



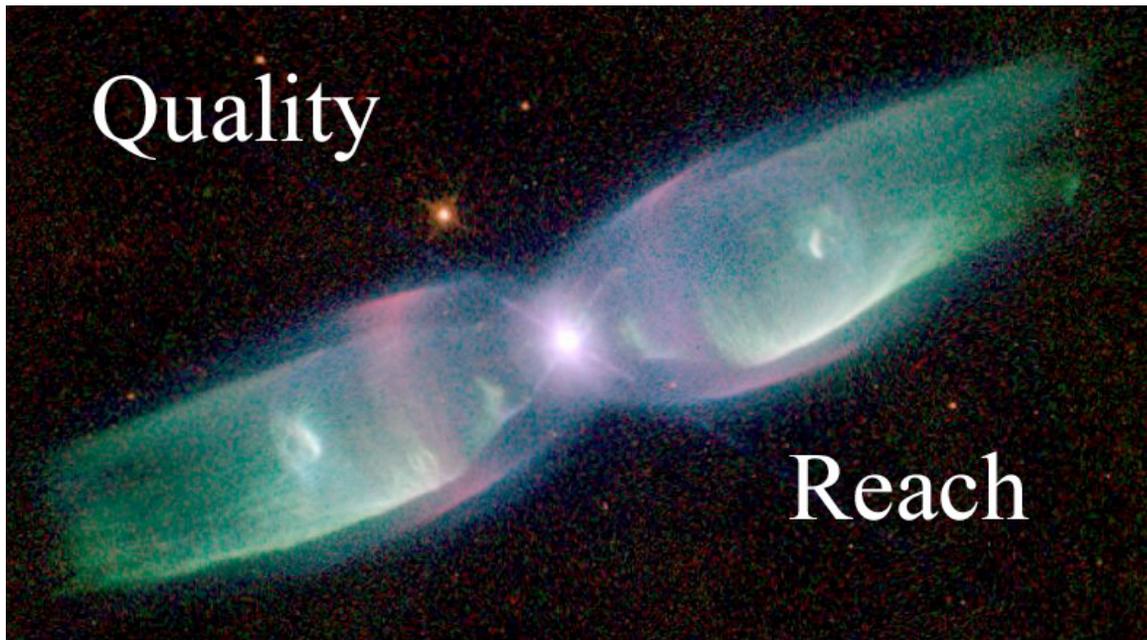
NSDL Reflections Project

Comments on this essay? Please visit
nsdlreflections.wordpress.com

Reflections on NSDL

by Frank Wattenberg
Department of Mathematical Sciences
United States Military Academy
Revised: September 18, 2008

The picture below is one that I've used in virtually every essay or presentation I've given about NSDL. As you can see, this essay is no exception.



The Twin Jet Nebula

This picture captures three things about our initial vision of NSDL – our aspirations, truly astronomical in scope, and our two goals – making a substantive improvement in the quality of learning in the STEM disciplines and a substantive difference in the extent to which the very best learning reaches all students. The two goals together are what give the vision its character – there are many pockets of very high quality STEM learning and the reach of STEM learning is huge but sadly the reach of the very best STEM learning has been too limited.

It is no secret that many of us are disappointed in the results of a project that at its beginning was so promising and had such huge potential. The shortcomings of NSDL can, for the most part, be traced to two related, foreseeable and, indeed, foreseen problems.

- When I was at NSF we had to avoid the use of the word “facility” and we had to work within NSF’s usual grant solicitation mode. Anything else would have led in the best case to delays of ten or more years. But, NSDL, is very much a facility for research and education and building it through the usual grant solicitation mode without acknowledging that reality is like building a house through many individual grants – you might wind up with 5 beautiful kitchens but no bathrooms. And, that is exactly what happened. Moreover, grantsmanship is essentially competitive and one result was friction among some of NSDL’s most enthusiastic and creative supporters.
- NSDL has never had a good sustainability plan. This stems, in part, from our use of the word “library.” The word has many wonderful connotations. Think of how many biographies begin with an ode to the lion statues on the steps of the New York City Public Library; think of the role libraries and librarians play in defending our rights to write and read the most controversial views on the most controversial topics; and think of the way that card catalogs and librarians have helped us find and choose among the millions of available volumes. Ultimately, however, the word “library” is misleading in two very important ways.
 - Many people believe that the word “library” is somehow related to the word “free” but libraries are expensive. When NSDL was first getting started I remember doing a quick calculation, dividing the budget of the Arlington Virginia Public Library by the population of Arlington. I don’t remember the result but through the magic of Google and the Internet I just found the FY 2007 budget of the city-funded Arlington, Texas public library system (\$6,703,585) and the 2006 population of Arlington (364,300). A quick division yields \$18.40. Post-secondary enrollment in the United States is roughly 17,500,000, so a comparable budget might be on the order of \$300,000,000. More significantly, the production of most of the books found in libraries is supported by bookstore sales. This is significant for two reasons – first, the best content in NSDL is likely to be purely digital and not supported by a separate market – and, second, content in NSDL could replace some textbooks. In short, NSDL merits very substantial funding, well beyond current levels.
 - The care and feeding of the contents of digital libraries are closer to the care and feeding of the animals found in zoos than to that of the books found in traditional libraries. This is especially true for the highly interactive contents that are most likely to have a real impact on the twin goals of quality and reach – for example, few of us are able to read floppy disks and many of us spend inordinate amounts of time rewriting old programs in new languages. For this reason, the name National Digital Science Zoo would have been better than National Digital Science Library.

