

# FINDABILITY IS MOVING EDUCATION TO THE NET

---

What you want to learn or teach will show up  
for free in your hand, and in context

---

## INTRODUCTION

Today schools still rely mostly on printed paper and closed digital resources while strongly cautioning their teachers and students in using the open Net . For the ten-plus years since the Net began expanding from a small education resources network into multiple sectors and eventually mammoth information repositories, educators have been concerned about the challenge of finding the best learning resources within the vast open content. This perceived challenge has been a main reason that education has kept online materials at arms length.

Meanwhile, along with the much-lamented educational “junk” that poured into the Net, arguably most of the knowledge the academy exists to teach went online where these digital resources are open to all, free—and are usually better kept up-to-date and more compelling than in their printed counterparts.

Net native methods have emerged that make online content findable. Outside of education, powerful tools have been developed that could direct students and teachers to the best content for precisely what they are trying to locate. This white paper describes those tools—forming the new findability discipline—and outlines how educators can help make findability work for learning. When the process moves to mobile, you will find what you want to learn is in your hand.

## FIVE STAGES OF NET CONTENT:

- 1) DIGITIZED, 2) MOVED AND OPENED ONLINE, 3) SEARCHED FOR,
- 4) OPTIMIZED FOR SEARCH, 5) FINDABLE.

## EDUCATION HAS SO FAR ONLY PARTIALLY ENTERED ANY OF THEM

This white paper traces the steps Net content has gone through that have laid the foundation for findable education, explains what educators can do to build findability, and sketches the findable future. Findable education is fruit for learning of the Web 2.0 era. It promises no less than significantly higher levels of learning for the youngest global generation and those that follow.

### STAGE ONE: CONTENT WAS DIGITIZED

As the postcard from the 1904 St. Louis World's Fair reminds us, establishing physical centers of knowledge to attract people who wanted to learn is an ancient tradition. From the small schoolhouse to local and regional colleges, to the great universities and grand libraries, the purpose for them was to bring visitors to the location. In the biggest of these visions, the Greeks labored for centuries on their library at Alexandria to which they dreamed people could one day come to study every book ever written. That vision is about to be realized, but there will be no physical place. Everything known by humankind will be at the beck and call of every person on earth—findability will bring the knowledge to the student.

The commercial sector is now taking the final steps of a decades long process that has led to findability. The first step began in the second half of the twentieth century: methods were devised by which almost any kind of information could be represented by sequences of zeroes and ones. Numbers, text, images, sound, colors—all could be stored and transmitted in any medium that handled digital material (the zeroes and ones). Vast amounts of information that had once been printed on paper was keyed into machines that dealt with it digitally. The information was stored on punched tape, in computer memory, on tape reels and magnetic cards, and floppy disks.

The education industry has participated in digitization only minimally. The knowledge it delivered stayed mostly in teachers' memory and on the printed pages of textbooks. A November 7, 2005 article in the New Yorker summed up part of the financial scope of educational print: "College students now spend more than five billion dollars a year on textbooks, while states spend another four billion on books for elementary and high-school students."

In the early 1990s, a digital technology appeared that looked for a while like it had great promise for pedagogical



Figure 2 - CDs have digital data, but it cannot interact between CDs

resources. The technology was the presentation of interactive materials on compact disks, like the ones in figure 2. Digitized tutorials could be inserted into computers on the CDs. Students could interact with the tutorials to learn. But there was a fundamental limitation to the CDs: they could not interact with each other. The limited digital material on a single CD seemed paltry, and remained outside of the pedagogical mainstream.

Then the Net began to grow and small, isolated CD digital spaces were replaced with a huge and growing digital universe where anything could connect to anything else. While other sectors—commerce, culture, media, government, (porn, of course)—pored content into the Net, education stood aside mostly, not participating very much. When education did go online, most of what it did kept the traditional educational concept of coming to a place. In what came to be called distance learning, educational institutions created connections to students away from their physical classrooms, but the learning still was coming from the institution. The distant learning model does not open the content for learning online. The content remains inside virtual walls of ivy in university intranets and paid-for subscription services to closed digital libraries.



