THE LAWS OF THE JUNGLE

I'm ashamed to say I sometimes empty the dishwasher of all its contents except for those in the silverware basket, hoping that my wife will do it. It's a miracle our marriage has lasted so long.

I don't mind putting away the plates because I have a technique. I swoop in from the top and take as many of the dinner plates as possible, then pick out the salad plates, and continue on, from largest to smallest. That's not because I'm an obsessive-compulsive who has to put things away by size order. If anything, it's reality that's obsessive-compulsive. The bigger plates insist on sticking up higher than the smaller plates, and thus are easier to take out first. The silverware, on the other hand, refuses to play along. It mobs up in the dishwasher's basket, unsorted and defiant. I grab fistfuls and have to go through them item by item. The silverware basket is too damn miscellaneous.

Of course, it's only a problem because we insist on "rescuing" the silverware from its miscellaneousness, dividing it among the designated sectors of the silverware drawer. If we could leave it as miscellaneous, we'd just dump it into the drawer and be done with it. That, in fact, was our Silverware Maintenance Policy when I was living with three other college students. But we paid for our cavalier refusal to split the silverware lump every time we went to set the table. We would paw through the drawer, picking out pieces we needed. If we're going to use utensils to eat—a nicety we sometimes skipped in my college days—then one way or another, we have to take the silverware out of its state of pure miscellany. In college, we postponed that moment until we were ready to eat. As responsible married adults, we now sort the silverware in advance of its use so that when we need to set the table, the silverware is tucked away in its little beds, waiting for us.

Each strategy has its place in the world of atoms. The college-student approach works best when you have only a few types of silverware. Once you've gotten flatware as a wedding gift and you begin to add to your collection, you find you don't just have knives, forks, teaspoons, and soup spoons. You have butter knives, steak knives, paring knives, and service knives. You have salad forks, regular forks, and forks for eating shrimp. Then you get older and have more silverware for the family, for guests, and for super-special guests. And if you're lucky enough, like some of us, to marry an Orthodox Jew, you have separate sets for meals with meat and meals with dairy, and double that for Passover. At that point, making one giant heap of all your silverware while cranking up Jefferson Airplane on the hi-fi no longer seems like such a good idea. If you don't sort after each meal, your silverware world falls into chaos.

In the digital age, computers have become demonically good at sorting through gigantic, complex piles of information. Crate and Barrel's online catalog has fifty types of place settings, each with two different forks, two different spoons, and a knife, yet there you can find a pierced serving spoon that goes with the genuine-pakka-wood-handled flatware faster than you can grab a bunch of soup spoons from your perfectly ordered silverware drawer. And that means we college students had it right. We were just ahead of our time. The best digital strategy is to dump everything into one large miscellaneous pile and leave it to the machines to find exactly the table settings we need for tonight's dinner.
A BIG CAN OF WORMS

"Alison Lukes et Cie is Washington's premiere closet consultant, personal shopper and stylist." Her Web site shows her sitting in front of a closet in which dozens of women's shoes on shelf after shelf are arranged by color. She is young. She is stylishly dressed. She is pretty. She is smiling proudly. And why not? Ms. Lukes et Cie wrestle one of the last remaining pockets of household miscellany into order. Closets crammed with junk shrink back in terror when they hear her dainty tread. Chaos itself quails before her label gun.

What is our problem with the miscellaneous, anyway?

At its heart, the miscellaneous is a set of things that have nothing in common. Of course, that "nothing" is relative since the utensils in your kitchen's miscellaneous drawer all have a use in preparing and eating food, all are physical objects, and all are smaller than the drawer itself. Likewise, the miscellaneous section of a greeting-card display does not hold bassoons or riding lawn mowers. Nevertheless, within some particular domain, the miscellaneous gathers things that are unlike whatever sits next to them.

Sometimes we like that. Sometimes, as with college housemates, sorting on the way in takes more effort than sorting on the way out. And there can be positive benefits to miscellany. Enlightened human resources directors will tell you that workplace diversity isn't just a matter of equity. PepsiCo says that about an eighth of the company's revenue growth in 2004 came from new products "inspired by diversity efforts." Putting unlike things together also works for Oscar-winning film editor Walter Murch. When he was editing Jarhead, he filled the walls of his studio with jumbled photos of the five thousand separate shots in the movie. "It makes images collide with each other in very opportune ways," he said.

Nevertheless, when Murch is done with a project and is filing away the photos, he undoubtedly wants to sort and order them. In the first and second orders, mixing things up may be great for creativity, but for refining them, it's a disaster. As one of the "Slob Sisters" writes in their popular book Sidetracked Home Executives: "On that fateful June day, I was in my new home, lying on the living-room floor, surrounded by 157 Belkis' moving boxes—all marked Miscellaneous." That's not a good thing.

