Proper citation:
Hwa A. Lim, "Einformatics — a green-collar career", *Symbiosis*, June, 2008, pp. xx–yy.

Einformatics — A Green-Collar Career

By

Dr. Hwa A. Lim, Ph.D., MBA, MA, B.Sc. (Hons.), ARCS

Silicon Valley, San Jose, California, USA.

hal@dtrends.com

₪ "The biggest mistake that you can make is to believe that you are working for somebody else. Job security is gone. The driving force of a career must come from the individual. Remember: Jobs are owned by the company, you own your career! ₪

- Earl Nightingale (1921–1989).

---

1 Einformatics

Informatics started in the second half of the 20th century, and bioinformatics in the last quarter of the same century. We are now anticipating Einformatics to be a dominating force in the 21st century.

In the 1980s, when Hal first initiated the field of bioinformatics—a field that is now a part of university curriculum—it was to be called "Dinformatics," where the 'd' may stand for 'data,' 'database,' 'data mining,' 'digital,' 'DNA,' 'discovery,' 'diseases,' 'drugs' or even 'dotcoms.' But later it became "bioinformatics," "cheminformatics," and variations thereof.[1–8]

Now by the same rationale we are defining a new discipline, "Einformatics," where the 'e' may stand for 'efficiency,' 'energy,' 'environment,' 'ethanol,' or even the 'e' in the electronic world. Some people may want to call this "eco-friendly informatics," but this is a mouthful, and a shortened version "ecoinformatics" will sound too much like "informatics of economics." Others may prefer "Greeninformatics," which is also fine, because the issues of being green are precisely an aspect that Einformatics will address. With this understanding, in fact, the 'e' may also stand for 'economics' since the economics of "being green" is a central issue in Einformatics. Indeed, to stretch further, 'e' may also stand for 'earth' since Einformatics will be addressing global issues such as global warming, carbon neutral or carbon negative. But then we are getting carried too far afield. For all our intents and purposes of discussing this new discipline, we shall just call it Einformatics.

Informatics arises from information technology (IT); similarly bioinformatics arises from biotechnology. In an analogous fashion, Einformatics derives from Energy Technology (ET), or Green Technology. Again, here we shall stay with Energy Technology because it sounds more whimsical, having been made popular by "extra-terrestrial" (ET).

1.1 À la Bioinformatics

Since its humble beginning, the definition of "bioinformatics" has metamorphosized from the original—and in many ways, naïve—definition of "data collection, analysis and dissemination" to a current more encompassing definition found in literature. Its domain of application has widened from being a special niche tool to that of an essential corporate technology. The scope has also widened from a laboratory-based tool to an integrated corporate infrastructure.[9]

For years after its introduction, a recurring issue that regularly popped up is the definition of bioinformatics, particularly from individuals who were trying to get into the field. Most lamented that there are too many definitions of bioinformatics. This was a consequence of the expanding scope of the subject. As both an enabling and enabled discipline, it has been defined differently depending on the domain of expertise of the person who is giving the definition: A computer scientist will give one definition, a biologist another, a biotechnologist a different version, and an individual from a pharmaceutical company will provide yet another variant…

Each definition is as good as the other; this is just the nature of the beast. The important point to note is that bioinformatics is an enabling and an enabled discipline. As such it will never replace the bench work and wet lab experiments of the biological, biochemical, health and clinical sciences. It only helps the areas it is being applied in to

- Eliminate unlikely candidates (such as in drug target discovery);
- Interrelate data and information (such as in analysis);
- Extrapolate into regime inaccessible by experiments (such as in cases not possible with current state of the art of the technology);
- Study cases that will be unethical to do (such as those studies that will be too invasive to the human body);
- Etc.

Thus it would be unwise, particularly for bioinformaticists, to just write the best computer software, or to just integrate the most sophisticated packages to churn out numbers and beautiful graphics. The bottom line is still the biology, the biochemistry, and the healthcare system. Nothing, not the best software, nor the most sophisticated package, shall replace real life systems.

