large majority of our ancillary goals, as fully documented in our NEH 2010-11 compliance matrix (see Appendices). We also established a variety of new image capture and image processing, data management, and workflow processes during this phase. As a result, although LEAP focuses on a new set of activities, the project builds on and extends the lessons learned and collaborative relationships established during the previous, NEH-funded period.

Accomplishments and Future Objectives. A number of factors distinguish *Livingstone Online*. We have collected all our data -- images, transcriptions, and metadata -- to meet international archival standards and ensure interoperability and long-term digital data preservation. We have extensively documented and disseminated our results so that other project teams can learn from and build on our work.²⁴ We have also prioritized rigorous scholarship and digital innovation, highlighted by the recent inclusion of the *Livingstone Spectral Imaging Project* website in NINES,²⁵ the leading peer-reviewed digital resource aggregator for nineteenth-century literary and cultural studies. Finally, we have established a solid record of building institutional partnerships and multidisciplinary, collaborative relationships.

Currently, the fact that Livingstone's manuscripts are scattered across archives and private collections around the world impedes far-ranging study of Livingstone's texts. The work of the *Livingstone Online* team is enabling unified study of Livingstone's rich manuscript legacy; LEAP will continue this process. From May 2011 to June 2012, the *Livingstone Spectral Imaging Project* has averaged 270 hits per week, with a one week high of 2,428, and approximately 15,650 users total during this one-year period.²⁶ Although current-use statistics are not available for *Livingstone Online*, the most recent data (summer 2009) indicates that about 3,500 users were accessing the site per month at that time.

Our long-term goal is to enhance access to and preserve all of Livingstone's primary manuscripts and to expand the contextual materials available through our site. With LEAP, *Livingstone Online* will support the written archive with standardized digital access to an integrated data set and ultimately present a more comprehensive picture of Livingstone, his life, and his times.

C. Methodology and Standards

LEAP will provide a uniform set of digital materials presented on the web for both (1) immediate reading, review, and manipulation in an enhanced web interface and (2) download as raw archival files. This dual provision of content will enable a wide spectrum of use and give the freedom to users to direct their own use of the materials.

Three streams of digital content are included in the LEAP proposal:

- Digital images and transcriptions already part of the *Livingstone Online* digital archive (2,200 folia)
- Digital images to be created from the materials held by the National Library of Scotland (1,000 folia)
- Digital images and transcriptions to be created from the material held by the David Livingstone Centre (2,500 folia)

²⁴ NEH program officers have previously singled out this aspect of our work for commendation.

²⁵ <http://www.nines.org/>

²⁶ The Appendices include emails sent by a representative sample of site visitors.

Each of these streams has a slightly different digital acquisition workflow prior to being integrated as part of *Livingstone Online* and permanently hosted by the UCLA Digital Library. To ensure this integration meets the needs of program stakeholders and end users in providing quality data on schedule and within budget, we will implement accepted program management and system integration practices developed and refined from those used in industry and in previous digitization programs. This will include program planning and documentation, integrated scheduling, and regular program communications.

I. Digital Acquisition

National Library of Scotland (NLS)

Preservation and Access to the Original Materials. All NLS manuscript collections of Livingstone material are freely accessible to the public. In terms of storage, the collection is packaged in archival quality folders and boxes, and stored in the National Library of Scotland's George IV Bridge building in the Manuscripts strongrooms. The stack floors have full environmental controls aiming at compliance with BS 5454, along with fire protection, including a sprinkler system, 24-hour security and full disaster-planning procedures. The Library's buildings hold a five-star fire safety award from the British Safety Council. The NLS manuscripts consist of loose and bound manuscript pages of varying sizes. The physical condition of the relevant NLS Livingstone collection is stable and any required conservation work has been carried out.

Digitization. Number of folia: 1,000. Digitization of relevant materials from the NLS has been or will be carried out in one of the library's reprographic studios, using equipment approved for use with NLS collections by the NLS Preservation & Conservation Manager. Imaging has been or will be done by library staff who have received training on handling original documents and who have experience working with all of the library's special collections.

David Livingstone Centre (DLC)

Preservation and Access to the Original Materials. All DLC manuscript collections of Livingstone material are freely accessible to the public, although patrons must schedule their viewing with staff in advance of their visit. The DLC materials are held in a store environment that is regulated by conservation heating and monitors that measure relative humidity and other conditions. The DLC manuscripts consist of loose and bound pages of varying sizes. The physical condition of the DLC materials ranges in quality and, given the number of documents in question, it is not possible to make any general statements. The materials vary from manuscripts in excellent condition to those that require minimal or more substantial conservation. The lists of DLC manuscripts to be digitized in the Appendices include information related to the condition of some manuscripts; a full assessment will be conducted as part of the proposed project. It is our intention to fund conservation of these documents through a future phase of the project.

Digitization. Number of folia: 2,500. Digitization of relevant DLC materials will be carried out by the Centre for Heritage Imaging and Collection Care (CHICC) in the John Rylands Library at the University of Manchester. The John Rylands Library is a Grade 1 listed building, but has a modern extension which contains state-of-the-art storage facilities. CHICC has a dedicated, lockable area within a climate-controlled area in the new building for storage and digitization. Access to this secure area is controlled by proximity card and keycode. Each item will be assessed for suitability for digitization by CHICC conservators, who will brief the photographers before imaging commences. Each folio of bound or loose items will be individually photographed, rectos then versos, by photographers who are experienced in heritage photography and trained in conservation handling. Images will be produced using a Phase One digital camera and back. CHICC's cameras range from the P25 to the top-of-the range 80 megapixel IQ180. The bound manuscripts will be supported on a Manfred Mayer Traveller's Conservation Cradle and illuminated with cold, LED lighting. Flat items will be digitized on a copystand. Files will be named

according to LEAP standards. Images will be supplied in TIFF and JPEG formats on hard drive or by FTP.

