OCTOBER 2019 VOL. 1 ISSUE 4

COMPUTATIONAL AND DATA SCIENCES

THE OFFICIAL NEWSLETTER OF IISc BANGALORE, CDS DEPARTMENT

cds.iisc.ac.in

The Department of Computational and Data Science presents the fourth issue of the first volume of the Newsletter series! Through this newsletter, we hope to give you a glimpse into our little world- what we do, where our interests lie and what we hope to achieve. The Department of Computational and Data Sciences (CDS) is an interdisciplinary engineering department covering the broad research areas of computational science and engineering, and scalable computer & data systems.

Our journey began in 1970, when the Computer Centre of the institute was established as a central computing facility. In 1990, it became Supercomputer Education and Research Centre (SERC), providing computing facility to the faculty and students of the Institute. A master's degree programme M.Tech in Computational Science, one of the first of its kind in India, was started by SERC in 1999. Therefore about 16 years later, CDS was formed in December 2015 from the academic wing of SERC.

From the Chair's Desk

It is our pleasure to share the activities of the department through this newsletter. The goal of this publication is to highlight our research activities, achievements of faculty, students and project staff and new initiatives.

The main motive of this news digest is to reach out to the world outside the campus including student , faculty, corporate and any individual who is interested in computational and data sciences. Ultimately, we hope to build a highly connected, a truly inter-disciplinary society and, of course, help us to network.

Faculty Interview

- with Dr. Jayant Haritsa

Dr. Jayant Haritsa
Senior Professor

Q1. Why did you choose DBMS as your field of research ?

My original interest was not in databases, but in performance modelling. I took several courses during my PhD covering both analytical and simulation-based modeling. My advisor was an expert in simulation and was working with database faculties and they realized that simulation is the only viable technique with which the deep-rooted dependencies within the database engine can be modeled. So my advisors applied this technique in a variety of database environments, and this is how I got into database systems. My thesis was an organic combination of performance modeling and database management, but over the last two decades I have moved on to hardcore database management problems such as query optimization and data-mining.

Q2. What are the research problems that you are working on and what are the challenges that are currently in your field ?

So for the past decade we are looking at how the internals of current database engines operate. In India most of the industries as well as academia are working at the application level where they treat the engine as a black box and are building their products around it. I usually explain it using the analogy of a driver vs a mechanical engineer. A driver has limited understanding of the working of the engine and is mostly concerned by what they can do given that engine. Whereas, we work like the mechanical engineers and are concerned by the performance of the engine itself. We developed a tool in 2005 called PICASSO and were able to show in visual format how currentdatabase engines are misbehaving and this fact was, surprisingly enough, largely unknown in the industry. We called the software PICASSO because the pictorial results resembled cubist paintings, an art genre co-founded by Pablo Picasso. This tool showed for the first time the brittle underside of database engines. I presented these results in Norway in front of both database vendors and users, and they were taken aback by the results. So, several of these vendors have used our tool and hired our lab's alumni in their companies to address these issues. Moreover, we were also able to develop the tools to fix these problems and our work was widely appreciated throughout the industry. Our software PICASSO also got the best software award in the 2010 edition of VLDB, a very prestigious conference in our field.

We also demonstrate that database engines are not robust and a small change in parameter values can make the performance swing widely. This is an inherent brittleness and people were trying to work around this brittleness for a long time with little success. We showed that these approaches won't work and designed new execution strategies that for the first time give provable performance guarantees to the user.

Q3. You did your Masters and PhD from Wisconsin Madison and did your post doc from University of Maryland. What motivated you to return back to India and join IISc ?

Many people, including my advisor advocated me against taking this step since they thought that my global visibility would disappear after returning to India. However some of my peers and I decided to take it as a challenge to come back to India and still be able to publish high quality papers. The system in the USA was like a well oiled machine whereas ours wasn't so smooth in early 90s, which is even before the software boom in India. So we decided to take the challenge and make the research environment more conducive in India. This gamble did pay off when we published two papers in 1996 in the Sigmod conference, these were among the first papers from India to get published in that conference. The program committee was surprised by the 560012 pincode! The feeling that we were able to break some barriers that existed in those times and were able to create these publications is pretty special to me even today

Q4. Your career in academia has spanned three decades. In this time how have you seen the culture of the academia and academia as a whole to evolve ?

A troubling trend in academia that has sprouted is the obsession with number of publications. If you look at earlier times, people like Michael Stonebraker, who later went on to win the prestigious Tuning Award, got a faculty job in UC Berkeley with no papers because people recognized his work. Today, people are obsessed with numbers like h-index, number of citations, rankings of institutions etc. This has transformed education from something which brings a positive transformation to the society to a sport. What I see many researchers doing is what I can term as scalar research, where you do incremental addition to the field. What people should aim is to be a vector and change the direction of the field. In my work I consciously aim not to do scalar research. People should choose a challenging problem and enjoy the process of solving that problem. I believe we should look up to Venki Ramakrishnan as the ideal scientist from whom we can learn these positive traits. I think that a student should be judged not by his resume alone but the overall transformation i.e from where he started and where he is now. There cannot be a quantitative assessment of academic performance that can do justice and we should judge the person qualitatively. We should compare the output relative to input, that is what we should measure.

Q5. What are your interests outside academics ?

I have an abiding interest in Carnatic classical music and it is usually playing in the background even when I am at work.

Q6. What books have you cherished reading the most ?

I have especially liked "The Name Of The Rose" and “Foucault's Pendulum” by Umberto Eco, a famous Italian writer with fine sensibilities. My childhood favourite is "Les Misérables" by Victor Hugo because of its broad social canvas. Apart from them I like books of Sir Arthur Conan Doyle, especially his historical romances about "Sir Nigel".

Q7. What is your advice for young researchers and students who are at the dawn of their career ?

Treat it like a test match, it takes patience and can be frustrating at times. But the key is that can you engage the problem in a creative manner? Can we look at a problem with a different perspective and come up with a novel solution ? In this matter as well, I have a lot of admiration for Venki Ramakrishnan. My advice for research students is that don't worry about the people around you, researchers should be goats, not sheep

The crucial thing to analyse is "does this problem interest me, do I think about it at night ?". Only when you ponder over it again and again, you will have that moment of epiphany. Moreover, if you don't like the problem, change it, otherwise you cannot come up with novel ideas.

Visiting Faculty Interview

- with Dr. Ananth Grama

Dr. Ananth Grama
Professor

Q1. Can you share your background with us ?

I was born in beautiful Bangalore many moons back, at St. Martha's Hospital! I spent my formative years (ages 0 - 8) in Pilani, Rajasthan, and then Roorkee, UP (ages 8 - 21). I have spent my entire life on college campuses (my father was a professor as well as albeit of Civil Engineering). I got my undergraduate degree from the University of Roorkee (now called IIT-Roorkee), my MS from Wayne State University in Michigan, and my Ph.D. from the University of Minnesota and all in Computer Science/ Engineering. Upon completion of my Ph.D., I moved to Purdue in 1996 and have been here since. I was an Assistant Professor from 1996 - 2001, Associate Professor from 2001-06, Professor from 2006 onwards, and the Samuel Conte Professor from 2017 onwards. A fun factoid about Prof. Samuel Conte - I learned my Numerical Analysis in my undergraduate years at Roorkee from the book by Conte and deBoor (around 1987), got to meet him at Purdue around 1996-97, and have the good fortune of carrying the Chair in his name now.