Figure 1 - A palace of knowledge at the 1904 St. Louis World's Fair

## STAGE TWO: CONTENT OPENED ONLINE

### The closed content and the open Internet

In the mid-1990s digital content across the sectors of human knowledge and enterprise was pouring into the Internet. Content could be kept closed if you wanted to limit it only to certain people. Banks are an example of institutions that did not want much of their online information openly available—and neither did their customers. Businesses created “intranets” where their confidential information was used by their employees, with little connection to the open Net.

It is not the closed sectors of online information that dominates the public conscious. We usually think of the online world as the vast amounts of digital content that is out in the open creating a huge and rapidly growing global commons of information. It is this now gigantic ocean that is referred to in this white paper as the Net.

It is fair to generalize that most of the content that established education has digitized since the mid-1990s has been kept closed. On the other hand, a lot of digital knowledge

***Caveat: If it isn't open no one is going to find it.***

Throughout this white paper—and in considering the searching and optimizing of educational resources—it is the open, global commons called herein the “Net” that is being considered. Intranets and/or any other form of walled off digital content are outside the increasingly findable Net ecology discussed in these pages.

resources that are excellent for learning have been put online by individual professors and other experts, museums, laboratories, governments, associations, and the media. Beginning with the open courseware project at MIT, a movement for open educational resources (OER) has been joined by dozens of universities across the world. Although education has been positively impacted by these open resources, they have not led in efforts to make online materials findable. The story that begins the online focus on finding content within the Internet takes place first in commerce.

### The first efforts to find content online relied on pre-digital methods.

In the early 1990s, rows of shelves at book stores like Barnes & Noble filled and then overflowed with books full of links to look up on the Net. There were thick volumes for all sorts



Figure 3 - In April 1997, this was the entire Yahoo! homepage.

of subjects. For example, a book focusing on links for children would offer the links in chapters by categories. Each link would have its title printed, with a brief review of what the website's content was, and then its printed online address. You would take your book to your desktop computer, open your browser, and copy the URL from your book, typing it into the Web address box on your screen. No wonder short URLs were highly prized!! One of the worst flaws of the book system for finding links was the printed references went out of date at warp speed.

Online listings soon grew into the a better way to find content in the open Internet. The Yahoo! homepage in Figure 3 displays an index. This Yahoo! Web directory vintage 1997 is a simple hierarchy of subject listings structured like the Dewey Decimal System or any other brick and mortar library system for listing of resources. The links listed in this type of index were evaluated by humans, who chose the categories where they would put the links. Paid placement of links soon skewed the objectivity of some of the indexes.

Yahoo! and other online indexes were a terrific improvement over the printed books of links in your lap. Soon, though, as content continued to cascade online, a better way was needed to find what you were looking for on the Net.

## STAGE THREE: FINDING CONTENT GOT HARDER WITH THE EXPLOSIVE GROWTH OF THE INTERNET AND SEARCH BECAME BIG BUSINESS

### Search engines are born and raised in cyberspace.

As the top down approach of indexing methods transplanted from the paper world became less and less effective for finding what visitors wanted in the proliferating Net, search methods native to networks began to be teased out and harnessed.

In the early 1990s Archie and Gopher, with programs Veronica and Jughead, were developed to search indexes. Yahoo! continued to thrive as a seemingly ever-expanding Web directory—and as part of a portal that grew and grew under the Yahoo! mantle.

Methods modelled on the paper world were joined by a parade of what are now called search engines—native creatures

### ***Necessity: Search engines were invented because paper world ways broke down.***

With online materials multiplying exponentially, a method different from what had been used to catalog printed resources was needed.

of the Net that use software robots called spiders to crawl the Web searching for a words submitted by users. The use of top down hierarchies was replaced by considering the meaning of the smallest pieces of text.