Quality Assurance for Images. A quality assurance plan will be established and implemented to ensure quality data and products are provided to users, with methods of tracking performance. This will build on standardized operations across the system to minimize risk and opportunity for error. Training, documentation, and quality control will be addressed in the program and systems integration plans. Data quality will be measured, reported, and acted upon to ensure production quality from data creation through access with objective quality standards based on user feedback.

II. Data Integration and Management

File Formats for Images. For each object the following will be delivered to *Livingstone Online*: (1) all the master TIFF images of the object, and (2) an XML manifest file that completely documents the digital objects, including technical metadata, catalog data (for content identification), and (3) MD5 checksums of each image and other data file belonging to the object's "information package." Each set of data will be validated and verified before delivery to *Livingstone Online*. For all folia, we will create uncompressed TIFF images of at least 300dpi for archival purposes, derivative full-resolution JPEG2000 images for zooming and panning, compressed JPEG web images at 1800 pixels on their longest edge, and thumbnails at 190 pixels on their longest edge. TIFF images will conform with version 6.0 of the TIFF specification. JPEG2000 and JPEG images will also comply with current standards for those file types.

Textual Conversion. Text Encoding Initiative (TEI) P5 XML²⁷ transcriptions will be in diplomatic format with three levels of textual production recorded: original text written by Livingstone, corrections made by Livingstone, and corrections made by present editorial team. All transcriptions will be done manually by project scholars. We will use TEI customizations to more tightly constrain the recommendations of the TEI (e.g. with a more limited set of classifications), to document extensions to the TEI, and to provide a single source for not only documenting local project encoding guidelines, but also generating a variety of schema formats for validating the XML files produced. One deliverable from the project will be a TEI ODD (one document does-it-all) customization documenting a rigorous schema for the *Livingstone Online* materials against which all new materials can be validated. The conversion of 2005-10 *Livingstone Online* materials from the (now dated) TEI P4 format to TEI P5 XML will ensure longevity, as this will extensively document their relationship to a particular version of the TEI and make any future format migrations a much simpler and automatable process. A retrospective conversion of these legacy files to validate them as a single format will also introduce encoding improvements where these are able to be automated across the whole corpus of materials.

Quality Assurance for Transcriptions. Quality will be ensured by producing and encoding transcriptions in three stages prior to online release: 1) initial transcription and TEI encoding will be done by a single individual, 2) a second individual will review the transcription, 3) the second individual or a third will review the TEI encoding. All of the transcribing and proofreading will be done by scholars or by students working under the scholars' direction. After online release, a fourth stage of proofreading will be done.

Metadata. For all digitized works, the project team will provide descriptive metadata adhering to the MODS standard, and mapped to Dublin Core for metadata harvesting purposes. Metadata will also be included in the <teiHeader> of the TEI/XML transcription files. The UCLA Digital Library usually populates the following elements in MODS as a minimal record: titleInfo.title, identifier, name (to record a creator and the holding repository), originInfo.date.dateCreated, language, typeOfResource, physicalDescription, and accessCondition. The creation of descriptive metadata will be overseen by the

²⁷ See <http://www.tei-c.org/index.xml> and <http://www.tei-c.org/Guidelines/P5/>

UCLA Library's Cataloging and Metadata Center. At this point, we do not intend to provide subject analysis on the item-level, but we will provide access to the materials based on the text-encoding, which will include place names, personal and corporate names, and dates. See, for example, the search options provided within the *Livingstone Spectral Image Project*.²⁸ We will build from this search example to provide a browse list as well.

The *Livingstone Online* materials are rich in textual content, which is ideal for search engines. We will enhance access to these materials by updating search engine optimization for the site. The increase in materials and creation of uniform data through LEAP will also improve access via search engines. In addition, the *Livingstone Online* site as a whole will be submitted to NINES for inclusion. NINES uses RDF for resource discovery, so we will also use RDF to further disseminate the *Livingstone Online* materials. In addition, all our metadata will be harvestable through the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH). We are particularly excited about the richness embedded in the TEI transcriptions, which can be exploited from these files and exported as linked open data in RDF. LEAP will leverage development of recommended outputs to standard public ontologies such as the CIDOC-CRM where feasible. Such output can be aggregated by relevant digital repositories or cooperative networks such as NINES.

III. Online Preservation of and Access to All Materials through LEAP

Images. This project will create a uniform system of open access for the images, transcriptions, and metadata via HTTP. All data will be organized to allow predictable and regular access of images, image metadata, and transcriptions. For each folio, TIFF and JPEG images, available technical metadata, and transcriptions will be provided. TEI P5 XML files will provide detailed content metadata, transcriptions, and links to corresponding images. All data files will be accompanied with MD5 checksums to ensure data integrity. Metadata will be stored with the images, as well as in the TIFF headers. At a minimum, metadata will conform to the Dublin Core (DC) Metadata Element Set (International Standards Organization Standard Number 15836 -- ISO 15836), which describes resources with a set of core elements for effective discovery and retrieval. As a result, users will easily be able to draw on our entire core data collection for use in other projects not linked to *Livingstone Online*. The *Livingstone Online* primary data archive will provide all the raw files (TIFF and JPEG images, XML transcriptions, relevant metadata) upon which our website is based.

Transcriptions. The *Livingstone Online* primary data archive will offer users a complete package of transcriptions integrated with images as well as documentation and full metadata. The transcriptions will be offered online in multiple formats: the original TEI/XML file and derivative XHTML, PDF, and ePub files. Our standard output will emphasize side-by-side viewing of the transcription and the digital image. The derivative files will be created using both open source XSLT conversions provided by the TEI and custom XSLT conversions developed by the UCLA Digital Library in consultation with James Cummings.

