Scientific zero to one: Some common properties of highly-influential papers

Mohammad Mahdi Jahani Yekta
Department of Computer Science,
Stanford University
Gates Computer Science Building,
353 Jane Stanford Way, Stanford, CA 94305, USA
e-mail: m_mahdi_jahani@yahoo.com

ABSTRACT

Based on an intensive study of the most popular papers of the approximately 6,000 top-1% highly-cited (HC) researchers 2018 announced by Clarivate Analytics, the most HC papers in different fields of science in the interval 2001-11 reported by Web-of-Science (Thomson- Reuters), the top 100 most-HC ones ever; reported in 2014 by the same institute, and also the scientific backgrounds and career trends of the involved authors, we first analyze several major properties common in most such highly-influential papers that add something fundamentally new to their corresponding fields, taking them from a "zero" state in some domain to "one". These properties include importance of the paper topic in essence, firsthand approach to the problem it focuses on, notable improvement of the solution it offers for the problem compared to its existing counterparts, monopoly in some aspect, durability, proper timing, knowledgeable research team and availability of the required resources, effective presentation to the right audience, and; finally, perseverance. We next propose soft citation; an upgrade to the current citation standards, and then a problem-oriented interactive research model as a transformative complement for the state-of-the-art academic publishing system, that would expedite this zero to one transition.

Keywords: Highly-influential papers; Highly-cited papers; Scientific breakthrough; Research models; Academic publishing systems.

INTRODUCTION

Setting aside the graduate students and university professors who publish only to graduate and to receive career promotions or become tenure-tracked respectively, there exist scientists who do not believe in the publish-or-perish culture, and aim to make major breakthroughs and noticeable influence. For this, some knowledge about properties highly-influential (HI) papers generally have in common would naturally be helpful. This paper, in the first place, aims to study some of the most essentials of such properties.

Before anything else, there is a need to make clear what is meant by the term "highly-influential" used in this paper repeatedly, and its difference with other phrases that may be used in similar contexts for a merely "good" research. Similar to the categorization Terence Tao; the well-known mathematician and winner of the 2012 Fields medal, proposes for good mathematics (Tao 2007), good research may also refer in general to a piece of scientific study having one or more of a number of desired properties. These include good (1) problem-solving, a research that solves a major problem; (2) technique, a smart usage of existing methods or devising of new tools; (3) theory, that unifies a body of scattered results in a common framework; (4) insight, providing a deeper understanding of a principle or phenomenon; (5) discovery, noticing an unknown or unexpected fact; (6) application, in major problems in the same area or probably other fields of science; (7) exposition; an informative survey of an important topic; (8) pedagogy, a writing or presentation approach that helps others learn the subject more effectively; (9) vision, demonstrating a program for an organized long-run research; (10) taste, a research that is intrinsically of great interest for a broad range of scientists, while helps solve big problems as well; (11) public-relations, effective presentation of findings to non-specialists; (12) meta-research, advances in the foundations, philosophy, or practice of the topic; and (13) rigor; containing details of the steps taken in the work. While except for rigor the rest of these items are of course more or less subjective, this list of 13 properties can naturally be extended with even more subjective properties such as the research being for instance (14) beautiful, presenting results nice to state and intuit, but maybe hard to prove; (15) elegant, deriving difficult results shortly and with minimum effort; (16) creative, meaning to be highly original; (17) useful, to be used widely and repeatedly in future works; (18) strong, that derives a strong conclusion from relatively weak assumptions; (19) deep, reaching something beyond the reach of elementary tools or generic approaches; (20) intuitive, being sensible from a natural perspective; and (21) definitive, to remove ambiguities in a field. It is of course true that these categories do not have sharp boundaries, and overlap. Also, a research excellent based on either of these properties will likely score high in some others as well. The remarkable phenomenon here; anyway, is that a highly influential research in any of these “local” aspects, corresponding to only one dimension of strength, tends to spark more such research in many of the other aspects as well, which can effectively discriminate between these and merely good researches. In fact, a highly influential work is not just an isolated strong work, but is instead a piece contributing exceptionally to the progress of a bigger “global” and long-term scientific story, and the interaction and orchestration of its parts with each other.

This article, in the first place, analyzes the major properties most highly influential papers have in common. This analysis is based on a thorough study of the most popular papers of the approximately 6,000 top-1% highly-cited (HC) researchers 2018 (Clarivate Analytics 2018) announced by Clarivate Analytics, the most-HC papers in different fields of science in the interval 2001-2011 reported by Web of Science (Thomson-Reuters), the top 100 most HC ones ever, reported in 2014 by it as well, along with a thorough investigation of the scientific careers of the involved authors using their university home-pages, Scopus, Google Scholar, and other such reliable venues. The analysis is not merely another informetrical work to be put on the thousands of the already existing ones, but instead a visionary and insight-giving one looking to the scene of research from a different and fresh point of view, from where it is seen as a “business” in the much broader context of the whole society as the ultimate consumer of the corresponding scientific products. From such a perspective, a highly-influential research not only has to have high scientific standards, but has also to fit well to the playground and be a differentiated product in the research “market” as well. In this framework, the spotted properties are much the same as those of the most successful businesses studied in the book “Zero to one: notes on startups, or how to build the future” Masters and Thiel (2014), formed from a computer science course instructed in 2012 by Peter Thiel, the founder of PayPal and one of the early investors of Facebook, at Stanford University. In this book, which has been the New York Times number 1 best seller in 2014 as well, Thiel divides businesses to two main categories: those only copying an existing business probably with some minor modifications, and the others which create and offer something essentially new. In his phrasing, the first group take the world from 1 to n, while the latter take it from 0 to 1. Being a nice conveyer of the underlying difference, this interpretation has been used in the title of this paper as well.

The next contributions in this work are proposing (a) the idea of soft citation, which is an upgrade to the current traditional citation standard in which many essential information about the referred document is missing, and (b) a problem-oriented interactive research model as a transformative complement for the state of the art scholarly publishing system, that can facilitate and expedite this zero to one transition to a great extent.

OBJECTIVES AND METHOD

The objectives of this work are to:
(a) study some of the most essentials properties highly influential papers have in common;
(b) propose the idea of soft citation, which is an upgrade to the current traditional citation standard in which many essential information about the referred document is missing; and
(c) propose a problem-oriented interactive research model as a transformative complement for the state of the art scholarly publishing system, that can facilitate and expedite this zero-to-one transition to a great extent.

The method for the first objective has been to investigate a huge number of resources stated in the introduction, including many websites and papers, to check whether the nine properties Thiel states for highly successful businesses hold for highly influential scientific papers as well. The papers most deeply scrutinized for this purpose have been selected based on a thorough study of the most popular ones of the approximately 6,000 top-1% highly-cited (HC) researchers 2018 (Clarivate Analytics 2018) announced by Clarivate Analytics, the most-HC papers in different fields of science in the interval 2001-2011 reported by Web-of-Science (Thomson-Reuters), and the top 100 most HC ones ever reported in 2014, again by it. A dominant majority of these papers have received more than 3,000 citations to date.

CITATION COUNTS REVISITED

It is a rather common practice in academia to think as a highly-influential research simply as one that would receive thousands of citations. While such an achievement of course proves a work to be HI, it should be noted that citation count is by no means a perfect estimator neither for the level of this influence nor for its nature. The realities behind these counts, some of which discussed in Van Noorden, Maher and Nuzzo (2014) in the framework of a review of the top 100 most highly-cited papers ever, are indeed much different from what most academicians may think of. These 100, that have been reported in 2014 by Web of Science, do not; e.g., contain many of the world’s most famous and absolutely highly-influential works, including those of Albert Einstein. On the other hand, Oliver Lowry, the late first author of the most HC paper ever (Lowry et al. 1951) (no. of citations till Nov. 2018 > 305, 000) that describes an assay to determine the amount of protein in a solution, mentions in a writing in 1977 that “Although I really know it is not a great paper . . . I secretly get a kick out of the response” Van Noorden, Maher and Nuzzo (2014, p. 1).

Most of the papers in this top-100 list are in fact on experimental methods or softwares that have become popular in their fields (Van Noorden, Maher and Nuzzo 2014). This shows, in contrast to what a considerable portion of researchers think of, that the vast majority of most HC papers have achieved this honor not because of a rarely reachable and extremely exceptional intellectual breakthrough, but simply for their frequent usage in the daily scientific activities of other researchers. They in fact cite these sorts of papers almost only to make clear the grounds and framework in which their research has been carried on (Van Noorden, Maher and Nuzzo 2014). On the other hand, Einstein's papers and other works of this class are such outstanding breakthroughs that soon become famous, go into textbooks, and get known to almost everyone in the field. There would therefore remain no need for them to be formally cited when referred to - it suffices for them to be merely hinted at in the text. This would naturally cause them to receive far less citation counts than they are eligible for, which may eventually prevent their authors from climbing the lists of highly-cited researchers (Van Noorden, Maher and Nuzzo 2014). Some other facts affect the neutral interpretation of these sorts of rankings as well. Biologists, for example, citing one another's work more frequently than physicists (Van Noorden, Maher and Nuzzo 2014); and the number of publications in different fields not being the same, which are of course generally tried to be relaxed and compensated for in such lists via a variety of balancing methodologies.

This paper, as mentioned previously, is not going to get drown in such informetrical discussions anyways. It will instead focus on "how" major scientific breakthroughs have emerged, and; as stated earlier, use citation counts only as a measure to give an initial rough estimate about their order of popularity. Much more realistic and sensible evaluations about their long-run impact on the whole body of science are left to be made by expert judgments over time. This would also be more elaborated on in a later section of this paper.

MAJOR COMMON PROPERTIES OF HIGHLY-INFLUENTIAL PAPERS

Based on an in-depth analysis of the most popular papers, scientific careers, and the chain of events leading to the ultimate success of the approximately 6,000 Clarivate HC researchers 2018, the major properties more or less common in their most HC papers are as studied and described in the following sections, which include (1) high importance of the subject in essence; (2) firsthand approach to the problem; (3) 10X improvement of the offered solution; (4) monopoly; (5) durability; (6) proper timing; (7) research team, atmosphere, resources; (8) effective presentation to the right audience; and finally (9) deep interest and belief of the researchers in their work, and their perseverance on it. The citation counts reported in this paper for the HC papers, which are based on Google Scholar records in 2018; when the first draft of the work was being prepared, are indicated in square brackets when the corresponding articles are referred to.

High Importance of the Subject in Essence

The first property lying at the core of the properties of a highly-influential piece of science is the importance of its topic over a wide range of disciplines. Although the level of this importance is most realizable by the field's specialists, it is generally somewhat sensible to others as well. For instance, the density functional theory (DFT) which is the most highly-cited concept in physical sciences (Van Noorden, Maher and Nuzzo 2014) is a well-clarifying example of such topics which covers 12 entries of the 2014 top-100 list (accompanied to the online version of Van Noorden, Maher and Nuzzo 2014), including 2

(Lee, Yang, and Parr 1988) [81,450] and Becke (1993) [85,510]) of its top 10. DFT is mainly an approximation that "makes impossible mathematics easy", says Feliciano Giustino; a materials physicist at the University of Oxford (Van Noorden, Maher and Nuzzo (2014), p. 4). It makes the study of electronic behavior in a silicon crystal that would require analyzing 10^{21} tera-bytes of data possible to be conducted using only a few hundred kilobytes of samples. The wide importance of such a topic is clearly needless of being discussed.

On the opposite side, contributions made to topics that are relatively marginal and with limited importance would likely not be much recognized, irrespective of their level of novelty and other positive factors. A moderately original work on an important topic is generally expected to receive much more notice than a highly creative one in a less important subject.