Likewise, in classification systems, an overstuffed miscellaneous category can be a sign that the system isn't using all the relevant information. If I cluster elephants with spotted owls, the basis of the clustering—the category's Aristotelian definition—adds the information that both are endangered species. But if I throw them both into the miscellaneous category along with dachshunds and crickets, I've buried that relationship. So we should be suspicious (as Stephen Jay Gould brilliantly pointed out) when a taxonomic system divides a domain into two major lumps that are wildly uneven—as Linnaeus's classification of animals into vertebrates and invertebrates did. He divided the vertebrates (which we now know includes forty thousand species) into four subcategories, and the invertebrates (about a million species) into just two: the rather well-defined category of insects and the undifferentiated mass of creepy-crawly-swimmies—everything from earthworms to clams and jellyfish—which he called "worms" (Vermes in Latin). Linnaeus didn't know that his system was so skewed—he thought there were fewer than fifteen thousand species in total—but he nevertheless paid inordinate attention to the animals that, like him, have backbones.

It isn't obvious why Linnaeus pushed so many species into the Vermes bucket. While it's true that the simpler the organisms, the fewer ways to differentiate them, Linnaeus broke plants down in intricate detail, even though they are less complex than worms. Perhaps it's simply that Linnaeus loved plants and just didn't care much for worms. Whatever the reason, he left too much information hidden in that miscellaneous bucket.

It was Jean-Baptiste Lamarck—unjustly remembered primarily for being wrong about how giraffes got long necks—who not only sorted out Linnaeus's worms but changed the basic shape of Linnaeus's tree.
As Gould recounts the story, Lamarck loved invertebrates so much that when he was almost fifty, he was appointed professor of insects and worms at the Muséum National d'Histoire Naturelle. Lamarck called *Vermes* "a kind of chaos where very disparate objects have been united together" and, starting in 1793, he began to draw distinctions. When, in 1802, he split the Annelida—earthworms and the like—from intestinal and other host-based worms (Linnaeus’s decision not to delve deeply into this category will strike many of us as increasingly understandable), *Vermes* was left with creatures not obviously more complex than the category of sea urchins supposedly below it. Rooting around in the bucket, Gould claims, Lamarck came to realize that life could not be ordered in a single line, from least complex to most, as Linnaeus had.

There are two different lessons we could draw from Lamarck's correction of Linnaeus's system. We might say that the miscellaneous category should make us wary because it hides information waiting for a Lamarck to come along and split the lump in useful ways. Or we might say that Linnaeus was not miscellaneous enough. Sure, Lamarck discovered important distinctions among worms. But every time you organize matters in one way, you are disordered them in others. Sorting my dessert recipes into cakes, cookies, and pies obscures their carbohydrate order.

The basic fact that order often hides more than it reveals has sometimes itself been hidden within the art and science of organizing our world. We have been like the proverbial seven blind men feeling the elephant, except unlike the narrator of the story, we've had to pick our favorite blind man. Lamarck's division of *Vermes* indeed reveals relationships Linnaeus missed, but a fisherman would divide the bucket still differently based on which wriggling creatures desirable fish are likely to snap at. Lamarck's and the fisherman's divisions both have merits, but if it's a first-order bucket, we can divide it only one way, just as Staples has to sheve its printer inks one way and not another. In the second order, we have the flexibility to organize physical metadata in a few ways—library catalog cards sorted by subject, author, and title—but not much beyond that or the catalog gets too big to be usable.

These physical limitations on how we have organized information have not only limited our vision, they have also given the people who control the organization of information more power than those who create the information. Editors are more powerful than reporters, and communication syndicates are more powerful than editors because they get to decide what to bring to the surface and what to ignore.

At least in the first and second orders of order. In the third order, bits rule. And so does the miscellaneous.

**TAGGING LEAVES**

There's something glorious about a well-crafted, treelike structure of information, even if it sometimes borders on the absurd.

The list of categories in the International Press Telecommunications Council list of "NewsCodes"—a Dewey-like system for categorizing news articles—in spots accidentally reads like headlines that encapsulate a story: "Financing and stock offering," "Government contract," "Global expansion," "Insider trading." Such accidents of meaning happen, especially when you're trying to cover every topic about which a newspaper might want to write. The list's real peculiarity is how uneven its specificity is. While under "Cinema" there is a single entry ("film festival"), under "Sailing" there are seven sorts of dinghy races, including "one man dinghy (4.57m sq mainsail)."