Digital Library Infrastructure. The UCLA Digital Library is in the process of migrating content from its current locally built digital library repository and access system to a digital library framework based on robust open source components. The key components of this digital library framework, "Islandora," are: Fedora, a digital asset management system particularly well-suited for digital library content and widely used in the cultural heritage community; and Drupal, a content management system used as the back-end for at least 1.5% of all websites worldwide. At UCLA the Islandora framework is used to create large, searchable collections of digital assets of all types and is domain agnostic in terms of the type of content it can steward. It has a highly modular architecture with a number of key features, including support for any file type via the Fedora repository system and a flexible faceted search driven by Apache Solr software.

²⁸ <http://livingstone.library.ucla.edu/1871diary/search.htm>

The Islandora framework provides several useful components that will ensure the *Livingstone Online* project reaches a wide audience and is easy to use. These include:

- A modular “solution pack” framework for defining specific data models and associated behaviors.
- Support for any XML metadata standard.
- Many out-of-the-box metadata services, such as a “formbuilder” module for XML-based data entry and editing, and support for a wide variety of metadata standards and Open Archives Initiative metadata harvesting.
- Support for semantic ontologies and the creation of relationships between objects.
- Microservice-based workflows for automating the transformation of assets.
- Editorial workflows for approving submissions to the repository.

Standardization on Fedora for the repository layer and Drupal for the discovery layer allows us to leverage open-source activity in these open source communities, which in turn enhances the ability of the UCLA Library to provide sustainable services in the long term.

Standards for the *Livingstone Online* Primary Data Archive. The *Livingstone Online* primary data archive will offer users a complete package of transcriptions and images with documentation and full metadata, which will be self-documenting and autonomous. This will be a full digital archive of transcriptions and images of all *Livingstone Online* materials -- both those previously collected and those produced by the proposed project -- that meets international library and archival standards. The archive will be based on the archive and metadata model used for the Archimedes Palimpsest²⁹ and the *Livingstone Spectral Image Project*,³⁰ and will be developed by Doug Emery, who also implemented the model in those two projects. A primary task of LEAP will be to integrate the heterogeneous image data from the several existing and coming image sources: existing *Livingstone Online* images, the spectral imaging project images, images now being made at the NLS, and DLC images to be captured under this project.

LEAP will also follow the archival model cited above by combining an open Creative Commons licensing policy, direct access to source data, and standard presentation. Each *Livingstone* document will have:

- 1) At least three image derivatives for each folio: archival TIFFs, and web and thumbnail JPEGs, all in compliance with project standards;
- 2) Accompanying technical metadata for the images, including available capture information, image dimensions, and checksums provided in an XML manifest;
- 3) A TEI P5 XML file containing transcription, cataloging metadata, and links to each of the provided images for each folio.

The archive will also provide supplemental material including guides for use of the data set, details of the cataloging and transcription methods, and any XML schemata or style sheets needed to work with the data. As noted below, permission will be secured from stakeholders to migrate image data from the old *Livingstone Online* site³¹ hosted by University College London to the new home at UCLA. In cases where permission is not secured, the archive will document the location of remotely hosted images and provide full URLs to those images, transcriptions, and other relevant navigational metadata needed to make access to all *Livingstone* materials as uniform as possible across the archive.

²⁹ <http://archimedespalimpsest.net/>

³⁰ http://livingstone.library.ucla.edu/livingstone_archive/

³¹ <http://www.livingstoneonline.ucl.ac.uk/>

The integration and storage of digital images, transcriptions, and associated catalog, imaging, and processing metadata as a useful capability in *Livingstone Online* will require the establishment of a technical infrastructure that includes a content management system for storage, archiving, and access. This will be used to host transcriptions, digital derivative images, and other contextual materials and project information for broader access. For digital files of critical objects, *Livingstone Online* will establish a DOI (Digital Object Identifier) that will ensure access to each resource through a permanent URL. The UCLA Digital Library will lead this effort of conversion in collaboration with Wisnicki and with consultation on migration of site data/functionality, site enhancements (including integrating and exploiting TEI markup and metadata), open development, bug-testing, etc. provided by Cummings.

IV. Tools and Applications to Facilitate Access Online

IIPImage. Users who wish to engage our core data collection will probably want to work with a variety of images, ranging from high resolution TIFF and JPEG2000 images to 12 to 16 "raw" spectral images plus a variety of "processed" images.³² They will need to review and manipulate these images rapidly online, rather than downloading them one by one. IIPImage³³ is the ideal server system for meeting these viewing and manipulation requirements. The IIPImage system, which is open source, allows web-based streamed viewing and zooming of high-resolution images and is both fast and bandwidth-efficient with low processor and memory requirements. The system accepts both TIFF or JPEG2000 images. Most importantly, because streaming is tile-based, users can view, navigate, and zoom around images online and in real-time.

ImageJ and Paleo. We would also like the proposed *Livingstone Online* study environment to enable users to modify images online. Such functionality is necessary for two reasons: 1) Modification, enhancement, and creation of new images by users offers the potential of recovering or accentuating original manuscript features (for instance, faded text, staining, tearing, etc.). 2) *Livingstone Online's* core digital collection is distinguished by the inclusion of both color images of all manuscripts and raw and processed spectral images of select manuscripts (see IIPImage description, above) that users may want to modify further. As a result, we will integrate ImageJ³⁴ and Paleo,³⁵ an ImageJ toolbox developed partly through the NEH-funded *Livingstone Spectral Imaging Project*, into our proposed study environment in order to meet these user-directed image processing needs. ImageJ is a public domain image processing program that can run as an online applet that enables users to display, edit, analyze, process, and print images in many formats, including TIFFs and JPEGs. It provides standard image processing functions (e.g., contrast manipulation, sharpening, smoothing, etc.) as well as more advanced processing capabilities and -- memory permitting -- supports simultaneous modification of any number of images. Paleo is open-source and consists of a set of image processing tools implemented using both macro and plugin facilities designed to make ImageJ more user-friendly.

