

# The Thesaurus Linguae Graecae®: Specialised Historical Content for a Niche Audience

University of California, Irvine, California, USA

www.tlg.uci.edu



Online resources are often described and evaluated in terms of their ability to serve vast amounts of diverse content to wide audiences, but well-targeted, specialised digital projects can have a profound effect on an academic discipline. The Thesaurus Linguae Graecae® (TLG®), a digital corpus of over 12,000 works of Greek literature ranging from the ancient era to the modern age, has proven its value to scholars and has been able to convert that value into a sustainability model that incorporates multiple revenue streams. The resource is targeted towards academic classicists and medievalists, who rely on it as the only comprehensive body of historical Greek-language works available online; it also offers a small Open Access selection of canonical Greek works for use by a wider audience. The project, which is hosted at the University of California, Irvine, depends on three main revenue streams: subscription fees, direct financial support from the university and a project endowment. The endowment was originally intended to supplement the other two revenue streams, but the project's goal now is for the fund to some day cover all of the ongoing costs for the TLG. This case study looks at some of the questions facing the TLG and outlines the broader implications for other resources with highly specialised content: How does such a project build an audience and keep users excited and engaged? What characteristics make a project a strong candidate for a subscription model? And how do the leaders of the TLG envision their resource – and its funding – evolving in the future?

## Introduction

The quest to publish and preserve the entirety of Greek literature has a rich history, stretching back to the efforts of humanist scholars and printers in 16th-century Europe. Only in the past four decades, however, has technology provided a solution to this problem, through a unique digital humanities project hosted at the University of California's Irvine campus. The Thesaurus Linguae Graecae® (TLG®) is an online compendium of Greek-



language works, ranging from Homer to the 15th century CE. The original goal of the project was to create an electronic resource of all ancient Greek texts; having now largely accomplished that, the project leaders have expanded the historical scope of the TLG to include texts from the Byzantine and modern periods. The digitised texts are displayed in Greek font (rather than as page scans) and are searchable by standard criteria – author, title, date of composition – as well as by specialised fields such as author epithet, geographical origin and publication information. In addition, users can search the entire TLG database using either Greek or Latin transliteration. The site is updated two to three times a year, with between 20 and 30 newly digitised works added each time.

The TLG was initially funded not with a grant from a charitable foundation but through a donation from a graduate student. In 1972, Marianne McDonald, a PhD candidate in classics at the University of California, Irvine (UCI), gave \$1 million to start the project at the university, where it was led by classics professor Theodore Brunner for its first 25 years.<sup>1</sup> Over the past four decades, the project's computer-ready texts have migrated formats, from magnetised computer tape, to CD-ROM, to online access. Today, the TLG has nearly 1,000 subscriptions from higher education institutions and another 1,000 from individual

This case study was researched and written by Matthew Loy as part of the Ithaka Case Studies in Sustainability project.

<sup>1</sup> McDonald, now Professor of Theatre and Classics at the University of California, San Diego, is well known for her beneficence; her gift-giving was itself the subject of a general-interest case study. See Jerold Panas, *Mega Gifts: Who Gives Them, Who Gets Them* (Medfield, MA: Emerson and Church, 1995).

scholars.<sup>2</sup> The project is overseen by Professor Maria Pantelia of the UCI Classics Department and Assistant Director Betsy Shanor. At the time of Pantelia's appointment to the directorship, the university promoted the TLG as 'the largest humanities-based electronic resource ever created in the United States'.<sup>3</sup>

## Sustainability model

### Goals and strategy

The TLG has two chief strategies for covering its direct costs: drawing revenue from multiple streams (including an endowment and five-year-long subscription contracts) and keeping costs low through careful prioritising and selective outsourcing of digitisation projects.

*The project currently pays a digitisation firm in China that charges \$1.58 per 1,000 keystrokes. These cost savings allow the project to focus resources on editing and correcting the digitised works.*

### Costs

The TLG's greatest costs are salaries and benefits for the project's nine staff members (7.15 FTE), which account for about \$400,000 annually, or 80% of the TLG's overall budget of approximately \$500,000.<sup>4</sup> Part of the director's salary is paid by UCI through its direct cash contribution to the TLG; the university counts Pantelia's leadership of the TLG as one-half of her appointment, and teaching two courses per year (alongside research and other faculty duties) as the other half. (The half of her appointment related to teaching and standard faculty research and administration is not part of the TLG's budget.) The project's assistant director is employed full time and is also partially paid by UCI. In addition, the TLG's budget covers the salaries of three programmers (two at full time and one at 40%), one full-time researcher to oversee text and data entry, one part-time researcher and one graduate student researcher (each at 50%) and an in-house distribution and licensing specialist (at 75%).