The Getty Art and Architecture Thesaurus is even more ambitious than NewsCodes. "The aim was to classify the material world," says Joseph Busch, the project director who oversaw its construction. "Quite a modest project," he adds, playfully. The thesaurus was created to enable institutions to find and share information about their contents, a boon for curators and scholars. The experts who accomplished this wonder of the well-organized world divided 128,000
terms into seven top-level categories. Search for “apple corer” and you’ll find that it’s part of the Corers category, which, through several shoots and twigs, eventually attaches to a main branch of the tree:

corners
culinary tools for extracting

culinary equipment for preparing and cooking food
culinary equipment
equipment by context
equipment
tools and equipment
furnishings and equipment
objects

Because the designers wanted to provide ways to describe artworks that depict scenes in which things happen, the thesaurus includes an Activities facet at the same level as the Object facet. Under Activities, you’ll find the subcategories Standing, Sitting, Whispering, Archery, and Torture, all on the same level and hanging from the same branch. Classifications make strange bedfellows.

The Getty thesaurus provides what’s called a controlled vocabulary, so curators can classify their holdings without having to make arbitrary decisions about whether to describe a maritime painting by J. M. W. Turner as depicting ships or boats. Standardization makes it easier to retrieve information: If you know the vocabulary (or if you browse the tree) you don’t have to guess whether to use the word ship or boat when looking for that Turner painting. The fact that the Getty’s terms are arranged into trees also helps avoid ambiguity, because the system knows that the word rock is a stone in the Objects branch, but rock is what Whistler’s mother does in the Activities branch.

The Getty thesaurus is a mighty tree but, like all such projects, it can strive for comprehensiveness only by reducing the richness of what it’s comprehending. This is the nature of organizational trees, for they are built on single relationships applied over and over again: “B is a type of A,” or perhaps “B reports to A” or “B is the child of A.” No matter the relationship explained by the branch, it is almost certainly too simple to capture all of the relationships and complexities of its subject.

Usually we know that. We understand that a genealogical tree expresses only the path of DNA through time and that it tells us nothing about the emotional ties among the children and parents. We understand that the Getty thesaurus is intended as a convenience for curators, not as a comprehensive guide to how the universe works. Even so, when we draw a map of knowledge, it is all too easy to assume that knowledge is a territory that can be subjugated by applying a rigorous and relentless methodology.

The Getty project was only practical because it was undertaken by a single organization that could make the hard decisions. This explains the problem with ladies’ pants. Browse online at the Gap and you’ll find pants, jeans, and capris. Capris are split into cropped pants and cropped jeans, with some messy overlap between the two sets. J.Crew has pants, loungewear, denim, and suiting, and splits the pants and denim groups into different leg cuts—matchstick, hip-slung, bootcut, boy jeans, slouch, city, favorites. Anthropologie has pants divided into wide leg, slim leg, trousers, denim, short pants, petite, and tall. Even browsing all the pants doesn’t necessarily get you all pants: The Gap seems to put capris in the gap between pants and shorts. It doesn’t get any better with skirts: Anthropologie does graphic, short, straight, fluffy, and petite skirts, whereas the others have short, long, and suit. And all the stores have sale sections, of course. If there were a controlled vocabulary and a standard tree of pants, shoppers could know how to browse every store, confident they’re not missing the capris. It hasn’t happened not because ladies’ pants are more complex than the set of everything that might show up in an artwork but because there is no single entity with the Getty’s standing to declare a standard classification scheme. Classification is a power struggle—it is political—because the first two orders of order require that there be a winner.
The third order takes the territory subjugated by classification and liberates it. Instead of forcing it into categories, it tags it. Tagging lets a user of online resources—Web pages, photos—add a word or two to them so she can find them again later. The basic idea has been around for decades, but one particular site, Delicious.com (also spelled “delicious”), gave it a twist that sparked a new round of interest. Joshua Schachter, Delicious’s creator, calls it an “amplification system for your memory of Web sites.” On its most basic level, Delicious is a bookmarking site that lets you list Web pages you may want to go back to, especially if your list no longer fits comfortably in your browser’s bookmark menu. To help you find the sites you’ve bookmarked, Delicious lets you attach whatever words you want to them. If it’s a page about San Francisco, you might tag it “San Francisco,” “SF,” “my hometown,” or “4-syllable cities.” Tags let you remember things your way. When you want to find sites that talk about San Francisco, you just click on the list of tags displayed on your own page at Delicious.com and it shows you a list of all the sites you’ve given that tag.

Tagging grew out of a very personal need. Schachter’s own list of Web addresses had grown to twenty thousand, many of which he wanted to share with friends. So he built a site, now defunct, called Muxway, where friends could see the sites he’d listed. But because his friends were finding interesting sites, too, Schachter opened up Muxway so others could contribute sites they’d found. In 2003, Schachter, who was working as a financial analyst during the day, took what he’d learned from Muxway and built Delicious. Until he sold it to Yahoo! a few years later, he ran it from his apartment.

Tagging was the most important feature Schachter added to Delicious. The idea goes back to his original list of twenty thousand Web addresses. Just eight lines into the list, Schachter had annotated a site’s Web address with “#math.” By using a hash mark (#) to flag tags, Schachter could easily search the list specifically for them. At Delicious, of course, users don’t have to type a # when they want to create a tag; when you put a bookmark onto your Delicious page, a simple form pops up.

