Ali H. Sayed | IEEE Signal Processing Society President | ali.sayed@epfl.ch

Big Ideas or Big Data?

Many ask me what signal processing should be doing in the age of big data. My answer is clear: signal processing should continue to generate big ideas. Big ideas for big data.

Our discipline has always advanced ingenious methods and theories, irrespective of the size of the data: *small* or *big*. Many of these ideas permeate disciplines far and wide, ranging from imaging to video; speech processing to coding and communications, forensics, security, and privacy; and also social media, machine learning, and data science.

Keep in mind that the qualification *big* is relative to the technology of the times. When we were trying to save large files onto floppy disks of limited capacity in the early 1980s or when we were trying to run complex simulations on a memory-constrained desktop computer from that same time period, we were also dealing with *big* challenges. When researchers were attempting to compute large-scale discrete Fourier transforms two decades earlier in the 1960s, they were also faced with a *big* data problem for their time.

The *size* of the data has, of course, grown exponentially fast in recent years with the spread of cloud resources, social media, and online powerhouses such as Google, Facebook, Twitter, and Amazon. Many new and interesting research challenges are driven by this reality. Signal processing met similar challenges in the past with vigor and is already contributing to the new big challenges with remarkable strides. However, at times, we are victimized by our own successes. We have perfected our art so well that our contributions often go unnoticed to our detriment. As we like to say, signal processing has become the invisible intelligence behind our modern digital revolution, or "the science behind our digital life." Our algorithms power everything from cell phones, to spacecraft, robotic explorers, imaging devices, automated systems, computing machines, as well as recommender systems. You name a technology, and you will find some signal processing concept or solution driving its functionality.

We are a singular discipline. The training we provide to our students and engineers is special. We prepare them with a unique gift to modeling, problem solving, and optimization. We know how to extract the essence of a problem. We are creative, inventive, innovative, and, above all, we like challenges. We also do not shy away from venturing confidently into other fields. I recall the time when my former Ph.D. advisor at Stanford University, Thomas Kailath—a giant in our field, an IEEE Medal of Honor recipient, and a U.S. National Medal of Science honoree by President Barack Obama—launched into a large project on semiconductor manufacturing in the early 1990s, which was a field well outside his domain of expertise. He ended up making award-winning contributions and also cofounded a successful company. That is the spirit of signal processing.

Many other disciplines can benefit from our expertise and way of thinking. If they look the other way, they will often rediscover our methods at added expense and wasted resources. So stand up and be proud of your association with the IEEE Signal Processing Society. It is no wonder that we were the very first IEEE Society founded back in 1948 and are commemorating our 70th anniversary in 2018. We are also one of the largest, ranked fourth in size among close to 40 other IEEE Societies. It also makes us proud to know that one our Society's recent former presidents and distinguished colleague, José Moura, will be leading the IEEE in 2019.

A glimpse of history

Signal processing and data have always had an intertwined history. Many of our tools and concepts have been motivated by data-driven inquiries, even from the remote past. Let us take a tour of some historical contributions. The list given next is, of course, biased by my own preferences and is not intended to be exhaustive given the space limitations. It is only meant to illustrate how signal processing has always been at the center of promoting big ideas for data.

Let us start with Gauss himself who, in 1795 at the young age of 18, was fitting lines and hyperplanes to astronomical data and invented the least-squares criterion for regression analysis. He even devised the recursive least-squares solution to address what was a *big* data problem for him at the time: he had to avoid tedious repeated calculations by hand as more observational data became available. What a wonderful big idea for a data-driven problem! Of course, Gauss had many other big ideas.

de Moivre (1733), Laplace (1812), and Lyapunov (1901) worked on the central

Digital Object Identifier 10.1109/MSP.2018.2791378
Date of publication: 7 March 2018

limit theorem. The theorem deals with the limiting distribution of averages of *large* amounts of data. The result is also related to the law of *large* numbers, which even has the qualification *large* in its name. Again, big ideas motivated by *large* data problems.

Our own Fourier (1807), whose tools are at the core of our discipline, developed the phenomenal Fourier representation for signals. It is meant to transform data from one domain to another to facilitate the extraction and visualization of information. A big transformative idea for data.

Bayes (circa the mid-1750s) and Laplace (1774) appear to have independently discovered the Bayes' rule, which updates probabilities conditioned on data observations. The rule forms the backbone of much of statistical signal analysis, Bayes' classifiers, naive classifiers, and Bayesian networks. A big idea for data-driven inference.