<sup>2</sup> This number includes a small number of subscriptions to the TLG's pre-internet incarnation as a CD-ROM (on which more later in this case study). These subscriptions are slowly being phased out.

<sup>3</sup> 'Native of Greece to Oversee UC Irvine's Pioneering Computer Data Bank of Ancient Greek Texts'. *Today@UCI* [30 October 1996], [www.today.uci.edu/news/release\\_detail.asp?key=234](http://www.today.uci.edu/news/release_detail.asp?key=234)

<sup>4</sup> All budget figures and estimates were provided by the project leaders. For further detail on the financial data presented in this report, please see Appendix B: Summary of revenues and costs.

The TLG enjoys considerable cost savings by outsourcing elements of the digitisation process. The project has a long tradition of farming out its data entry, starting with the production of computer tapes in the 1970s; as Pantelia jokingly likes to say, 'We invented out-sourcing!'<sup>5</sup> The project currently pays a digitisation firm in China that charges \$1.58 per 1,000 keystrokes. These cost savings allow the project to focus resources on editing and correcting the digitised works. First, the project's researchers locate a text they wish to digitise, which can often be difficult given the requests they receive from users, which are often for obscure texts. Then, the researchers photocopy the volume and ship the copy to the Chinese firm, where the text is hand-keyed to create an electronic file. TLG staff still must devote considerable time to editing and correcting the resulting file.<sup>6</sup> A software program developed in-house scans the text for irregularities, after which the graduate student and the full-time researcher check the text a second and third time, often comparing the text to existing print sources, reviews and relevant bibliographies to correct errors and flag ambiguities. Any particularly difficult editing questions are then passed on to Pantelia, who has the final jurisdiction. According to Pantelia, the TLG text may have up to 100 changes per page compared to the printed original.

### Revenues

The project takes in approximately \$500,000 per year from a combination of three sources: institutional and individual subscriptions (which account for approximately half of the project's total income), investment returns from the project's endowment and direct financial support from UCI. Aside from the donations and grants which provided start-up funding for the project, subscriptions are the longest-running source of funding for the project; UCI began providing grant support in 1987, and the endowment was founded in 1992.

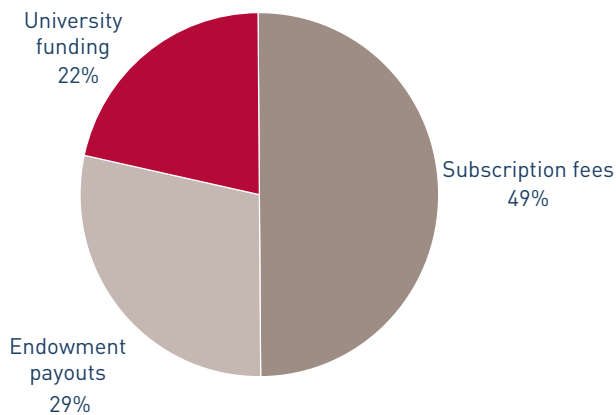
*Subscriptions.* The TLG began offering subscriptions in the mid-1980s with a CD-ROM version of the corpus, which was licensed for individual or campus single-workstation use. The release of the subscription CD-ROM product was concurrent with a time when the project was struggling with financial sustainability. 'It [was] daunting,' Shanor says, recalling years in the 1980s which included four staff members being laid off and uncertainty about the long-term survival of the resource. She credits the subscription fees not only for generating revenue but for smoothing the year-to-year income of the project: 'We don't have the highs and lows of fundraising...We now have a steady stream of income.'

Individual subscriptions to the TLG cost \$100 for one year or \$400 for five years. Institutional subscriptions can be purchased for discrete numbers of computer terminals or for unlimited campus-wide use; in either case, the institutional subscriptions are available for five-year periods only. A five-year subscription for up to three terminals on a campus costs \$1,500. Approximately 60% of institutional subscribers, according to Shanor, opt for the unlimited campus-wide access plan. Most institutional subscriptions for campus-wide access range from \$3,500 to \$10,000 for a five-year period, although multi-campus institutions are charged higher fees.