Instead of using tags, Schachter could have set up Delicious so that users create folders into which they drag Web addresses, much like the typical Internet browser’s bookmarks and like our computer desktops. But folders have a big disadvantage over tags: An item can go in only one folder, just as a physical book can go on only one shelf of a library. True, advanced computer users know that they can create what Windows calls a “shortcut,” which allows you to put links to a file into multiple folders, but it’s a time-consuming process that can quickly clutter a desktop. If you want to file a page about Aruba under “Aruba,” “Caribbean,” “beach,” “vacation,” “snorkeling,” “trips,” “too expensive,” and “daydreams,” you’d have to make a folder for each term. At Delicious, you’d simply type in those terms when you bookmark the page. Each of those tags then shows up in the tag list on your Delicious page, and clicking on any one of them assembles a list of all the pages you’ve tagged with that word. You can also find all the Web sites you’ve tagged with both “beach” and “vacation,” which would exclude the pages about Greenland you tagged as “vacation” and “extreme.”

Think how different this is from the Getty thesaurus. Rather than using a standard set of categories defined by experts, at Delicious—and the many sites that followed its lead—each person creates her own categories in the form of tags. The Getty’s categories are carefully nested, creating a well-ordered tree that would have made Dewey and Aristotle proud. At Delicious, the relationships among the tags are much messier. For example, in a traditional tree, an object can be on only one branch. At Delicious, tagging a Web address with multiple tags in effect puts it on many branches. Yet despite the lack of a well-organized scheme of categories, Delicious can make a list of twenty thousand Web addresses thoroughly usable.

That was Schachter’s first insight: Tags work as a way for individuals to remember and find pages. His second was understanding the power of making people’s lists public. At Delicious, you can not only find all the bookmarks you’ve tagged as “San Francisco,” you can also find all the bookmarks anyone else at Delicious has tagged “San
Francisco”—or “San Francisco” and “restaurant,” or you can add “vegetarian,” “Chinese,” and “cheap” to focus your search even more narrowly. Every time you check back at Delicious and click on a tag, you’ll see the latest pages to which other people have applied the tag. It’s like having a world of people with similar interests out scouring the Web for pages that you’ll find interesting, relevant to your work, or simply delightful.

These tag streams—digitally assembled lists of pages that share a tag applied by people who may not even know one another—can be immensely useful if you need to follow the latest ideas and developments on a particular topic. If you’re an industrial chemist, the tag stream of pages people have tagged “polymer” is likely to turn up information you would have otherwise missed. You could even subscribe to a tag feed, so that a daily list of new pages tagged “polymer” is automatically sent to your email in-box or to software—called an “aggregator”—designed to handle feeds. As you get into the habit, you may find yourself thinking that tagging a particular page “poly-styrene” as well as “polymer” might help other chemists find and benefit from the page. Indeed, it’s becoming common at technical conferences for the organizers to recommend that attendees tag their conference-related blog posts, photos, and online articles with a tag specific to that conference—“etech2006” or “poptech07”—so they can all be easily found by those using tag search sites such as Technorati.com. Because tagging is such an easy way to share knowledge, some companies, including IBM, are setting up their own internal Delicious sites so that research is shared within the company borders.

If you could take a top-down look at the tags at Delicious, you wouldn’t see a tree. In fact, it would far more like the floor of a forest in autumn. There are millions of tagged bookmarks at Delicious, each with an average of two tags, and over half a million different tags. Printed out, those tags would be as orderly as confetti. But in the third order of order, the messiness of miscellaneous information doesn’t reduce its utility. For example, users have uploaded over 225 million photos onto Flickr, the photo-sharing site—and are currently adding about 900,000 per day—and have applied 5.7 million different tags a total of 540 million times. Yet if you search at Flickr for photos tagged “Capri,” it neatly divides them into photos of the island of Capri and of the Ford Capri by analyzing the tags people have applied. (Apparently, not enough people are photographing their pants for a pants cluster to emerge.) The clusters are surprisingly accurate given that they’re based on nothing but the photos’ tags. It turns out that the bigger the mess—more tags, and more tags per photo—the more accurate is Flickr’s analysis. Other techniques are being developed for sorting the leaves into useful piles, including better photo recognition, bottom-up taxonomies, and even games—Google lets two strangers tag a photo simultaneously until they come up with the same word.

But we’ll never be done making sense of these piles of information. Because tags are created by ordinary people using words that are meaningful to them, there will always be ambiguity. Is “SF” San Francisco, San Fernando, or Sally Field? That ambiguity can be a problem if you have to find absolutely every resource available. But if you’re at Flickr to browse photos of San Francisco because you’re planning to go as a tourist, it won’t really matter if some of the more than 680,000 pictures tagged “San Francisco” are actually pictures of the San Francisco in Guatemala or if you miss a few thousand photos of the Golden Gate because they were tagged “SF.” The ambiguity may even introduce us to other San Franciscans we want to visit.