1.2 Now the Enabling Einformatics

Einformaticists already exist. Examples are the superb global warming modelers in NASA, Jet Propulsion Lab (JPL) and at various universities, or the group of tireless people, including the Intergovernmental Panel on Climate Change (IPCC), with former U.S. vice president Al Gore as its messenger. The latter two are co-laureates of the 2007 Nobel Peace prize. These people have been working in Einformatics, without even knowing it.

A definition of bioinformatics is the study of information content and information flow in biological systems and processes. Einformatics can be defined, in a manner analogous to bioinformatics, as the study of the information content and information flow in energy systems and processes. So what is the big deal? Historically, every discipline started as seemingly disparate areas until a coordinated effort to integrate the fragments into one with a common, central focus, either by combining different disciplines or using a common tool. No discipline ever started from vacuum; even Sir Isaac Newton sat on the "shoulders of giants."

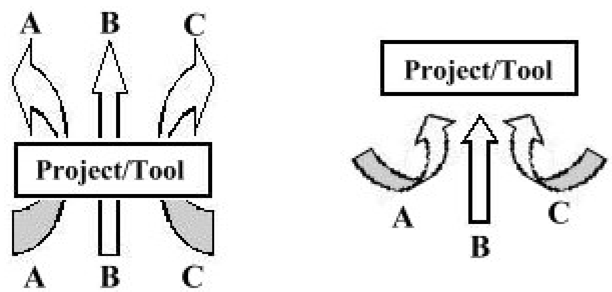


Figure 1. There are at least two models in which different disciplines can work together. In the first, the disciplines come together, work on a project or use a common a tool, and then they return to their respective disciplines. In the second, the disciplines come together to be part of a common project.

In the modern days of megaprojects and grand challenges, there are at least two ways different disciplines may be brought together, and often times, to create another discipline. In the first, experts from different disciplines come together to solve a problem and then they return to their respective disciplines. This is common in an academic setting. For example, a high-performance computer institute at a university, in which departmental faculty members are encouraged to work with research scientists at the institute to solve a common problem using the high-performance computer (the tool). At the end of the day, departmental faculty members return to their respective departments, from which they are seeking tenure or in which they already have tenure.

In the second, experts from different disciplines are brought together to widen the perspectives of solving a grand challenge problem. This is more common in a national laboratory setting. Two excellent examples are the Manhattan Project in which the Los Alamos National Laboratory recruited to best brains of diverse backgrounds to create the first atomic bomb, and the international genome consortium.

Bioinformatics operates equally well in either mode, so does Einformatics. In these modes, bioinformatics took almost a decade to go mainstream. How soon Einformatics will go mainstream will depend on how quickly policy-makers enact carbon tax or equivalent laws, how quickly the dwindling energy resources are affecting economic growth, and the impact of pollution on the environment. The 2007 Nobel Peace Prize does help bring awareness to the public, so do the greensayers. But unlike the decade-long incubation period of bioinformatics, it is anticipated that Einformatics will go mainstream in a relatively short period, perhaps in 2–3 years, or in the next generation of university graduates.

1.3 University Curriculum

From the time bioinformatics was introduced in the 1980s until about the year 2000, students and young researchers interested in getting into bioinformatics constantly asked questions. Most of the questions are of a common nature: what do I (or my children) have to take in college to become a bioinformaticist?

It is anticipated that new Einformatics entrants will ask similar questions. First note that Einformatics, just like bioinformatics, is an interdisciplinary and transdisciplinary field:

- It is interdisciplinary because it involves a number of subject areas.
- It is transdisciplinary because experts in peripheral fields can apply new breakthroughs—
    a. In their fields of expertise into the field, that is, in an enabled way;
    b. In the field into their fields of expertise, that is, in an enabling way.