In addition, Computer Science, the discipline and the community, is responsible for some of the successes of NSDL and its greatest failures. In the beginning much of the formal CS focus was on searching, metadata, reusability, portals, and harvesting. The underlying theme was that the resources were there and the problem was helping users find them. NSDL was just another “domain” for CS. Interestingly, the nonacademic CS community, has made the successes of the academic NSDL CS community less relevant while the academic NSDL CS community’s framing the problem as a matter of “harvesting” a crop that was just lying around waiting to be harvested doomed the effort from the beginning. Metadata efforts have had less impact than we expected; reusability, while incredibly important, has not posted many successes; and NSDL’s portal does not seem to have the traffic to generate support. Meanwhile Google’s successes – including, searching, realizing the potential of advertising revenue, cloud computing, and mash-ups – have upended the view of the Web we all held five and ten years ago. In the Army we often use the term OBE (overcome by events). Much of the academic CS NSDL effort has been overcome by events – rendered largely irrelevant.

NSDL can still achieve the goals we had in mind at the beginning but the focus must always be on the “books” – that is, the content – of the library. Quality does not come cheap and maintaining highly interactive quality is not automatic. We have spent a great deal of time and money on NSDL and now is a good time to review what we have done and chart a course for the future.

I propose a review process based on the work of a series of tiger teams whose members are drawn from a wide community and which have tightly focused charges and three month time lines. The tiger teams should be supported by NSF funds. As you will see below the tiger team charges will collectively be quite broad.

The tiger teams themselves should choose the membership of an editorial committee that will synthesize the reports of the tiger teams and their recommendations. The editorial committee should also have a three month timeline that would start with a meeting at which the full membership of the individual tiger teams presents their reports and end at a meeting during which the full membership of the individual tiger teams discusses and approves the final report and its recommendations. The work of the editorial committee should also be supported by NSF funds.

The Tiger Team Topics and Charges

- **Long term “business plan.”** This is the most important and the most controversial topic. Without a realistic long term business plan based on unconstrained and realistic discussion, NSDL simply cannot achieve its goals. The tiger team should address (but not necessarily decide among) at least four alternatives or combinations of alternatives.
 - Large scale long term government funding.

- Self-financing based on a combination of subscription fees and advertising revenue. The key to widespread access is that a single modest cost subscription should provide access to the entire NSDL collection.
- Support by a consortium possibly including government, educational institutions, and private industry.
- An endowed foundation with a substantial initial endowment.

The membership of this tiger team should be very broad and include, for example, someone from the leadership of the Digital Promise project. Note that there are serious public policy issues that must be addressed including independence of content from political pressure and widespread accessibility. Must NSDL be free to users? Should it be international? Can and should it remain an NSF project? What is its relationship to other projects, like Digital promise?

- **Accessibility, reliability and maintainability of interactive content.** This is the core of the “digital zoo” metaphor. We can still read books placed on the shelves of traditional libraries hundreds of years ago and, equally importantly, everyone can read the books on the New Books shelves. We can even take two books written at the same time from different shelves and read them together. Digitally, however, we live in a world inspired by the story of the Tower of Babel and its aftermath. Our hardware and software are not forward compatible, not backward compatible, and not even sideways compatible. The tiger team process is unlikely to solve this problem. It can, however, address the following issues.
 - The over-arching issue is replacing the production and maintenance functions assumed by traditional publishers for traditional books in this new context of highly interactive digital content. This tiger team needs to address this issue directly. The discussion should include, for example, the level at which such functions should reside. With the current structure some functions might reside most naturally at the NSDL level, others at the Pathways level, and still others at the collections level. In traditional publishing authors do what they do best, write, and then publishers take over the responsibility for production and maintenance. In NSDL we often rely on the authors for production and maintenance, a job they often have little interest in, and less experience with, and a job that conflicts with their creation of new content.
 - Standards and “best practices” for interactive content. This issue is closely related to the first one. There is considerable tension now between the burden of standards and best practices on what essentially are individual creator/producers and the accessibility, reliability, and maintainability of their products. NSDL discussions often include a distinction between free-flowing developmental collections and the “permanent” collections. In current practice, the frontier style of the developmental process has not lead to accessible, reliable, and maintainable permanent collections with substantial interactive content. We probably need two sets of standards and “best practices” – one set for creation and development and another for the

“permanent collections” – and we need effective mechanisms for transitioning developmental content to production-quality, maintainable “permanent” content.