Meandre. Meandre³⁶ provides a software environment for *Livingstone Online* that will offer users a variety of tools that can be tailored to work specifically with Livingstone files and for Livingstone users. Meandre is an open-source data-flow environment that allows users to execute text analysis. Meandre provides the infrastructure for using a variety of open-source tools, including Mallet, Google maps, Protovis Network Graph, and Simile Timeline. Meandre leverages these analytical capabilities so that users can explore the Livingstone corpus in order to make scalable discoveries across the corpus and from document to document. The creators of Meandre fully support our proposed project and, additionally,

³² See http://livingstone.library.ucla.edu/1871diary/spectral_imaging1.htm for more on images produced by the *Livingstone Spectral Imaging Project*.

³³ <http://iipimage.sourceforge.net/>

³⁴ <http://rsbweb.nih.gov/ij/>

³⁵ http://livingstone.library.ucla.edu/1871diary/documents/04_Equipoise_Toolbox_Christens-Barry.zip

³⁶ <http://seasr.org/meandre/documentation/>

have offered, first, to provide access to their existing infrastructure and their latest components for the duration of the project and, second, to arrange an initial training for our proposal participants (see G. Garnett's letter of support in Appendices). With Meandre users will be able to gain new understandings of connections between people and places using OpenNLP Entities to Protovis Network Graph, which visualizes syntactical connections using a network graph. Users can mine dates from the Livingstone letters, chart them onto a Simile timeline, and then overlay additional data about those dates onto the same graphic. The geographic elements of Livingstone's travels can be mapped to Google maps or the emotions in various letters can be compared using sentiment tracking. Meandre houses these tools and more as data flows within the Meandre environment, rather than attempting to build each tool separately from different sources.³⁷

SharedCanvas. In addition to robust manipulation and analysis of core Livingstone image and transcription data, we would also like to enable users to record and annotate the results of their work with our digital tools in a way that will be interoperable with other digital projects. The SharedCanvas data model³⁸ offers an ideal approach to meeting these needs. SharedCanvas allows users to develop multimedia, multicomponent digital canvases. Such canvases allow complex rendering and study of digitized objects, including multiple resources from a single source such as *Livingstone Online* or the integration of resources from different servers and repositories. So, for instance, for study of a single Livingstone manuscript page, users might combine a series of spectral images, transcriptions, annotations, textual visualizations, maps, and outputs from tools like Meandre or ImageJ. Alternately, items digitized by *Livingstone Online* might be brought together on a single canvas with non-Livingstone items from other digitized collections, and this canvas, in turn, might then be annotated by a community of users. As a result, SharedCanvas will give users the opportunity to work with all our core data and with the outputs of tools embedded in our site, but will also enable them to use the results of such research to a variety of other contexts. Our integration of SharedCanvas will be supported by one of its developers, Ben Albritton, who will be a member of our advisory board.

D. Sustainability of *Livingstone Online* and Digital Content Created by LEAP

Livingstone Online, a virtual archive of Livingstone materials, is a highly collaborative endeavor. Materials have been gathered from a wide range of repositories. Therefore, *Livingstone Online* relies on permission agreements with these repositories to make the materials available on the Internet. *Livingstone Online* core image data falls into two categories:

- 1) data publicly released or resulting from the 2005-10 project phase
- 2) data publicly released or resulting from the 2010-12 and the LEAP 2013-16 phases

For the *Livingstone Spectral Imaging Project* and for LEAP, both the David Livingstone Centre and the National Library of Scotland have granted permission to *Livingstone Online* to digitize and make freely available online their relevant Livingstone manuscript collections. All permissions are granted under a Creative Commons Attribution-NonCommercial license (see Appendices for permission letters). Industry best practices will be employed to make the transition from the current *Livingstone Online* site and its addresses for text content, TEI transcriptions, and image data to the new *Livingstone Online* site. All current URLs at <http://www.livingstoneonline.ucl.ac.uk/>, including those for images captured in 2005-2010, will remain valid until data has been moved to UCLA. University College London (UCL) has agreed to continue hosting the 2005-10 materials to ensure their long-term viability. When and as necessary permission has been received to transfer the images, a strategy will be developed for connecting users to the data at their new location, whether by HTTP redirection or other means, again, according to

³⁷ See the Appendices for an example of Meandre use with TEI-encoded transcriptions.

³⁸ <http://www.shared-canvas.org/>

industry best practices. In the case that permission is not secured for a given group of images, the new UCLA site will direct users to the correct location of files at UCL.

All digital assets and services developed during the grant period will be integrated into the data management and preservation standards of the UCLA Digital Library. Digital objects will be assigned Archival Resource Keys (ARK)³⁹ and be replicated in data centers on both the Los Angeles and Berkeley campuses of the University. In the past UCLA Library digital objects have been deposited with the UC Digital Preservation Repository,⁴⁰ but this “one size fits all” service is being unbundled by the University of California Curation Center⁴¹ into a set of more flexible micro-services,⁴² including the Merritt repository service.⁴³ The UCLA Library is in the process of adopting this more flexible preservation strategy and anticipates adding the Merritt repository as an additional preservation layer to the current Los Angeles/Berkeley replicated storage and redundant backup strategy. Content deposited with Merritt will in turn be replicated in diverse geographical locations by the Curation Center.

The UCLA Digital Library works with the UC Curation Center to provide a full set of services to support digital preservation of content. These services include the assignment of globally unique identifiers to all digital content; redundant and geographically diverse digital storage; verification of bit-level integrity of content files; inventory, indexing and search services; and transformation services (forward migration of file formats when needed). These services together provide the UCLA Digital Library with the ability to provide various levels of preservation services, from simple bit-level preservation (i.e. what you put in you can get back), to a full digital preservation service that includes replication, integrity services, and forward migration of file formats.