Evaluation of the real importance of a topic to work on, maybe for months or even years, is an extremely quaint matter to be approached by a researcher. It is highly probable for a researcher to get misled by some local factors that make the topic seem to be important, while it not being as such at all from a global and unbiased perspective. As a first rule-of-thumb guide in this context, and based on the analysis of the most HC papers in different fields of science in the interval 2001-2011 reported by Web of Science, the five main classes of topics that have been popular in this interval are discussed in the following sub-sections.

a) Softwares

Softwares have been the most frequent HC research outputs in 2001-2011. The most top-cited article in agricultural and biological sciences in this interval for example is Tamura et al. (2007) [31,620], that presents MEGA-4, the fourth version of a free software for conducting statistical analysis of molecular evolution and for constructing phylogenetic trees. The paper by Pettersen et al. (2004) [chemistry, 18,750] also offers a software for interactive visualization and analysis of molecular structures and related data, including density maps, supramolecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles. The first place in the field of energy is for Allison et al. (2006) [5,480] written by 44 authors from across the globe, which presents a software for simulation of the passage of particles through and interacting with matter. The paper by Tzourio-Mazoyer et al. (2002) [neuroscience, 9,700] provides a freeware for interface with the statistical parametric mapping package of Ashburner and Friston (1999) [1,960] to be used by neuroimaging researchers. It is noting that softwares form a considerable portion of the top-100 papers in Van Noorden, Maher and Nuzzo (2014) as well.

b) Methodologies and algorithms

Papers that propose powerful algorithms or methodologies applicable to a wide range of problems are the next group that have received the highest citation counts. In arts and humanities, the first rank in 2001-2011 is for McCall (2008) [5,410] on inter-sectionality (or inter-sectionalism), which is the study of intersections between forms or systems of oppression, domination, or discrimination. This article discusses the complexity of studying this subject, and presents new methods for this purpose. McCall believes that the reason for his paper becoming highly-cited is the fact that while qualitative research tends to dominate the study of inter-sectionality, this work has helped justify research in more quantitative areas in the field. The paper by Lowe (2004) (computer science) with the eye-catching number of 50,900 citations proposes an algorithm in computer vision to match images in a novel manner. The 2004 work of Bendtsen et al. (2004) [immunology and micro-biology, 7,030] improves the then most popular method for prediction of classically secreted proteins; SignalP-version 2, which is based on neural network and hidden Markov model algorithms. In materials science, Soler et al. (2002) [9,650] develops and implements a self-consistent density functional method. The article by Deb et al. (2002) [mathematics, 27,110] suggested a non-dominated sorting genetic (NSG) multi-objective evolutionary algorithm (MOEA) named NSGA-II, which relaxed the three main drawbacks of its other MOE counterparts algorithms to a great extent. In psychology, Braun and Clarke (2006) [54,320] argues that thematic analysis offers an accessible and theoretically-flexible approach to analyzing of qualitative data, provides clear guidelines for researchers who are about to start thematic analysis or to conduct it more rigorously, and re-introduces it as a useful and flexible method for qualitative research in and beyond psychology.

To see the reason for the popularity of methodologies and algorithms, it is necessary to flashback to the decade before the 2000's. When five out of 15 top mathematical scientists cited in the 1990's turned out to be among Stanford's statistics professors, Bradley Efron; the third of these five, highlighted the key reason for this to be their work focused on useful methodology (see the 14th Feb. 2001 news at https://news.stanford.edu). He mentions that one of the nice recent turns in the field of statistics, since 1970, has been a greater interest in algorithms that help scientists with difficult data problems. He believed this to be applicable to most scientists, as the computer age has allowed much more ambitious data-collection methods.

Many other exceptionally HC papers other than the first ranks above fall in this class as well. For instance, after the tutorial paper by Arulampalam et al. (2002) [11,860], the second most-cited paper of *IEEE Transactions on Signal Processing*; is the 2006 paper by Aharon, Elad, and Bruckstein (2006) that presents an algorithm for designing over-complete dictionaries for sparse representation of signals, and has collected 7,930 citations.

c) Foundational research

Papers that have made the foundations for mega-trends in research form another class of highly-influential works. For instance, the paper by Li et al. (2008) [7,550]; being the most-cited work in chemical engineering in 2001-2011, has opened numerous opportunities for using of graphene (an allotrope (form) of carbon consisting of a single layer of atoms arranged in a hexagonal lattice) in many technological applications via the low-cost way it develops for large-scale production of aqueous grapheme dispersions. The article by Kreuer (2001) [chemistry, 2,950] presents a semi-quantitative connection between morphology (micro-structure) and transport (proton conductivity, water transport) of fuel cell membranes (hydrocarbon versus perfluorosulfonic acid - PFSA), from which a wide range of related research in fuel cells and other electro-chemical applications has span off. Another exceptional work from this class is the 2010 Nobel Prize winning paper (Geim and Novoselov 2010) [engineering, 33,110] of A. K. Geim and K. S. Novoselov, which in combination with Novoselov et al. (2004) [46,370] provides the laying foundations for ghraphene research. With 8,070 citations, the most-cited paper in environmental sciences in this interval; Kolpin et al. (2002), also falls in this category. This paper, which measures concentrations of 95 OWCs (organic wastewater contaminants) in water samples from a network of 139 streams across 30 US states during 1999 and 2000, has been the first national-scale study of such compounds conducted in US that has strongly triggered research on contaminants of emerging concern (CEC's) in this country as well. The article by Vargo and Lusch (2004) [economics, econometrics, and finance, 13,830] founds the grounds for a marketing paradigm in which service provision, rather than goods, is fundamental to economic exchange.

d) Molecular and cell-level medical research

Another group of highly-influential papers are those in molecular and cell-level medical research. The most-HC 2001-11 articles Writing Group (2002) [medicine, 3,490], Livak and Schmittgen (2001) [biochemistry, genetics, and molecular biology, 93,760], Dominici et al. (2006) [pharmacology, toxicology and pharmaceutics, 11,580], and Gronthos et al. (2002) [dentistry, 1,970] can be categorized in this group. Another informative package of works from this class are those of Guido Kroemer, an expert in immunology, molecular biology, and pharmacology, and one the Clarivate’s 24 HC researchers 2018 in three ESI (Essential Science Indicators) fields (Clarivate Analytics 2018). Kroemer's breakthroughs in immuno-pharmacology of cell death, mitochondrial cell death control, mechanisms of HIV-1-induced apoptosis, crosstalk between lethal, stress-adaptive and metabolic pathways in aging and disease, immunogenic cell death for optimal anti-cancer chemotherapy, clinical research concentrated on patient-relevant bio-markers and successful clinical trials, and also for reviews and position papers in *Annual Reviews* and *Nature Reviews*, have earned him a total of 214,900 citation counts along with an impressive h-index of 219.

Papers in areas such as biological techniques, bioinformatics, and phylogenetics; being some branches from this group, have been among the ones most-cited ever as well (Van Noorden, Maher and Nuzzo 2014).

e) Tutorials and reviews

Tutorials and reviews are the next category of papers widely seen in the highly-influential literature. This is partly because researchers tend to study some good reviews on their subjects at the very beginning of their work to prevent reinventing the wheel. One of the first sources of inspiration for Geim's works that led to his 2010 Nobel Prize 30 years later, for instance, was an old 1981 review paper (Dresselhaus and Dresselhaus 1981), in which it was indicated that even after many decades graphite had been a yet little-understood material, especially in terms of its electronic properties (Geim 2011). In this framework, the most-HC 2001-11 paper (Venkatesh et al. 2003) (business, management, and accounting, 23,470) is a review. The 2002 paper by Albert and Barabási (2002) (physics and astronomy, 21,200) reviews the then-recent advances in the field of complex networks focusing on the statistical mechanics of network topology and dynamics, and discusses the corresponding main models and analytical tools. Besides being categorized in the foundational class studied earlier, the highly-cited work of Geim and Novoselov (2010) is of course also a review paper to be referred to here again. As stated previously, a notable portion of Guido Kroemer's scientific fame is for his excellent reviews and position papers in *Annual Reviews* and *Nature Reviews* as well.

Another pedagogical example well-demonstrating the attention a good review may catch is the citation profile of Vinod Kumar Gupta; another member of the 24 hinted earlier, whose areas of research lies in the cross-section of ion selective electrodes, chemical sensors, and wastewater treatment by developing low cost absorbents. The most highly-cited paper of Gupta is the 2009 review article (Gupta et al. 2009) with 2,620 citations to date, which is about three times that of his second HC work being 860. Except for this most HC paper, his other citation counts fall in a rather gradual pattern, conveying the message that what a deep difference between an exceptional review and one’s other regular works, even being of rather high-qualities, might be.

The profile of Frank B. Hu; another scientist from the 24, contains a highly-cited review as well. The 2nd HC paper of Hu, whose research focuses on diet-lifestyle, metabolic, and genetic determinants of obesity, type 2 diabetes, and cardiovascular disease, is the 2006 paper (Malik, Schulze, and Hu 2006) [2,660] which reviews the literature on the relation of sugar-sweetened beverages and overweight in the long interval of 1966-2005. As mentioned earlier, the most HC paper in *IEEE Trans. Signal Process* (Arulampalam et al. 2002) [11,860], is also a tutorial.

Firsthand Approach to the Problem

The second property of almost all highly-cited papers is firsthand approach to the problem they study. For this, out-of-the-box thinking is crucial. Firsthand ideas for influential research can almost never be inspired from already hot topics, where; surprisingly, most researchers look for around. A hot topic is simply a consequence of years of transformative research by some pioneering scientists leading to breakthroughs that have then attracted an extremely large number of other researchers, along with their overwhelming volume of further work on the subject. It would therefore be far from expectation for one to succeed to add something essentially new to such a topic. A subject becoming hot in fact somehow implies that the time for contributing anything noticeable to it is almost over. Geim (2011) confirms this in his Nobel lecture, where he states that his first publishing experience on thin films taught him that introducing a new experimental system is generally more rewarding than trying to find new phenomena within crowded areas. The chances of success are much higher where the field is new (Geim 2011). He also adds there that while in the early 2000's when nono-tubes were at their peak of glory, although nice results in this field triggered him to enter it, he was “too late”, and needed to find a different perspective “away from the stampede”(p. 5).

The faulty cycle of spending over-justifiable time on literature review rather on searching for new phenomena is also well-described by Geim. As he states, many researchers fail to live up to their promise because instead of spending their time for the latter, they waste it doing the first (Geim 2011). After months of literature search, they then, as explained above, come to the conclusion that everything they had planned for has been done before. They thus see no remaining reason to try their own ideas, and; consequently, begin another literature search (Geim 2011). This is while spotting essentially-new scientific directions is much the same as starting of highly successful businesses, where the next Bill Gates will not build an operating system, the next Larry Page or Sergey Brin will not make a search engine, and the next Mark Zuckerberg will not create a social network, as mentioned by Thiel in the preface of Masters and Thiel (2014). If one is copying these people, he is not learning from them (Masters and Thiel 2014).

Great researches have their very first roots in other, and sometimes seemingly unrelated sources instead. In fact, the old-fashioned hall of fame still has value (Van Noorden, Maher and Nuzzo 2014). To make exciting advances, HC researchers generally rely on relatively unsung papers to conduct new works based on Van Noorden, Maher and Nuzzo (2014). One reason for this is that local circumstances as well as new technologies offer a reasonable chance for old failed ideas to work unpredictably well at the time being (Geim 2011). Another reason has roots in the process of creativity in humans. Every glossy idea is almost always shaped on previous knowledge. The differentiating factor is “where” one looks for this knowledge in. Considering the countless brilliant talents across the globe, if the pool of literature on a hot topic is selected for this purpose, the inspired idea is unlikely to be a firsthand one not have been discovered by anyone else before. The conditions; on the other hand, dramatically change if the corresponding knowledge would be obtained from a less-crowded area. An idea inspired from there can then be checked with the recent literature to find the best spot for it to be stuck to the currently-trending scientific subjects for its utmost influence.