Q2. Tell us about your lab and your research field ?

My lab works in broad areas of parallel and distributed computations and their applications in scientific and engineering domains. The core expertise in our lab is model, method, and software development. In these investigations, we work closely with scientists to deliver complete software solutions. As an example, in one project, we developed an entire software suite called PuReMD (Purdue Reactive Molecular Dynamics), working closely with Prof. Bill Goddard's group at CalTech. Goddard lab developed the ReaxFF force field and my lab developed the algorithms, solvers, and parallel software for ReaxFF.

In yet another project, we work closely with Prof. Shankar Subramaniam's lab at University of California, San Diego, to develop models, methods, and statistical techniques for solving problems in Systems Biology. More recently, we have been working on problems in analyzing single cell RNASeq data to identify cell types, their markers, underlying regulatory mechanisms, and associated drug targets. In yet another project, we work with Prof. Doug LaCount's lab at Purdue to analyze interactions of viral proteins (in particular, Ebola) with human proteins to identify pathways of infection and new targets for drugs. Beyond these, my lab is also a part of the Center for Science of Information - a Science and Technology Center of the National Science Foundation. The mission of this Center is to develop new models, methods, and applications of the Science of Information, incorporating elements of space, time, semantics, resource constraints, and other elements into traditional metrics of information. Finally, my lab also focuses on more traditional problems in Computer Science - investigating problems of resource allocation and fault tolerance in distributed/ parallel systems, novel solvers for optimization problems in machine learning, and use of machine learning techniques in verification of autonomous systems.

Q3. What motivated you to choose this field?

It was not a deliberate decision at any point of my life - merely a sequence of local decisions that led me here. I learned to program relatively young in life (on punch cards, no less), joined the Computer Science program for my undergraduate, since it was most in demand, and simply made all of the obvious decisions that led me to this point in my career. As to what motivates the kind of problems I work on, I have always been intrigued by things I do not understand and thus my very broad intellectual interests.

Q4. How has your experience been in Purdue? How is it different from academic environment in India ?

I have never been a faculty at an institution in India (before my Visiting Appointment at IISc), so I cannot speak with any authority on academic environment in India.

At Purdue, faculty experience complete freedom in how they conduct research and what they work on. It is not abnormal for a faculty in Computer Science to be working on identifying drug targets for Ebola!

The workplace hierarchy is very flat - I do not have a boss, nor do I get to boss anyone else. That is not to say there is no pressure to perform and excel. We are constantly under pressure (largely self-imposed) to publish and get grants. These are indeed highly competitive - requiring one to be at the top of their game/ field and to constantly innovate and push. Overall, it is an immensely rewarding experience - I would do it just the same, if I had a do-over.

Q5. Why did you decide to join CDS as a visiting faculty ? What courses will you be teaching ?

I used to visit IISc quite often, and have been tremendously impressed by the depth and stature of faculty at IISc. I have also had the good fortune of sitting through multiple graduate student symposia at IISc - the quality of work and students at IISc is comparable to the best, anywhere around the world. Given my frequent visits, interactions, and deep respect for faculty and students, it was only natural that I seek the honor of a visiting faculty position at IISc.

Q6. What are your interests outside academia?

Sports, politics, traveling, adventure. I have played tennis and ping pong for over 40 years, and take every opportunity to play (although I don't get very many these days). I also ski and golf as often as I can (which is not too often).

My work does afford me the luxury of traveling to distant places - which I enjoy greatly. I also enjoy long drives (in this context, for instance, a drive from my house to New York is about 1000 miles).

I follow markets, the economy, and politics world-wide, as well as a wide variety of sports (American Football, Formula 1).

Q7. What is your advice for students starting to pursue research at CDS?

Think big, try to solve important problems, do not fall into the 10% trap (improving someone else's result by 10%). We are fortunate to work in an area (computing) that is a critical tool for just about every major scientific, social, and engineering investigation. Strive to make impact, do things no one else has done before.

Alumni Interview

- with Dr. Jaya Prakash (currently Assistant Professor IAP, IISc)

Q1. Can you share the background of your work in MIG with us?

My work was on developing signal processing based methods in the context of medical imaging/clinical applications with an intent to reduce burden on radiologists and surgeons. Photoacoustic imaging (PAI) is nowadays considered as a big boon to radiologists, as it could provide molecular contrast (of important biomolecules like lipids, proteins, hemoglobin, water, etc) at acoustic resolution non-invasively. Providing molecular information regarding lipids, hemoglobin, and proteins has direct impact in cancer, cardiovascular, and neuroimaging applications. PAI involves shining nanosecond pulsed laser light on biological samples, these biomolecules will then absorb the shined light causing a small rise in temperature and emits sound waves due to thermoelastic expansion property of the tissue.The emitted sound wave is detected in a tomographic fashion using ultrasound transducers and is used to reconstruct the exact location and concentration of these biomolecules in the biological medium. PAI problem involves inverting a wave equation which becomes a highly ill-posed problem in limited data situation/noisy environment, requiring inclusion of regularization term to obtain stable solution. Limited-data cases often arise due to physical limitations, faster acquisition, and limited number of transducer elements. Choice of the regularization parameter affects the quality of the reconstructed images, therefore my work involved developing a computational feasible method to automatically choose the regularization, thereby enabling radiologists to get an unbiased image faster, which they could analyze for diagnosis.

More importantly, the radiologist can now stop worrying about tuning the reconstruction parameter and focus on the provided reconstruction to confirm biological/clinical hypothesis.

Secondly, I was also involved in a project to enable real-time non-invasive therapy using high-intensity focused ultrasound (HIFU).HIFU has been used clinically for kidney stones treatment, breast cancer treatments and surgeons are trying to explore HIFU for other therapy protocols. HIFU involves focusing low-frequency ultrasonics bursts on cancerous tumors, this focused ultrasound will heat the cancerous region causing tumor control/shrinkage. The process of HIFU involves increasing the temperature in the biological medium, hence it is imperative for us to accurately recover the temperature maps over larger three-dimensional volumes to avoid killing of normal cells and focusing the sound energy onto cancerous tumors. To obtain temperature maps, people have used Magnetic Resonance Imaging, together with HIFU, popularly known as MRgHIFU.MRI is typically a slow imaging modality due to point by point acquisition, hence performing accurate reconstruction with highly under-sampled data is desired. To this end, I was involved in developing a compressive sensing based approach for accurately reconstructing the temperature maps in MRgHIFU with 16x undersampling. Further these compressive sensing algorithms were also parallelized on graphics processing unit (GPU) environment to enable real-time temperature imaging, which is much needed in the context of MRgHIFU for surgeons.

Q2. You have done your Masters and Phd from IISc, have done more than 3 years of Post Doc and are now a faculty in IISc. How different has your experience been in these different roles ?