Locally all *Livingstone* digital objects produced through the proposed project and through the 2010-12 NEH-funded phase of development (images, TEI/XML, MODS/DC metadata etc.) will be hosted by the UCLA Library’s Islandora instance, consisting of the Fedora repository layer and a Drupal-based presentation layer. Existing *Livingstone Online* files will be available either on the UCLA’s Islandora instance, or linked to from that site, depending on granted stakeholder permission to transfer the data from UCL. UCLA Library is working with the Islandora development team at the University of Prince Edward Island and DiscoveryGarden Inc. to unite the Merritt repository with the Islandora framework, in order to integrate Merritt-based preservation services more directly into the digital object ingest and publication workflow. We anticipate that this Merritt integration will have been achieved by the conclusion of the grant period and that all *Livingstone* digital content will be deposited with Merritt for preservation purposes.

All the digital content created during the LEAP as well as that resulting from the 2005-10 and 2010-12 project phases will be prepared to conform to UCLA Library standards and appropriate national and international standards. From that point forward, digital content will be sustained and further developed, consistent with the Library’s mission to develop and provide access to first class collections -- both physical and digital -- to support research and instruction of the highest order. Additionally, all digital content, where permission has been provided, will become a permanent part of the UCLA Digital Library collections. Finally, the use of open international standards in metadata, images, and transcription data files will help to ensure that the future process of forward migration is simplified through the provision of machine-processable metadata. That all data produced through the 2010-12 and LEAP phases will also be made freely available for download under an open Creative Commons license aids their long-term preservation and will encourage the dissemination of multiple copies of the material. It is our hope that

³⁹ <https://confluence.ucop.edu/display/Curation/ARK>

⁴⁰ <http://www.cdlib.org/services/uc3/dpr.htm>

⁴¹ <http://www.cdlib.org/services/uc3/>

⁴² <https://confluence.ucop.edu/display/Curation/Home>

⁴³ <https://merritt.cdlib.org/>

LEAP will also result in discussion of and/or plans for our core data collection being mirrored by a suitable repository in Scotland.

E. Dissemination of Project Results

The results of the *Livingstone Spectral Imaging Project* have drawn sustained, worldwide interest from 2010 to 2012⁴⁴ and laid the foundation for a range of new initiatives, including invitations to take part in David Livingstone bicentenary exhibits (2013) at the National Museum of Scotland, the National Library of Scotland, and the David Livingstone Centre, Scotland. By commencing LEAP during 2013, we will be able to build on the academic and popular excitement generated by the bicentenary. Our dissemination and promotion strategies will follow the highly successful, international dissemination model established by the *Livingstone Spectral Imaging Project*.⁴⁵ We will target both scholars and the general public, and emphasize continuous project visibility through a combination of dissemination strategies, such as press releases, academic announcements, educational exhibits, and solicited peer reviews in journals. We will put out a series of press releases to mark significant project milestones. We will also reach out personally to journalists with whom we have previously collaborated in promoting the *Livingstone Spectral Imaging Project*.

In addition, our previous efforts have laid the groundwork for an outreach program for schoolchildren. The Livingstone's Diary Outreach Program (LDO), a pilot version of which we will fund through LEAP, will adapt the broad humanities and scientific themes of *Livingstone Online* and the *Livingstone Spectral Imaging Project* for the classroom and deliver a multidisciplinary teaching program for students aged 10-14. The pilot will be run in collaboration between *Livingstone Online*, the David Livingstone Centre, Blantyre, and seven Blantyre primary and secondary schools. Ultimately, LDO will result in the production of a collection of digitized teaching resources that will be made freely available online through *Livingstone Online* and that will be offered for the websites of other relevant archival and education institutions. In future phases of the program, which fall outside of the present grant application, we will broaden the program's national scope in the UK and implement a pilot program for US schools (see Appendices for a longer program description).

We will give conference presentations and lectures to academic audiences, and mark project milestones by directly contacting specialists interested in our work and by distributing academic announcements to listservs such as DLF-Announce, Victoria, Humanist, H-Africa, H-Slavery, H-Travel, TEI-L, and Sharp-L. As the proposed project matures, we will submit the integrated *Livingstone Online* site to cooperative networks and digital resource aggregators, including NINES. We will solicit formal reviews of the completed site from a multidisciplinary range of journals, including *Digital Humanities Quarterly*, *Romanticism and Victorianism on the Net*, *Scholarly Editing*, and *African Research & Documentation*.

F. Work Plan

LEAP will be a three-year project (May 2013 – April 2016). We have broken down our activities into work periods of six months. Overall project planning will be conducted by Toth (project manager), and managed and developed with Wisnicki (PI) and McAulay (UCLA lead). For the activities listed below, a lead person is designated with * as necessary, although almost all tasks will require a group of staff to complete. Dissemination and documentation will be ongoing. For communication, the team will use in-person meetings, teleconferences, and networked communications.

⁴⁴ See Appendices for a full list of the print and broadcast media press coverage.

⁴⁵ http://livingstone.library.ucla.edu/1871diary/analysis_dissemination.htm

Activities (Note: “Scholars” = Ball, Lawrence*, Simpson, Ward, Wisnicki*.)

May 2013 - Oct. 2013 (Kickoff, Digitization, Data Analysis, Migration, Transcription, Outreach)

- Detailed project plan and integrated schedule completed (McAulay, Toth*, Wisnicki)
- Meetings in the UK with imaging, encoding, and repository staff (Emery, Toth, Wisnicki)
- Digitization of materials from DLC (2,500 pages) (DLC, Manchester, Toth*)
- Digitization of materials from NLS (1,000 pages) (Emery, NLS*, Toth)
- Analysis of 2005-10 and 2010-12 TEI encoding (Cummings*, McAulay, Wisnicki)
- Evaluation of existing image and transcription files for archive creation (Emery, Wisnicki)
- Migration of pre-2013 materials; begin development of site and streamlined functionalities through Islandora (Ghorpade*, McAulay, Ward, Wisnicki*)
- Test transcription and encoding of Field Diary I and select DLC letters (Scholars, Cummings)
- Preparation of outreach teacher packs; first phase of outreach (DLC, Simpson*)