This type of out-of-the-box thinking has been the very first step in almost all exceptionally popular scientific works. The discovery of diamagnetic levitation that had previously earned Michael Berry and Geim the 2000 Ig Nobel Prize; e.g., was triggered by the extremistical act of Geim pouring water inside the electromagnet of his lab when it was working at its maximum power (Geim 2011). Although he confesses about this to have undoubtedly been unscientific and unprofessional (Geim 2011), the point here is of course his underlying firsthand approach to the matter. This experiment also taught him that poking in directions far away from one's immediate area of expertise could lead to interesting results, even if the initial ideas were extremely basic (Geim 2011). Based on his opinion, serious work usually requires many months of digging through irrelevant literature without any clear idea in sight, which would eventually give a researcher a "feeling" - rather than an idea - about what might be interesting to explore. A researcher will next give it a try, which would normally fail. The most difficult step in this process is to decide whether to continue on the subject or start thinking on something else (Geim 2011). Geim's 2003 work on Gecko tape (Geim et al. 2003) [1,350] published in *Nature Materials* ; for instance is itself a result of his maybe-accidentally; as he hints, reading of the 2000 paper (Autumn et al. 2000) [2,340], which describes the mechanism behind the amazing climbing ability of geckos (Geim 2011). All put together, the main conclusion of his Nobel lecture in this context is that poking in "new" directions, even randomly, is more rewarding than is generally perceived (Geim 2011).

a) Graphene research, compressed sensing, and network science

To provide a deeper intuition about the crucial position of firsthand approaches in successful researches, in what follows are some details how this has led to booms in the three fields of graphene research, compressed sensing, and network science, being from the different areas of experimental, theoretical, and mixed (including both experimental and theoretical aspects) sciences, respectively.

Graphene research: The history of Geim's Nobel Prize-winning contributions on graphene, for instance, themselves root back to 1859 and the work of the British chemist Benjamin Brodie (Brodie 1859). Studies on the topic by different scientists progressed in the next decades. In 1962, Ulrich Hofmann, Hanns-Peter Boehm, and their other colleagues looked for the thinnest possible fragments of reduced graphite oxide and identified some of them as mono-layers (Boehm et al. 1962). Tremendous volumes of further research was then conducted in the field as well, but what ultimately made the big graphene boom was the firsthand golden difference the 2004 *Science* paper (Novoselov et al. 2004) [46,370] of Novoselov, Geim, and their other co-authors had with all its predecessors. While all the existing works were only observational experiments that observed ultra-thin graphitic films and sometimes even mono-layers without reporting any of graphene's distinguishing properties, what started the 2004 trend was its electronic properties studied and the related measurements presented in this paper.

Compressed Sensing (compressive sampling): The firsthand approach paradigm is not only limited to experimental scientific advances, but also is seen in some highly theoretical breakthroughs. An excellent example for this is Compressed Sensing (CS), a topic in computer science and signal processing that emerged in early 2000's which studies reconstruction of signals such as high resolution camera images using only a few registered (captured) pixels. In fact, while it has been traditionally thought; and the reader may already think of as well, that to compress a mega-pixel digital camera image (for instance, one in PNG format) for one with a still good resolution but having only hundreds of kilobytes of size (in a format such as JPEG), it has first to be captured and then compressed (with one of the available softwares). What CS says is that why should in essence millions of pixels be registered and saved initially when most of them are to be discarded later in compression? In answer, it proposes an alternative imaging method that captures (senses / samples) and saves only some hundreds of kilos of specifically chosen pixels from the original scene at the beginning, and then deploys them offline to reconstruct a high-resolution image of it via a mathematical algorithm. The notably lower number of pixels needed to be registered initially would be a great advantage in applications such as medical imaging, where it considerably reduces the duration the patient is exposed to radiation.

The theoretical foundations of this topic root back to as early as the 18th century with the work of Prony (1795), then to the 20th century with papers of Caratheodory (1907; 1911), and then to the 2000's by contributions such as Vetterli, Marziliano, and Blu (2001) [1,050]. The main firsthand approach to the problem; however, was the replacement of the conventional l_2-norm (a mathematical gauge for evaluation of the size of a signal[1]) for measurement of the reconstruction error with l_0, and then approximation of the latter with l_1, proposed by the works of Donoho; a 2018 HC computer scientist, Candes, Tao (both being 2018 HC mathematicians), and Romberg around 2006 in papers described in the following paragraphs, which collected thousands of citations for them.

A brief glance at David L. Donoho's scientific background on the topic could be particularly instructive. Being a professor of statistics at Stanford, he is the single author of the 2006 paper (Donoho 2006) [23,550] which formally proposed the theory of compressed sensing for the first time, and is now the most highly-cited paper in *IEEE Transactions on Information Theory*. Donoho states in his autobiography, presented on 23 September 2013 in Hong Kong in a Shaw Prize event (see www.shawprize.org), that he has been aware of the l_1-norm approach since 1978 when he was a Bachelor of Science student at Princeton, where his undergraduate thesis advisor (John Tukey - inventor of the Fast Fourier Transform and the words "bit" and "software") deployed it for purposes such as equation fitting. Part of the art of Donoho, therefore, has been spotting of the perfectly-fit context of signal and image reconstruction to apply this technique on, which ultimately led to the above foundational paper. Another notable portion of the impressive popularity of this work has also absolutely to do with its proper timing. It was presented right at the very beginning of the big-data era, in which efficient data processing has become a serious issue of concern among scientists, calling for methodologies such as compressed sensing.

Analysis of the professional career of Emmanuel Candes in the CS framework is also pedagogical. Candes received his PhD in statistics and non-linear approximation theory in 1998 under the supervision of Donoho. His thesis was concerned with highly theoretical mathematical methods for signal and image processing, which of course have countless areas of emergence in the non-mathematical practical world. This firsthand strategy, which was rarely followed in mathematics departments at the time, finally resulted in the following papers - Candes and Tao (2005) [6,600], Candes, Romberg, and Tao (2006)

---

[1] For a vector $\boldsymbol{v} = (v_0, v_1, \ldots, v_{N-1})$, the l_p-norm is defined as $(\sum_{i=0}^{N-1} |v_i|^p)^{1/p}$.

[14,680]; Candes and Tao (2006) [6,680]; Candes, Romberg, and Tao (2006) [6,450], and Candes and Wakin (2008) [8,470], as some of the main foundations of CS when he was a professor of computational mathematics at Caltech. The paper Candes, Romberg, and Tao (2006) is the third most highly-cited article in *Transactions on Information Theory* as well. Considering its deep mathematical concept, Compressed Sensing could probably never be proposed by any electrical or computer engineer or computer scientist as the main readers of this journal, and *IEEE Signal Processing Magazine* where the other four of this five works have been published in.

Achieving of such eye-catching successes by the above two formerly professor-student scientists is of course not accidental. Besides many other factors, this has strong reasons particularly in the research habit in the Statistics department at Stanford. Proposing methods that non-statisticians are not generally aware of, to solve problems in their diverse fields of research, is the firsthand approach behind all these works indeed. Besides the papers of Donoho and Candes discussed earlier, this paradigm is also seen for instance in those of Jianqing Fan, a highly-cited professor of statistics at Princeton and another former Ph.D. student (1989 - Berkeley) of Donoho as well. Fan has applied statistical methods to solve a wide range of problems in finance, economics, risk management, machine learning, computational biology, and biostatistics, that have led to popular publications such as Fan and Li (2001) [5,630] and Fan (2018) [5,880].

Network Science: Another insight-providing case which is worth being scrutinized here and in some next parts throughout this paper as well, is emergence of the whole field of network science, a discipline that started to shape around 2000 whose subject is intricate networks and interactions of their elements with each other Barabási (2016) [1,130]. One of the pioneers in this area is Albert-L. Barabási, a professor at Northeastern University, whose arrow of progress in the field contains many instructive events. After his first network science paper (Barabási 1996) that ties the two disciplines of physics and computer science by demonstration of equivalence of the network problems studied in the two, Barabási faced years-long failures Barabási (2016). His second manuscript, “Dynamics of random networks: connectivity and first order phase transitions”, was inspired by his imagination of the remarkable complexity of the electric, telephone, and Internet cables under the Manhattan's pavement Barabási (2016) - an issue that almost all people were aware of, but became the inspiration source for a scientist only when looked at from his different and non-daily perspective. Although this manuscript was of course rejected in the 1995-1997 interval by four journals and never got published thereafter either, the challenges in the process of getting this work published thought Barabási that besides selecting of an original research topic, success calls for a firsthand approach to the main body of the work as well Barabási (2016). He was in fact convinced to stop his conventional graph-theoretical approach to network science, and to follow an original path instead. For him, this was to continue to get inspired from the real-world in the main body as well - a skill that physicists; including him, are generally good at Barabási (2016).

In this framework, Barabási decided to use maps of real networks in his research. At that time (1996) the World Wide Web (WWW) had just started to boost, and a smart idea was indeed to select it as the subject network. His trials to obtain third parties data for this purpose however, failed. What eventually motivated him to retry was again a firsthand approach, this time inspired by Isaac Asimov’s science fiction trilogy, on the basis of which he decided to formulate a set of equations to predict the future of a system as complex as the society Barabási (2016). He finally succeeded to overcome his problem with the WWW maps via a web-crawler designed by H. Jeong, a post-doctoral researcher graduated from the top Seoul National University of Korea, who had joined his group in 1998 Barabási (2016). The data showed that the node degrees in the WWW had a power law distribution. While in almost all random graph and social network literature of the time the probability density function (pdf) considered for such networks was the Poisson distribution, this discovery was undoubtedly a true breakthrough. This was reported in the now highly-cited paper (Albert, Jeong, and Barabasi 1999) [5,440] published in *Nature*.

The next big discovery of his team was inspired from analyzing the highly unrelated networks formed by (a) the wiring diagram of an IBM computer chip provided to them by Jay Brockman; a computer science professor, now at the University of Notre Dame; (b) a map of the power grid from Duncan Watts, a sociologist currently at Microsoft Research; and (c) Hollywood actors database by Brett Tjaden, also a professor of computer science now at James Madison University. Despite being highly unrelated, tails of the degree distributions of all these networks showed a power law pattern similar to that of the WWW as well. This was discovered by R. Albert, a Ph.D. student of Barabási whose paper on granular media had made the cover of a *Nature* issue a couple of years before Barabási (2016). Inspired from this firsthand unifying pattern in such highly scattered domains, they proposed a model explaining this behavior based on the two principles of growth and preferential attachment in the 1999 *Science* article (Barabási and Albert 1999) [33,480], which became the most highly-cited paper of physical sciences a decade later.

b) Incomparable attractions of pioneering and span-off researches

To highlight from yet another instructive perspective the far distance between the attractions of pioneering works and the ones span-off from them later even by the same researcher, consider the papers of P.M. Ajayan, another of the 24 introduced earlier and a pioneer in the development of carbon nano-tubes, depicted in Figure 1. His most-cited paper is the 1992 3-page *Nature* article (Ebbesen and Ajayan 1992) [4,220] that reports the synthesis of graphitic nanotubes in gram quantities. His 2nd HC paper is another breakthrough (Ajayan 1999) [3,690] published in *Chemical Reviews*, solving some problems in effective dispersion of individual Carbon nano-tubes in polymer matrices. These two documents; as shown in Figure 1, have on the order of two times the third highest citation count (2,020) of the same author, being for Gao et al. (2009).

This phenomenon is seen in the majority of other top-cited research profiles as well. For Zhenan Bao, another member of the 24 with 83,860 total citations; her major scientific impacts root back to her works as a postdoc in the very early development of the first all-plastic transistor or organic field-effect transistors in Bell Labs and Lucent Technologies, which then continued by studies on the related areas of organic semiconductor and carbon nano-tubes. Her 2008 ACS Nano paper (Becerril et al. 2008) as one of the corresponding outputs which is also her most-cited work, has received 2,790 citations. Her second top-cited contribution is the 1996 paper (Bao, Dodabalapur and Lovinger 1996) [2,020], and after her 3rd (Lipomi et al. 2011) [1,780] and 4th (Mannsfeld et al. 2010) [1,560] HC works, her 5th-rank article (Crone et al. 2000) [1,380] has received, however, only on the order of half the citation count of the first.