Dr. Jaya Prakash
Assistant Professor

Starting from master's all the way to my current role, I have always been learning. The excitement to learn and acquire knowledge/skill is what keeps me motivated. During my master's and PhD at IISc, I was exposed to the most fundamental aspect of today's world, i.e. computing. I enjoyed attending courses at IISc, starting from Linear Algebra, Numerical Methods to High Performance Computing, and Digital Image Processing. The best part of the institute is in allowing students to take courses in any department. These subjects were foundational and enabled me to venture into topics of my interest. Even though the coursework is rigorous, all the faculties at the institute were accommodative and are willing to help you. My master's and PhD thesis topics groomed me to use these optimization schemes in real-world situations; obviously the most satisfactory part was that these real-world situations were associated with improving human health.

After my IISc stint, I moved to take up a post-doctoral research position in Germany. During this period, I started to learn various aspects of experimental research, interacting with PhD students from different parts of the world and mentoring them was truly a memorable/learning experience.

My post-doctoral research involved interactions with biologists and clinicians on a daily basis, these interactions made me realize the importance of cross-communication across disciplines and importance of experimental research. During my tenure at Germany, I was involved in developing a new kind of spectrometer, a new kind of bio-compatible photoacoustic probe, different biological imaging applications like non-small cell lung cancer (NSCLC) model, atherosclerotic model, and asthma model alongside data analytics. Further, I started venturing into the area of theranostics (therapy+diagnostics) with light. I was actively working on performing photothermal and photodynamic therapy alongside multispectral photoacoustic imaging. All these experimental exposure alongside the computational/theoretical rigour of IISc, enabled me to setup a multidisciplinary lab focusing on spectroscopy, imaging and theranostics at IISc.

As a faculty of IISc, I have started teaching Biomedical Optics and Spectroscopy course, obviously I am enjoying teaching my first course which overlaps a lot with my research work. Apart from research and teaching, my current role requires myself to be a good manager on both the scientific and administrative front, which is again a good learning experience. Apart from these learning, the institute provides utmost freedom to perform your academic activity, which is the best part of my job. My role has now become to guide students and solve important problems in the healthcare sector. I am already seeing loads of challenges in achieving these goals, however these hurdles will seamlessly disappear when you realise that your innovative work will help someone. Last but not the least, during this journey I have met loads of wonderful people and obviously my time was well spent only because of having these nice people beside me both inside and outside the lab

Q3. What motivated you to join academia?

Academic job involves performing teaching, and research related activities. Academia is a place which allows and encourages everyone to study and think freely. This aspect of thinking freely (without any strings attached) is what motivated me to pursue an academic career.

In my opinion, teaching helps in improving one's institution and understanding of the subject, which can directly help in pursuing research ideas in a thought provoking manner. Teaching involves sharing acquired knowledge during research and inspire students to become academics/entrepreneurs, hence research and teaching are interwound (one cannot exist without the other). More importantly, an academic is given a wonderful opportunity to work with people who are very young, training/exposing them to the cutting edge research activity is obviously joyful. Motivating these young minds to become torch-bearers of your area of research is what pushed me to take up an academic career. I would like to quote Nelson Mandela 'Education is the most powerful weapon which you can use to change the world', these words have always influenced me in realizing the importance of educating students and help in nation building.

Secondly, my role as an assistant professor requires me to wear different hats like a teacher, researcher, manager, promoter, volunteer etc. Each of these roles require different kind of skill set, hence becoming academic will develop once overall personality. Further, as an academic, I can always update myself based on the current trends (both scientific and market), and this never-ending learning process is what I was looking for. Most university campuses are pleasant as you find almost everything within the campus and this helps in keeping the motivation to get off the bed everyday to reach your office. Last but not the least, you get summer vacations in universities along with freedom to schedule your time as per your priority, which is very rare to find in any other role.

Q4. What will be your research area now ?

I have started the Frontiers in Imaging, Spectroscopy, and Theranostics (FIST) Lab at the Department of Instrumentation and Applied Physics (IAP) in IISc. FIST lab will focus on developing novel systems for improved sensing, diagnosis, and therapeutic approaches, targeted towards biological and medical theranostics. FIST lab aims to achieve non-invasive investigation of physiological and molecular processes in tissues at unprecedented depth to resolution ratio. Specific focus will be on developing multispectral photoacoustic and fluorescence imaging systems to probe biological media. Low-cost photoacoustic systems will be realized using frequency-domain approach. Development of frequency domain photoacoustic system will involve innovative projects on theoretical, computational and experimental fronts. These systems will be developed to understand life-threatening diseases like atherosclerosis, and cancer in mice models. On the fundamental aspect, light scattering happens in many real-life situations, like while driving a foggy day, taking a pic during foggy morning, or shining life on tissues. Light scattering is one of the limiting factors for imaging biological tissue with light. Light scattering is the main reason for limited resolution and limited penetration depth. On this front, my lab would focus on developing imaging methods that could overcome the light scattering issues while probing biological media.

Q5. What advice would you give to students starting their research at CDS? Please answer this only as an alumni and senior to current students.

My advice is very simple, just try to be a good student (always; this role should never change in life) and give it your fullest on whatever endeavor you are trying to achieve.

Once you have a deep belief that you want to achieve something, this deep belief along with hardwork and dedication will realize your goals. Just remember, the world is not a fair place, hence don't bother about anything/anyone, my advice is to set your priorities and work according to your responsible priority without anyone's intervention in your life. Most important, particularly for research students, you might not get the results you are expecting, just don't hesitate to ask for help and persevere until you solve the problem. This attitude towards life will take you very far, and keep your surroundings happy.

Last but not the least, life is really small; just enjoy every moment with your friends and family. IISc is the best place to find great friends, and is a pleasant campus. Just enjoy every moment in IISc (be in courses, hostels, coffee time, movie time, gymkhana, birthday celebrations, etc), you will realize the beauty of the institute only when you leave the institute. If possible, try helping others and set your work goals to improve the standard of living of humankind.

Good Luck Juniors!!!

Science in the Golden Shackles of its Imaginary Impacts

by Prof. Murugesan Venkatapathi

(originally published in June, 2014)

The impact of science on our daily lives is ubiquitous, but tracing bits of this impact to individual works of science with reasonable certainty is becoming impossible in most cases. Nevertheless, considering that we spend a notable part of our financial commons (or GDP) on scientific endeavors, such a microscopic estimate of the impact by each scientific work seems unavoidable. Even in the case of a work in fundamental sciences that is far away from any immediate use, an estimate of its impact on our knowledge is quite pertinent. As a first approximation, measuring impacts of science have been relegated to the quantification of citation-impact by a scientific work. Such measures assume that the citations received by a scientific work are unbiased pointers to its real impact. Further the citation-impact of works are cumulatively considered for estimating impacts of larger aggregates such as journals, scientific institutions and the careers of individual scientists. Moreover, monetizing both scientific discoveries (by patents) and the access to scientific publications exacerbates this necessity of micro-estimating scientific impacts. The argument that technology is a great end-use of science but not the primary motivation is unappealing even to most scientists today. Hence, use of citation-impacts to justify the quantum of funding to scientific organizations and in the evaluation of scientific competence of countries to individuals is the order of the day. This has begun a debate on the pitfalls of such conclusions using an emphasis on citation-impacts [1-3].