Nov. 2013 - Apr. 2014 (Archive/Site Design, ODD, Transcription, Outreach)

- Preliminary archive design; test transfer of archival data to UCLA (Emery*, McAulay)
- Ingest test transfer of archival data into Islandora (Emery, Ghorpade*, McAulay)
- Continue and conclude development of site and streamlined functionalities; web release of new site (Ghorpade*, McAulay, Ward, Wisnicki*)
- Development of TEI P5 ODD customization and generated schema (Cummings*, Wisnicki)
- Initial conversion of legacy transcriptions to agreed TEI P5 schema (Cummings)
- Finalizing of TEI P5 ODD customization documentation (Cummings*, McAulay)
- Transcription and encoding of Field Diaries II, III, and IV and Unyanyembe Journal (Scholars)
- TEI Encoding review of Field Diaries and Unyanyembe Journals transcriptions (Cummings)
- Second phase of pilot for outreach program (DLC, Simpson*)

May 2014 - Oct. 2014 (Archive Delivery, Transcription, Experimentation, Outreach)

- In-person meeting in the UK with imaging and repository staff (Wisnicki)
- Beginning archive preparation and delivery of existing image data (Emery*, Wisnicki)
- Ingest of incoming data through Islandora (Emery, Ghorpade*, McAulay, Wisnicki)
- Proofreading and quality assurance analysis of converted legacy transcriptions (Cummings, Wisnicki)
- Transcription and encoding of Field Diaries V, VI, and VII and Unyanyembe Journal (Scholars)
- TEI Encoding review of Field Diaries and Unyanyembe Journals transcriptions (Cummings)
- Integrate and experiment with IIPImage, ImageJ, and Paleo (Christens-Barry, Ghorpade*, McAulay, Ward)
- Outreach report and resources released on site; design US pilot (Simpson*, Ward*, Wisnicki)

Nov. 2014 - Apr. 2015 (Archive Delivery, Transcription, Experimentation)

- Ongoing archive preparation and delivery of image data (Emery*, Wisnicki)
- Ongoing ingest of data into Islandora (Emery, Ghorpade*, McAulay, Wisnicki)
- Transcription and encoding of Field Diaries X, XIV, and XV and Unyanyembe Journal (Scholars)
- TEI Encoding review of Field Diaries and Unyanyembe Journals transcriptions (Cummings)
- Research for introductory and critical materials for Diaries and Journal (Scholars)
- Conclude work with IIPImage, ImageJ, and Paleo; produce technology report (Ghorpade*, McAulay*, Ward)

May 2015 - Oct. 2015 (Archive Delivery, Transcription, Critical Materials, Experimentation)

- In-person meeting in UK with stakeholders and assess progress (Wisnicki)
- Ongoing archive preparation and delivery of image data (Emery*, Wisnicki)
- Ongoing ingest of data into Islandora (Emery, Ghorpade*, McAulay, Wisnicki)
- Transcription and encoding of Field Diaries XVI and XVII and Unyanyembe Journal (Scholars)
- Begin writing introductory and critical materials for Diaries and Journal (Scholars)
- Incorporate and experiment with Meandre and SharedCanvas (Ghorpade*, Ward, Wisnicki)

Nov. 2015 - Apr. 2016 (Archive Delivery, Critical Materials, Experimentation)

- Final archive preparation and delivery of image data (Emery*, Wisnicki)
- Final ingest of data into Islandora (Ghorpade*, McAulay, Emery, Wisnicki)
- Continue and conclude writing introductory and critical materials for Diaries and Journal (Scholars)
- Incorporate and experiment with Meandre and SharedCanvas; produce technology report (Ghorpade*, McAulay*, Ward, Wisnicki)

Deliverables

- Fully functional, enhanced, and integrated *Livingstone Online* website with open and integrated access to images and text transcriptions of circa 5700 pages of Livingstone manuscript materials;
- An open raw data archive of all the archival TIFF images, metadata, and TEI/XML; web-accessible and open for download;
- A scholar-focused set of web-based tools integrated into *Livingstone Online* for image manipulation, textual analysis, and interoperable ;
- A customized TEI schema for Livingstone Online with associated documentation (ODD);
- Full project documentation; technology report on integration of tools and applications;
- Livingstone Diary Outreach Program (LDO) digital teacher packs and documentation for site.

G. LEAP Staff

Our **staff** consists of *Livingstone Online* core team members (Ball, Lawrence, Simpson, Wisnicki), individuals with specific project responsibilities who have participated in previous data access projects including the *Livingstone Spectral Imaging Project* (Christens-Barry, Cummings, Davison, Emery, Ghorpade, McAulay, Toth), and new collaborators recruited for LEAP (Horning, Ward).

Ball, Heather F. Morgan Library & Museum. Performed transcription and proofing of Livingstone's writings for the *Livingstone Spectral Imaging Project* and *Livingstone Online*; created content for the projects' websites and oversaw the work of other assistants; created, collated, and integrated folio-level catalog records for digitized manuscripts into the Morgan's online catalog, CORSAIR. *LEAP duties*: Collaborate in transcribing, encoding, and researching on Livingstone manuscripts. Continue as *Livingstone Online* head research assistant. *Time commitment*: 1,400 hours.

Christens-Barry, Bill. Equipoise Imaging, LLC. Has participated as core scientific member in numerous cultural heritage imaging projects and is inventor of the LED-based narrowband illumination technique used in the field. Developed the Paleo image toolbox. *LEAP duties*: Consultant on Paleo development and modification to match processing, annotation, and metadata requirements of LEAP; serve on advisory board (see below). *Time commitment*: 80 hours.