- Institutional mechanisms for accessibility, reliability, and maintenance. This is one place NSF could take immediate action even before the tiger team process is complete. This tiger team should recommend specific mechanisms – for example, one mechanism might focus on Java and certify and maintain Java-based interactive components for permanent collections. Another mechanism might focus on shareable software components. Another mechanism might focus on Open Source or cloud alternatives to expensive proprietary components, like *Mathematica*. A sufficiently robust long term business plan might also, for example, include access to proprietary components for all NSDL subscribers.
- Maintainability and access issues associated with game quality modeling and simulation.

The membership of this tiger team should include people with experience in the commercial world with production and maintenance issues for digital content. It should also include people on the cutting edge of game quality modeling and simulation – for example, from the Federation of American Scientists – and from innovators in cloud computing.

- **Quality assurance and broad authorship.** One of the primary drivers behind the idea of NSDL was the uneven quality of materials on the Web. This tiger team should address that issue. In particular, it should consider two somewhat different approaches and make recommendations for how both might contribute to the NSDL mission.
 - The traditional refereeing approach used by professional societies.
 - Dynamic broad community authorship, as exemplified by the Wikipedia.

Membership on this tiger team should include people active in professional societies and active Wikinomicists

- **Market Research.** The most striking metric leading to the disappointment many of us feel with NSDL is simple – it is not being used. Our colleagues often don’t know about it and we, ourselves, rarely use it except when we are working specifically on NSDL itself. My personal belief is that this is due to the fact that NSDL has not successfully addressed the key goals of quality, accessibility, and maintainability and that the largely CS goals it has addressed have proved either largely irrelevant (e.g., harvesting and metadata) or have been (e.g., searching and portals) more successfully addressed by Google and others. The successes of NSDL are primarily at the collection level and many of its collections are widely used. We need to know more about what NSDL should do to address the real needs of STEM education. This tiger team should at a minimum

- Do a meta analysis of existing “market research.”
- Convene at least one national meeting and hold subject area sessions at meetings in different disciplines. Project Kaleidoscope meetings are another natural venue for such sessions. The membership of this tiger team should include representatives from professional societies who are able to convene such sessions within the tiger team timeline.
- **Technical challenges and opportunities.** This is the CS tiger team. It should look at the title subjects in the context of the real needs of the “market.” Among the things it should look at are the Horizon Report produced by the New Media Consortium and EDUCAUSE. (Google “New Media Consortium Horizon Report”.) This tiger team’s membership should include computer scientists from the academic and nonacademic communities, people active in the production of the Horizon Report, and STEM educators.
- **Potential collaborations.** One of my personal disappointments is the very limited collaboration that has been built around the NSDL. This has led, in large part, to the morselization of NSDL that has endangered its future and compromised its present usefulness. My belief is that the key to NSDL’s success is that it is viewed as a large bundle of resources – with one single subscription and with a structure that mirrors the multidisciplinary structure of modern STEM. Although these questions are largely “business plan” issues they are sufficiently important to merit their own tiger team. This tiger team should specifically address the following potential collaborations and its membership should be drawn from national leaders in each area.
 - Collaboration among professional societies. Many collections are maintained by individual professional societies whose finances are limited and who are afraid to take bold steps toward a true NSDL. Such steps could provide each society’s members with access to the digital collections of all societies. The current situation harms professional societies, their members, and both research and education.
 - Collaboration among educational institutions. One possible business plan would involve an NSDL supported and maintained by a consortium of colleges and universities whose students pay access fees to the entire bundle and whose faculty, supported through NSDL, would create the content. Such a consortium would easily have the resources necessary for production and maintenance.
 - Collaboration between NSDL and major private sector players, like Google.

If this author could choose an overall chairman for this tiger team-based review and a convener for this tiger team it would be someone like Hal Varian.

This has been a rather ambitious proposal for reviewing NSDL and for setting a new course but this is a project that is worth the effort. Many of us involved at the beginning did not yet have grandchildren but were looking forward to telling our grandchildren that we were part of the NSDL. I now have four grandchildren and they have no idea what NSDL is. Nonetheless, NSDL can do wonderful things. We can still realize the promise illustrated by the Twin Jet Nebula.