Early on webmasters submitted their pages to search engines for indexing as a way to cause items on their pages to be reported to users. From the beginning of search methods, webmasters looked for ways to entice the engines to report their webpages high up on search engine result pages. It was a small step from doing that to coming up with ways to embellish what the crawling spiders looked at, making the search engine reports to users even more favorable. This embellishment has become known as search engine optimization (SEO).

### Google orchestrates PageRank

In the late 1990s, founders Larry Page and Sergey Brin invented a new approach to link analysis that has caused their search engine Google to become one of the largest and most influential enterprises in history. The inventors built on the nature of the Net, moving ways to find content ever farther from predigital methodology. Their core concept is defined on the Google.com technology page as follows:



Figure 4 - In April 1999, Google was in beta.

***“PageRank Explained:*** PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page’s value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at considerably more than the sheer volume of votes, or links a page receives; for example, it also analyzes the page that casts the vote. Votes cast by pages that are themselves “important” weigh more heavily and help to make other pages “important.” Using these and other factors, Google provides its views on pages’ relative importance.

“Of course, important pages mean nothing to you if they don’t match your query. So, Google combines PageRank with sophisticated text-matching techniques to find pages that are both important and relevant to your search. Google goes far beyond the number of times a term appears on a page and examines dozens of aspects of the page’s content (and the content of the pages linking to it) to determine if it’s a good match for your query.”

Google made search native to the network ecology and has continued for a decade to refine the science of search.

### Education has pushed for search skills

One way educators have responded to the Net is by expecting teachers and students to become expert searchers. This is an opposite approach to what commerce did in seeing to it that its customers could find what they are looking for. If the commerce sector had taken the sort of approach education did, sellers would be teaching the customers how to find products to buy.

Instead, the sellers gussy up their products with things that appeal to search engines so the products appear in front of potential online customers. A major way this is done is that the online commerce sector goes to great length and expense to get its products to appear at the top of search engine result pages. Doing that is called search engine optimization (SEO).

There has been no noticeable comparable effort by education to put quality online content in front of students.

## STAGE FOUR: THE COMMERCE SECTOR REALIZED THAT OPTIMIZING CONTENT FOR SEARCH ENGINES MAKES SALES AND MONEY, AND SEO WAS BORN

Techniques have become highly-developed for optimizing content to meet search engine PageRank criteria.

Search engine optimization (SEO) is formed around the value of page rankings. The definition given by the specialists at a leader in the field for business websites, SEOmoz.org, is: "SEO is the active practice of optimizing a Web site by improving internal and external aspects in order to increase the traffic the site receives from search engines."

***Duh!: Search engines will rank our Web pages higher if our pages have things search engines like.***

As search engine algorithms have grown more sophisticated so have the optimization methods for Web pages developed by product marketers.

The optimizing does things to a Web page that entice robot software spiders to crawl through the page and then to report in ways that boost the page's rank on their home search engine results pages.

SEO is not rocket science, the pros like to say, but it requires tweaking carefully many aspects of content that is placed online.

Figure 5 is an illustration from a series of expert articles describing *Search Engine Ranking Factors* on the SEOmoz.org



Figure 5 A Data Key for Search Engine Ranking Factors  
<http://www.seomoz.org/article/search-ranking-factors>

website. Explained are the Top 10 Positive Factors, another 10 Most Controversial Factors, and the Top 5 Negative Factors. The devil is definitely in the details for SEO. Tending to these details has been discovered to be very profitable for commercial websites.

The two big methods of SEO are using keywords and acquiring links to the Web page being optimized.

### SEO: Keywords and tags

Search engine spiders read text. Putting the text that identifies a webpage's contents into positions on the webpage where the spiders will find and read them is the core work of SEO. This has become a highly developed process. The main steps:

1. Selecting general and specific keywords that will attract your target audience.
2. Placing keywords into text and html in titles, as headings and tags, link labels, anchor text, file names, attributes, captions, first lines of paragraphs, within the URL, etc.
3. Adding the keywords as tags in meta data and as blog post tags, etc.