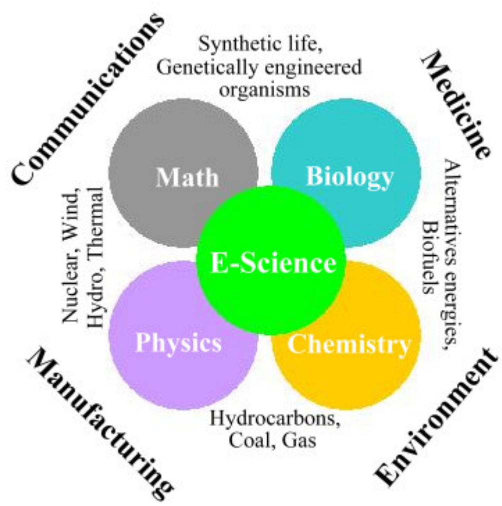


Figure 2. Energy science (E-science) is an interdisciplinary and transdisciplinary discipline. Shown above is an example of how the interfaces of basic sciences form E-science, and some of its applications.

Just like bioinformatics, depending on the background of the experts, they each will likely offer a different definition of Einformatics. Each definition will be as good as the other. The important point to note is that Einformatics is an enabling/enabled discipline. As such it will never replace the bench work and wet lab experiments of energy systems; it only helps the areas in which it is being applied to

- Eliminate unlikely candidates (such as which feedstock to use);
- Interrelate data and information (such as in analysis of Arctic ice thaw, carbon calculus);
- Extrapolate into regime inaccessible by experiments (such as in nuclear reactor technology);
- Elucidate long-term repercussions (such as global warming impact);

- Analyze energetic (such as that of corn ethanol, sugar ethanol);
- Etc.

Similar to bioinformaticists, it would be unwise for Einformaticists to just write the best computer software, or to just integrate the most sophisticated packages to churn out numbers and beautiful graphics. The bottom line is still the biology, chemistry, physics, engineering, ecology, geology, environment, resource allocation, and economical impact.

1.4 E-technology and E-companies

An indication that a new field has taken off—for example, from E-science to E-technology to E-commerce—is when lawyers (attorneys), accountants, recruiters, publicists, and others such as investors are beginning to show up at pertinent scientific meetings. There are already signs of such presence and mingling in energy technology (ET).

In the Internet boom, the mantra was to change the world and get rich quick; this time, given the size and scope of the energy market, the idea is to change the world and get even richer—but somewhat more slowly. But starting a clean technology firm is not like starting a low entry-barrier Internet company, or an online do-it-yourself (DIY) company; scientific credibility is the primary currency. One of the credentials will be Einformatics certifications to certify green-collar technicians and professionals. How the new technology will be adopted is a discourse in itself.[10]

2 Green-Collar Einformaticists

The Y2K problem was the "millennium bug." Just prior to the new millennium, the only country that had enough software programmers to adjust all these computers so they would not go haywire, and do it at a reasonable price, was India and a few Southeast Asian countries (such as the Philippines). In retrospective, it was this huge operation that launched today's Indian lead in outsourcing industry. Wipro Technologies and Infosys Technologies come to mind immediately.

With a carbon tax or cap-and-trade legislation looming, more and more companies are expected to become carbon-neutral, or even carbon-negative. This is going to create the next big global business transformation, which is going to require tons of software, programming and back-room management to measure each company's carbon footprint—monitor the various emissions-reduction and offsetting measures—on an ongoing basis.

Now some experts are predicting that this will create an opportunity bigger than Y2K. Similar to Y2K, we shall call this the problem of E2n4E. "E2n4E" stands for "efficiency to and for the environment." This includes all the energy programming and monitoring that thousands of global companies are going to be undertaking in the early 21st century to either become carbon neutral or far more energy efficient than they are today.

In this globalized world, a low-cost brainpower will be the beneficiary of this new wave. But a notable distinction is that Y2K was a deadline imposed by the calendar, and therefore it had a huge ability to concentrate the mind. It became a drop-dead date for everyone. Making a company carbon neutral or negative is not a date; it is an inevitability.