<sup>5</sup> Unless otherwise noted, all quotations from staff members and other individuals knowledgeable or associated with the TLG are drawn from interviews conducted as part of this case study during December 2008 and January 2009. A full list of interviewees is included in Appendix A.

<sup>6</sup> Unlike some digitisation projects, the TLG digitisation process does not involve scanning the pages of the documents to mount digital facsimiles, but rather involves creating a highly accurate transcription of the original text.

TLG's Sources of Revenue



Shanor credits the subscription fees for smoothing the year-to-year income of the project: 'We don't have the highs and lows of fundraising... We now have a steady stream of income.'

The TLG considers several factors when determining the price of an initial five-year subscription. Some are typical of online resources (eg total student enrolment), while others are keyed to the pool of potential users in classics departments, including the number of classics faculty members and the highest degree offered by the department. (The TLG collects this information from schools through a pre-subscription questionnaire.) Pantelia and Shanor use this information to place the school in one of three tiers of pricing, based on an estimate of how much the TLG will be used on the campus. At the time of renewal, the actual usage on each campus is taken into consideration (in addition to the factors listed above) in determining the subsequent five-year renewal subscription fee.

The subscriptions are priced in five-year increments in part to save time and money on billing services. It is possible that these costs could be decreased further if billing services were outsourced, but the project leaders want to retain quality control over this function. Subscribers usually pay the full five-year subscription fee up front, but they have the option of making annual payments, and may cancel at any time to receive a refund for the unused portion of the subscription fee.

**Endowment.** In the early 1990s, it became clear to Theodore Brunner, the project's first director, that the long-term survival of the TLG was in jeopardy. Brunner believed that the five-year subscription fees for the CD-ROM product were insufficient to sustain the TLG, but he worried that raising the subscription

price would put the product out of reach of individual subscribers and less well-resourced institutions. Eager to ensure the long-term viability of the TLG, Brunner submitted a challenge-grant proposal to the National Endowment for the Humanities (NEH) to start building an endowment. Although the project had received approximately \$7 million in funding up to this point, the bulk of its private and government grants were for digitisation and data-creation activities, not long-term maintenance, preservation or upgrades; the endowment was intended to supplement, rather than replace, subscription fees and university support.<sup>7</sup>

The endowment fund was built from private donations raised alongside a \$500,000 three-to-one matching-funds Challenge Grant from the NEH from 1993–1999, as well as a matching grant from the Andrew W. Mellon Foundation. The endowment is managed by the Office of the President of the University of California system, alongside the university's endowment as a whole.

The current project leader, Pantelia, would like the TLG to be completely Open Access at some point in the future. She estimates that the endowment would need to grow to \$8–10 million to meet her goal of making the resource freely available. Although there is no firm time frame for a concentrated capital campaign, she plans to focus fundraising efforts on private donors, rather than seeking larger commitments from charitable foundations or academic libraries. She feels that the first endowment-building phase demonstrated strong support from donors, and the project's \$7 million in prior funding included two major gifts from individual classicists.

In an interview at the end of 2008, Pantelia estimated the pre-recession value of the endowment at \$3 million. As of January 2009, the impact of the recent economic downturn on the project's endowment was still not clear to the project leaders. Although Pantelia and Shanor were uncertain how significant the short- and long-term effects will be for their fund, they expressed confidence that the TLG is on a relatively firm financial footing. 'Because our endowment is not our only source of income,' Pantelia said, 'so far, we are okay.' For now, they plan to keep costs low and defer any hiring (in accordance with a university-wide hiring freeze).

**University support.** In 2008–2009, UCI will provide the project with approximately \$108,000 in direct funding. (From the project's inception, the university had made in-kind contributions of office space and accounting and payroll services, as well as providing direct support by paying some salaries.) Shanor acknowledges that there is some inherent uncertainty in the level of ongoing university support, since UCI's endowment is also being affected by the broader economic slowdown.<sup>8</sup> This has already impacted these payments: the 2008–2009 figure already includes a 10% budget cut over the previous year.