Tagging is one way the miscellaneous is coming into its own, but it’s not the only way. Objects that used to be organized by individuals or institutions are rapidly becoming available to us free of their old structures. Online music sites aggregate the world’s music and let us access it in any order we want, as if it all resides on an unthinkably large jukebox. Wikipedia, the grassroots encyclopedia, does the same for encyclopedia entries. New online services let biologists refer to species without having to locate them in an often-contentious tree of life. While eBay turns the world of garage sales into a miscellaneous pile, Amazon does it for books, as does the University of Pennsylvania’s PennTags project. The IBM consultant database does it for
potential team members. What may be considered the twenty-first century’s largest media company, Google, does it for Web pages. Dabble.com does it for videos. We are rapidly miscellanizing our world, breaking things out of their old organizational structures, and enabling individuals to sort and order them on the fly.

This goes far beyond simply organizing your information so you can find it again. It can change how a business works.

The British Broadcasting Corporation, known for years as “Auntie” because of its prim image, is tearing itself apart so that it can better accomplish its mission of bringing news and entertainment to British subjects. Since it began, in 1922, the BBC’s content has been organized into programs, schedules, and channels. Today, the channels, like U.S. stations, are justifiably possessive of the shows over which they’ve labored. But as the millennium turned, the BBC realized that the ability to deliver radio (and eventually television) programs over the Internet meant the audience would no longer behave according to the BBC’s schedule or way of organizing itself.

The system the BBC had to wrangle was a classic second-order monolith. Sarah Hayes, head of media asset management, and her crew work in a light-filled, airy space in the BBC’s busy headquarters, managing access to goods in an industrial warehouse kept five miles away. File a request and just as soon as someone can schlepp out to the right spot on the fifty-five miles of shelves, it’ll be put on a shuttle van and delivered to you—the very definition of what “instantaneous” isn’t. So in 1999 the BBC started spending approximately $100 million a year to preserve and to convert its archived material to modern, digital formats.

To take advantage of third-order means of finding information, the BBC began a long and complex process of turning its layout of stations and schedules into a miscellaneous pile of programs. This breakdown of the traditional ways of organizing content affects every aspect of the BBC’s business, from how it compensates its channels to their licensing agreements with producers and artists who thought they were going to control when and how often the programs were going to “air.” But the BBC realized that changing the rules of broad-

casting enables their viewers to get more value from the content the BBC produces. People want to be able to listen to or watch programs whenever they want. When listeners are trying to find, say, a jazz performance, they don’t care if the program originated on BBC Radio 1 or Radio 4. In the digital world, channels make more sense to the creators of the information than to the users of it. The audience can also find programs long buried in the BBC archives and watch them when they want to. And not only watch: The BBC has also been working on clearing the rights for programs so viewers can use portions to create new works of scholarship and creativity. It is slow and expensive work, and the BBC’s progress has been uneven, in part due to changes in management. But this radical “mixing it up” of programs—both by untethering them from their broadcast schedules and by making them available for reuse—sharply increases the BBC’s value to its customers, which is precisely its mission, and a goal for every business.

MISCELLANEOUS FROM A TO Z

On paper, it sounds like a terrible idea. Build an encyclopedia by letting anyone create or edit an article, even anonymously. Yet four years after its launch at the beginning of 2001, Wikipedia had more people reading its pages than the New York Times’ Web site did. By the middle of 2006, Wikipedia boasted over a million articles in its English edition, with more than a hundred editions in other languages. The traditional sources of authoritative knowledge have begun to pay close attention to the new kid on the block, and not only to its content. Traditionally, the articles in a work that size would be carefully arranged. But Wikipedia’s organization is as bottom-up as its content.

The Encyclopaedia Britannica does not have the luxury of being as thoroughly miscellaneous as Wikipedia. If we’re looking for the Britannica’s article on elephants, we count on being able to open the volume with the E stamped on its spine and page through alphabetized entries until we get to the one we want. If we’re feeling
adventurous, we can check out the carefully planned cross references at the end of the article. Or we can go to Mortimer Adler’s *Propædia* to find a family of Adler-approved concepts related to elephants. Either way, we are able to find information in the *Britannica* precisely because it isn’t miscellaneous.

At Wikipedia, there are no volumes—not even digital representations of volumes—to thumb through. There is an alphabetical listing of the topics, but it’s poorly done—Mortimer Jerome Adler is listed under the *M*s—probably because the listing is rarely used. There are tens of thousands of entries for each letter, on average. That’s a lot of riffling, whereas with eight keystrokes and a press of the Enter button, you could have searched for *elephant* and found the article about pachyderms instantly. At the top of the elephant article, there’s a link to a page that lists all the other articles in Wikipedia you might have meant to find when you typed *elephant* into the search box: a film by Gus Van Sant, an album by the White Stripes, a World War II German antitank vehicle, a brand of beer, or the 105th chapter of the Koran. Wikipedia reminds us that even a word as simple as *elephant* has a touch of the miscellaneous about it.