Asia, given its track records with Y2K, is poised to get a lot of this work. To tackle E2n4E requires Einformatics.

2.1 E2n4E Winning Strategy—From IT to ET

When Y2K came along, some companies responded tactically, doing only the minimum reprogramming to keep their computers operational when the millennium rolled over. Others approached it more strategically. These latter companies went from seeing information technology (IT) as a cost to looking for ways to make money from it—through data mining and using better information to cross-sell products, reduce cycle times by introducing new services and to manage inventories more efficiently.[11]

Similarly, the key to winning E2n4E business will be showing big global companies, such as Google and Yahoo, how becoming more energy efficient, carbon neutral or carbon negative is not a new cost they have to assume to improve their brand or satisfy regulators, but the switch can actually be a strategic move that makes money and gives them an edge on the competition.

To that end, they will use Einformatics to reduce material costs, simplify logistics, drive down electricity charges and shorten supply chains. As they start to do this, it will require a lot of data management, which companies will want to do as cheaply as possible. It will be India and Southeast Asia again—and any other emerging low-cost brainpower nations—for backroom work.

IBM seems to be moving into this space, too. Big Blue knows that even if Indian companies do a lot of the backroom work, there will be lots of front-end jobs nearer the customers. A new, good and stable-growth career has emerged—green consultants, green designers, green builders, green this and green that—are all going to be in huge demand in the foreseeable future. And if these green (like new and inexperienced) green labor entrants can speak English, a little Hindi, or Chinese, they will likely climb the new corporate ladder faster.[12]

2.2 Writing on the Wall

As is now known, Y2K set India on the track for information technology (IT) outsourcing; the 2008 Summer Olympics might just be the E2n4E launch pad for the Chinese energy technology (ET) and Einformatics outsourcing, though the dragon is somewhat bruised by the May 2008 Wenchuan earthquake.

There are a few reasons why India has been so successful in IT. Cheaper labor aside, the Indians speak English, or Inglish. However you want to call it, they can read, write and converse understandable Inglish with no major problem. Thus they have a huge advantage over other Asian countries. Another reason is that India is 12 hours ahead of Silicon Valley, California. Thus, while the technocrats and entrepreneurs in Silicon Valley are asleep, it is workday in India. In this sense, the IT world never sleeps.

China and Southeast Asian countries also enjoy the advantage of cheap labor. They are 15 hours ahead of Silicon Valley, or 12 hours ahead of Wall Street. They thus enjoy the same "time" advantage as India. As for language, the Chinese is slightly behind. But because of the 2008 Summer Olympics, there is a craze to learn English (some native English speakers like to say Chinglish). The Chinese are picking up fast, and this language disadvantage is rapidly disappearing. Returning scholars and Chinese diaspora overseas are also helping.

English, Inglish, Chinglish, Singlish (for Singapore), no problem. Einformatics transcends accent-barrier. Asia and Southeast are poised to get the bulk of the E2n4E work, and become a hotbed of Einformatics. And if done properly, some day the world may be speaking *x*glish where "*x*" may be "In," "Chin," or "Sin", depending on which country eventually prevails as the dominant Einformatics player. This is the world of "plurality" and the "majority dictates the norm."

And best of all, Malaysia is Chinese, Indians, and Malaysians all in one. The country is strategically situated at the corridor of Asia, geologically located in a haven sheltered from earthquakes and typhoons, and politically very stable—a haven for Einformatics. The country is multilingual, speaking besides fluent English—that is, perfect English, Inglish, Chinglish, Singlish or Minglish—Bahasa Malaysia, Hindi and Chinese. It is a true haven for Einformatics.