Let us start with the broad agreements among the scientific community on this issue:

1. Measuring impacts is necessary
2. Citations earned by a scientific work have a positive correlation with its actual technological and scientific impacts
3. Citation based indicators are far from perfect primarily due to uncertainty in the relationship between real and citation impacts (notwithstanding any advanced processing of citation data).

The strong disagreements arise from the effects of (3), especially its effect on the way we do science in the long term [1-5]. While many believe that despite its limitations the current strong emphasis on such indicators has been fruitful, many others argue that its premature use in decision-making severely stifles science due to fundamental deficiencies of the citation system and its indicators. In this article, I point to the large lacunae in our rudimentary citation system that makes quantification of real impacts unreliable except in restricted cases. These points are valid irrespective of the statistical metrics used in processing the citation data. Next I point to specific practices in the current system of scientific publication that can multiply these negative consequences into a vicious runaway cycle in the long term. This discussion offers suggestions that can make the measurement of real impacts more accurate and also help increase the signal-to-noise ratio of scientific publications. I argue that these improvements in the systems of citation and publishing are vital, and should receive strong support of the scientific community irrespective of which side of the above argument one submits to.

Does every citation indicate an identical impact? Does this fallacy result in a folly?

When one attempts to derive metrics for scientific impacts from citations, the following issues should be pondered.

A] Grades of citation: A citation earned by a scientific work indicates any of the 3 kinds of contributions to the citing publication. The first more notable kind is a contribution to the methods used in the citing work; the second is a relevant work with comparable/contradictory results; and the third is a related work used to highlight either the historical antecedent or the contemporary significance of citing work. The first kind is enumerated in the methods and introductory sections of a manuscript, whereas the second type is typically found in the introduction and results/discussion. The third kind is limited to the introductory section of a manuscript. It is thus natural to require that citations are distinguished based on their graded relevance to work as the difference in the real impacts to the citing work may be separable by orders of magnitude. On an average less than 20% of the references of a typical manuscript are unique indispensable citations, more so in the applied areas of science.

B] Methods matter: Also to be noted is that the current practices of highly visible journals (as described in the next section) explicitly discourage a detailed description/verification of methods, to be replaced by longer introductory sections and more plots of the results. The questionable justification is that today many of the methods are eventually repeated in the prolific publications of increments, and also, they do not appeal to a wider readership. The above factors introduce a large bias against manuscripts describing new essential analytical/experimental methods that are fundamental and general.

C] Citations can be inherited: Even before the era of search engines, it was showed that indicators like citations have had the characteristics of a greedy propagator (i.e.) the effect of rich getting richer [6], making advisors/co-authors at graduate school a significant causal factor in the citation-impacts of later works of a scientist. This effect has increased subsequently with the internet age and also introduces bias against a scientist working in multiple scientific areas; while largely favoring incremental publishing on a problem to saturation as it can garner higher hits in a search engine.

D] More the authors more the merrier: One of the most glaring faults in the current indicators is that total citation-impacts earned by a publication are not shared by authors, but instead is duplicated to each of them (i.e.) the sum citation-impact attributed to authors is not conserved by the citation-impact of the publication!

E] Quality of citations: Recently, there has been an effort to include the apparent quality of a citing publication in the determining the impact of a work. In principle, this can be done using the citation data provided the pitfalls A, C and D are sufficiently addressed. If these are allowed to linger, impact indicators based on advanced data processing techniques can only enlarge those lacunae.

Conflicts of interest: Science Vs the Journal

Monetizing the scientific publications has resulted in a necessity to make journals highly visible. It is in the interest of scientific community that parochial interests do not trump the larger interests of science. Unfortunately, a high standard of science does not necessarily have a notable correlation with a wide readership (that is needed for high visibility and citation impacts). Large increases in doctoral students and the number of publications along with this need for journals to be distinguished have severely stressed the peer-review process. Introduction of full-time editorial staff to screen manuscripts before peer-review is a result of this need. A non-practicing scientist is employed in a journal (for decades together resulting in entrenched interests); primarily to screen submitted manuscripts for maximizing the future citation-impact of the journal. Naturally, they are well trained to distinguish the apparently good manuscripts from the average ones, but more importantly they mimic the non-expert wide readership they seek for the journal. Typically each one of them is expected to peruse and make decisions on a few thousand manuscripts in a year. Such decisions are not scientifically justified but more importantly, it ensures that scientific merit plays a minor role in comparison to the significance perceived by a non-expert [7].

It is a system that is designed to publish manuscripts that are appealing to even the people who may not understand the contents of the manuscript sufficiently. Based on a superficial understanding, a vicious cycle of inflation in publications on any subject along with its citation-impacts can result, and this seems to satisfy the false premise of an increasing quality and quantity of science. There is also an explosion of literary/algebraic embellishments in publications appealing to such editorial staff and the larger readership, naturally at the cost of our understanding in the science. In many cases, peer-reviews in these journals have been relegated to the opinions of the peers on the appropriateness of a manuscript to the journal; many a times shifting the focus unscientifically from ‘what is being said’ to ‘who is saying it’.

Above all, the above practice and negative consequences have been justified based on the imaginary impacts enumerated by the citations accrued to journals. But the actual signal-to-noise ratios in science may have drastically fallen. Leaving aside this opinion on the difference between real and citation impacts, one should at least take note of the ability of highly visible journals to accommodate the most cited publications [3]. The correlation of the most highly cited papers to the highly cited journals (in all areas) was moderate before 1960 and did climb until the dawn of the internet age (~ 1990). Subsequently, this has taken a sharp downward trend recently (~2002) clearly showing that the publication practices to ensure high visibility run counter to accommodating the most excellent scientific works of our times. Systems optimized for high throughput and higher averages naturally have trouble in accommodating the most original works. Also too much specialization of journals is counterproductive as well; where duplication of scientific knowledge and vocabulary slows down the actual scientific progress despite an increase of citations.

Finally, a specific example of the uncoupling of citation-impacts and real impacts is tempting here. The seminal paper of Pines and Bohm [8] on collective excitations of free electrons in a metal (called plasmons today) has earned ~ 700 citations in sixty years. Not surprisingly, even invited opinions on the use of plasmonics that were published in highly visible journals have attracted more than 5000 citations in just the last decade; which in naiveté would signal an impact almost hundred times stronger. The remedies for the large lacunae in journal publishing practices are mostly well-known. An effort to limit unbridled monetizing of the access to scientific publications has already begun. This should be followed by a double blind peer-review process that puts emphasis on the scientific rigor and simplicity of a solution to the problem as this has become a dire need.

