CLIVE THOMPSON for BUSINESS, *Wired,* DATE OF PUBLICATION: 06.25.07

FOR CERTAIN TASKS, THE CORTEX STILL BEATS THE CPU

Which is prettier? A picture of a black cat sleeping on a pillow or one of a curly-haired brunette woman in a miniskirt? I've got only a few seconds to decide. I vote for the cat. I'm sitting in a laboratory at Carnegie Mellon University, playing Matchin', a computer game developed by Luis von Ahn. In the game, two players — von Ahn and I, seated at different terminals — watch as pairs of pictures swiped off the Internet flash up on our screens. Our goal is to pick the one we think both of us will find more attractive, not necessarily the one we personally prefer. This requires a sort of mindmeld, and it doesn't always work: Von Ahn picks the girl in the miniskirt instead of the cat. We've got one minute to process as many pictures as we can, so we race on frantically, evaluating photos in an instant. Soon we hit a groove: We both say that a picture of a peacock is prettier than one of a picnic, a baby is lovelier than a tombstone, a wedding couple beats a field of wheat. Then the game is suddenly over, and we get our score: We agreed 70 percent of the time. Pretty good, but not enough to hit the high-score tables.

"Man," laughs von Ahn. "You picked some weird stuff!"

Luis von Ahn's new games pair random players to solve a computing problem. Because the two players get points when their answers match, the accuracy — and fun quotient — increases. To try them, go to [www.gwap.com](http://www.gwap.com/).

1) Matchin' Players are shown the same pair of images, then each tries to pick the one they'll both agree is more attractive. Creates a database of images searchable by aesthetic value, a task no algorithm can perform.

2) Babble Two English-speaking players are shown a sentence in a foreign language that neither of them speak. A list of possible English meanings appears below each word. Players try to agree upon a set of English words that forms the most coherent sentence. Translates foreign text into English without requiring anyone fluent in both languages.

3) InTune Players listen to the same audioclip and then try to come up with the same phrase to characterize it. Tags sounds with searchable descriptive text.

4) Squigl Two players are shown the same picture and a word describing an element within the image (e.g., a picture of a dog and the word "leash"). They each draw a border around the element. Produces a set of pictures with their internal components tagged — terrific for very specific image searches.

5) Verbosity One player is given a word, and the other tries to guess that word by completing phrases such as "It is near a \_\_\_\_" or "It is a type of \_\_\_\_." The first player answers "true" or "false" but can't use the word itself. Creates a database of commonsense knowledge describing the objects.

It's an oddly enjoyable game. But Matchin' is also a covert experiment in artificial intelligence. Every time players agree on a picture, it's tagged as prettier. Von Ahn, a 28-year-old professor of computer science at Carnegie Mellon, will put the game online this summer, and as thousands of people play it, his database of 100,000 photos will be imbued with something quintessentially human: an aesthetic sensibility, encoded as a ranking of attractiveness.

The game basically tricks humans into teaching computers what constitutes prettiness. If enough people play Matchin'— and von Ahn's previous games have garnered millions of play-hours — it could eventually rate the appeal of every image on the Internet. Google could incorporate the ratings into its search engine, so you could search specifically for "beautiful" pictures of houses, people, or landscapes.

"People are good at figuring out what's attractive, and computers are good at quickly searching and finding," von Ahn says. "You put them together, and bang!"

This is "human computation," the art of using massive groups of networked human minds to solve problems that computers cannot. Ask a machine to point to a picture of a bird or pick out a particular voice in a crowd, and it usually fails. But even the most dim-witted human can do this easily. Von Ahn has realized that our normal view of the human-computer relationship can be inverted. Most of us assume computers make people smarter. He sees people as a way to make computers smarter.

Odds are you've already benefited from von Ahn's work. Like when you type in one of those stretched and skewed words before getting access to a Yahoo email account or the Ticketmaster store. That's a Captcha, which von Ahn developed in 2000 to thwart spambots. Or there's von Ahn's picture-labeling games, which have lured thousands of bored Web surfers into tagging 300,000 photos online — doing it so effectively that Google bought his idea last year to improve its Image Search engine.