Pantelia notes that the university began providing this \$100,000+ segment of direct funding for the project as a condition of an NEH grant and has continued to do so since then. The project was founded only seven years after the Irvine campus was chartered in 1965, so it has a long history at the university, and Pantelia

<sup>7</sup> 'The resultant \$2,000,000 [of the planned endowment] is to be established as a permanent TLG endowment, the annual yield of which – combined with user fee revenues and UCI contributions – would sustain TLG operations on a permanent basis.' See 'NEH Challenge Grant', *Thesaurus Linguae Graecae Newsletter*, 20 (May 1992), 3.

<sup>8</sup> Between 31 December 2007 and 30 September 2008, the University of California system's endowment dropped from \$6.7 billion to \$5.7 billion. For more on this, see Tanya Schevitz, 'UC's Endowment Plunged \$1 Billion', *San Francisco Chronicle* (13 November 2008), B-3. Available at: [www.sfgate.com/cgi-bin/article.cgi?file=c/a/2008/11/13/BA781436M4.DTL](http://www.sfgate.com/cgi-bin/article.cgi?file=c/a/2008/11/13/BA781436M4.DTL)

believes this may have helped the TLG win continuous support from administrators. Still, administrators ‘come and go’, she points out, so she considers reaching out to new campus officials and talking to them about the TLG an important part of her job.

## Key factors influencing the success of the sustainability model

### Understanding value to the users

The audience of the TLG is very different from the audiences for other online academic resources such as the Stanford Encyclopedia of Philosophy, which appeals to a wider range of users than the academic philosophers for whom it was originally intended, or the UK National Archives’ projects to digitise genealogical records, which are popular with the public at large. Formal researchers are the primary audience for the TLG – Pantelia told us she often hears from scholars who say they could not do their research without access to an online compendium of texts – so project outreach focuses on them. Pantelia promotes the resource at professional meetings such as the annual conference of the American Philological Association, and forwards news of site updates to relevant blogs and listservs. ‘Because we are directed at a particular audience,’ she says, ‘it’s not a question of letting people know that we exist, but of informing people about what we are doing.’ Outreach to secondary audiences – high school and beginning undergraduate students, casual readers and others – is minimal. In addition, there is a non-academic audience of individual subscribers (some of them in Greece) who use the resource for pleasure reading rather than for academic research.

*...the TLG has attracted its audience in part because of a first-mover advantage: it is the only complete, centralised compendium of ancient and Byzantine Greek texts on the internet...*

Beyond the specialised nature of the content and its applicability to research needs, the TLG has attracted its audience in part because of a first-mover advantage: it is the only complete, centralised compendium of ancient and Byzantine Greek texts on the internet, with approximately 105 million words. Its major online counterpart, the Open Access Perseus Digital Library hosted by Tufts University, holds only 8.1 million words of Greek texts, and it aims to serve high school and introductory-level undergraduate students rather than scholars.<sup>9</sup> Because the TLG is such a unique resource, it seems well positioned to charge a subscription fee – while classics scholars are a very small audience, the TLG’s content and search capabilities are crucial to their work.

<sup>9</sup> Perseus Digital Library, Greek and Roman Materials, ‘Word counts by language’, [www.perseus.tufts.edu/hopper/collection.jsp?collection=Perseus:collection:G-reco-Roman](http://www.perseus.tufts.edu/hopper/collection.jsp?collection=Perseus:collection:G-reco-Roman)

*To gauge the needs of these users, the project leaders conduct occasional user surveys, and they respond to user suggestions.*

To gauge the needs of these users, the project leaders conduct occasional user surveys, and they respond to user suggestions. The project’s staff members prioritise functionality suggestions if the change would benefit a large number of users. If the upgrade seems like a great deal of work for a low impact, staff members try to work around the problem individually with the user (for example, by doing a one-off extraction of specialised statistical data, rather than programming a new advanced search function to the site). When a user requests that a text be digitised and added to the TLG collection, the project leaders move ahead with the request ‘99% of the time,’ according to Pantelia; in rare cases, the correct volume cannot be found or the text exists in a very old edition and cannot be digitised easily.