Even if you use Wikipedia’s alphabetical index, the pages are not really in alphabetical order. In fact, a Wikipedia article isn’t a single object. Although an article’s Web page looks unified to the reader, as with many pages on the Web, its text, graphics, and formatting rules are each stored separately and are pulled together only when a user requests a page by clicking on a link. If you search for *elephant* at the Wikipedia site, it’s probably the computer named Vincent (after Vincent of Beauvais, a Dominican priest who compiled an encyclopedia with 3,718 chapters in the thirteenth century) that comes up with the list of articles that use the word. If you click on the link to the main article, this sends a request to another computer, which checks to see if that article was recently requested by someone else; if so, a copy of that page is kept ready to go and a third computer—perhaps the one named Will Durant, after the historian of philosophy—simply sends the page you’re looking for. If not, Wikipedia sets about constructing the page for you. It randomly looks at one of the half dozen computers (including one named after Mortimer Adler) that store the complete text of the current articles in Wikipedia. Wikipedia then looks on Bacon (named after the philosopher Sir Francis Bacon) or one of the other computers that store the graphics, and passes both the text and the graphics to one of the dozens of computers that do nothing but assemble contents into Web pages based on templates. The finished page is then passed to your computer, where you see a text-and-graphics page about elephants.

Another level down, Wikipedia, like all computer applications, is even more miscellaneous. The computer may decide to store any single element of an article—say, the text or a photo of an elephant—in discontinuous sectors of a hard drive in order to fit the most data onto the drive and to optimize the time it takes to retrieve all those bits. That’s why when I asked Brion Vibber, the chief technical officer of the Wikipedia organization, where the text information for the elephant article is actually stored, he replied, in the chat room we were in:

<brion> god only knows.
<brion> On the disk somewheres

A shame-faced admission of an appalling ignorance? Not at all. The gap between how we access information and how the computer accesses it is at the heart of the revolution in knowledge. Because computers store information in ways that have nothing to do with how we want it presented to us, we are freed from having to organize the original information the way we eventually want to get at it. The bits and pieces of Wikipedia are, in effect, an enormous reserve of miscellaneous information that can be assembled in precisely the ways we need at precisely the moment we need it. That’s true all the way through Wikipedia, from the microscopic bits stored on the hard drives to the finished articles we read.

At the top level of this hodgepodge of bits, images, text, articles, and ideas, something remarkable happens. The million articles in English are not arranged alphabetically. They are not put into a
Dewey-like categorization scheme. There is no controlled vocabulary. There is no usable overview. Yet this enormous miscellany gets organized richly and in tremendous detail. How it happens would have driven Mortimer Adler over the brink: Wikipedia articles are packed with hyperlinks created by anyone who takes the time to add one. No qualifications are required, and no expertise is needed beyond knowing that to link the word *elephant* in an article to its entry in Wikipedia, you type “[[elephant]]”. In some entries, almost every second word is linked to another article. Together these links constitute a web of knowledge, communally constructed, ever shifting, and frequently extraordinarily useful.

Wikipedia's hyperlinked web, like the Web itself, does not look like a tree. It is a far, far more complex structure. But its shape, freed from the two dimensions of paper, better represents the wild diversity of human interests and insight.

NEW PROPERTIES, NEW STRATEGIES, NEW KNOWLEDGE

College students' silverware drawers, Delicious, Flickr, the BBC, and Wikipedia are miscellaneous in different ways, except for one thing: How their content is actually arranged does not determine how that content can and will be arranged by their users. In some cases—Wikipedia, for example—no one even knows exactly where the raw contents are. These examples are miscellaneous because users don't need to know the inner organization, because that inner order doesn't result in a preferred order of use, and because users have wide flexibility to order the pieces as they want, even and especially in unanticipated ways. This means that the miscellaneous enables all of the information contained in the set to be discovered over time.

But this also means the miscellaneous doesn't much resemble our traditional view of knowledge. Knowledge, we've thought, has four characteristics, two of them modeled on properties of reality and two on properties of political regimes.

As we've seen, the first characteristic of traditional knowledge is that just as there is one reality, there is one knowledge, the same for all. If two people have contradictory ideas about something factual, we think they can't both be right. This is because we've assumed knowledge is an accurate representation of reality, and the real world cannot be self-contradictory. We treat ideas that dispute this view of knowledge with disdain. We label them "relativism" and imagine them to be the devil's work, we sneer at them as "postmodern" and assume that it's just a bunch of French pseudointellectual gibberish, or we say "whatever" as a license to stop thinking.