### SEO: Link baiting

The SEO folks say that when one webpage links to another, it is "giving juice" to the page it links to, or "giving love." The juice and love are important because the search engines look for link quantity, link quality, and link relevancy. The more juice and love the spiders report, the higher the webpage is ranked on the search engine report page.

Good SEO is to: link out; place links in content areas; have a wide variety of linking practice; put energy into authority links: deep link; create topic-specific SEOed landing pages.

Using closely guarded secret algorithms, search engines place links on their results pages. In his 2008 book from New Riders, *Building Findable Websites*, Aaron Walter provides this list of what pleases search engines:

### *The Deepest Desires of Search Engines:*

- Content that is naturally keyword rich (not stuffed) and valuable to readers
- Content that is visible to search engine spiders with no barriers that may prevent a full indexing of pages
- Content that communicates a clear informational hierarchy so spiders can understand what the page is about
- Content that loads quickly so spiders can index it efficiently
- Links to your site from reputable sources so they can determine the reputation of your site
- Honest content that isn't trying to trick the search engine
- More content than code to mark up the page
- Clean, meaningful URLs with keywords in them if possible
- Domains that have been around for a while

## STAGE FIVE: NETWORK EFFECTS (WEB 2.0) ARE MORPHING SEARCH INTO FINDABILITY

The pros tend to say, that it is actually search engine marketing (SEM) that they do, not just SEO. The word “findability” is beginning to solve this confusion.

It gets confusing because the newer kinds of marketing activities are not about search engines and are less about optimizing content so spiders can find it than about making it findable for human beings within a network ecology. Here are four methods that have recently joined with SEO in the SEM toolkit:

### Blogs

A blog is a means to flow little webpages that are targeted to a topic and SEOed for that topic into the open Net.

### Syndication

RSS (really simple syndication) allows individuals to receive notice of new content when it is posted on webpages to which they subscribe. RSS newsreaders are a timely way to find new online materials selectively.

### Viral

This method is to spread a message by virtual word-of-mouth using the open Net. Email has been a powerful viral marketing tool for several years. The recent emergence of social media has opened a vast new environment for optimizing findability of material distributed by word-of-mouth.

### Social media and communities

The marketing world is now scrambling to make consumer products findable in the exploding online social sphere. In early 2008 the photo collection website flicker.com is a top target for taking a product to customers. The hugely successful bookseller Amazon.com is using individual books it sells as nodes—with the potential to build a customer community around each one.

Marketers are focused on online communities at myspace, facebook, xanga, blackplanet, asianavenue, mothersclick, nascar, and many more. Twittering has caught on. YouTube is

***Realization: When webpages are SEOed by the same attributes (like the same keyword) the webpages will find each other.***

In the open Net the effect of network laws is for patterns to emerge among related nodes—the SEOed nodes are webpages.



Figure 6 - A slide from Gordon McLeod's keynote at the March 2008 Search Engine Strategies conference.

dominant. Social sites thrive: Digg, Stumbleupon, del.icio.us, Redit, and Skirt. Large on the online marketing horizon are virtual worlds, with leaders like: Barbiegirl.com, neopets.com, SecondLfe, and the MMOGs (massively multiplayer online games). In each of these, making products findable is a key step to selling those products.

### Search Engines are putting related content on a webpage.

An interesting new challenge is a swing away from optimizing one bit of content so someone using a search engine will find that bit. As the title of Figure 6 indicates, search is becoming a content platform. One of the things that means is that bits of content that are related to each other are interfaced together.

Gordon McLeod is president of The Wall Street Journal (WSJ) Digital Network. In a keynote address at the Search Engine Strategies 2008 conference in New York City, McLeod described the very recent move of the WSJ through the steps described thus far in this white paper. Although there has been an online WSJ for many years, it has been open only to paid subscribers. With the content closed, there was no relevancy to SEO because the spiders could not get through the WSJ firewall to crawl the webpages of the online editions.

In January 2008 WSJ.com partially dropped its subscription barrier. Figure 6 is a slide from McLeod's keynote. It illustrates how content generated by search is input to pages of the open WSJ online pages. The result is a step into the future where choice and context are displayed to online visitors. The search engine finds content and places it on a page. The possibilities of this capability for education are stunning. Pieces of content relevant to each other can find each other. Context is created.