Last winter, von Ahn was awarded a $500,000 MacArthur genius grant, and in April he received another $200,000 as one of Microsoft's New Faculty fellows. This summer he's putting the money to good use, launching five new games that will identify sounds, give computers commonsense logic, and even help scanners perfect their optical character recognition.

"Captchas would be enough for most people to build their entire career on," marvels Josh Benaloh, a cryptographer at Microsoft who hired von Ahn as a summer intern just three years ago. "Luis keeps coming up with new stuff."

Von Ahn wants to harvest every idle moment in our lives and turn it to productive use. If the world's computer solitaire players could be coaxed into enjoying a game that contributed to solving a computing problem, he calculates, it would produce billions of man-hours of labor each year.

"Basically, I want to make all of humanity more efficient by exploiting the human cycles that get wasted," he tells me over lunch at a diner near his lab in Pittsburgh. "As humanity goes online, it's becoming an extremely advanced, large-scale processing unit."

I suddenly realize how very weird the world appears from inside von Ahn's head. He gestures around the restaurant. I see businessmen eating hamburgers, a couple chatting over coffee, waitresses hustling to fill orders; von Ahn sees idle processors waiting to be put to use.

"Together," he says, "we're the biggest supercomputer in the world."

New Captchas Help Digitize the World's Libraries

Captchas thwart spambots. Now reCaptchas are helping digitize the world's libraries.



Captcha A regular Captcha helps keep bots off of Web sites. You read the skewed word and type it in. If you get it right, you've proved you're a human and are allowed access.



reCaptchas The next-gen test presents two skewed words bisected by a line, making machine deciphering even harder. Both are taken from the Internet Archive's project to scan public-domain books. One word is known to the computer; the other couldn't be read by the Archive's scanners, so when you type it in you're doing a tiny bit of work for the project.

The son of two doctors in Guatemala City — "They said to me, ‘Son, you can do anything you want, but don't be a doctor'" — von Ahn grew up speaking fluent English and geeking out on Commodore computers. In 2000, he enrolled in the computer science graduate program at Carnegie Mellon, where he became fascinated by cryptography. It wasn't the programming that intrigued him so much as the paranoid psychology: He loved thinking about the ways people break security systems and especially the problem of how to verify that someone is who they say they are.

Making sure that people are real, as it turned out, was precisely the problem that confronted Yahoo. Bots were running wild on the company's site, spamming chat rooms and email accounts, and Udi Manber, the company's head of technology at the time, couldn't figure out how to stop them. Manber discussed the problem with Manuel Blum, a computer science professor at Carnegie Mellon and von Ahn's PhD adviser. To keep the bots from setting up bogus accounts, Yahoo needed a Turing test, a challenge that humans could pass but computers couldn't.

Visual recognition was just such a task. Blum and von Ahn set about designing a test. Von Ahn wrote a program that generates four random letters and numbers, distorts them, and places them on a fuzzy background. Type in the four characters correctly and you're in. Von Ahn dubbed his invention the Completely Automated Public Turing Test to Tell Computers and Humans Apart — or Captcha — and when Yahoo began using it, the spambot problem was greatly reduced. Soon, Ticketmaster was using Captchas to keep scalper-bots at bay, and AOL and Microsoft adopted them to protect their email systems.

Von Ahn got to thinking: If people could so easily recognize pictures of letters and numbers, could he coax them to use this ability to identify and label the vast number of images on the Web? It wouldn't be easy. If he simply showed people pictures and asked them to type in a label, odds are they would screw it up — using an incorrect or weird word. So he began thinking like a crypto grapher. Perhaps he could get strangers online to verify each other's output. If two people agreed on the same word to describe an image, each would be holding the other to account.

The way to do it, he realized, was as a game. It would pull images off the Web, then randomly pair two players from around the world. They would be shown the same images, then each would type in as many words as they could to describe those images, hoping to hit upon the same ones as their anonymous partner. They'd get 50 points for each match, and two and a half minutes to earn as many points as possible. Von Ahn suspected that whenever the players agreed on a word — "meadow" to describe a tree-lined clearing, for example — they would be choosing a highly accurate label for the picture.