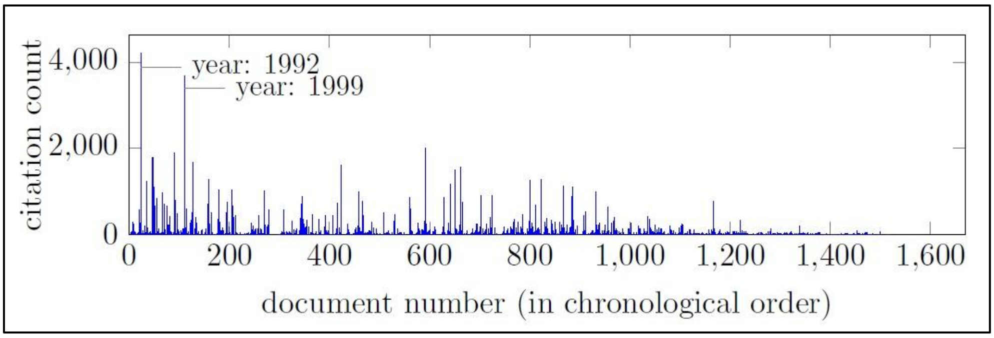

Figure 1: Citation counts of P.M. Ajayan's papers based on Google Scholar (1,517 documents)

Yi Cui in the same list also works on nano-scale properties of materials, where he has made pioneering early contributions to nano-wires when being a PhD student at Harvard. Similar to the counts of Ajayan and Bao, Cui's citations shows a highly-skewed pattern as well, with their first (Cui et al. 2001) and 5th (Huang et al. 2001) ranks having notably distant citation counts of 6,600 and 2,660 respectively.

For Hongjie Dai in the list, whose field is carbon nano-tubes as well and has presented breakthrough researches on NIR-II dyes and plasmonic gold (pGOLD), although the citation pattern falls with a much lower rate, the first and 5th highest citation counts being for Thess et al. (1996) [7,260] and Liu et al. (1998) [3,960] respectively are also still considerably spaced.

10X-Improvement Solution

Contributions to important topics, even via firsthand approaches, will not be much recognized unless if they would offer something being on the order of 10X better compared to their existing alternatives. A three or even 5X improvement may not likely be enough. This is because such levels would be depreciated by some other, generally unpredicted, factors that will emerge later maybe when the solution is put into practice. For instance, a computer algorithm that has been proved to be 5X faster than its existing counterparts mathematically, may eventually turn out to outperform them by only; say 0.2 times, due to the limited computational power and numerical precisions of the state of the art hardwares. No one would of course probably switch from a well-established and working algorithm to a new one for only such a negligible excellence, and the discovery will thus fail. The improvement level has therefore to be especially taken care of.

A good case for a work that offers such a 10X improvement solution (or even more than) is the 1991 4-page *Nature* paper (O'regan and Gr¨atzel 1991) [29,570] of Michael Graetzel; also in the 24, who works in the areas of photonics and interfaces. This breakthrough article of Graetzel describes a photo-voltaic cell created from low to medium-purity materials through low-cost processes which also exhibits a commercially realistic energy conversion. While each of these three properties alone would likely boom a product in the energy industry, it is needless to explain about the popularity their synergy would bring about.

a) Complementarity (inter-disciplinarity) instead of substitution

Offering a 10X improvement in a classic field, where thousands of scientists have worked on well-known topics for decades through more or less the same approaches, is generally far from expectation. And this is the reason for inter-disciplinary domains becoming more popular instead. This is much like the path successful businesses pave. Thiel says in this context that such businesses do not try to compete to substitute already-existing products with their owns, maybe with some minor improvements. Such competitions, in his view, consume a startup's time, energy, and other resources without paying off much. Successful startups, instead, tend to complement what already exists by what they themselves excel in some sort of multi-disciplinarity (Masters and Thiel 2014).

Lots of good case studies for the tremendous power of interdisciplinary research in production of highly improved and upgraded science can be provided. Among them, are the major progressions of network science which were mainly due to the synergy between mathematicians with expertise in graph theory, sociologists and their social network approaches, the universality concept of physicists, observations in biology, algorithms in computer science, and engineering Barabási (2016). The software offered in Allison et al. (2006) is also another cross-field work that has found applications in high energy physics, astrophysics and space science, medical physics, radiation protection, astronomy, and engineering. It has also been frequently cited by articles in medicine, where the toolkit has been used to track the effect of materials on the human body (Tendeiro et al. 2014). Mc.Call's paper (McCall 2008) is a trans-disciplinary one as well. It has been highly-cited not only in the field of arts, humanities, and social sciences in the framework of gender-related psychology, ethnic identity, and feminism, but also in areas such as business management in the contexts of women's careers in business (Jyrkinen 2014), workplace diversity (Kokot 2014), and women's leadership skills development (Levac 2013), or in environmental sciences in frameworks such as gender-related climate change (Carr and Thompson 2014), and gender migration patterns (Ju´l´usd´ottir, Skaptadottir, and Karlsdottir 2013). Mc.Call explicitly refers to the interdisciplinary nature of this work as one of the reasons for its high popularity as well. Besides its main fields of engineering and mathematics, the paper (Lowe 2004) discussed in the same part has also been frequently used in areas as diverse as health for organ imaging (Wu et al. 2014), decision sciences (Liu, Liu, and Sun 2014), and even social sciences to track the processing and interpretation of images such as street signs by humans (Laeng et al. 2014). The article by Kreuer (2001) has been popular in as various fields as chemistry, materials science, chemical engineering, and energy. The work by Novoselov et al. (2004) has found applications in materials sciences, chemistry, energy, pharmacology, and computer science as well, and has been used even in social sciences in papers such as those by Arora et al. (2013), Lehman (2011), and Pham and Fayerberg (2011).

Monopoly

A monopole work is referred to one that dominates a big portion of its corresponding region of applicability, or in business terms its market, because of some particular aspects that distinct it from its counterparts. For research, this is equivalent to the solution offered in the paper being an extremely superior option for readers (customers) to consider, when compared to the other ones in its class. To be influential, a paper has to be sufficiently monopole.

The quaint point here is right spotting of the target market. Inspired by Thiel's example in Masters and Thiel (2014), suppose, for instance, a research in the field of solar panel systems offering a solution for 10 Gigawatts of power generation capacity. If it would be proposed in a US-context with 67 GW of solar energy production capacity (based on the 2018 statistics of the U.S. Energy Information Administration - EIA), the 14.92% of the market it intends for is of course a good portion to excite the readers. But if the relevant context is that of the entire globe with over 500 GW of solar production capacity, the share of the paper's solution will drop down to 2%, seriously decreasing its importance. Even worse would be if the target is global renewable energies in general with an amount of 2,350 GW (based on the 2019 statistics of the International Renewable Energy Agency), shrinking the work's share further to 0.42%. And finally, if the research is to be presented in the framework of global power generation in any form with 18,300 GW of capacity, it is of course already knocked out.

Monopoly in research can be achieved in many different aspects. The most common form of it is perhaps the scientific one obtained by the new doors the idea, solution, or results presented in the paper open, the technological barriers they remove, and the fresh visions they sketch. What made the compressed sensing theory discussed previously monopole; for instance, is the minimal number of pixels it captures at the shooting (sampling) stage, and reconstruction of the actual desired image offline at the far end of the process. This has brought about monopoly for CS in applications such as medical imaging, where minimal exposure of the subject to the imaging equipment is crucial. In such applications neither of the existing compression competitors such as JPEG, wavelet-based methods, etc., have intrinsically anything to offer in front of it.

A research can also have scale monopoly. A case with a high explanatory value here has to do with the two big population-based prospective cohort epidemiologic studies of Albert Hofman, a scientist in the aforementioned 24, performed in Rotterdam - the Netherlands, whose data collections started in 1990 and 2002 respectively. As stated in Hofman's Harvard web-page, what they both have in common is that they target multiple common diseases, have a very extensive and state-of-the-art assessment of the putative determinants of these diseases, and employ as much as possible new technologies to be applied in the setting of epidemiologic population studies. The wide span of these projects being a big city, their long duration, multiplicity of the diseases they cover, their extensive assessment, and the overwhelmingly vast and modern technologies deployed in the studies have made them being monopole in scale, being rather impossible; at least in the short term, for others to run ones similar to them. This monopoly has led to a rich sequence of highly-cited papers such as Bots et al. (1997) [2,780], Hak et al. (2000) [1,730], and Heeringa et al. (2006) [1,720] to refer to a few. The paper by Kolpin et al. (2002) discussed earlier enjoys a similar form of monopoly as well.

Besides the two forms of monopoly, another important type of it that sometimes helps a paper stick is that of its authors. A paper by a well-known and highly-cited, or in brief; monopole, author will probably face fewer obstacles at the first step from the Editor and reviewers of the journals to which it is submitted, and will also likely catch the eyes of more prospective readers later as well. The impressive popularity of the paper by Albert and Barabási (2002) [21,200] for instance, which formalized the intellectual foundations of network science and became the most-cited paper in physics and astronomy later, is probably related in part to the monopoly they were enjoying after their outstanding previous pioneering contributions to network science to be explained later.

Monopoly may boost the progress of a whole scientific field as well. Three years after publication of Albert and Barabási (2002), monopoly of the studies conducted on networks in those decades led to the specific term "Network Science" officially assigned to it in the 2005 National Research Council report published by the US National Academies, which then convinced the US government to invest tremendous amounts of money on it as well

Barabási (2016). Today, highly recognized global academic publishers such as Cambridge and Oxford University Presses as well as the IEEE publish specific journals in the field, and many conferences on the topic are also held annually across the globe.

Durability

Another central property of exceptionally highly-cited papers is that they have remained attractive for scientists over long times to be read and cited; i.e., they have been durable. A fine intuition about the central role of this property can be obtained by scrutinizing the citation profiles of HC scientists in fields such as computer science or electrical and communications engineering, in which a widely seen phenomenon is the seldom existence of more than one or two papers having some thousands of citation counts. Researchers in these areas have generally become highly-cited by publishing hundreds of papers, each having only some hundreds of citations instead. For Jinde Cao in the 24 with 44,380 total citations, the profile of his 10 most top-cited papers has been illustrated in Figure 2. It is seen that in contrast to those of Ajayan, Bao, Cui, and most of the other 23 researchers in the corresponding list, Cao's citations consists of notably lower counts falling with a much lower rate as well, representing a continuous chain of good, planned, and sequenced researches which have not; however, been much durable. The reason for this non-durability seems to be the fast pace of technology in such fields, which soon makes most purely-technological researches in them to a great extent obsolete, preventing them from receiving thousands-scale citations. This is while, as can be seen in the interactive plots in the online version of Van Noorden, Maher and Nuzzo (2014), durability of the top-100 papers studied there has led to their continuous earning of hundreds or sometimes thousands of citations yearly, resulting in their ultimate tens of thousands of overall citations records.

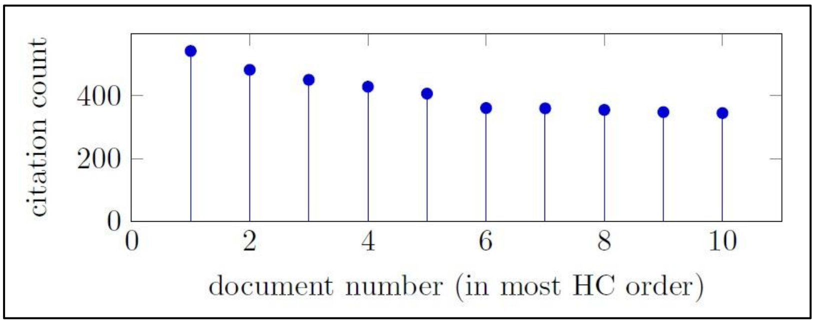

Figure 2: Citation counts of Jinde Cao's first 10 highly-cited papers

Durable research in the technological fields can be conducted almost only in areas dealing with long-lasting general-purpose theories, algorithms, and methodologies rather than; as stated earlier, only temporarily-popular purely technological results. A set of instructive examples for such works that have received thousands of citations are those of G.E. Hinton; another 2018 HC researcher in computer science and a winner of the 2019 Turing award, and his co-authors on the process of learning in neural networks. The 1986 paper (Rumelhart, Hinton, and Williams 1986) [17,300] that boomed the popularity of the back propagation algorithm in training of neural networks is, for instance, one of them. Another exceptional and also rather recent one is the 2012 paper (Krizhevsky, Sutskever, and Hinton 2012) which has attracted 39,970 citations. Ground-breaking durable works of Hinton and his collaborators in the learning process of neural networks and related topics have earned him 278,920 overall citation counts, demonstrating the crucial position of this parameter in a successful research plan in technology. As further high-quality works by others on the same durable topic that have been considerably influential, reference can be made to Huang et al. (2004, 2006, 2012) [2,680, 7,060, and 3,210] by G.-B. Huang; another 2018 HC computer scientist, and co-authors. Good examples of durable general-purpose theories, algorithms, and methodologies in engineering from 2018 HC computer scientists also include the works of Thomas J.R. Hughes and Dervis Karaboga in computational mechanics and numerical optimization respectively, and many other outstanding ones as well.