## EDUCATION CONTENT HAS PRETTY MUCH WATCHED FROM OUTSIDE

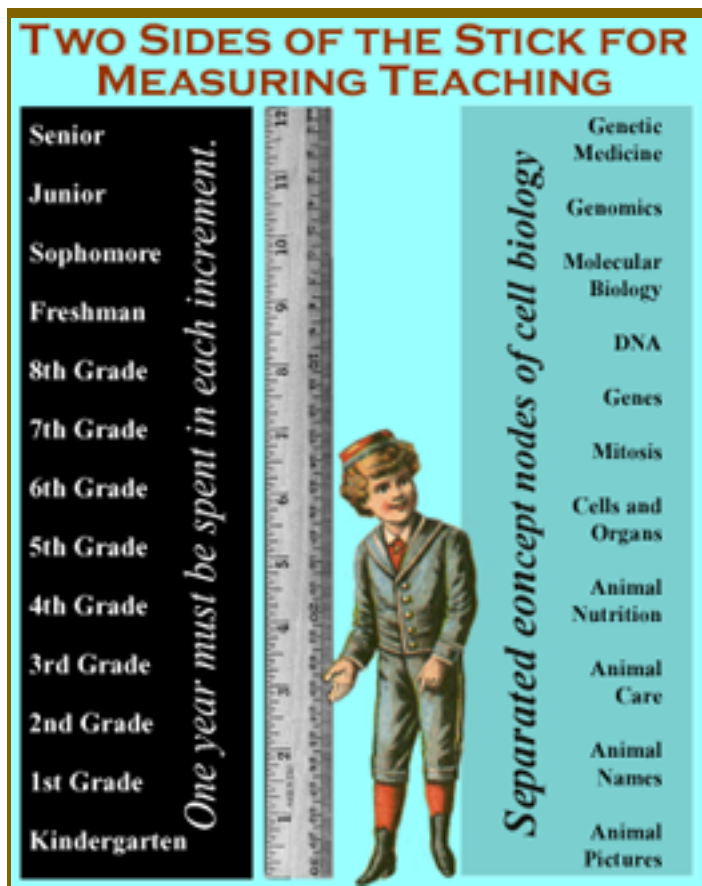


Figure 7 - The distribution of students across geography and the need to deliver education resources printed in books led to dividing subjects into relatively few units boxed into hierarchies of subjects and grades.

Over the ten-plus years that the commercial sector has pioneered methods for dealing with online content, education not done much:

### 1) DIGITIZED

Education-produced resources remain largely print-based. The open publishing movement is at work to change this. Interestingly, because open textbook and other Net publishing are coming at this late date, they can engage the optimization techniques as they go online—benefitting from years of method development in the commercial and social sectors.

### 2) MOVED AND OPENED ONLINE

Many education resources that are digitized remain closed to the global Net. But this situation is changing. A movement for open educational resources (OER) is making real progress in encouraging universities, libraries, and other academic repositories to open their courses and archives on the Net. In addition to content from academic institutions, vast amounts of excellent learning materials are hosted online by experts, museums, laboratories, governments, associations, and the media.

### 3) SEARCHED FOR

Education has made efforts to train teachers to search the Net, maintaining caution in using educational resources that are open in the Net and found by search engines.

### 4) OPTIMIZED FOR SEARCH

Very little has been done to optimize educational resources for search engines. Educators have yet to undertake what commercial marketers have done: optimize educational resources for search engines. Education is not noticeably active in networking methods in blogging to optimize content, RSS distribution of learning materials, viral distributions, and/or optimizing for the social networks that have captured the student generation.

Since so little educational content is even yet digitized, educators are in a position to start from digital scratch in creating findable resources for the Net.