Cummings, James. University of Oxford, IT Services, Manager of InfoDev (Research Support and Data Solutions). Specialist on text encoding and markup languages. Chair and elected member of the Text Encoding Initiative (TEI) Technical Council. *LEAP duties*: Design and implementation of TEI

conformant ODD customization for all existing and future *Livingstone Online* materials. Legacy Data migration from TEI P4 to customized TEI P5 schema. TEI, data migration, and XSLT consultation. Proofreading encoding choices for new transcriptions. *Time Commitment*: 240 hours.

Davison, Stephen. Head, UCLA Digital Library. Leads the UCLA Digital Library Program and oversees its collection development, technical infrastructure, and project management. Has served as the DLP project lead on several grant-funded projects, including three NEH-funded projects: two phases of the Strachwitz Frontera Collection of Mexican and Mexican American Sound Recordings⁴⁶ and the Caro Minasian Persian and Arabic Manuscripts.⁴⁷ *LEAP duties*: Provide administrative oversight of UCLA Library's participation in the project, supervise project manager McAulay and project programmer Ghorpade, manage the project budget, and provide progress and budgeting reports to Library management and the UCLA Office of Contracts and Grants Administration. *Time commitment*: 3% FTE.

Emery, Doug. Emery IT. Data Manager for the Archimedes Palimpsest Project, Syriac Palimpsest Project, Sinai Palimpsests Project, and *Livingstone Spectral Imaging Project* as well as the Walters Art Museum's several NEH-funded manuscript imaging projects. *LEAP duties*: Design and assemble public data archive, delivery of collected data to UCLA. *Time commitment*: 250 hours.

Ghorpade, Parinita. Project Programmer, UCLA Digital Library. Programmer/Analyst III responsible for digital library development, especially XSLT, Java, and Solr programming. Was the programmer on the NEH-funded Caro Minasian project (see above) and has done all the programming to support the *Livingstone Spectral Imaging Project*. *LEAP duties*: Responsible for all programming and data management for the project. Write code for data management, data ingest, access, delivery services, metadata mapping services, and metadata remediation. *Time commitment*: 20% FTE.

Horning, Claudia. Metadata specialist, UCLA Library Cataloging and Metadata Center (CMC). Head of the Metadata Section within the CMC; oversees metadata creation for digital library projects. *LEAP duties*: Provide all metadata advice for the project, including advising on metadata standards to adopt; the level of detail needed to support discovery, management, and preservation services; and guidelines on the implementation of metadata standards. Work with the Project Manager to ensure high quality metadata creation and mediation, if needed. Ensure consistent application of metadata guidelines across the entire collection. *Time commitment*: 5% FTE.

Lawrence, Christopher. Professor Emeritus of the History of Medicine, University College London. Internationally recognised historian of medicine and science. Has published and edited numerous books, including volumes of historical correspondence. The original founder of *Livingstone Online* and now co-director. *LEAP duties*: Co-direct overall development of *Livingstone Online* with A. Wisnicki; write essays that contextualize primary Livingstone manuscripts; review transcriptions of Livingstone manuscript materials. *Time commitment*: 1,560 hours.

McAulay, Lisa. Librarian for Digital Collection Development, UCLA Digital Library Program. Oversees the development and execution of digital projects, including the St. Gall Monastery Virtual Library,⁴⁸ the metadata and text encoding of the UCLA Encyclopedia of Egyptology, and the *Livingstone Spectral Imaging Project*. Manages project content models, ingest procedures, quality control, and TEI/XML activities. *LEAP duties*: Project Manager for UCLA Library's participation in the project. Coordinates activities of the Project Programmer and Metadata Librarian; supervises Student Assistants. Principal point of contact with project partners. Responsible for all aspects of UCLA participation in the project. Provides progress reports as needed. *Time commitment*: 25% FTE.

⁴⁶ <http://frontera.library.ucla.edu/>

⁴⁷ <http://minasian.library.ucla.edu/>

⁴⁸ <http://www.stgallplan.org/>

Simpson, Kate. Graduate Student, Edinburgh Napier University. Has been a research assistant for *Livingstone Online* and the *Livingstone Spectral Imaging Project* since 2010. Researches imperial adventure fiction and exploration in Africa; works as a librarian within her university; created metadata at the NLS for the International Mission Photography Archive.⁴⁹ *LEAP duties:* Collaborate in transcribing, encoding, and researching on Livingstone manuscripts; guide the outreach project for schools in collaboration with staff from the DLC and local Scottish teachers. *Time commitment:* 1,200 hours.

Toth, Michael B. President, R.B. Toth Associates. Develops, refines, and implements new technologies and standardized work processes for efficient digitization of cultural objects. Over 25 years of experience with systems integration and program management for cultural and scientific institutions. *LEAP duties:* Planning, scheduling and management for the overall project; Systems Integration of the technologies, data and work processes of the participating institutions and personnel, in particular digitization and encoding. *Time Commitment:* 240 hours total.

Ward, Megan. Assistant Professor of English, Point Park University. Expertise in Victorian literature and the periodical press, the history of the novel, and early artificial intelligence. Recently attended an intensive course on the Meandre Workbench at the Digital Humanities Summer Institute at the University of Victoria. Has over ten years of teaching experience. *LEAP duties:* Collaborate in transcribing, encoding, and researching on Livingstone manuscripts. Oversee the development of manuscript tools, especially the Meandre workbench. Plan U.S. pilot of Livingstone Diary Outreach Program (LDO). *Time commitment:* 1,560 hours.

Wisnicki, Adrian S. (Principal Investigator). Assistant Professor of English, Indiana University of Pennsylvania. Internationally-recognized specialist with publications and lectures on Victorian literature, nineteenth-century African history, and British exploration, colonial, and travel literature focused on Africa. Serves as project director for the NEH-funded, peer-reviewed *Livingstone Spectral Imaging Project*. *LEAP duties:* Co-direct overall development of *Livingstone Online* with C. Lawrence; review LEAP results and progress; liaise with major stakeholders; supervise and contribute to transcription, encoding, and research on Livingstone manuscripts; document LEAP activities as public record of project history; disseminate project results. *Time Commitment:* 2,250 hours.