To investigate quantitatively the axial role of durability in scientific papers becoming widely popular, let us again consider the 100 most top-cited papers, whose corresponding yearly citation patterns; as stated previously, can be found in the online version of Van Noorden, Maher and Nuzzo (2014). Taking the generic rate of 100 annual citations for these papers as the threshold of being referred to as yet actively cited, those of their decisive majority of 80 ones are still above this level. Their first; Lowry et al. (1951) [> 305, 000], for instance, has enjoyed more than 10,000 citations per year around 1985 for almost a decade, as depicted in Figure 3. For the remaining 20, having the well-known bell-shaped citation curve, the time spans during which they had been as such are also notably wide. It is observed from Figure 4 that their 19th top-cited paper (Southern 1975) [35,420] for instance has enjoyed about 35 years of active citations. As seen in Table 1, most of these 20 works have received citations actively for at least approximately 30 years, three (numbers 32, 38, and 81 - in bold) even for 45.

All put together, the long-lasting popularity of these 100 and many other more recent papers is because of their content proving over years to be still live and useful - their durability.

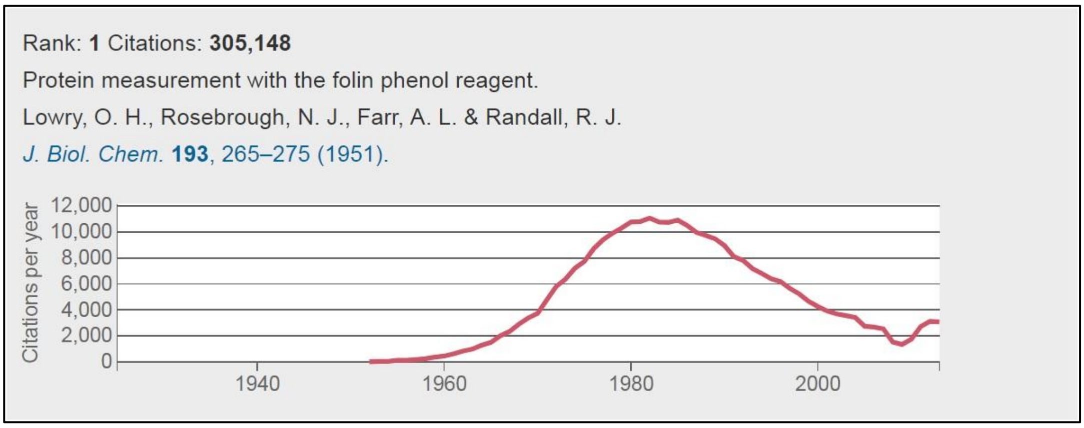

Figure 3: Yearly citations pattern of the most top-cited paper ever.
A majority of 80 of the top-100 papers are still being actively cited
(source: Van Noorden, Maher and Nuzzo 2014).

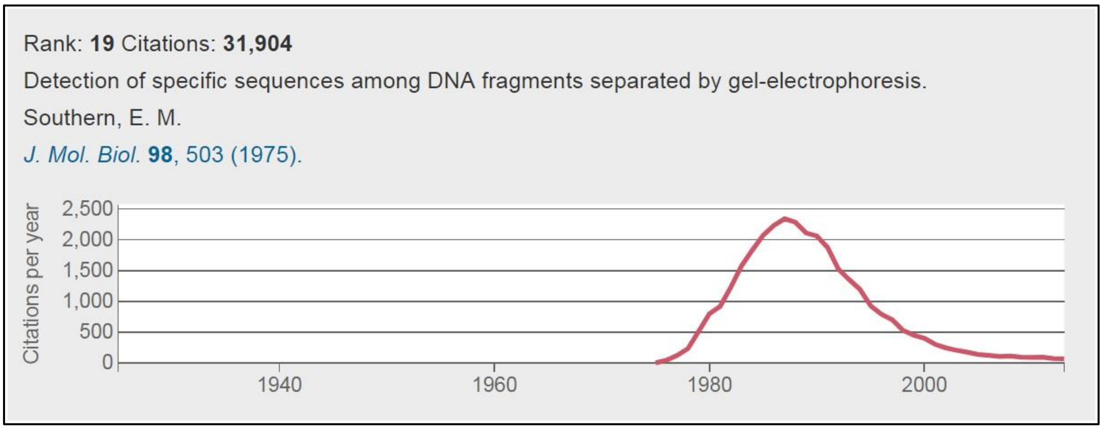

Figure 4: Yearly citations pattern of the 19th top-cited paper ever.
Although not being actively cited any longer, the minority of 20 of the top-100 papers ever have enjoyed this in the past for decades
(source: Van Noorden, Maher and Nuzzo 2014).

Table 1: Durations of being actively cited for the 20 papers out of the first 100 top-cited ones ever which are now less-actively cited.

Paper No.	19	30	31	**32**	**38**	40	44	50	52	63	67	74	75	78	80	**81**	82	87	94	95
Duration (years)	35	35	30	**45**	**45**	25	40	40	35	25	30	30	20	30	30	**45**	20	25	30	20

Proper Timing

Excellent research sticks only at its right time. Paper no. 11 of the top-100, (Kaplan and Meier 1958) [55,670], authored by the two statisticians Edward Kaplan and Paul Meier, for instance, is on a method to find survival patterns for a population, which introduced the Kaplan-Meier estimate. This paper; however, received almost no citations till 1970, as seen in Figure 5, when advances in computational power made the method deployable by non-specialists (Van Noorden, Maher and Nuzzo 2014). In another case, the idea of Bland-Altman plot for visualizing how well two measurement methods agree, which is the subject of the list's no. 29, Bland and Altman (1986) [42,930], by Martin Bland and Douglas Altman, was proposed by another statistician 14 years earlier as well (Van Noorden, Maher and Nuzzo 2014). The key for Bland and Altman (1986) to attract the citations however, was the accessible way of presentation it enjoyed as a result of the available computational power at the time (Van Noorden, Maher and Nuzzo 2014). This advantage has been likely among the success reasons of many recent heavily-cited works as well. Compressed sensing with its computationally-intensive algorithms, for instance, would unlikely have had the chance to receive such tremendous attentions if proposed 10-15 years earlier.

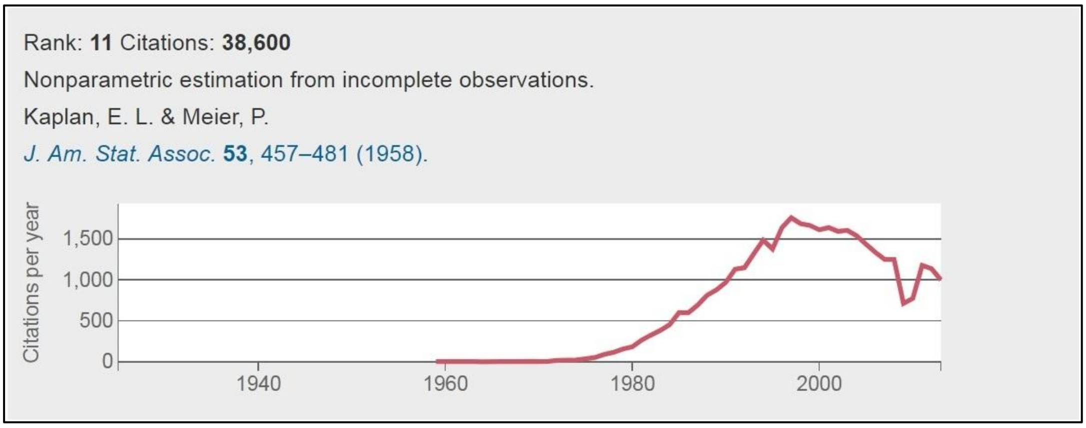

Figure 5: Effect of proper timing.
This paper received almost no citations till 1970, when advances in computational power made its proposed method deployable by non-specialists
(source: Van Noorden, Maher and Nuzzo 2014).

Proper timing is not limited to cases benefited from the endowments of computational power. For Barabási ’s first network science papers discussed earlier, proper timing was related to his works being fine-tuned with the Internet boom. Many other such cases can also be noticed by careful analysis of technological, social, economic, and similar trends. Kumar for example believes that one reason for softwares (Tamura et al. 2007) [31,620] continually becoming more important is the emergence of the age of genomics, when increasingly larger numbers of scientists use them to analyze their data. About Allison et al. (2006) [5,480], Asai (one of the co-authors) emphasises in this framework on the rapid expansion of the areas of high energy, nuclear and accelerator physics, and medical, space, and material sciences, where their software may have applications.

Research Team, Atmosphere and Resources

The late 1999 Nobel laureate, Ahmed H. Zewail, states in answer to a question about a formula for managing discovery making, that first, and most important, are the people involved. Second, that an atmosphere of intellectual exchange is of paramount importance for ideas to crystallize. And third, that without resources little can be achieved, no matter how creative the mind (Zewail 2010).

Properties of the research team are far more influential on the produced outputs than what is typically thought of. A recent study (Wu, Wang, and Evans 2019) shows that the output’s type, being developmental or disruptive, is closely related to the team’s size being large or small respectively Effects of the mind, knowledge, expertise, and creativity of the team members are also of course needless to be much elaborated on. In Barabási’s case, for instance, his trials for obtaining maps of real networks, as stated previously, were fruitless until Jeong’s expertise in computer science helped the group map out the WWW. Without this, as Barabási confesses in Barabási (2016), he might have never discovered the scale-free property. R. Albert’s mathematical skills were also crucial for developing of the corresponding theories Barabási (2016). Barabási states in the context of the importance of teamwork that without their collaboration “all I could produce was a string of ideas and failures” (Barabási 2016, p. 16). Applying of network theoretic methods to cell biology could also only be sparked in the group by the expertise of someone as Zoltan Oltvai, a medical doctor and researcher at Northwestern University.

Besides the team members, the atmosphere and environment it works in also affects its output quality considerably. A provocating case in this framework is related to Geim. A few months after he had wanted his new PhD student Da Jiang in 2002 to make as thin as possible graphite films via polishing, he realized that instead of highly-oriented pyrolytic graphite (HOPG), as was intended, he had mistakenly given Jiang high-density graphite to work on instead. Even if this mistake had not occurred, his intended process towards making the film was, as he states that of "polishing a mountain to get one grain of sand" (Geim 2011, p. 5). What prevented this inefficient approach from being continued was only the unbelievably simple alternative of Oleg Shklyarevskii, a senior fellow from Ukraine and an expert in scanning tunneling microscopy (STM) working nearby, who brought over to them a piece of cellotape just fished out from a litter bin with graphite flakes attached to it (Geim 2011). HOPG is the standard reference sample for STM, and a fresh surface of graphite is normally prepared by removing the top layer with a sticky tape. Although they had used this technique for years, they had surprisingly never looked carefully about what is thrown away with the tape.

Despite Shklyarevskii not being among the co-authors of Novoselov et al. (2004) [46,370], the impact of his hint in the ultimate publication of the paper has been history-changing. In August 2004, and before this article had been published yet, Philip Kim's group had submitted the 2005 paper (Zhang et al. 2005) [500], describing the electronic properties of ultra-thin (down to ∼35 layers) graphite platelets following the same approach as that of Novoselov et al. (2004). The difference of the two papers is mainly in the thickness of the studied devices. The similarities, on the other hand, are fascinating. A good sense of these similarities can be obtained by noting that after the deployment of the Scotch tape technique of Geim by Kim's group and their studying of mono-layers in early 2005, the two papers (Novoselov et al. 2005 [18,200] and Zhang et al. 2005 [12,590]) describing the observation of Dirac fermions in mono-layer graphene, submitted independently by the two groups to *Nature*, both got accepted, and appeared back to back in the journal. Besides this rare event, which also explicitly shows the reproducibility of the results in the two, the quality of Kim's works have been to such a high level that Geim has frequently expressed his terms of honor if he could share the 2010 Nobel Prize with him as well (Geim 2011). What made Geim's works more visible was indeed the Scotch tape technique of Shklyarevskii; the paper (Novoselov et al. 2004) officially introduced to the graphene research community for the first time, along with the many new experiments it made possible. This lucidly shows the importance of the research environment and people working around. Many more examples about the synergistic effects in great research teams can also be seen in other successful works.