### 5) FINDABLE

The two images on this page contrast the past that education has been reluctant to move beyond and the future formation of what is known by humankind on the Net. The matrix shown in tiny part in Figure 8 is a network, where laws operate that will make learning content findable. Educators can speed the process and harness it sooner by focusing now on its fascinating and enormously promising challenge.

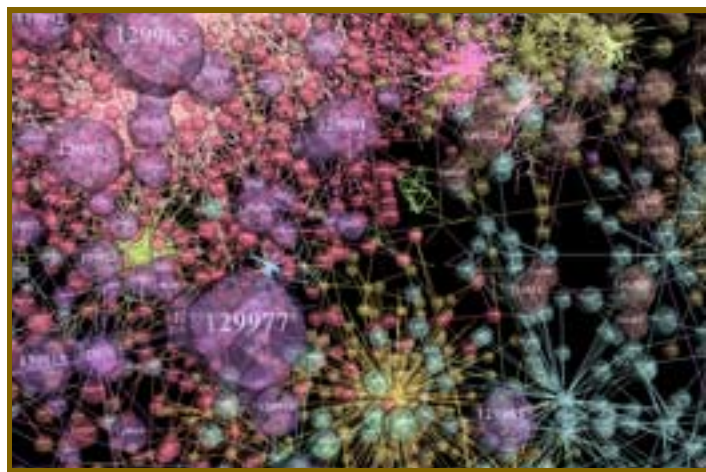


Figure 8 -- This image from the School of Engineering, University of Tokyo<sup>1</sup> is a representation of a tiny Net moment. The image barely suggests the actual complexity and dynamics of the Net. To platform information, every node (round dot) can connect with any and every other node in the entire open Net. Our minds are not equipped to visualize that reality even if it were possible to depict in on a paper, screen, or even in a 3-dimensional model. How can we expect to find anything at all? The answer may turn out to be that what we seek may find us.

1 <http://www.race.u-footfootnoc.jp/~uchida/blogdata/>

## A GLIMPSE OF EDUCATIONAL RESOURCES IN THE FINDABLE NET FUTURE

It is beyond the scope of this white paper to argue that educators should put their minds and energies toward making educational resources findable online. The following sketches very briefly what that process would be, and highlights what we are beginning to understand could happen when educational resources are part of the findable Net future.

The obvious first step is for educators to apply to educational content the findability tools developed in the commerce sector.

These steps are set out in the preceding pages. Educational content needs to be digitized and put into the open Net. Educational resources that remain nondigital and offline will atrophy in the vigorous environment of content digitization of the present and future.

The concept of searching for Net content is morphing into techniques of engaging findable emergent knowledge and ideas. Educators can optimize content for findability, utilizing their academic expertise and authority. Educators can also participate in optimizing educational resources for use in social networks and connecting the global commons through syndication and viral cognitive connectivity.

The steps described are not just in the development pipe. They are available now. The Animal Bytes page in Figure 9 is rich in ideas but probably not fully SEOed for education. Each animal name and fact is potentially a node of information that can make that page findable for the individual fact. The Animal Bytes page does not have blocks of information that are placed on the page and kept up-to-date as the WSJ is doing in Figure 10, where the search index is making content findable and placing it adjacent to related context on a page—creating context. The context is a glimpse the findable future.

### Defining Findability

The simplest way to justify using the word findability to label the next phase on online learning is to note that denotes a switch from burdening teachers and learners with search and causes what they are looking for to be located. Education has been stalled for a decade by the search obstacle that has kept teachers and students away from the Net.

Aaron Walter, author of *Building Findable Websites*,<sup>1</sup> writes that: “Findability is the broader discipline that unites all strategies to help your audience find what they seek.” That definition is a good and useful one to guide builders of Net content. Educators will do well to respect the findability discipline for what

1 Aaron Walter, *Building Findable Websites: Web Standards, SEO, and Beyond*. New Riders 2008, p.2..



Figure 9 - Animal Bytes at the San Diego Zoo has rich educational content, but it is all static. There are no changing feeds.



Figure 10 - This portion of Figure 6 shows two search engine feeds that keep the content up-to-date as it changes on other websites. The latest information is kept findable on pages remote from its source.

they put online.