Fast-forward to modern times. The fast Fourier transform (FFT) is another example of an algorithm driven by challenges posed by data size. Its modern version is due to Cooley and Tukey (1965); the latter received the IEEE Medal of Honor and was also a U.S. National Medal of Science honoree. Their algorithm revolutionized our field, and FFT processors have become common components in many modern electronic devices. Even Gauss had a role to play here, having proposed an early version of the algorithm some 160 years before, again motivated by a data-driven problem while trying to fit astronomical data into trigonometric polynomials. A big idea for a data-driven problem.

Closer to the core of statistical signal processing, both Kolmogorov (1939) and Wiener (1942) laid out the foundations of modern statistical signal analysis and optimal prediction methods. Wiener was a recipient of the U.S. National Medal of Science. Their theories taught us how to extract information optimally from data, leading to further refinements by Wiener's student Levinson (1947) and, more dramatically, by Kalman (1960), who was also an IEEE Medal of Honor recipient and a U.S. National Medal of Science honoree. The Kalman filter is applied across so many domains, including in financial analysis from market data. Kalman's work was an outgrowth of the model-based approach to system theory advanced by Zadeh (1954), another IEEE Medal of Honor recipient and the founder of the field of fuzzy logic. The concept of a recursive solution from streaming data was a novelty in Kalman's filter; the same concept is commonplace today in most online learning techniques. Again, big ideas for recursive inference from data.

Cauchy (1847) early on, and Robbins and Monro (1951) a century later, developed the powerful gradient descent method for root finding, which is also recursive in nature. Their techniques have grown to motivate huge advances in stochastic approximation theory. Notable contributions that followed include the work by Rosenblatt (1957) on the Perceptron algorithm for single-layer networks and the impactful delta rule by Widrow and Hoff (1960), widely known as the LMS algorithm in our discipline. Subsequent work on multilayer neural networks grew out of the desire to increase the approximation power of single-layer networks, culminating with the backpropagation method of Werbos (1974). Many of these techniques form the backbone of modern learning algorithms. Again, big ideas for recursive online learning.

Shannon (1948), an IEEE Medal of Honor recipient, a U.S. National Medal of Science honoree, and a modern giant, contributed fundamental insights to data representation, sampling, coding, and communications. His concepts of entropy and information measure help quantify the amount of uncertainty in data and are used, among other areas, in the design of decision trees for classification purposes. Nyquist (1928), also a winner of the IEEE Medal of Honor, contributed to the understanding of data representations as well. Big ideas for data sampling and data manipulation.

Bellman (1957), a towering system-theorist and an IEEE Medal of Honor winner, introduced dynamic programming and the notion of the curse of dimensionality, both of which are core underpinnings of many results in learning theory, reinforcement learning, and the theory of Markov decision processes. Viterbi's algorithm (1967) is one notable example of a dynamic programming solution, which has revolutionized communications and has also found applications in hidden Markov models widely used in speech recognition today. Viterbi is also an IEEE Medal of Honor winner and a recipient of the U.S. National Medal of Science. Big ideas for conquering complex data problems by dividing them into simpler problems.

Kernel methods, building on foundational results by Mercer (1909) and Aronszajn (1950), have found widespread applications in learning theory since the mid-1960s with the introduction of the kernel Perceptron algorithm. They have also been widely used in estimation theory by Parzen (1962), Kailath (1971), and others. Again, a big idea for learning from data.

Pearson and Fisher established the modern field of mathematical statistical signal analysis with the introduction of methods such as principal component analysis by Pearson (1901) and maximum likelihood and linear discriminant analysis by Fisher (1912, 1922, and 1936). These methods are at the core of statistical signal processing. Pearson (1894 and 1896) also had one of the earliest studies of fitting a mixture of Gaussian models to biological data. Mixture models have now become an important tool in modern learning algorithms. Big ideas for data-driven inference.

Markov (1913) introduced the formalism of Markov chains, which is widely used today as a powerful modeling tool in a variety of fields including word and speech recognition, handwriting recognition, spam filtering, gene analysis, and web search. Markov chains are also used in Google's PageRank algorithm. Markov's motivation was to study letter patterns in texts. He laboriously went through thousands of letters in a classical Russian novel and counted pairs of vowels, consonants, vowel followed by a consonant, and consonant followed by a vowel. A *big* data problem for his time. These were great ideas (and an example of great patience) for data-driven inquiries.

The list goes on with many modern day and ongoing contributions by signal processing experts to network science, distributed processing, compressed sensing, wireless communications, array processing, optimization, multiagent systems, computational imaging, speech processing, forensics, etc.

Just count the number of IEEE Medal of Honor recipients and U.S. National Medal of Science honorees in the previous remarks. You are in good company.

SP