### Adding value to the historical content

Because the TLG deals in historical texts rather than original content, there is some risk that the resource might lose appeal for users if it were to consist only of a static canon of texts, or if a mass digitisation project like Google Books were to build a rival corpus. To maintain the value of the resource to users – which, in turn, drives continued subscription renewals – the project team digitises new texts, edits these texts meticulously and layers on search functionality and specialised types of research unique to this content. Because nearly all the available works from the ancient period have been digitised, the project has widened its scope to include texts from the Byzantine period and the 16th and 17th centuries, and it will continue to expand forward in time as texts move out of copyright and financial resources allow for digitisation. The project’s leaders hope that the addition of these later works will draw new scholars to the TLG.

Still, Pantelia believes that tools are vital to the future of the resource; merely adding new works will not be sufficient to guarantee the TLG’s survival. ‘Digitisation is slowing down,’ Pantelia said. ‘The question now is infrastructure: how can we organise data and offer it to the world? And what tools will we use?’ While these are also open questions for future development of the resource, the previous attention paid to the TLG’s search functionality reflects the staff’s concern with the user experience. The search engine supports the Latin and Greek alphabets and includes a lemmatised search function through which a user can enter a dictionary word form and receive a myriad of different word forms that would otherwise have to be searched for one at a time. Several other search features specific to ancient and Byzantine Greek texts are available, including the ability to search by the degree to which the texts were corrected by their original editors (‘diplomatic’ and ‘vernacular’ editions). Pantelia believes that this functionality is a significant part of the resource’s value to its subscribers, and that this partially insulates the project from future competition with mass-digitisation projects like Google Books: ‘Humanities collections are very idiosyncratic... We’ll always need specialists to look at the particularities.’



### Consistent leadership

The project has retained leaders and staff for a strikingly long time. Over its nearly four decades of operation, the TLG has had only two directors: Theodore Brunner, who founded the project in the 1970s and led the TLG until his retirement from academia in 1997, and Pantelia, who has taken charge of the project since then. In addition, Shanor has been the assistant director of the TLG for 28 years, and the TLG's head informational technology specialists, Nishad Prakash and Nick Nicholas, have been with the project for ten years.

Shanor credits the dedication of Brunner and Pantelia for the continued existence of the project, highlighting the need for committed leadership for similar digital resources. 'These projects really become part of you. [For] Professor Brunner, [the TLG] became his personality; he really wanted it to survive. I think you need a leader who is that committed to make these things survive, and Professor Pantelia is like that, too.' And Pantelia emphasises the commitment that these projects demand: although UCI allocates half of her time to the TLG and half to teaching, research and university service, she notes that leading the resource is effectively a 'very demanding' full-time job.

The TLG also has an advisory board that offers advice to the project leaders on fundraising, the scholarly direction of the TLG (including future additions) and legal issues. Its members are mostly scholars of the classics and Byzantine studies, and two of the members are appointed by the American Philological Association. The board does not formally set subscription prices (those decisions rest with Pantelia and Shanor), but they do make general pricing recommendations – for example, that subscription fees should correlate with institution size and intensity of usage.

### Open Access versus gated content

The project's leaders have occasionally faced criticism for charging subscription fees for cultural heritage content; some believe that these works should be freely available. Pantelia and Shanor both spoke of their concern on this point. 'Sometimes you feel guilty charging [for the content],' Shanor commented, 'but you just have to...How do we sustain and improve [the TLG] if we don't have an income?' Still, Pantelia's ultimate goal for the project is to grow the endowment to the point that the entire resource can be offered free of charge and sustain itself through annual payouts from investments.

For now, the TLG serves a wider, non-subscribing readership through an Open Access selection of canonical texts by 68 writers (including Homer, Plato, Sophocles and Xenophon). These works are familiar to a wider readership and readily available in print editions, so there is little danger that this side project will cannibalise subscription revenue. Pantelia intends this portion of the site for beginner-level undergraduates, as well as for secondary students whose high schools cannot afford a subscription to the TLG. She hopes the Open Access selection will meet the needs of a non-scholarly audience, advance the goals of sharing Greek literature and promoting the classics profession, and partially address criticism from Open Access advocates.

### Benefits and challenges

A significant benefit of the TLG's sustainability model is its reliance on multiple streams of funding, which can lessen the impact of a drop in any one stream. Thus, the TLG can rely on its subscription revenue and internal university funding if endowment payouts drop.<sup>10</sup> Conversely, when the market is on an upswing, some excess funds can be used to reinvest in the resource and fund upgrades. The long history of the TLG demonstrates the critical importance of having sufficient resources for periodic upgrades and data migrations to meet evolving technology standards.