Peter Morville, the philosopher of findability spent three full pages in his book *Ambient Findability*<sup>2</sup> on his “Definition” section for findability. Morville blogs on the subject at Findability.org. Before highlighting a few of the characteristics of a world that lies ahead where educational resources are findable, the scene is well set by this conclusion to Peter Morville’s definition of findability.

2 Peter Morville. *Ambient Findability*. O’Reilly, 2005, pp. 4-6.

Of course, the user experience is increasingly *out of control*, as wireless devices inject new interfaces and affordances into an already complex network ecology. How do we design for mobility? How do we create good experiences when we can't predict context of use? Will our users be in the office or in the bathtub? What is their bandwidth and screen size? The variables will only multiply as ubicomp transforms the Web into both interface and infrastructure for an ambient Internet of objects we can barely imagine.

*am-bi-ent* *adj*

- a. Surrounding; encircling: *e.g., ambient sound*
- b. Completely enveloping

Ambient findability describes a fast emerging world where we can find anyone or anything from anywhere at anytime. We're not there yet, but we're headed in the right direction. Information is in the air, literally. And it changes our minds, physically. Most importantly, findability invests freedom in the individual. As the Web challenges mass media with the media of the masses, we will enjoy an unprecedented ability to select our sources and choose our news. In my opinion, findability is going ambient, just in time.<sup>3</sup>

Education's wait and see attitude toward engaging the Net is made obsolete by both Aarron Walter's practical definition of optimizing Net content and Peter Morville's future vision.

### Glimpses of educational resources findability—a few random observations and conjectures.

- Most existing educational resources could attract more users by applying the commercial SEO methods to their Web pages.
- Education resources are content-rich. This makes applying SEO to most learning websites easy by comparison to many commercial products where content has to be contrived to attract search engine spiders. Learning subject content usually has a lot of text that usually includes words that serve as keywords to optimize the webpage containing them.
- Educational resources will experience two kinds of findability:
  1. What one seeks to learn becomes findable in the Net.
  2. Findable learning materials find and connect to related ideas, linking to each other: findability creates context.
- Cognitive optimization will be both the major practical technique that builds the global learning commons and the source of its authority. By tagging and linking the Net resources they respect, academic experts create findable patterns that underlie online educational resources of the future. (Actually, this has been happening in a small way for years.)

<sup>3</sup> Ibid.

- When experts optimize online academic knowledge so that it can be found, they also activate network emergence causing related knowledge to cluster and be vetted by the authorities on the subjects that emerge.
- Findability assures that the long tail of academic subjects is not cut off. Students can pursue their curiosity limitlessly.
- Findability will be different for pedagogical resources and cognitive resources. The pedagogy will be optimized by human picking and choosing. Regular SEO works very well for pedagogy: courses, curricula, lesson plans.
- Cognitive content has the extra cluster power of the network inherent in the meaning of the subjects themselves. Emergence of patterns of knowledge about math, history, physics and the other academic subjects will facilitate findability through network laws.

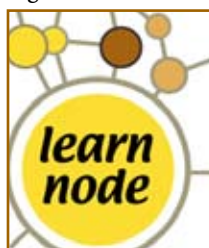
### Findability is moving education to the Net

The threshold has been crossed. We now live in the time of findability. Education will appear here soon and your mobile device will put what you want to learn or teach in your hand for free and in context.



Figure 9 -- The above network image is from a movie<sup>1</sup> by the same University of Tokyo laboratory that created Figure 8. By watching the movie you can get an idea of how information connects within the Net.

<sup>1</sup> [http://www.race.u-tokyo.ac.jp/~uchida/blogdata/dataset1\\_evolutioning.m1v](http://www.race.u-tokyo.ac.jp/~uchida/blogdata/dataset1_evolutioning.m1v)



The LearnNodes.com project builds the global learning commons by connecting links to what experts know. Blog posts that are SEOed are used. The learn node approach is a way you can contribute your expertise and authority to the emerging findability on the NET.