Second, we've assumed that just as reality is not ambiguous, neither is knowledge. If something isn't clear to us, then we haven't understood it. We may not be 100 percent certain about whether the Nile or the Amazon is the longest river, but we're confident one is. Conversely, if there's no possibility of certainty—"Which tastes better, beets or radishes?"—we say it isn't a matter of knowledge at all.

Third, because knowledge is as big as reality, no one person can comprehend it. So we need people who will act as filters, using their education, experience, and clear thinking. We call them experts and we give them clipboards. They keep bad information away from us and provide us with the very best information.

Fourth, experts achieve their position by working their way up through social institutions. The people in these institutions are doing their best to be honest and helpful, but until humans achieve divinity, our organizations will inevitably be subject to corrupting influences. Which groups get funded can determine what a society believes, and funding is often granted by people who know less than the experts: The fate of a DNA research center may rest with congresspeople who can't tell a ribosome from a trombone.

The way we've organized knowledge has been largely determined by these four properties of knowledge. We've tried to settle on a single, comprehensive framework for knowledge, with categories so clear and comprehensive that experts can put each thing in its proper place. Institutions grew to maintain the knowledge framework. Their ability to certify experts and to vouch for knowledge made them powerful and, sometimes, rich. So when the miscellaneous shakes our certainty in the nature of knowledge, more than
the future of the card catalog is at stake. Because a third-order miscellaneous is digital, not physical, we no longer have to agree on a single framework. Things have their places, not a single place. We get to create our own categories, ones that suit our way of thinking. Experts can be helpful, but in the age of the miscellaneous they and their institutions are no longer in charge of our ideas.

These are big changes, but perhaps the most urgent one is this: Over the course of the millennia, we’ve developed sophisticated methods and processes for developing, communicating, and preserving knowledge. We have major institutions—serious contributors to our culture and our economy—devoted to those tasks. We’re good at it. Now we have to invent new ways appropriate to the new shape of knowledge. We are doing so at a pace unparalleled in our history.

Four new strategic principles are emerging, severing the ties between the way we organize physical objects and ideas.

**Filter on the way out, not on the way in.** A friend of mine who worked at the *Harvard Business Review* tells amusing stories about the “slush pile,” the unsolicited manuscripts that arrive every day. The *Harvard Business Review* is a sober journal of research and ideas, yet people submit poetry, short stories, andarty photographs. My friend’s job was to go through the slush pile to see what, if anything, was worth passing along for serious consideration. She was a gatekeeper, a filterer, doing a job that makes sense when the economics and physics of paper force us to make decisions about what knowledge we will publish and thus preserve. We rely on experts such as my friend to spare us from having to wade through the slush pile on our own.

But when anyone can publish at the press of a button, the social role of gatekeepers changes. For example, from the outside the “blogosphere” looks like a self-indulgent pool of slush that wouldn’t get past the usual publishing filters. While the economics of publishing ensures that most blogs indeed wouldn’t be let through the gates, the aggregate value of all the blogs in the “long tail” (to use the term Chris Anderson made popular in his book of that name)—each perhaps of interest to only a few people—is incalculable. This is an inversion of the old model. In a world of parsimonious access to paper, filters increase the value of what’s available by excluding the slush. But in the third order, where there’s an abundance of access to an abundance of resources, filtering on the way in decreases the value of that abundance by ruling out items that might be of great value to a few people. Filtering on the way out, on the other hand, increases the value of the abundance by locating what’s of value to a particular person at a particular moment. For example, a physics professor at McGill University, Bob Rutledge, started an electronic bulletin board that posts new findings for *any* astronomy research as soon as it can be summarized. Rutledge doesn’t apply criteria to decide for the reader whether the research is important enough to be included (though only active, professional astronomers can register to post to the site). It’s up to each reader to be the filterer. Similarly, the Public Library of Science’s biology journal, a peer-reviewed but free online resource, started PLoS One in November 2006. “The idea is to take the editorializing out of the peer review process,” says Hemai Parthasarathy, the managing editor. So long as a paper is “sound,” it will be published. If it’s good science, *someone* may find it useful. So long as the user has good tools for finding what she needs—and this is a task many are working on—filtering on the way out vastly increases our shared potential for knowledge.

**Put each leaf on as many branches as possible.** In the real world, a leaf can hang from only one branch. In the first order of organization, there’s no way around that limitation. In the second order, most cataloging systems have provisions for listing books under more than one heading, but the physicality of the second order still usually demands that one branch be picked as the primary one, and there is a limit on the number of secondary listings.

In the third order, however, it’s to our advantage to hang information from as many branches as possible. If you get a new Casio digital camera to sell in your online store, you'll want to list it under as many categories as you can think of, including cameras, travel gear, Casio products, graduation gifts, new items, sale items, and perhaps
even sports equipment. Hanging a leaf on multiple branches makes it more findable by customers. Unlike in the second order, this doesn’t make your e-store disorganized or messy. It makes it more usable . . . and more profitable.