Von Ahn cobbled the game together in a week — "crappy, totally terrible code," he admits — and threw it online. He dubbed it The ESP Game and emailed the URL to a few friends. Within days it was Slashdotted, whereupon his server nearly crashed under the load of new players. Astonished, von Ahn watched for the next four months as 13,000 players produced 1.3 million labels for some 300,000 images — with a few hardcore fans clocking more than 50 hours of play. "It's like crack," as one player complained in an email to von Ahn.

The labels his players generated were far more accurate than what other image-search technologies produced. Most search engines are limited to sniffing out words associated with a picture, such as the name given to the image, words in the page around it, or links pointing to it. That's inherently imprecise: When von Ahn recently searched for "dog" on Google, a third of the pictures showed no dogs at all. When he queried the ESP database, almost all the results contained canines. Better yet, players often generated labels that were subtle and nuanced. A search for "funny" found a picture of Ronald McDonald being hauled away by police and one of Queen Elizabeth picking her nose.

In December 2005, von Ahn demo'd his game at Google. After the presentation, Sergey Brin and Larry Page approached him. "They stayed for the whole speech, which apparently they never do, and then came up and said, ‘Hey, let's commercialize this,'" von Ahn recalls. Four months later they had licensed the game, and in August 2006 they debuted it as the Google Image Labeler, quietly using it to make the company's database of images better and smarter.

The day I met up with von Ahn, he and his team of 10 students were in their lab working on a set of new games they're launching in July. The goal, as with ESP, is to harness human labor to do the important work for which computers are ill suited — and to make the process so entertaining that people are happy to do it for free. But there's one big problem: Some of the games just aren't fun to play.

Von Ahn rocks back in a big leather chair, frowning faintly as Edith Law, a 27-year-old student, plays InTune with Severin Hacker, an exchange student from Switzerland. In the game, two players listen to a sound — pulled from the Freesound Project, a European database of noises — and try to generate the same word to describe it in just a few seconds. The goal is to label the sounds so they're searchable, not only by obvious descriptive words ("guitars," "waterfalls") but also by emotional responses ("sounds that make you happy").

The problem is, the sounds are so obscure that no one in the room can figure out what the heck they are. The first one is vaguely like a sci-fi laser… maybe. The next is a perplexing low drone. Then it's a bouquet of echoey chirping that might have been a bunch of angry crickets. Law cocks her head toward the laptop speaker and looks completely baffled.

"OK," von Ahn says abruptly after an excruciating minute or two. "This isn't working." Nobody's going to play a game that asks them to decipher impenetrable noise. "If we want people to enjoy the game," he tells his team, "we need to be playing them clips of things they care about, things that are fun — like pop music. We need, like, Britney Spears!"

"I don't have any," sighs Law.

"Well, if it isn't like that, it won't be fun, and no one will play it," he replies.

This is the problem faced by all of von Ahn's human computation projects. People will contribute their brainpower, but only if they're given an enjoyable, time-killing experience in exchange. Play is the unexpected glue that lashes human brains together into a global overmind. So to build a good human-computation project, you can't merely be a scientist; you also need to be a videogame designer.

This, as it turns out, is a significant hurdle, because few academics are trained in game design. Von Ahn tries to find students who have done it themselves: Law worked for Ubisoft, and others have programmed their own indie Flash games. "Game design is a funny thing," von Ahn admits. "There are people out there who are really good at it, but it's not clear that they can teach it. It's a very intuitive process. It's an art."

His team has generated dozens of concepts in the past year, and this summer von Ahn will launch five of them on the Web (see "Serious Play," this page). The others, however, failed this crucial test — they just weren't entertaining enough.

Which raises a question: Are von Ahn's games inherently limited to simple challenges? If the overmind will process only pleasant tasks, how many deep problems can it really tackle?

Most projects that harness human processing power rely on a different motivator: money. Many companies use Amazon's Mechanical Turk to farm out informational piecework, usually for about 10 cents a task. These are often quick-hit recognition jobs that computers can't perform. PriceGrabber .com uses thousands of Web surfers to update its catalog by hunting down and in putting descriptive information on products. A mapping firm called Geospatial Vision hires crowds to pore over satellite pictures of cities and identify tiny features like lampposts and road signs. "There are 6 billion human beings in the world. We could do a lot of stuff if we all work together," says Peter Cohen, director of Mechanical Turk.