References:

1. www.ascb.org/SFdeclaration.html.
2. Luís A. Nunes Amaral, “Measuring Impact: Scientists must find a way to estimate the seemingly immeasurable impact of their research efforts,” The Scientist (Opinion), February 24, 2014.
3. George A. Lozano, Vincent Larivière and Yves Gingras, “The weakening relationship between the Impact Factor and papers’ citations in the digital age,” Journal of the American Society for Information Science and Technology 63, 2140–2145 (2012).
4. Richard Naftalin, “Rethinking Scientific Evaluation: Asymmetry in the Research Excellence Framework in the U.K. is a threat to basic medical sciences within British medical schools”, The Scientist (Opinion), July 16, 2013.
5. Orion Penner, Raj K. Pan, Alexander M. Petersen, Kimmo Kaski, and Santo Fortunato, “On the Predictability of Future Impact in Science”, Scientific Reports 3, 3052 (2013).
6. Matthew J. Salganik, Peter Sheridan Dodds, Duncan J. Watts, “Experimental Study of Inequality and unpredictability in an artificial cultural market,” Science 311, 854-856 (2006).
7. Steen RG, Casadevall A, Fang FC, “Why Has the Number of Scientific Retractions Increased?,” PLoS ONE 8(7): e68397. doi:10.1371/journal.pone.0068397 (2013).
8. D. Pines and D. Bohm, “A collective description of electron interactions:I and II” Physical Review 82, 625-634 (1951); Physical Review 85, 338-353 (1952).
9. Douglas N. Arnold and Kristine K. Fowler, “Nefarious Numbers,” arXiv:1010.0278 (2010).

Faculty Startup

From labs and tweed jackets with elbow patches to board rooms and black berry suits, here are some of our faculties who bridged big data with big businesses..

ರೆನೆಲೈಫ್ ಪ್ರೈವೇಟ್ ಲಿಮಿಟೆಡ್
ReneLife Private Limited
Accelerating Life Sciences

ReneLife is a technology startup, launched as a faculty enterprise of the Indian Institute of Science, in the area of Life Sciences. It was co-founded by Professor Debnath Pal and Professor S. K. Nandy. ReneLife is devoted to providing cutting-edge hardware-software co-designed accelerated solutions.

It has made a start by developing ReneGENE - a proprietary core technology, for accelerated accurate alignment of short reads from Next Generation Sequencing (NGS) platforms through massively parallel and scalable realizations.

Professor Pal holds a position of full Professor and heads the Biomolecular Computation Lab in our department. His research interests are in the areas of Computational Biology, Bioinformatics, Omics and Systems Biology. He is noted for his contributions in the area of protein structure, function, interaction and dynamics. He is an INSA Young Scientist and NASI-SCOPUS Young Scientist awardee, and a member of several professional and academic bodies. He has over 70 journal publications and a patent to his credit.

Interview with Professor Pal

Tell us something about your startup.

It is a Life Science startup aiming to bring in new software and hardware-software co-designed technologies in precision medicine. The aim is to improve cost, time, and quality of analysis that feeds into precision medicine pipeline. Since a large part of the difficulties pertain to big biology data in genomics, our current focus is on genome analysis, where our products have undergone several rounds of trials and we hope to launch the same within a year. Our long-term goal is to create a product portfolio that will traverse all aspect of analysis contributing to improve precision medicine.

What makes your idea unique in Indian /global environment (your USP)?

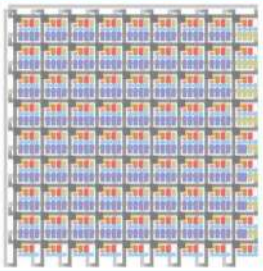
Our basic motto is to improve precision in precision medicine. Current big data challenges impel applications to make compromises to save time and cost that affect the final outcomes. We intend to avoid that by bringing hardware-software co-designed application incorporating accelerator technology (GPGPU, FPGA, ASIC).

What motivated you to start your own company?

As basic science experts we do a lot of research to develop computational methods/algorithms that improve biological analysis. But at some point you realize that several of these have translational potential and are commercially exploitable. Luckily, Prof. Nandy (expert in hardware acceleration and our startup co-founder) inspired an understanding that many of the big data impediments in biology are solvable through accelerator technology. This is expected to bring down time, and cost without compromise on quality. This was our main premise to take a startup plunge.

What advice would you give to young aspiring students, who want to open their startup?

Technology startup is a challenging effort and needs grit and determination since there is also competition in the market. It does not matter whatever technology you have developed in the lab, to bring it to the market it must comply with the expected standards which are basically defined by your competitors. Therefore, for commercial viability you may need to do many more tasks that you may have originally thought as not necessary. Patience and perseverance are therefore key to success and bring investments that takes your venture to the next level.



Morphing Machines
REDEFINE – Reconfigurable Silicon Cores

Morphing Machines Pvt Ltd is a closely held Indian fabless semiconductor IP products and solutions company launched from the Technology Entrepreneurship initiative of the Indian Institute of Science at Bangalore. Morphing Machines is focused on creating semiconductor IPs centered around its path-breaking reconfigurable silicon core technology and massively parallel processor platform named REDEFINE (TM).

A variety of high-performance silicon solutions for various application domains have been developed at Morphing Machines in the areas of cryptography, data security, avionics, high-definition image and video processing, numerical computation, neural networks and deep learning, and other areas. It was founded in 2006 by Professor S. K. Nandy and a group of distinguished IIT and IISc alumni with decades of rich research and industry experience at leading global computer technology corporations. In 2011, Morphing Machines featured as one of the four global semiconductor start-ups in the Cool Vendors 2011 report of Gartner Research.

Professor Nandy holds a position of full Professor and heads Computer Aided Design Lab in CDS. His research interests are in areas of High Performance Embedded Systems on a Chip, VLSI architectures for Reconfigurable Systems on Chip, and Architectures and Compiling Techniques for Heterogeneous Many Core Systems. He has over 170 publications in International Journals, and Proceedings of International Conferences, and 5 patents. He is a valued IEEE member for 22 years.

Kenome

Kenome was founded in late 2017 by Professor Partha Talukdar. Professor Talukdar is a faculty member in the Department of Computational and Data Sciences at the Indian Institute of Science (IISc), Bangalore (India), where he is also the founding director of the Machine and Language Learning (MALL) Lab. Professor Talukdar is one of the architects of the NELL system at CMU, USA.

The name Kenome is a play on the phrase knowledge genome. Just as genome encodes the genetic material in an organism, Kenome's AI aims to capture and unify the knowledge in an enterprise and harness it for improved decision making.

Interview with Professor Talukdar

Tell us something about your startup.

Businesses as part of their day to day work generate lots of unstructured data, which is data that is written in natural language, like the presentations, manuals, communication among employees, communication with vendors, communication with customers, maintenance and research notes. All of it runs over natural language data. Structured data on the other hand is like a relational database, like sales data. While most of the decision-making by businesses is done based on the structured data, we believe there is lots of value to unlock from this unstructured data. That is what Kenome's mission is - to help enterprises make sense of this heterogeneous data, part of which is structured, but a large volume of it is unstructured. According to various estimates, 80-90% data in an enterprise is actually unstructured.

Our secret sauce in doing that is through Knowledge Graphs, which is also the area of research for MALL lab. Companies like Google have been using Knowledge Graphs doing crude web search, Amazon has been building a Product Knowledge Graph, but the exploration has been limited to these frontier companies, who are at the forefront of AI. Even worldwide, these kinds of explorations are in their nascent stage, and we want to bring more focus on it. So, through Kenome, we are looking to democratize those benefits, and bring knowledge about how to handle unstructured data to enterprizes.

What advice would you give to young aspiring students, who want to open their startup?