References

This work is based on a forthcoming book:

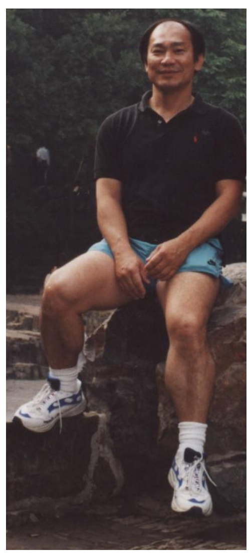
Hwa A. Lim, *Eco-friendily Yours: Energy, biofuels, green and clean technologies*, (World Scientific Publishing Co., Hackensack, New Jersey, USA, 2008).

And the following work:

1. Hwa A. Lim, "Bioinformatics and cheminformatics in the drug discovery cycle", In: Lecture Notes in Computer Science, Bioinformatics, R. Hofestaedt, T. Lengauer, M. Loffler and D. Schomburg (eds.), (Springer, Heidelberg, Germany, 1997), pp. 30–43.
2. Hwa A. Lim, *Genetically Yours: Bioinforming, biopharming and biofarming*, (World Scientific Publishing Co., New Jersey, USA, 2002), 417 pages.
3. Roger James, "The market size of bioinformatics", *Drugs and Markets Bulletin*, August 2001.
4. N Goodman, "Biological data becomes computer literate: new advances in bioinformatics", *Current Opinion in Biotechnology*, 13, 2002, pp. 68–71.
5. Ayon Roy, "Where chips and cells collide: bioinformatics", *SiliconIndia*, August 2002.
6. Anne Cordon, and Donna Messersmith, "Bioinformatics, virtual labs, and the human genome project", In: *Proceedings of the 23rd Workshop/Conference of the Association for Biology Laboratory Education*

(ABLE), M.A. O'Donnell, (ed.), (University of Toronto, Canada, 2002), 392 pages.
7. Shiva Kanaujia, "Bioinformatics and Internet: new paradigm to disciplines and information technology", *SRELS Journal of Information Management*, 41(1), March 2004, Paper E, pp. 43–55.
8. Corrado Priami, and Paola Quaglia, "Modelling the dynamics of biosystems", *Briefings in Bioinformatics*, 5(3), 2004, pp. 259–269.
9. Hwa A. Lim, "Bioinformatics, nanotechnology and SARS", *Symbiosis*, October 2003, pp. 26-29.
10. Hwa A. Lim, "The business of innovation", *Symbiosis*, April 2008, pp. 6–9.
11. Thomas L. Friedman, "The dawn of E2K in India", *The New York Times*, November 7, 2007.
12. Hwa A. Lim, "Technology adoption", *Symbiosis*, April 2008.

About The Author



Dr. Hwa A. Lim, on a trip in Asia. Though resident in the U.S., he can speak excellent *x*glish. (Source: HAL Archive).

Dr. Hwa A. Lim obtained his Ph.D. (science), M.A. (science), and MBA (strategy and business laws) from United States, his B.Sc. (Hons.), ARCS from Imperial College of Science, Technology & Medicine, University of London. He is sometimes also known as "The Father of Bioinformatics." Most of his early work on bioinformatics was performed at a U.S. Department of Energy (DoE) supported supercomputer institute, where he was program director, and tenured state-line faculty. Hal has served as a bioinformatics expert for the United Nations, a review panelist for the U.S. National Cancer Institute, and as an expert consultant for McKinsey, Prudential, VAXA, Eli Lilly and Company, Monsanto and Company, Dupont CR&D, and Robertson Stephens.

His career started with short stints at the Strong Memorial Hospital, New York, then at the Laboratory for Laser Energetics (LLE) at the University of Rochester, and later computational work using computers at the John von Newmann Center at Princeton University. In 1995, he advanced his career to California after having been at Florida State University as a tenured faculty and program director for eight years. In May 1996, he was on the BioMass Panel organized by the American Association for the Advancement of Science (AAAS) at Stanford University.

Hal currently resides in Santa Clara, "The Heart of Silicon Valley"^{SM}, California, USA, which has one of the highest concentrations of high-tech companies in the world.