*Because the TLG's content is unique and considered indispensable to a core group of faculty, its subscription revenue is somewhat insulated from economic downturns...*

Because the TLG's content is unique and considered indispensable to a core group of faculty, its subscription revenue is somewhat insulated from economic downturns; academic librarians who make purchasing decisions based on faculty needs might drop other, non-core resources before the TLG. At the same time, the TLG's leadership recognises the importance of augmenting the value of the content to its users by investing in the development of tools and features.

The TLG is exceptional in that its host university manages the endowment and provides significant direct funding, and the project has taken steps to keep spending low on project marketing and billing for subscriptions. At the same time, the costs of editing and quality control are relatively high because the project targets scholars with exacting needs.

The five-year subscription period decreases billing costs, but it also hampers the project's ability to adjust prices quickly. The TLG's staff credits the subscription cycle for helping them achieve a more stable flow of income. But the initial five-year

<sup>10</sup> See the case study on the Stanford Encyclopedia of Philosophy elsewhere in this report for a detailed analysis of the benefits and challenges of the endowment funding model, many of which apply to the TLG's endowment as well.

cost of an institutional subscription is based on an *estimate* of usage, leaving a 60-month period during which the staff can do nothing to correct an underestimate of an institution's usage or respond to external market forces. Similarly, endowment funding can limit flexibility in responding to market opportunities.

## Broader implications for other projects

*Specialised niche resources, even with small audiences, can be good candidates for a subscription model if the resource is of high enough value to its users.* The TLG can charge subscription fees to supplement its other revenue streams in part because the content is well suited to a subscription package. Its content is irreplaceable to scholars working in this field, and they, in turn, are willing to ask academic libraries to subscribe or to purchase individual subscriptions themselves. At the same time, the content is so specialised that there is no other comprehensive digital library of primary sources to which scholars of ancient and Byzantine Greek can turn.

*Developing a strong relationship with users is as important as providing valuable content.* Although the TLG has a targeted core audience for its content, the project leaders do not take their users' loyalty for granted. They invest resources to keep content offerings from stagnating by developing search functionality, and by digitising new texts from later periods. Because the TLG deals in out-of-copyright works, the project's subscription revenues would potentially be vulnerable if a content aggregator were to digitise a duplicate corpus. Such a scenario seems unlikely now, given the heavy workload involved in digitising and editing classical texts, which are of interest to a relatively small audience and thus not commercially attractive. But the digital landscape is changing quickly, and it is difficult to foresee what content for-profit and not-for-profit digital resources might some day want to incorporate. For that reason, maintaining the loyalty of core users by responding to digitisation requests and adding value through new tools and search functionality may act as an insulator against competition for smaller, niche projects with relatively limited resources.

*Multiple revenue sources may reduce risk and offer greater financial stability.* The benefit of being able to depend on an alternative revenue source if a primary source falters seems obvious – and the TLG's endowment helps supplement its subscription model in just that way. That said, the development of multiple revenue streams can be difficult to achieve, as it may require significant staff time, expertise and pre-existing infrastructure. (Indeed, there is a debate in the not-for-profit community over the desirability of diverse revenue streams.<sup>11</sup>) Still, the TLG's three-pronged funding approach may serve it well during the current economic downturn.

<sup>11</sup> See William Foster and Gail Fine, 'How Nonprofits Get Really Big', *Stanford Social Innovation Review* (Spring 2007), 46–55. The authors, both at the Bridgespan Group, isolated financial data for 110 high-growth nonprofits and found that nearly 100 of these received more than 90 % of their funding from a single source. In a separate report on the same study, Bridgespan analysts hypothesise that this might be due to the relative expense and difficulty of pursuing multiple revenue streams (among other possible reasons). For those hypotheses, see William Foster, Ben Dixon, and Matt Hochstetler, 'In Search of Sustainable Funding: Is Diversity of Sources Really the Answer?' *The Nonprofit Quarterly* (Spring 2007), pp. 26–29.

*Cost management is vital to sustainability, even when a project's overall budget is modest.* Keeping direct costs low, as the TLG has done through digitisation outsourcing, a five-year billing cycle and close management of the project's budget, is also necessary for a niche resource that is highly dependent on its host institution and has limited prospects for exponential future subscriber growth.