First of all, I think that is in short supply. I feel like many students are not thinking in terms of startups, because getting the comfort of a well paying job in the security of a big company is a thing that most students aspire to. But I feel like there are alot more options beyond that. Kenome has been around for slightly over a year and a half, and one of the most satisfying parts of this journey so far has been seeing the ideas from research solve actual business problems, which make other people's lives better. Rather than just following directions, which has its own value, you should also seriously consider doing things on your own, because jobs will always be there. If you are graduating from IISc and getting a job today, you are going to get a job two years down the line too. And I think your value should actually increase in these two-three years.

One of the concerns that I hear from companies is that there is a big disconnect between the expertise and products. Students are not thinking in terms of product, while they may be competent in terms of the methodologies and background. While that's also super important, if you start your own company, you'll bridge that part. You'll learn how to translate all of these research ideas to actual products, what works and what doesn't work. So I'll say you'll be even more attractive to companies now that you have filled that missing part.

Student Corner

There are few things one never misses about college life - like mess food and living on a shoe string budget in cubby holes . Let's ask some people who left all the riches to come back to this spartan lifestyle.

APOORV UMANG SAXENA

I left my SWE job at Google in 2017. I wasn't getting the job satisfaction that I needed as a part of the GSuite team - my interests in machine learning were not being applied at work. In order to get a more fulfilling role (ie researcher, prof) either at a tech company or in academia I needed to pursue a higher degree. Thus I left my job and applied for MTech Research at CDS. Fortunately I got admission in the direct PhD program. It's been 1 year now, and I am 100% sure that I took the right decision.

SRAVANTI ADDEPALLI

I have been interested in research from my undergraduate days, and specifically, research at IISc has always been a dream. After my bachelors, I started my career at an R&D company in VLSI, where I got ample opportunities to explore various interesting areas, and hence I worked there for about a decade. My recent interest in Machine Learning and my professors at an ML school that I attended motivated me to revive my thoughts on pursuing a PhD, and that is what brings me back to school after so long! To all those starting off, I would say the earlier the better, as long as you have clarity on what you want to do. And to those who are working... well, it is never too late!

CHANDRASHEKAR M.A

I did my M.tech from NITK in 2010. Since I had an inherent interest to pursue teaching career, I joined an engineering college as lecturer. To fulfil my yearning to understand Indian philosophies, I joined an year long course to study books like Bhagavad-Gita. My desire to explore teaching field took me to train myself in high-school teaching and teach students in mathematics for a couple of years. As I taught and thereby learnt maths from various perspectives, my inner interest to dive deeper into study and research in math and computation surfaced and grew stronger, which brought met to CDS.

Tidbits for a Healthy Life!

Cell Phone Radiation

How cell phone radiation affects life around us?

First we start with what is radiation, the word 'radiation' sounds scary, but in reality it includes any type of electromagnetic wave. The electromagnetic spectrum is broken up into two parts based on whether small doses of radiation can cause harm: ionizing radiation and non-ionizing radiation. The ionizing radiation is harmful but in reality the mobile phone radiation is non-ionizing. So let's see what are the other sources of non-ionizing radiation. The non-ionizing radiation, which encompasses the vast majority of light we are exposed to: visible light from lightbulbs, infrared light from an oven and from people, gigahertz light from our wifi, megahertz light to/from our cell phones, and radio waves hitting our car radio. It has been said that non-ionizing radiation is not harmful in small doses because one photon does not have enough energy to ionize atoms and/or break apart molecules, but in very large doses, non-ionizing radiation can be harmful.

What can cell phone users do to reduce their exposure to cell phone radiation?

We can reserve the use of cell phones for shorter conversations or for the times when landline is not available, but if cell phone plays an important part in your life and your most of the time spent on the phone, use a device with hands-free technology, such as wired headsets, which places more distance between the phone and the head of the user. It has been found in the studies the exposures decline dramatically when cell phones are used hands-free.

At last, the future is going to be safer because 5G is on the verge of making mark. Currently cell phones use electromagnetic radiation in the microwave range, but next generation 5G cellular networks will use higher frequencies in or near the millimeter wave band, 24 to 52 GHz. That is safer and possibly less carcinogenic.

How to get a Sound Sleep

DO's

- Stick to a sleep schedule of the same bedtime and wake up time, even on the weekends.
- Use bright light to help manage your circadian rhythms.
- If you have trouble sleeping, avoid naps, especially in the afternoon.
- Exercise daily, Practice a relaxing bedtime ritual, Evaluate your room, and Sleep on a comfortable mattress and pillows.
- Avoid alcohol, cigarettes, and heavy meals in the evening.
- Wind down. Your body needs time to shift into sleep mode, so spend the last hour before bed doing a calming activity such as reading.

DON'Ts

- Don't use an electronic device such as a laptop or cell phone as it can make it hard to fall asleep, because the particular type of light emanating from the screens of these devices is activating to the brain.

Fact that is not so FUN

Just because you're not using your cell phone before bed doesn't mean that it can't harm your sleep: Keeping a mobile within reach can still disturb slumber, thanks to the chimes of late night texts, emails, calls, or calendar reminders. About 72 percent of children ages 6 to 17 sleep with at least one electronic device in their bedroom, which leads to getting less sleep on school nights compared with other kids and these kids are at higher risk of insomnia.

Research Highlights

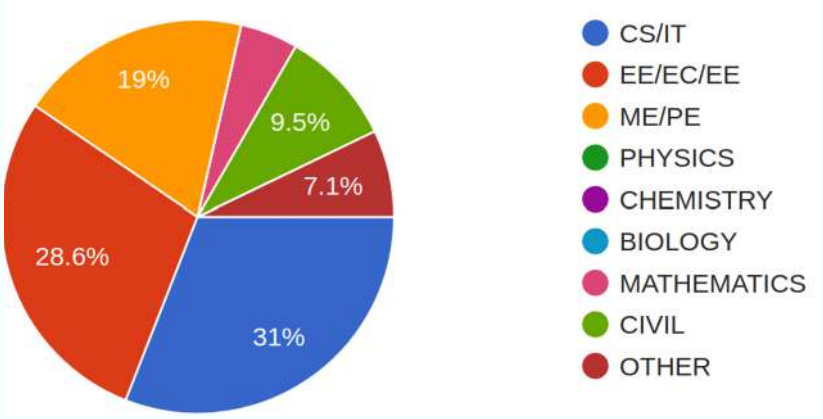
Sharing one's progress with the world is a moral duty of an academician. As Sir Isaac Newton once said, "If I have seen further it is by standing on the shoulders of Giants". Here are some contributions made by our department recently -