Infrastructural and other such equipment and resources are also undoubtedly among the other prerequisites for a successful research. One of the major reasons for breakthrough discoveries, especially experimental ones, emerging less frequently in developing countries; for instance, is indeed the limited access of scientists there to the required resources and tools compared to that of their peers in developed regions.

Effective Presentation to the Right Audience

The term "presentation" here refers to a broad range of issues beyond merely the good writing of a paper from scientific aspects. It is instead pointing to the "art" of making the work make sense to the readers, as its customers, not only by its scientific merits, but also based on aspects more or less similar to those businesses take care of in marketing and selling.

The first point to be emphasized here is that what readers have interest in are not necessarily those the authors may think of beforehand. While a researcher in a technical field may think, for instance, that a sophisticated work on a hard topic containing lots of theories and formulas may catch the readers' eyes; they may prove later to be more intended towards a less- or non-technical - but highly instructive - paper in the field instead. As Tao (2008) states a month after he modernized his 1999 UCLA mathematical web page "What's new?" to a blog with the same name at wordpress (terrytao.wordpress.com) in Feb. 2007, the most widely read and commented article there was a nontechnical article "Quantum Mechanics and Tomb Raider", which had "absolutely nothing" to do with his own mathematical work. Spotting readers' needs and interests are thus among the first steps of a successful research, which should be taken care of all the way long. The conducted work has to be packed and presented finally in a form easily "sellable" to them as well. Tao phrases in his blog in this framework, for instance, that "you should sell your paper in its introduction".

To see how the above concepts have been applied in practice, consider the article no. 41 of the 2014 top-100; which is a 1984 paper by Felsenstein describing how to apply some statistical methods to phylogenies (Felsenstein 1985) [37,890]. He says that the bootstrap concept (Efron 1979) [17,620] proposed in 1979 by Bradley Efron, the statistician at Stanford University quoted previously, was much more fundamental than his work. Applying it to a biological problem; however, will attract many more citations for the author (Van Noorden, Maher and Nuzzo 2014). This simply refers to spotting the proper audience for a work, which is a component of a well-planned presentation strategy. Part of Felsenstein's outstanding citation count also has roots in his decision to include all his findings in one paper rather than to publish them separately, as reflected in his statement "I was unable to go off and write four more papers on the same thing . . . I was too swamped to do that, not too principled" (Van Noorden, Maher and Nuzzo (2014), p. 3). This compact presentation strategy, though unintentional, has led to several quality ideas being offered in a unified pack. The result has made life noticeably easier for the readers, directing even more citations to the author's basket.

Becoming highly-cited in statistics by spotting the right audience has took place for many other papers as well. Most popular statistics papers in the top-100 are in fact, not at all those that have been most important to statisticians, as remarked by Stephen Stigler, a statistician and expert on the field's history at the University of Chicago in Illinois (Van Noorden, Maher and Nuzzo 2014). Stigler continues that they are instead the ones useful for practitioners from biomedical labs and elsewhere (Van Noorden, Maher and Nuzzo 2014). This philosophy seems to be also working for recent highly-cited works in the field as well. The compressed sensing papers of Donoho, Candes, Tao, and Romberg (referred to in Graphene research, compressed sensing, and network science) are a few confirming cases, for instance.

After spotting of the right audience, good organization and effective narration are the next major ingredients of an influential presentation. Switching back to network science, after the paper by Watts and Strogatz (1998) [38,690] on small networks was published in *Nature*, Barabási believed that the reason behind his second network science manuscript never getting published was not much related to its science, but instead mainly had to do with its presentation and framing Barabási (2016). There, he had asked questions interesting for physicists, while narrating the work for neuroscientists. This was while the work of Watts and Strogatz, rooted on the same ground of random networks, was well framed in sociology and the well-known subject of six degrees of separation there Barabási (2016). Having this in mind, after his team discovered the power law distribution of WWW links in 1998, he presented the finding in the same framework of six degrees with the smart title "Internet: Diameter of the Work Wide Web" (Albert, Jeong, and Barabasi 1999) [5,440], in which he explained that 6 is in fact 19 on the Web. He refers to this smart decision as a "Trojan horse" that helped their work on the topology of the WWW, which might had turned out to be too abstract to excite *Nature*'s Editors, finally get accepted Barabási (2016). The importance of a good presentation is also highlighted in Geim's Nobel lecture. He phrases there, when explaining about the excruciating publication process of their 2004 *Science* paper (Novoselov et al. 2004), that they continuously added data and "polished the presentation" (Geim 2011, p. 6).

Thiel says that people are always selling something, irrespective of what their job actually is, and even if they do so implicitly. An office secretary for example is selling her time along probably with her ICDL (International Computer Driving License) and public relation skills, for a monthly salary (Masters and Thiel 2014). The same is true for researchers. They are selling their intellectual and innovation capabilities in the form of scientific papers mainly for scientific credit, self-satisfaction, and other sorts of spiritual reasons, along with a typical level of financial remittance from their university/institute or project employer. Doing this successfully calls indeed for a minimum knowledge of skills such as marketing. Thiel adds in this framework that despite the importance of this issue, most people educated in engineering and similar disciplines, however, mistakenly think that the works they do are so great that do not need to be thought of from such perspectives (Masters and Thiel 2014). They in fact think that since for instance, gravity is a global and important component of the universe, a work on this subject will naturally be far bigger and more important than to need any marketing. And this will eventually lead to their failure sooner or later (Masters and Thiel 2014).

Another somewhat important ingredient of a good presentation package, besides the paper and its organization and wording, is post-publication exposure of the work to press, interviews, blogs, social networks, and other such media. Introducing of the research results via these sorts of media in a less-scientific, but understandable, language to the general public would make a sensible public awareness about the made advancement. The resulted circulation of the corresponding news in the society will have positive effects on the attention of the scientific community surrounding the work, in particular.

Deep Interest, Belief and Perseverance

Success, after all, is not achieved easily. Progress is made in a highly gradual and also nonlinear manner. Countless rises and falls would appear on the path. In science particularly, a researcher may sometimes feel that his efforts seem to be fruitless, and his peers are doing definitely better. These sorts of issues are, of course, inseparable companions of research, and should be learned how to be overcome by deep interest and belief in the work, and perseverance. As Geim (2011) states, perseverance allows one to progress much further.

Flashing back to Barabási again, for instance, his second manuscript was rejected by four journals during 1995-7, and, as mentioned earlier, was never published thereafter either, even to date. Its first submission in 1995 was declined by *Science* to be sent for review

Barabási (2016). It was next rejected by *Nature*. Its third and fourth submissions to *Physical Review Letters* and *Europhysics Letters* in 1996 and 1997 respectively led to the same result as well Barabási (2016). His deep interest in the subject however, did not let him get disappointed, and he started to continue after a while Barabási (2016). Similar to this unpublished manuscript, his paper (Barabási and Albert 1999) [33,480], which as mentioned earlier became the most highly-cited paper in physical sciences, was not at first even sent by *Science* for peer review either. His perseverance and belief in the work however, made him call the corresponding Editor and convince him about the importance of the work Barabási (2016). He surprisingly succeeded, and this opened doors for publication of this pioneering article.

Barabási's notable interest and passion finally persuaded him to completely switch from quantum dots, which was then his main field, to network science in 1999. Based on a decision he refers to as a "Catch 22" (a satirical novel by the American author Joseph Heller), he even returned the research grant awarded by the US National Science Foundation (NSF) received to work on a project in materials science that was only incremental in his view, to follow his dreams, and work, with no funds, on the network science problem he believed to be potentially transformative Barabási (2016). At the time, network science was not considered by any funding agency a proper field to fund. Sometime later, he noticed a call from the US Defense Advanced Research Projects Agency (DARPA) for technologies that would "allow the networks of the future to be resistant to attacks and continue to provide network services" (Barabási 2016, p. 14). While he was aware that addressees of this call are generally networking experts and computer scientists, he on the other hand believed, based on his deep knowledge about the essence of networks, that no fault-tolerant network could be designed without first understanding its topology and vulnerabilities Barabási (2016). His group thus started to prepare a related proposal. R. Albert simulated component failures by randomly removing nodes from the subject scale-free networks, and compared these failures to an attack, in which only the hubs are removed. The main result was that scale-free networks were highly resistant to random failures, but vulnerable to attacks. They added this to the proposal and submitted it Barabási (2016). Barabási's full confidence about the importance of this discovery convinced him not to wait for DARPA's funding to work on it, and to start right at the moment instead. They conducted some more research on the subject, and sent a derived manuscript to *Science*. This manuscript however, was not sent for review again, and calling the Editor for the second time did not work either Barabási (2016). They thus resubmitted it to *Nature*. A few months later, the DARPA project was rejected as well. The good news, however, was that *Nature* accepted the paper (Albert, Jeong, and Barabasi 2000) [8,260], and put it on its cover. Two years later, Albert and Barabási published their aforementioned paper (Albert and Barabási 2002) [21,200] formalizing the intellectual foundations of network science, which as stated earlier, then became the most cited paper in the fields of physics and astronomy.

In those years, and in research groups having belief and perseverance, the strong trend started anyways. The mathematics behind the scale-free property was improved and made exact in works such as Dorogovtsev, Mendes, and Samukhin (2000) [1,450], Krapivsky and Redner (2002), and Bollobás et al. (2001) [840]. Robustness was connected to percolation theory in Cohen et al. (2000) [2,240], which was then made exact in Bollobás and Riordan (2004) [310]. A continuum of papers such as Pastor-Satorras and Vespignani (2001) [5,200] next studied the many effects of the scale-free property on network behaviors. The field and its community gradually formed, and related university courses started to be instructed world-wide.

As the last point here, it is notable that unfortunate rejection of papers which would then be accepted elsewhere and turn out to become highly influential as well is not a rare accident at all. Besides those of Barabási referred to earlier, Geim's Nobel Prize-winning paper (Novoselov et al. 2004) [46,370] for instance, is another case for the misdeed of some referees making unprofessional comments on manuscripts without carefully analyzing them to spot their real value. This paper was first submitted to *Nature*, where it was rejected. It was rejected again as well after the further information requested by the referees had been added to it, with one referee's comments stating that the manuscript did "not constitute a sufficient scientific advance" (Geim 2011, p. 6). In Geim's view, these types of unprofessional reviews should not disappoint the authors. Their paper may also be prize-winning (Geim 2011).

SOFT CITATION AND PROBLEM-ORIENTED INTERACTIVE RESEARCH

After the discussions in the previous sections about different ingredients of an influential piece of science, another central issue to be elaborated on as well is that of the publishing systems through which these sciences are disseminated. Efficient publishing systems that broadcast research outputs in an agile and effective manner on one hand, and also make the assessment of their impact in their corresponding fields as straightforward as possible on the other, are inseparable parts of a well-organized research ecosystem indeed. The existing ones however, suffer from some serious drawbacks from these aspects resolving of which would greatly facilitate real "one"s coming out of "zero" much more smoothly. Here two updates are proposed; one for the currently used citation tradition and the other for the state-of-the-art publishing system, that seem to be good complements for their already existing counterparts in this regard.

Soft Citation: A Human-judged Complement to Quantitative Informetric Indices

Despite the deployment of citation counts throughout the discussions for providing a rough sense about the different levels of popularity of the papers referred to, it is in fact believed that evaluation of science is a highly subjective matter that cannot be reduced to such numerics. This type of indices seem more logical to be considered merely as first-estimate tools for this purpose instead. They should not be over-relied or over-expected.

Over-expectation, a major threat that has affected almost all IT-based technologies and methods, is the tendency to get more out of them than what they have been devised for and can actually offer. An illustrative example for this tendency, as Thiel well explores in Masters and Thiel (2014), is the one in computer science, and machine learning in particular, that tries to replace humans by computers in all tasks. This is however, not the right direction to follow. While computers are of course better than humans in computation, when it comes to cognitive tasks such as face recognition, even a child will definitely beat a super-computer with tens of thousands of CPUs. The right strategy in this framework is therefore complementarity rather than substitution, which has yielded outstanding results, for instance, in Palantir Technologies - a highly successful IT startup co-founded by Thiel, whose business is in areas such as health and banking. There, parts of the works that are best performed by computers are naturally assigned to them, leaving the portions humans are more effective in to be completed still by human resources (Masters and Thiel 2014).