*The potential importance of a host institution to a project's sustainability cannot be underestimated.* For small, high-value scholarly resources, nurturing a strong relationship with a host institution can be a key element of long-term sustainability. The TLG has been successful in this, winning significant direct financial support. Other projects may need to think about how they communicate value to their own host institutions, given that the true cost to the host institution includes the 'hidden' costs of in-kind contributions (such as office space and overheads). The value of these in-kind contributions may be much greater than the value of the direct funding.

“The TLG's project leaders do not believe that their current business model could support Open Access to the material.”

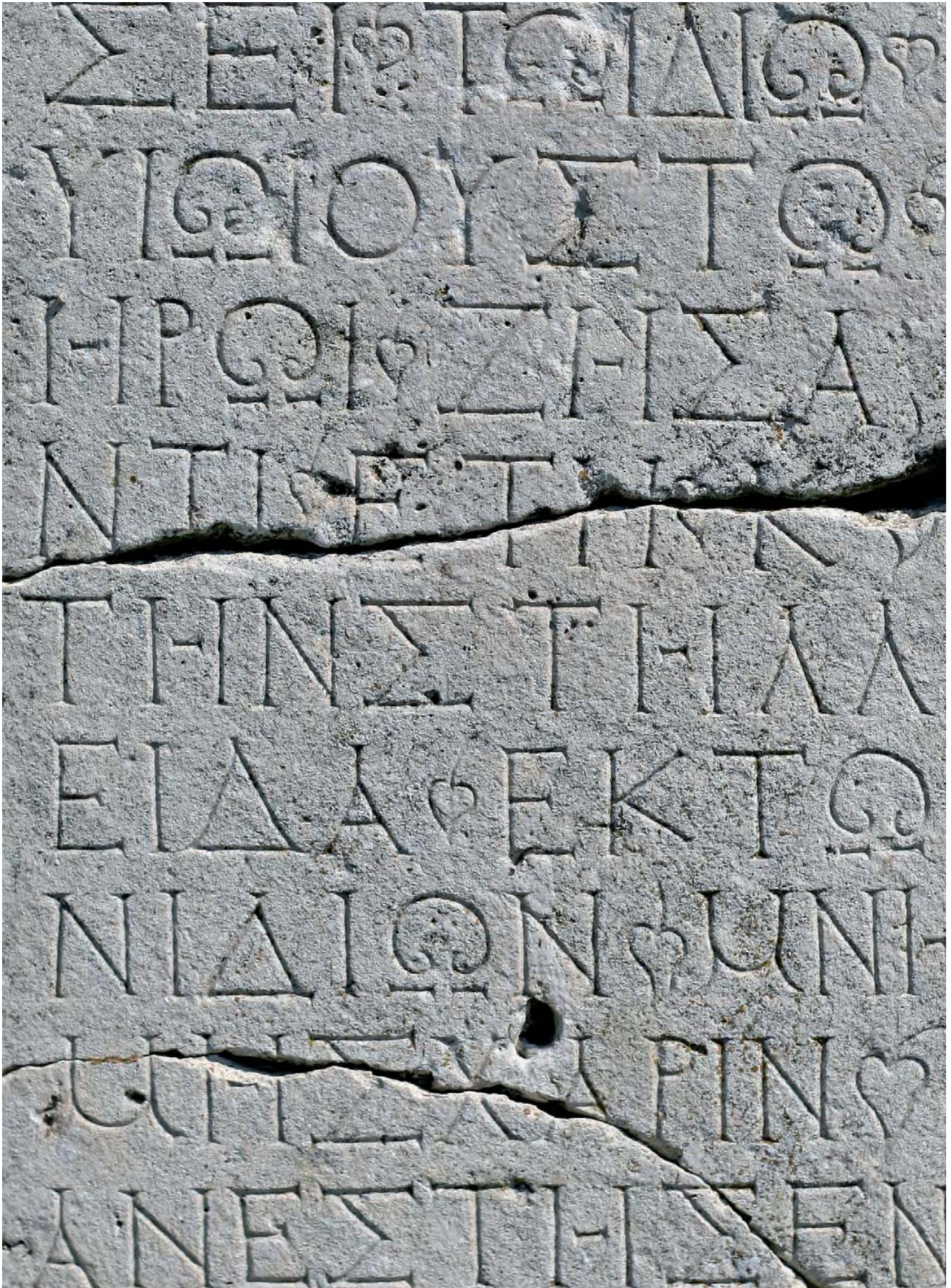
*A balance can be struck between free and gated content.* The TLG's project leaders do not believe that their current business model could support Open Access to the full body of TLG material – current endowment payouts and university support would be insufficient for the TLG's financial needs. However, they are able to serve up a substantial amount of free content, while keeping the bulk of the texts behind a subscription wall to generate needed revenue for sustainability.