The **institutional coordinators** will serve as our points of contact with the David Livingstone Centre (Karen Carruthers, Anne Martin) and the National Library of Scotland (David McClay, Alison Metcalfe). They will also work with LEAP staff to facilitate the digitization and archiving of relevant Livingstone manuscripts.

We will organize periodic progress reviews by the **advisory board**, but we will also call on the expertise of individual board members as relevant. The board will provide a range of expertise relevant to our project, including: David Livingstone scholarship (Lawrence Dritsas, University of Edinburgh); preservation and conservation science (Fenella France, Library of Congress); digitization (Bill Christens-Barry, Equipoise Imaging LLC); web development (Ben Albritton, Stanford University); data management (Steve Van Tuyl, Carnegie Mellon University); cultural heritage preservation (Alison Marsh, University of South Carolina); manuscript curation and access (William Noel, The Walters Art Museum and, from 5 Sept. 2012, University of Pennsylvania); descriptive and technical metadata (Dot Porter, Indiana University); and dissemination and outreach (Debbie Harrison, Birkbeck, University of London).

⁴⁹ <http://crcc.usc.edu/initiatives/imp/>

Bibliography of Works Cited in the Application Narrative

- Boucher, Maurice, ed. 1985. *Livingstone Letters, 1843 to 1872: David Livingstone correspondence in the Brenthurst Library, Johannesburg*. Houghton, South Africa: Brenthurst Press.
- Bridges, Roy C. 1987. "Nineteenth-Century East African Travel Records with an Appendix on 'Armchair Geographers' and Cartography." *Paideuma* 33: 179-96.
- Bridges, Roy C. 1998. "Explorers' Texts and the Problem of Reactions by Non-Literate Peoples: Some Nineteenth-Century East African Examples." *Studies in Travel Writing* 2: 65-84.
- Cashdan, Elizabeth. "Trade and Its Origins on the Botletli River, Botswana." *Journal of Anthropological Research* 43: 121-38.
- Clendennen, G. W., ed. 1976. *David Livingstone on the Zambezi: Letters to John Washington, 1861-1863*. Pasadena, CA: Munger Africana Library, California Institute of Technology.
- Clendennen, G.W., ed. 1992. *David Livingstone's Shire Journal, 1861-1864*. Aberdeen: Scottish Cultural Press.
- Clendennen, G. W., and I. C. Cunningham, eds. 1979. *David Livingstone: A Catalogue of Documents*. Edinburgh: National Library of Scotland.
- Cunningham, I. C., ed. 1985. *David Livingstone: A Catalogue of Documents: A Supplement*. Edinburgh: National Library of Scotland.
- Dritsas, Lawrence. 2005. "From Lake Nyassa to Philadelphia: A Geography of the Zambesi Expedition, 1858-64." *The British Journal for the History of Science* 38: 35-52.
- Du Toit, Andre. 1983. "No Chosen People: The Myth of the Calvinist Origins of Afrikaner Nationalism and Racial Ideology." *The American Historical Review* 88: 920-52.
- Endfield, Georgina H.; David J. Nash. 2002. "Drought, Dessiccation and Discourse: Missionary Correspondence and Nineteenth-Century Climate Change in Central Southern Africa." *The Geographical Journal* 168: 33-47.
- Flint, Lawrence. 2003. "State-Building in Central Southern Africa: Citizenship and Subjectivity in Barotseland and Caprivi." *The International Journal of African Historical Studies* 36: 393-428.
- Foskett, R., ed. 1964. *The Zambesi Doctors. David Livingstone's Letters to John Kirk, 1858-1872*. Edinburgh: University Press.
- Helly, Dorothy O. 1987. *Livingstone's Legacy: Horace Waller and Victorian Mythmaking*. Athens, OH: Ohio University Press.
- Holmes, Timothy. 1990. *David Livingstone: Letters & Documents 1841-1872: The Zambian Collection at the Livingstone Museum*. London: Currey.
- Koivunen, Leila. 2009. *Visualizing Africa in Nineteenth-Century British Travel Accounts*. New York: Routledge.
- Larner, A.J. 2004. "Ophthalmological Observations Made During the Mid-19th-Century European Encounter With Africa." *Arch Ophthalmol* 122: 267-72.
- McCracken, John. [1977] 2000. *Politics and Christianity in Malawi 1875-1940: The Impact of the Livingstonia Mission in the Northern Province*. Zomba, Malawi: Kachere Series.
- Pearce, G.D. 1985. *Bulletin of the British Mycological Society* 19: 39-50.
- Rijpma, Sjoerd. 1996. "Malnutrition in the history of tropical Africa." *Civilisations* 43: 45-63.
- Schapera, Isaac, ed. 1959. *David Livingstone, Family Letters, 1841-1856*. London: Chatto & Windus.
- Schapera, Isaac, ed. 1960. *David Livingstone, Private Journals, 1851-1853*. Berkeley: University of California Press.
- Schapera, Isaac, ed. 1961. *Livingstone's Missionary Correspondence, 1841-1856*. Berkeley: University of California Press.
- Schapera, Isaac, ed. 1963. *David Livingstone, African Journal, 1853-56*. London: Chatto & Windus.
- Schapera, Isaac, ed. 1975. *David Livingstone, South African Papers, 1849-1853*. Cape Town: Van Riebeeck Society.
- Sheriff, Abdul. 1987. *Slaves, Spices and Ivory in Zanzibar: Integration of an East African Commercial Empire into the World Economy, 1770-1873*. London: James Currey.
- Wisnicki, Adrian S. 2011-12. "Livingstone in 1871." In *Livingstone's 1871 Field Diary: A Multispectral Critical Edition*. Los Angeles: Livingstone Online and the UCLA Digital Library Program.
<http://livingstone.library.ucla.edu/1871diary/livingstone1.htm> (accessed 7 July 2012).