Over-expectation from IT platforms has affected the current state of the art citation system as well. In this system, informetrics analysts tend to judge paper qualities based mainly on the number of citations they have received via indices such as journal impact factor and h-index, and almost irrespective of the kind of the usage they have had in the referring document, the context they have been referred to, and other such qualitative aspects that cannot be indexed by machines essentially and have to be judged by human expertise. This system in fact suffers from two major drawbacks that degrade the on-the-fly usability of its output numerics and indices such as citation counts to a great extent. Their first, as already stated, is related to the use-cases being missing from the citation information. In the current tradition authors do not mention the type of usage the reference has had in their work, clarifying whether they have used its methodology, an idea it presents, or its pedagogical aspects. Sometimes they are even not citing the reference in a positive context at all, but only in the framework of reporting an error in or even to criticize it. On the other hand, setting aside the usage type, and also assuming that citation is done in a positive situation, levels of these usages have definitely a wide variation over the references as well. One may have been the main inspiration source of the referring work, without which it would have never been emerged, while the others might have had only marginal usages. In the current system however, all these references receive a rigid and non-flexible (hard) one unit of citation credit.

This system can be considerably upgraded through a simple transition towards a complement in which use-cases are included in citation information (authors' names, title of the work, publication date, etc.) as well, and citation credits are also assigned to the references softly from the whole (0,1] interval, discriminating their different levels of impact on the emergence of the referring paper. The upgraded standard would be in much more harmony with human perception of a work's real scientific influence.

Problem-oriented Interactive Research on a Global Scale

Academic science progresses in its current state of the art system through publication of research results in journals and conferences, to be studied and advanced further by other researchers somewhere, sometime, in future. This tradition, despite working for centuries, has its own drawbacks as well. The underlying progress is for instance highly ad hoc and, as an immediate result, rather slow. It often takes too long in it for a serious breakthrough to emerge. This status can be considerably improved through a wide range of measures, among the most effectives of which is the problem--oriented interactive research model proposed by the author in Jahani Yekta (2018). In the framework of this model, thousands of big problems whose importance have already been approved by numerous top scientists world-wide could be posted on a Web 2.0 platform accessible to all interested researchers across the globe. Instead of the old-fashioned tradition of scattered and sometimes nerve-scratching endeavors to spot a worthy topic to work on alone or at best in a small team, they can work in the new system on these problems in global-scale groups, and report their ideas, opinions, hints, comments, and every single piece of, even incremental, potential contributions towards their final solution interactively.

The proposed problem-oriented research system, if implemented professionally, would bring about many desired features. One of their firsts is that it can greatly push the peer review process towards an implicit, automatic, and on-the-fly model, performed by the other scientists who are already concentrated on the same problem and follow the progress of its solution up regularly. This completely eliminates false rejections such as the ones faced by Geim and Barabási regarding their respective papers (Novoselov et al. 2004;

Barabási and Albert 1999) hinted at previously. In the framework of peer review, this system, in contrast to its current traditional subject-oriented and explicitly-reviewed counterpart (Valcher 2015), would also be extremely sustainable, and can easily handle the ever-increasing volume of submissions.

In such a concentrated, crowd-based, lively, and agile publication system there is also no need for authors to wait for months for their full paper to be completed for submission. Every single piece of information or idea which may influence the progress of the solution can be presented right at the moment, and would be properly acknowledged and credited by the other involved researchers if appropriate. Contributions of each scientist in the final solution can be discriminated and measured exactly. The atmosphere would be highly live and competitive, which boosts ambitions for joining the game. Due to the large number of people concentrated on the problem, it will not enter an idle state, as frequently happens already for individual or small group researches. Energies would be directed towards impactful and essential works, and there would thus be almost no need for authors to wait for years to see the citation counts of their paper to evaluate how influential it was. Track of the hottest and the “flavor of the month” researches will be easy. They would simply be the ones with the largest numbers of followers and likes, much the same as what has become customary in blogs (Acero 2015). This is specially an outstanding feature for graduate students who are in the early stages of selecting their thesis topics. It would also be possible to rapidly register conjectures and call the world to work on them. Collaborations take place on a global scale rather than on laboratory or at most on departmental levels. History of the progress of a problem and people’s contributions would be well auto-documented. Citation of prior works would be done as straightforwardly as possible, and unnecessary and superficial ones, such as those based just on keywords, will decrease to a minimum. The research outlet will be at the utmost level of being open access. Also, citation-based metrics such as impact factor and h-index, for the measurement of the research quality would be replaced by more accurate and specific ones. This list of pros can of course be extended much further.

FINAL DISCUSSION

In this paper, we first studied some of the most essentials properties highly influential papers have in common. These properties should be considered by scholars when planning for a research they are going to spend months of time and lots of money and energy on. We then proposed the idea of soft citation, which is an upgrade to the current traditional citation standard in which many essential information about the referred document is missing. A problem-oriented interactive research model was finally proposed, as a transformative complement for the state of the art scholarly publishing system, that can facilitate and expedite scientific zero to one transitions to a great extent.

Based on what had been studied throughout this paper, it appears that a major and invaluable scientific direction to pave at the time being is further analysis of the different aspects of the interactive research platform discussed above, and to launch some small-scale pilot ones to see their performance in practice. Considering the optimistic results obtained from some ad hoc occasions in which such platforms have been simulated almost unintentionally (a question posed in 2006 by Tao in his aforementioned UCLA blog; for example, received a complete solution within days Tao 2008), it is expected that systematic and well-planned projects in this context would lead to much noticeable observations in the framework of taking zeros to ones. Such projects would be among our future lines of research.

ACKNOWLEDGEMENT

No grant from any public, commercial, or non-profit funding agency was offered for the conduct of this research.

REFERENCES

Acero, A. 2015. Should we experiment with new peer-review models? [president's message]. *IEEE Signal Processing Magazine*, Vol. 32, no. 5: 6.

Aharon, M., Elad, M., and Bruckstein, A. 2006, k–SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, Vol. 54, no. 11: 4311–4322.

Ajayan, P. M. 1999. Nanotubes from carbon. *Chemical Reviews*, Vol. 99, no. 7: 1787–1800.

Albert, R., and Barab´asi, A.-L. 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, Vol. 74, no. 1: 47.

Albert, R., Jeong, H., and Barab´asi, A.-L. 1999. Internet: Diameter of the World– Wide Web. *Nature*, Vol. 401, no. 6749: 130–131.

Albert, R., Jeong, H., and Barab´asi, A.-L. 2000. Error and attack tolerance of complex networks. *Nature*, Vol. 406, no. 6794: 378.

Allison, J. e. a.2006. Geant 4 developments and applications. *IEEE Transactions on Nuclear Science*, Vol. 53, no. 1: 270–278.

Arora, S. K., Youtie, J., Shapira, P., Gao, L., and Ma, T. 2013. Entry strategies in an emerging technology: a pilot web–based study of graphene firms. *Scientometrics*, Vol. 95, no. 3: 1189–1207.

Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. 2002, Feb. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, Vol. 50, no. 2: 174-188.

Ashburner, J., and Friston, K. J. 1999. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, Vol. 7, no. 4: 254–266.

Autumn, K., Liang, Y. A., Hsieh, S. T., Zesch, W., Chan, W. P., Kenny, T. W., et al. 2000. Adhesive force of a single gecko foot–hair. *Nature*, Vol. 405, no. 6787: 681.

Bao, Z., Dodabalapur, A., and Lovinger, A. J. 1996. Soluble and processable regioregular poly (3–hexylthiophene) for thin film field-effect transistor applications with high mobility. *Applied Physics Letters*, Vol. 69, no. 26: 4108–4110.

Barab´asi, A.-L. 1996. Invasion percolation and global optimization. *Physical Review Letters*, Vol. 76, no. 20: 3750.

Barab´asi, A.-L., and Albert, R. 1999. Emergence of scaling in random networks. *Science*, Vol. 286, no. 5439: 509–512.

Barab´asi, A.-L. 2016. *Network science*. Cambridge university press.

Becerril, H. A., Mao, J., Liu, Z., Stoltenberg, R. M., Bao, Z., and Chen, Y. 2008. Evaluation of solution–processed reduced graphene oxide films as transparent conductors. *ACS Nano*, Vol. 2, no. 3: 463–470.

Becke, A. D. 1993. Density–functional thermochemistry. III. the role of exact exchange. *The Journal of Chemical Physics*, Vol. 98, no. 7: 5648–5652.

Bendtsen, J. D., Nielsen, H., Heijne, G. von, and Brunak, S. 2004. Improved prediction of signal peptides: Signal P 3.0. *Journal of Molecular Biology*, Vol. 340, no. 4: 783–795.

Bland, J. M., and Altman, D. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, Vol. 327, no. 8476: 307–310.

Boehm, H., Clauss, A., Fischer, G., and Hofmann, U. 1962. The adsorption behavior of very thin carbon films. *Z. Anorg. Allg. Chem.*, Vol. 316, no. 3-4: 119–127.

Bollob´as, B., and Riordan, O. 2004. Robustness and vulnerability of scale–free random graphs. *Internet Mathematics*, Vol. 1, no. 1: 1–35.

Bollob´as, B., Riordan, O., Spencer, J., and Tusn´ady, G. 2001. The degree sequence of a scale–free random graph process. *Random Structures and Algorithms*, Vol. 18, no. 3: 279–290.

Bots, M. L., Hoes, A. W., Koudstaal, P. J., Hofman, A., and Grobbee, D. E. 1997. Common carotid intima–media thickness and risk of stroke and myocardial infarction: the Rotterdam Study. *Circulation*, Vol. 96, no. 5: 1432–1437.

Braun, V., and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, Vol. 3, no. 2: 77–101.

Brodie, B. C. 1859. XIII. on the atomic weight of graphite. *Philosophical Transactions of the Royal Society of London*, Vol. 149: 249–259.

Candes, E. J., Romberg, J., and Tao, T. 2006, Feb. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, Vol. 52, no. 2: 489–509.

Candes, E. J., Romberg, J. K., and Tao, T. 2006. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Math- ematical Sciences*, Vol. 59, no. 8: 1207–1223.

Candes, E. J., and Tao, T. 2005, Dec. Decoding by linear programming. *IEEE Transactions on Information Theory*, Vol. 51, no. 12: 4203–4215.

Candes, E. J., and Tao, T. 2006, Dec. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, Vol. 52, no. 12: 5406–5425.

Candes, E. J., and Wakin, M. B. 2008, Mar. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, Vol. 25, no. 2: 21–30.

Carath´eodory, C. 1907. Uber den variabilitatsbereich der koeffizienten von potenzreihen, die gegebene werte nicht annehmen. *Mathematische Annalen*, Vol. 64, no. 1: 95–115.

Carath´eodory, C. 1911. Uber den variabilitatsbereich der fourierschen kon stanten von positiven harmonischen funktionen. *Rendiconti Del Circolo Matematico di Palermo (1884–1940)*, Vol. 32, no. 1: 193–217.

Carr, E. R., and Thompson, M. C. 2014. Gender and climate change adaptation in agrarian settings: Current thinking, new directions, and research frontiers. *Geography Compass*, Vol. 8, no. 3: 182–197.

Clarivate Analytics 2018. *Highly Cited Researchers 2018 - Identifying top talent in the sciences and social sciences* (Tech. Rep.), Web of Science, Clarivate Analytics, http://clarivate.com.

Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. 2000. Resilience of the Internet to random breakdowns. *Physical Review Letters*, Vol. 85, no. 21: 4626.

Crone, B., Dodabalapur, A., Lin, Y.-Y., Filas, R., Bao, Z., LaDuca, A., et al. 2000. Large-scale complementary integrated circuits based on organic transistors. *Nature*, Vol. 403, no. 6769: 521.

Cui, Y., Wei, Q., Park, H., and Lieber, C. M. 2001. Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science*, Vol. 293, no. 5533: 1289–1292.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. 2002, Apr. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, Vol. 6, no. 2: 182–197.