*Host institutions can encourage digital projects by recognising their scholarly value.* At least in the US, tertiary institutions recognise the value of print scholarship in the tenure and promotion process. The impact of a digital project – particularly one that is not an e-monograph or online journal article – on a researcher's career is much less clear, despite the real benefits that the scholarly community realises from such work. In the case of the TLG, UCI includes the project management as part of a tenured position. This publicly recognises the scholarly value of the resource and provides a clearer picture of how the leadership of the resource will be handed on in the future: presumably, the leadership will be part of a future search to fill a tenured position.

## Appendix A: Interviewees

**Maria Pantelia**, Project Director, *Thesaurus Linguae Graecae* and Professor of Classics, University of California, Irvine, 17 December 2008 and 23 January 2009

**Betsy Shanor**, Assistant Director, *Thesaurus Linguae Graecae*, 15 December 2008 and 15 January 2009



## Appendix B: Summary of revenues and costs

### Thesaurus Linguae Graecae

Revenue Category	Description	Approx. amount
Subscription fees	Approximately 2,000 subscribing institutions and individuals	\$248,000
Endowment payouts*	Returns from the project's endowment	\$144,000
University support	Direct funds from the University of California, Irvine	\$108,000
<b>Total revenue</b>		<b>\$500,000</b>

\*Based on pre-recession endowment value.

Cost Category	Budgeted Costs			In-kind/volunteer contributions
	Description	FTE	Approx. cost	
<b>Personnel</b>				
Management	1 FT project manager & 1 FT assistant manager	2	partial	Half of management salaries paid by UCI
Content selection & production	1 FT researcher, 1 PT researcher, 1 PT graduate student researcher	2	yes	
Sales & Marketing	1 PT distribution and licensing specialist	0.75	yes	
Technology	2 FT & 1 PT programmers	2.4	yes	
<b>Total personnel costs</b>		7.15		<b>\$400,000</b>
<b>Non-personnel costs</b>			Included in budget?	
Administration & overhead	Furniture & office maintenance		partial	Office space provided by UCI; endowment management provided by UCI
Scanning, metadata, etc.	Approx. 20 works added every 6 mo; outsourced to China (\$1.58/1000 keystrokes)		yes	
Hosting & technology infrastructure	Project maintains its own servers and pays all software licences. Systems admin. provided by UCI (approx. \$450/mo)		yes	
Other				
<b>Total non-personnel costs</b>				<b>\$100,000</b>
<b>Total budgeted costs</b>				<b>\$500,000</b>

### Explanatory note

The information presented in this table is intended as a broad picture of revenues and costs associated with the project, not as a detailed financial report. The financial data, which are presented in the currency in which the project reported the information, were compiled as part of the interview process with project leaders and staff, and in some cases were supplemented with publicly available documents, such as annual reports. Project leaders were asked to review the information prior to publication. The column labelled 'Included in budget?' indicates whether or not the organisation includes that category of cost in its own definition of its budget. In many cases, the information was difficult for project leaders to provide because their institution does not record information in these categories, or because the project was combined with other projects in a larger department or unit. As a result, many of the figures are rounded or best estimates. Some leaders preferred not to offer figures at all, but suggested percentages instead. Frequently, certain types of costs are provided as in-kind contributions by the host institution. Although we did not attempt to place a value on these contributions, we felt it was important to highlight the significant role they play in many projects. Because of the variability in the way each institution estimated the various categories of revenues and costs, the information presented in the table is of limited value for detailed cross-project comparisons.

This case study was funded in part by:



**NATIONAL  
 ENDOWMENT  
 FOR THE  
 HUMANITIES**