**Everything is metadata and everything can be a label.** In a store, it’s easy to tell the labels from the goods they label, and in a library the books and their metadata are kept in separate rooms. But it’s not so clear online. If you can’t remember the name of one of Shakespeare’s plays, go to the search box at Google Book, type “Shakespeare tragedy,” and you’ll see a list of all of them. Click on, say, *King Lear* and you can read the full text, including the famous line, “How sharper than a serpent’s tooth it is to have a thankless child!” Now suppose you want to know where the quotation “How sharper than a serpent’s tooth” comes from. Type the phrase into the search box and Google Book will list *King Lear*. Simple, but in the first case you used Shakespeare’s name as metadata to find the contents of a book and in the second you used some of the contents of the book as metadata to find the author and title. In the miscellaneous order, the only distinction between metadata and data is that metadata is what you already know and data is what you’re trying to find out.

In the first two orders of order, we’ve had to think carefully about which metadata we’ll capture because the physical world limits the amount of metadata we can make available: A book’s catalog card has to hold far less information than does the book itself. In the third order, not only can every word in a book count as metadata, so can any of the sources that link to the book. If we want to help our customers or users find information, we’ll try to make as much of it usable as metadata as we can.

This not only makes sites easier to use, it vastly increases the leverage of knowledge. Think of what we can do with just the few words that fit on a second-order card or a label. Now that everything in the connected world can serve as metadata, knowledge is empowered beyond fathoming. We not only can find what we need based on what-
ever slight traces we have in our hand, we can see connections that would have escaped notice in the first two orders.

The power of the miscellaneous comes directly from the fact that in the third order, everything is connected and therefore everything is metadata.

**Give up control.** Build a tree and you surface information that might otherwise be hidden, just as Lamarck exposed information left hidden in Linnaeus’s miscellaneous category of worms. But a big pile of miscellaneous information contains relationships beyond reckoning. No one person or group is going to be able to organize it in all the useful ways, hanging all the leaves on all the branches where they might be hung. For example, iTunes shows users a branch that pulls together albums by a particular artist, but the millions of playlists that users have made there find relationships that the organizers of iTunes could not possibly have foreseen, from techno versions of children’s songs to tracks played at someone’s third wedding. iTunes simply cannot predict what people are going to be interested in, what a song is going to mean to them, and what connections they’re going to see. Some of the combinations will be of passing value to only one person, but other people may find their world changed by how a stranger has pulled together a set of songs to express a mood, an outlook, or an idea.

That’s why it’s so powerful to let users mix it up for themselves. Go into a real world clothing store and try pulling everything in your size off the racks and into a shopping cart so you can go through it in an orderly fashion. After all, that’s the rational way to proceed. Everything that’s not your size is just noise, a distraction. Yet, within ninety seconds you’ll be thrown out of the store and firmly asked not to return. On line, on the other hand, we just naturally expect to organize information our way, through tags, bookmarks, playlists, and weblogs. And then we add to the information that a site provides us by disagreeing with it in our own reviews. Users are now in charge of the organization of the information they browse. Of course, the owners of that information may still want to offer a prebuilt categorization, but
that is no longer the only—or best—one available. Put simply, the owners of information no longer own the organization of that information.

Control has already changed hands. The new rules of the information jungle are in effect, transforming the landscape in which we work, buy, learn, vote, and play.

SMART LEAVES

In 1948, two graduate students at the Drexel Institute of Technology in Philadelphia overheard the president of a local grocery chain asking a dean to sponsor research into how to read product information automatically. The students, Joseph Woodland and Bernard Silver, inspired by the dots and dashes of Morse code, came up with a set of straight lines much like the modern zebra-stripe bar codes, and in 1951 they unveiled a machine that could translate the bar codes back into numbers. It was the size of a desk, wrapped in black oilcloth, and used a 500-watt bulb as the light source. “It could cause eye damage,” Woodland recalled.

In 1966—four years after Silver died, at the age of thirty-eight—the idea went commercial when the National Association of Food Chains put out a call for automatic checkout machines to speed up checkout lines. The first was an RCA system installed at a Kroger store in Cincinnati in 1972, but to get real efficiency, bar codes would have to be put on the packages by the manufacturers, not the clerks working in the local stores. So the association established the Uniform Grocery Product Code, the grandparent of the Universal Product Code (UPC) standard we use today. In 1974, at a Marsh Supermarket in Troy, Ohio, the first working system successfully identified a ten-pack of Wrigley’s Juicy Fruit chewing gum that is now housed in the Smithsonian. In 1981, the U.S. Department of Defense required bar codes on all products it purchased and the UPC system went