Dominici, M., Le Blanc, K., Mueller, I., Slaper-Cortenbach, I., Marini, F., Krause, D., et al. 2006. Minimal criteria for defining multipotent mesenchymal stromal cells – the international society for cellular therapy position statement. *Cytotherapy*, Vol. 8, no. 4: 315–317.

Donoho, D. L. 2006, Apr. Compressed sensing. *IEEE Transactions on Information Theory*, Vol. 52, no. 4: 1289–1306.

Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. 2000. Structure of growing networks with preferential linking. *Physical Review Letters*, Vol. 85, no. 21: 4633.

Dresselhaus, M., and Dresselhaus, G. 1981. Intercalation compounds of graphite. *Advances in Physics*, Vol. 30, no. 2: 139–326.

Ebbesen, T., and Ajayan, P. 1992. Large-scale synthesis of carbon nanotubes. *Nature*, Vol. 358, no. 6383: 220.

Efron, B. 1979, Jan. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, Vol. 7, no. 1: 1–26.

Fan, J. 2018. *Local polynomial modeling and its applications: monographs on statistics and applied probability 66*. Routledge.

Fan, J., and Li, R. 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, Vol. 96, no. 456: 1348–1360.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, Vol. 39, no. 4: 783–791.

Gao, W., Alemany, L. B., Ci, L., and Ajayan, P. M. 2009. New insights into the structure and reduction of graphite oxide. *Nature Chemistry*, Vol. 1, no. 5: 403.

Geim, A., Dubonos, S., Grigorieva, I., Novoselov, K., Zhukov, A., and Shapoval, S. Y. 2003. Microfabricated adhesive mimicking gecko foot–hair. *Nature Materials*, Vol. 2, no. 7: 461.

Geim, A. K. 2011, July-September. Nobel lecture: Random walk to graphene. *Reviews of Modern Physics*, Vol. 83, no. 3: 851–862.

Geim, A. K., and Novoselov, K. S. 2010. The rise of graphene. In *Nanoscience and technology: A collection of reviews from nature journals* (pp. 11–19). World Scientific.

Gronthos, S., Brahim, J., Li, W., Fisher, L., Cherman, N., Boyde, A., et al. 2002. Stem cell properties of human dental pulp stem cells. *Journal of Dental Research*, Vol. 81, no. 8: 531–535.

Gupta, V., et al. 2009. Application of low–cost adsorbents for dye removal–a review. *Journal of Environmental Management*, Vol. 90, no. 8: 2313-2342.

Hak, L., Pols, H., Visser, T., Drexhage, H., Witteman, J., and Hofman, A. 2000. Subclinical hypothyroidism is an independent risk factor for atherosclerosis and myocardial infarction in elderly women: the Rotterdam Study. *Annals of Internal Medicine*.

Heeringa, J., Kuip, D. A. van der, Hofman, A., Kors, J. A., Herpen, G. van, Stricker, B. H. C., et al. 2006. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam Study. *European Heart Journal*, Vol. 27, no. 8: 949–953.

Huang, G.-B., Zhou, H., Ding, X., and Zhang, R. 2012. Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 42, no. 2: 513–529.

Huang, G.-B., Zhu, Q.-Y., and Siew, C.-K. 2006. Extreme learning machine: theory and applications. *Neurocomputing*, Vol. 70, no. 1-3: 489–501.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., et al. 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. *Neural Networks*, Vol. 2, 985–990.

Huang, Y., Duan, X., Cui, Y., Lauhon, L. J., Kim, K.-H., and Lieber, C. M. 2001. Logic gates and computation from assembled nanowire building blocks. *Science*, Vol. 294, no. 5545: 1313–1317.

Jahani Yekta, M. M. 2018, Aug. Problem–oriented interactive research with implicit peer review. *IEEE Control Systems Magazine*, Vol. 38, no. 4: 14.

Julıusdottir, M., Skaptadottir, U. D., and Karlsdottir, A. 2013. Gendered migration in turbulent times in iceland. *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography*, Vol. 67, no. 5: 266–275.

Jyrkinen, M. 2014. Women managers, careers and gendered ageism. *Scandinavian Journal of Management*, Vol. 30, no. 2: 175–185.

Kaplan, E. L., and Meier, P. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, Vol. 53, no. 282: 457–481.

Kokot, P. 2014. Structures and relationships: women partners' careers in Ger- many and the UK. *Accounting, Auditing, and Accountability Journal*, Vol. 27, no. 1: 48–72.

Kolpin, D. W., Furlong, E. T., Meyer, M. T., Thurman, E. M., Zaugg, S. D., Barber, L. B., et al. 2002. Pharmaceuticals, hormones, and other organic wastewater contaminants in us streams, 1999–2000: A national reconnaissance. *Environmental Science and Technology*, Vol. 36, no. 6: 1202–1211.

Krapivsky, P. L., and Redner, S. 2002. Statistics of changes in lead node in connectivity–driven networks. *Physical Review Letters*, Vol. 89, no. 25: 258703.

Kreuer, K. 2001. On the development of proton conducting polymer membranes for hydrogen and methanol fuel cells. *Journal of Membrane Science*, Vol. 185, no. 1: 29–39.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. 2012. ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).

Laeng, B., Bloem, I. M., DAscenzo, S., and Tommasi, L. 2014. Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, Vol. 131, no. 2: 263–283.

Lee, C., Yang, W., and Parr, R. G. 1988. Development of the colle-salvetti correlation–energy formula into a functional of the electron density. *Physical Review B*, Vol. 37, no. 2: 785.

Lehman, K. 2011. Reviews of science for science librarians: Graphene. *Science and Technology Libraries*, Vol. 30, no. 2: 132–142.

Levac, L. 2013. "Is this for real?" participatory research, intersectionality, and the development of leader and collective efficacy with young mothers. *Action Research*, Vol. 11, no. 4: 423–441.

Li, D., Mu¨ller, M. B., Gilje, S., Kaner, R. B., and Wallace, G. G. 2008. Processable aqueous dispersions of graphene nanosheets. *Nature Nanotechnology*, Vol. 3, no. 2: 101–105.

Lipomi, D. J., Vosgueritchian, M., Tee, B. C., Hellstrom, S. L., Lee, J. A., Fox, C. H., et al. 2011. Skin–like pressure and strain sensors based on transparent elastic films of carbon nanotubes. *Nature Nanotechnology*, Vol. 6, no. 12: 788.

Liu, H., Liu, Y., and Sun, F. 2014. Traffic sign recognition using group sparse coding. *Information Sciences*, Vol. 266, 75–89.

Liu, J., Rinzler, A. G., Dai, H., Hafner, J. H., Bradley, R. K., Boul, P. J., et al. 1998. Fullerene pipes. *Science*, Vol. 280, no. 5367: 1253–1256.

Livak, K. J., and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- ΔΔCT method. *Methods*, Vol. 25, no. 4: 402–408.

Lowe, D. G. 2004. Distinctive image features from scale–invariant keypoints. *International Journal of Computer Vision*, Vol. 60, no. 2: 91–110.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., Randall, R. J., et al. 1951. Protein measurement with the folin phenol reagent. *Journal of Biological Chemistry*, Vol. 193, 265–275.

Malik, V. S., Schulze, M. B., and Hu, F. B. 2006. Intake of sugar–sweetened beverages and weight gain: a systematic review. *The American Journal of Clinical Nutrition*, Vol. 84, no. 2: 274–288.

Mannsfeld, S. C., Tee, B. C., Stoltenberg, R. M., Chen, C. V. H., Barman, S., Muir, B. V., et al. 2010. Highly sensitive flexible pressure sensors with microstructured rubber dielectric layers. *Nature Materials*, Vol. 9, no. 10: 859.

Masters, B., and Thiel, P. 2014. *Zero to one: notes on startups, or how to build the future*. Random House.

McCall, L. 2008. The complexity of intersectionality. In *Intersectionality and beyond* (pp. 65–92). Routledge-Cavendish.

Novoselov, K. S., Geim, A. K., Morozov, S., Jiang, D., Katsnelson, M., Grigorieva, I., et al. 2005. Two-dimensional gas of massless dirac fermions in graphene. *Nature*, Vol. 438, no. 7065: 197.

Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., et al. 2004. Electric field effect in atomically thin carbon films. *Science*, Vol. 306, no. 5696: 666–669.

O'regan, B., and Gratzel, M. 1991. A low–cost, high–efficiency solar cell based on dye–sensitized colloidal TiO2 films. *Nature*, Vol. 353, no. 6346: 737.

Pastor-Satorras, R., and Vespignani, A. 2001. Epidemic spreading in scale–free networks. *Physical Review Letters*, Vol. 86, no. 14: 3200.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. 2004. UCSF Chimera – a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, Vol. 25, no. 13: 1605–1612.

Pham, C. H., and Fayerberg, R. 2011. Current trends in patenting graphene and graphene–based inventions. *Nanotech*. L. and Bus., Vol. 8, 10.

Prony, G. 1795. Essai experimental et analytique: sur les lois de la dilatabilit´e de fluides´elastique et sur celles de la force expansive de la vapeur de l´alkool´a diff´erentes temp´eratures. *J. Ecole Polytechnique Flor´eal Plairial* III, Vol. 1: 22.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning representations by back–propagating errors. *Nature*, Vol. 323: 533–536.

Soler, J. M., Artacho, E., Gale, J. D., Garc´ıa, A., Junquera, J., Ordej´on, P., et al. 2002. The SIESTA method for ab initio order-N materials simulation. *Journal of Physics: Condensed Matter*, Vol. 14, no. 11: 2745.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, Vol. 98, no. 3: 503–517.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. Mega4: molecular evolutionary genetics analysis (mega) software version 4.0. *Molecular biology and evolution*, Vol. 24, no. 8: 1596–1599.

Tao, T. 2007. What is good mathematics? *Bulletin of the American Mathematical Society*, Vol. 44, no. 4: 623–634.

Tao, T. 2008. *Structure and randomness: pages from year one of a mathematical blog*. American Mathematical Society Providence, RI.

Tendeiro, D., Lopes, G., Vieira, P., and Santos, J. P. 2014. Monte Carlo simulation of laser beams interaction with the human eye using Geant 4. *Biomedical Engineering Online*, Vol. 13, no. 1: 58.

Thess, A., Lee, R., Nikolaev, P., Dai, H., Petit, P., Robert, J., et al. 1996. Crystalline ropes of metallic carbon nanotubes. *Science*, Vol. 273, no. 5274: 483–487.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, Vol. 15, no. 1: 273–289.

Valcher, M. E. 2015. Peer review [president's message]. *IEEE Control Systems Magazine*, Vol. 35, no. 1: 14–16.

Van Noorden, R., Maher, B., and Nuzzo, R. 2014. The top 100 papers. *Nature News*, Vol. 514, no. 7524: 550–553.

Vargo, S. L., and Lusch, R. F. 2004. Evolving to a new dominant logic for marketing. *Journal of Marketing*, Vol. 68, no. 1: 1–17.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425–478.

Vetterli, M., Marziliano, P., and Blu, T. 2001. *Sampling signals with finite rate of innovation* (Tech. Rep.).

Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of small–world networks. *Nature*, Vol. 393, no. 6684: 440.

Women's Health Initiative Investigators, W. G. for the. 2002. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*, Vol. 288, no. 3: 321–333.

Wu, H., Wang, D., Shi, L., Wen, Z., and Ming, Z. 2014. Midsagittal plane extraction from brain images based on 3D SIFT. *Physics in Medicine and Biology*, Vol. 59, no. 6: 1367.

Wu, L., Wang, D., and Evans, J. A. 2019. Large teams develop and small teams disrupt science and technology. *Nature*, Vol. 566, no. 7744: 378.

Zewail, A. 2010. Curiouser and curiouser: managing discovery making. *Nature News*, Vol. 468, no. 7322: 347.

Zhang, Y., Small, J. P., Amori, M. E., and Kim, P. 2005. Electric field modulation of galvanomagnetic properties of mesoscopic graphite. *Physical Review Letters*, Vol. 94, no. 17: 176803.

Zhang, Y., Tan, Y.-W., Stormer, H. L., and Kim, P. 2005. Experimental observation of the quantum hall effect and Berry's phase in graphene. *Nature*, Vol. 438, no. 7065: 201.