- **Simulation of viscoelastic two-phase flows with insoluble surfactants**, Venkatesan, J., Padmanabhan, A., & Ganesan, S., *Journal of Non-Newtonian Fluid Mechanics (2019)*
- **Computational modeling of impinging viscoelastic droplets**, Venkatesan, Jagannath, and Sashikumaar Ganesan, *Journal of Non-Newtonian Fluid Mechanics (2019)*
- **How Deep Learning performs with Singularly Perturbed Problems?**, Yadav, Sangeeta, and Sashikumaar Ganesan, *IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE). IEEE, 2019.*
- **Predicting Gas Phase Entropy of Select Hydrocarbon Classes Through Specific Information Theoretical Molecular Descriptors**. Chandan Raychaudhury, Md. Imbesat Hassan Rizvi and Debnath Pal . *SAR and QSAR in Environmental Research (2019)*
- **Exploring the use of molecular dynamics in assessing protein variants for phenotypic alterations.** Aditi Garg and Debnath Pal. *Human Mutation Variation, Informatics, and Disease (2019)*
- **A proline insertion-deletion in the spike glycoprotein fusion peptide of mouse hepatitis virus strongly alters neuropathology.** Manmeet Singh, Abhinoy Kishore, Dibyajyoti Maity, Punnepalli Sunanda, Bankala Krishnarjuna, Sreeparna Vappala, Srinivasarao Raghothama, Lawrence C. Kenyon, Debnath Pal* and Jayasri Das Sarma. *Journal of Biological Chemistry (2019)*
- **Robust Initial Satellite Orbit Determination method using a Modified Kalman Filter**, Potu, Shirish, S. K. Anand, and Soumyendu Raha, *The Journal of Navigation (2019)*
- **Partitioning a reaction–diffusion ecological network for dynamic stability**. Kumar, Dinesh, Jatin Gupta, and Soumyendu Raha. *Proceedings of the Royal Society A (2019)*
- **A Systematic Approach for Acceleration of Matrix-Vector Operations in CGRA through Algorithm-Architecture Co-design,** Farhad Merchant, Tarun Vatwani, Anupam chattopadhyay, Soumyendu Raha, S. K. Nandy, Ranjani Narayan, Rainer Leupers, *IEEE VLSI Design (2019)*
- **Applying Modified Householder Transform to Kalman Filter,** Farhad Merchant, Tarun Vatwani, Anupam Chattopadhyay, Soumyendu Raha, S. K. Nandy, Ranjani Narayan, Rainer Leupers, *IEEE VLSI Design (2019)*
- **Strong Coupling of an Emitter with Absorbing Matter: A Regime for Enhancement of Light Emission**, Jain, Kritika, and Murugesan Venkatapathi. , *Physical Review Applied(2019)*
- **Fractional Regularization to Improve Photoacoustic Tomographic Image Reconstruction,** [Dween R. Sanny*, Jaya Prakash*], Sandeep K. Kalva, Manojit Pramanik, and Phaneendra K. Yalavarthy, *IEEE Transactions on Medical Imaging (2019)*
- **PA-Fuse: A Deep Supervised Approach for Fusion of Photoacoustic Images with Distinct Reconstruction Characteristics**, Navchetan Awasthi, K. Ram Prabhakar, Sandeep Kumar Kalva, Manojit Pramanik, R. Venkatesh Babu, *Biomedical Optics Express (2019)*
- **Modeling errors compensation with total least squares for limited data photoacoustic tomography**, Sreedevi Gutta, Manish Bhatt, Sandeep K. Kalva, Manojit Pramanik, and Phaneendra K. Yalavarthy, *IEEE Journal of Selected Topics in Quantum Electronics (Issue on biophotonics) (2019)*
- **Convolutional Neural Network based Robust Denoising of Low-Dose Computed Tomography Perfusion Maps**, [Sreedevi Gutta*, Venkata S. Kadimesetty*], Sriram Ganapathy, and Phaneendra K. Yalavarthy, *IEEE Transactions on Radiation and Plasma Medical Sciences (Special issue on machine learning methods for image processing and radiomics) (2019)*, [Among Top 25 popular articles for the months of August 2018, March 2019, & April 2019]
- **From Strings to Things: Knowledge-enabled VQA Model that can Read and Reason**. A. K. Singh, A. Mishra, S. Shekhar, A. Chakraborty, *International Conference on Computer Vision (ICCV), 2019*
- **FDA: Feature Disruptive Adversarial Attack**, Aditya G., Vivek B.S., and R. Venkatesh Babu, accepted in *ICCV 2019.*
- **GAN-Tree: An Incrementally Learned Hierarchical Generative Framework for Multi-modal Data Distributions**, Jogendra Nath Kundu, Maharshi G., Dakshit A., and R. Venkatesh Babu, *accepted in ICCV 2019.*

CDS Students at a Glance

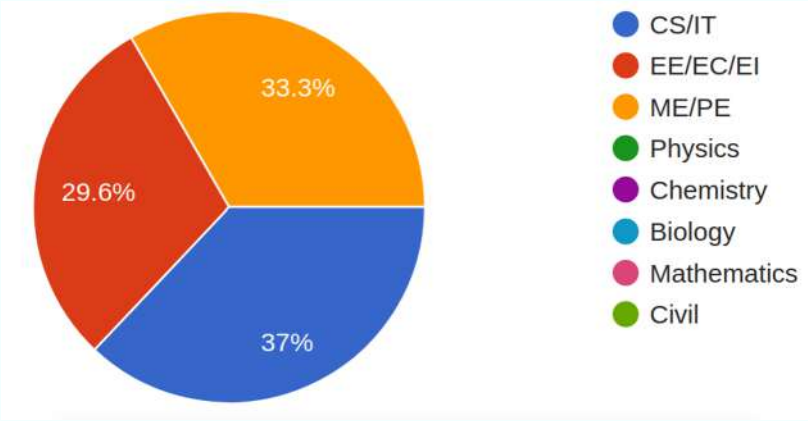
Although terms like "computational" and "data sciences" might seem overwhelming for students of some sciences, CDS is a truly interdisciplinary department with disciples of different disciplines and work experience. We have showcased this in the following charts.