Other companies have designed their own collaboration engines: Threadless sells T-shirts designed and rated by visitors to the Web site; Eli Lilly takes scientific problems that its in-house researchers can't crack and posts them online. There's a site that mashes up human computation and dating called I'm in Like With You. The idea has even been floated as a security concept: Jay Walker, the founder of Priceline.com, has argued that companies, and even the government, could hire online crowds to cheaply monitor security cameras from home.

Von Ahn has figured out how to get this labor — and tons of it — for free. But because it's so devilishly hard to make things fun, he's in a category by himself: No other researcher or company has successfully turned a collaborative project into a game. Two years ago, Bryan Russell, a graduate student at MIT, launched LabelMe, a project in which contributors draw outlines around objects in photos. The goal is to produce marked-up images that can be used to train visual- recognition software. Russell says he considered making it a game but ended up relying on the altruism of other researchers in his field. Boundary drawing is a tedious task, he says, and it's best performed by visual-recognition experts.

"We wanted high-quality labeling, and it's hard to get average people to do it well," Russell says. "I'm not sure you could make a game out of it."

Indeed, because nobody verifies the output from von Ahn's games, the results can occasionally raise eyebrows. Players in The ESP Game attached the label "Saddam" to a picture of the mustachioed actor Walter Matthau. A picture of George W. Bush was given the obvious labels "Bush," "George," and "president" — but also "dumb" and "yuck." Von Ahn defends the results; the point, he argues, is to create image descriptions that are meaningful to humans, and those labels certainly are. (Matthau did sort of look like Saddam Hussein, after all, and Bush… oh, never mind.) Wikipedia has to confront this all the time: Which produces better results — a small group of experts or a huge mob of amateurs?

Von Ahn realizes that some tasks are inherently unenjoyable — until you make them a game. People show up to play for two minutes and stay an hour. "I feel kind of bad sometimes about how I'm sucking them in," he jokes. Matchin' will offer players a side benefit. If they're intrigued by their partner's answers, they can click a button to introduce themselves, turning it into a dating game.

Some of his games may turn out to have deadly serious applications: This spring, von Ahn got a call from the Department of Homeland Security. He went to Washington to meet with DHS officials, and together they devised a game in which people are challenged to find dangerous objects in images of x-rayed baggage. The pictures would be fed from airport scanners, and players would act as a second set of eyes for overtaxed security employees. If enough players noticed something amiss, an alert would be triggered.

Von Ahn knows how kooky this sounds. "They're not trying to get rid of the screener in the airport," he says. "The idea is to provide assistance. It's a difficult task; maybe with 10 helpers, that guy can do a better job." To allay security and privacy concerns, only Homeland Security officials behind a government firewall will be allowed to play the game when it launches. Still, von Ahn was stunned that the department was willing to even entertain such a concept.

"They have some very intelligent people, and they're very forward-thinking. But I was like, ‘You're thinking about what?'"

Like most crypto freaks, von Ahn worries all the time about people breaking his systems. Cheating at human computation games would corrupt the data. And he's not necessarily being paranoid; hackers have relentlessly attacked von Ahn's creations in the past. Ever since Captchas hit the Net, spammers have tried, with some success, to bust them. Some have hired third-world workers to solve them by hand — ironically, a Mechanical Turk like approach. Others have offered free online porn to surfers willing to solve Captchas.

The stakes are high. Every day, thousands of spam blogs are created that threaten to corrupt search results, and companies like Ticketmaster lose consumer confidence when scalper-bots jump the queue for tickets. Von Ahn doesn't worry too much about the use of cheap overseas workers to defeat Captchas; solving each one manually takes too long to cause much damage. What concerns him is that spammers are developing algorithms to solve the puzzles.

So he's fighting back. In late May, von Ahn launched reCaptcha, a service that he believes is the toughest Captcha yet devised. ReCaptcha presents users with two stretched and skewed words, each bisected by a diagonal line. The line poses a particular problem for visual- recognition spambots, because machine vision has trouble with boundaries; if it can't figure out where one character ends and another begins, it can't separate the letters. Humans can do this easily, so we simply enter the two words and we're in.

But reCaptcha has an even sneakier — and more delightful — purpose. The words are pulled from the book-scanning project of the Internet Archive, a nonprofit project in San Francisco that aims to digitize millions of public-domain books and put them online for free. One of the two words in the test is the control word: The gatekeeper computer knows what it should be, so it's there to make sure the puzzle-solver is indeed human. But the other word is there for a different reason. The Archive's scanners are good, but some of the words are too smudgy for the software to decipher. The game takes the image of each smudgy word and puts it into reCaptcha. Each time someone completes a reCaptcha puzzle, they'll be doing a tiny bit of work — translating that difficult image into text, which von Ahn eventually feeds back into the Archive.

It is von Ahn's mania for efficiency taken to its logical extreme. Since people are going to be forced to solve Captchas every day, he figures, why not use that labor? "Every time somebody does one, they basically waste 10 seconds of their life," he says. "By inventing Captchas, I've essentially become this huge time waster. So the question is, can we get you to do work for those 10 seconds?"

Roughly 50 million Captchas are solved each day. If von Ahn can acquire just a fifth of those users, he'll have a stunning 30,000 daily man-hours of work at his disposal. It would constitute the world's fastest and most accurate character-recognition computer, processing 10 million words a day.

At that rate, the Archive's books become a trivial problem: Von Ahn could dispatch them in months. He's planning to launch a company — Games With a Purpose — to market the system, and he's already in talks with a major newspaperto digitize its 150-year back catalog. Beyond that, there are hundreds of law firms with equally vast scanning needs that von Ahn could service. He even envisions putting his massive virtual organization to work for banks. They currently pay full-time employees to look at each check before it's cashed, verifying that the numerical amount matches the amount written out in longhand. It's another task that only humans can do — but it could be farmed out to reCaptcha.

Von Ahn says it's not yet clear what the limits of such an enormous human computer would be. "We're still not thinking big enough," he tells me at one point. "If we have that many people all doing some little part, we could do something insanely huge for humanity."

"We'll never run out of things to digitize," he adds. Or human brains to do it.

Reading Questions: “For Certain Tasks, the Cortex Still Beats the CPU”

1. Wait a second. What is a cpu? If you have to quote a source, use a full bibliography.
2. Matchin’ is “a covert experiment in artificial intelligence.” What data is being associated with each picture that the computer could not generate on its own? P\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. The artificial intelligence part works because “the game… tricks humans into \_\_\_\_\_\_\_\_\_\_\_\_\_\_ computers” and “if enough people play Matchin’… [the computer] could rate the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of every image on the Internet.”
4. Computers are not good at “point[ing] to a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of a bird or pick[ing] out a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ in a crowd.”
5. Captcha was developed to “thwart \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.”
6. A Turing test is “a challenge that humans could \_\_\_\_\_\_\_\_\_\_\_\_\_ but \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.”
7. What does Captcha stand for?

C \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ A\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ P\_\_\_\_\_\_\_\_\_\_\_\_ T\_\_\_\_\_\_\_\_\_\_\_\_\_ T\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ to T\_\_\_\_\_\_\_\_\_\_\_\_ C\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

and H\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ A\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Computers don’t guess at descriptions of pictures well. “Most search engines are limited to…” what words associated with pictures? (list three sources of these words)
2. Matchin’ works with images. What does InTune work with? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. Why add Britney Spears, according to von Ahn?
4. What is missing in a lot of professionals working on computation/computer science projects? They lack experience in \_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ *(By the way, you will get experience in this, in this class!)*
5. List three other ways that crowd sourcing (using input from a lot of people) has helped companies or the government. (There are more than three mentioned in the paper.)
6. Labels like “Saddam,” “dumb” and “yuck,” are interpreted by some to be inaccurate, but Mrs. Frazier and Von Ahn see the great value of this. What is being communicated about pictures other than the literal description of what is there? *(This is not overt in the article, but the examples above are excellent examples – you’ll have to think about this answer, students!)*
7. What is Von Ahn concerned about? “spammers are developing \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.”
8. reCaptcha doesn’t just waste time, it also does useful work. Describe one of the ways reCaptcha knowledge/learning obtained from tons of human knowledge can help the world of business or literature.