Education background of CDS students

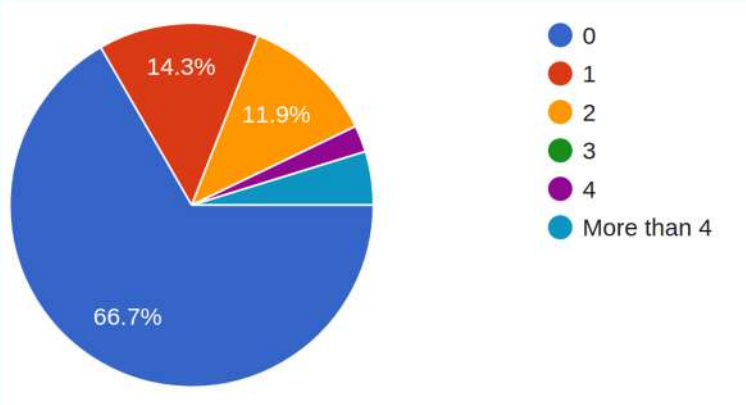


First Year students

Second Year students

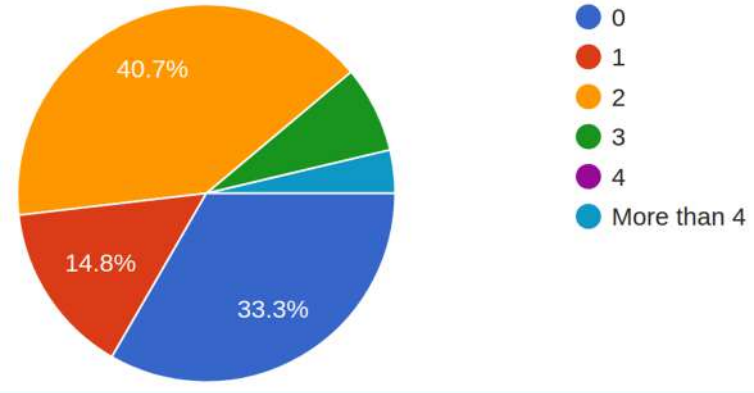


Work Experience (in years) of CDS students



First Year students

Second Year students



PhD Comics

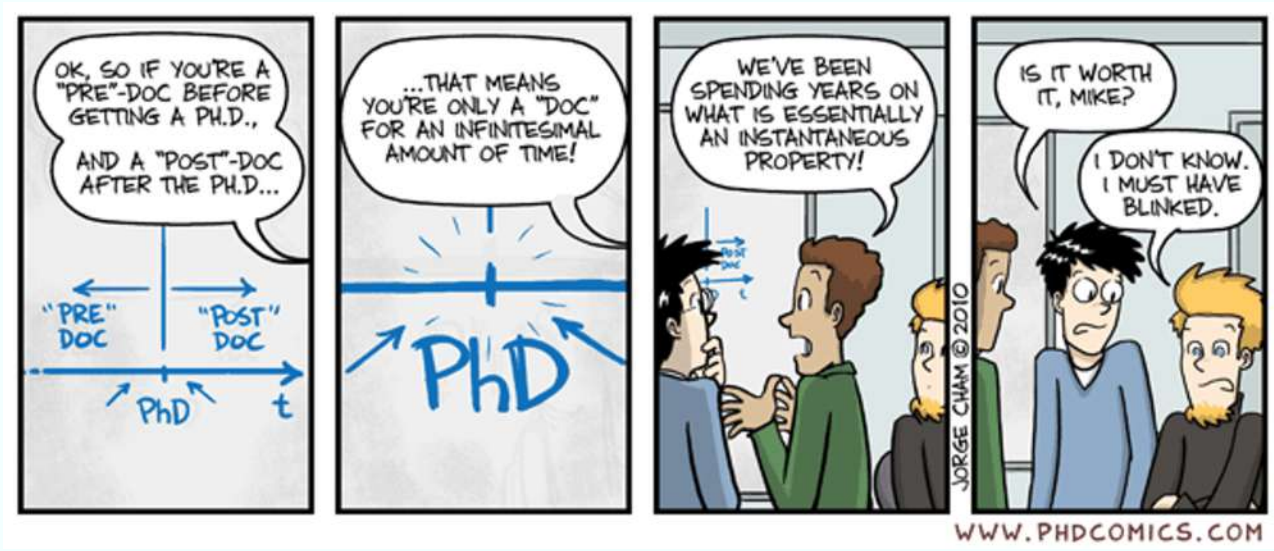
Just like atoms, life as a student in academia is not a piece of plum pudding (as Sir Joseph John Thomson would later realise for the former at least). Nowhere else do you celebrate success after failing thousands of times. The problems are pesky, approach is at times pesty and the solutions are more often plaguy. That being said, no student carries this burden alone and this ordeal is as ubiquitous as the dreams of winning a Nobel (or Turing). Here are some of these moments:-

"Pretty Sure professors are still called sensei in Japan"

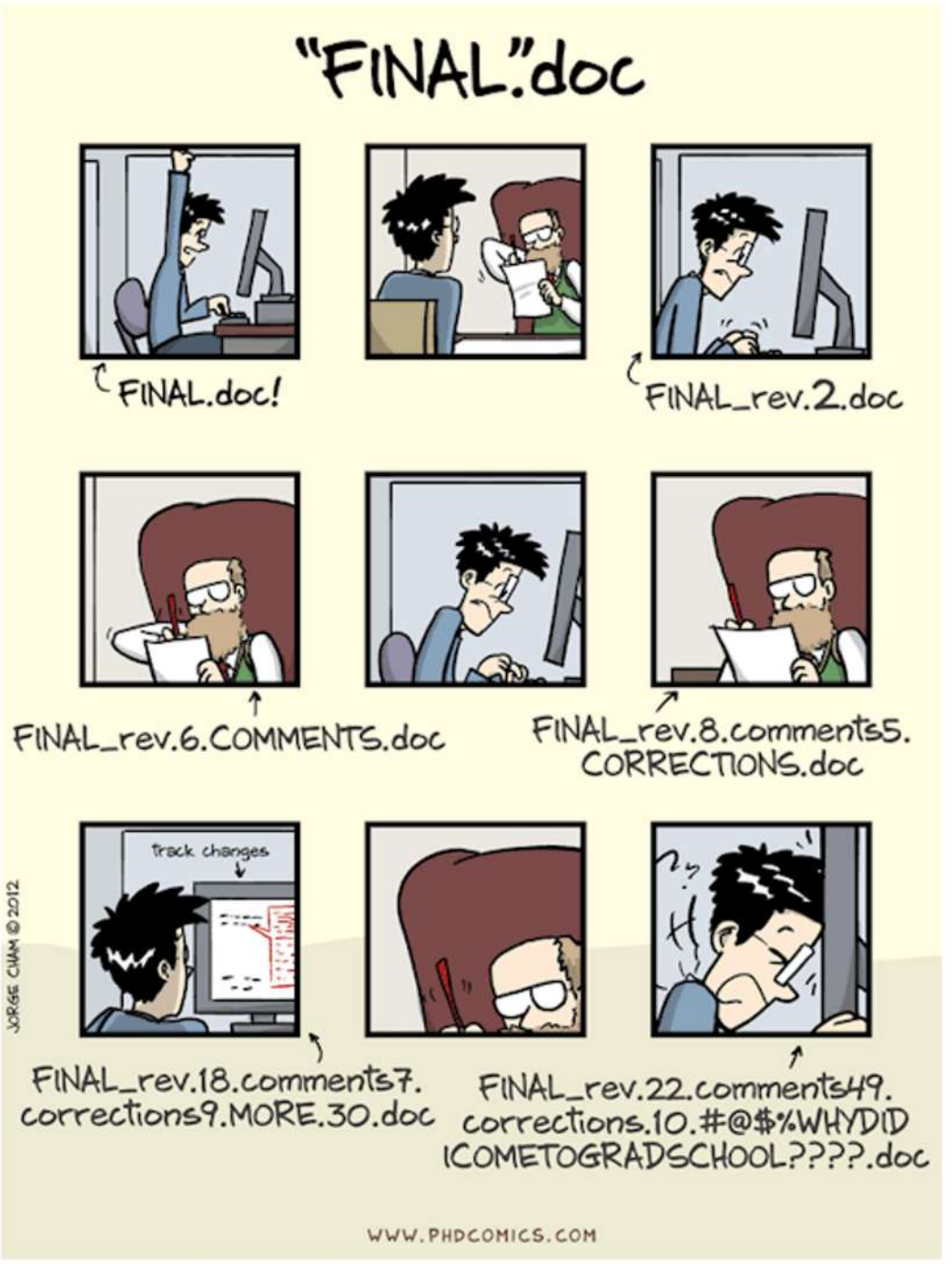
"Just a little push (actually a thousand of them)"

"Wry TAs are not born but are made such by their TAs"

"Does this mean a PhD has a half life of nearly ZERO seconds!!!!"



"The never ending pursuit of perfection"



Credits - Aditya Rastogi, Gaurav Oberoi, Ocima Kamboj, Somil Jain

Edited by Dr. Phaneendra Yalavarthy