Cite this paper as:

Al-Hidabi M.D.A., Teh P.L. (2019) Multiple Publications: The Main Reason for the Retraction of Papers in Computer Science. In: Arai K., Kapoor S., Bhatia R. (eds) Advances in Information and Communication Networks. FICC 2018. Advances in Intelligent Systems and Computing, vol 886. Springer, Cham.

Multiple Publications: The Main Reason for the Retraction of Papers in Computer Science

This paper intends to review the reasons for the retraction over the last decade. The paper particularly aims at reviewing these reasons with reference to computer science field to assist authors in comprehending the style of writing. To do that, a total of thirty-six retracted papers found on the Web of Science within Jan 2007 through July 2017 are explored. Given the retraction notices which are based on ten common reasons, this paper classifies the two main categories, namely random and non-random retraction. Retraction due to the duplication of publications scored the highest proportion of all other reasons reviewed.

I. INTRODUCTION

Over the years, Steen [1]; Wager and Williams [2]; Ferric C Fang [3]; M. L. Grieneisen and M. Zhang [4]; M. Zhang and M. L. Grieneisen [5]; Daniele Fanelli [6]; E Carafoli [7] Andrew B. Rosenkrantz [8]; and Philippe Mongeon [9] have highlighted that retraction has gradually increased [1]–[9]. Retraction is the act of taking out a paper from scientific publication. The retracted paper defined by the Web of Science (WoS) is "*An article that has been withdrawn by an author, institution, editor or a publisher because of an error of unsubstantiated data*"[1]. Ferric C Fang [3] reasons the Retractions could appear, due to misconduct, fraud, a scientific error,

---

1 Web of Science, "Retracted Paper: Web of Science - All Databases Help," Web of Science Website, 2017. [Online]. Available: https://images.webofknowledge.com/WOKRS522_2R1/help/WOK/hs_document_types.html.

plagiarism, duplicate publication, and so on [3]. It could also be because of faked peer reviews [10]. When this happens, editors will send a notice of concern in a bid for initiating a case of investigation in this respect [11].

Although there is a large proportion of retraction review available in the literature, there has not been much work focusing on the reasons for the act of retraction in computer science. In this paper, an attempt has been done to analyse all computer science retracted papers that were indexed in the WoS between Jan 2007 and July 2017. The pertinent works have taken diverse perspectives on different disciplines.

During the last decade, up to 31th of July 2017, the total number of retracted papers reached 36 in seven areas of computer science. Tentative results have shown that at this current stage, the highest number recorded and contributed to computer science disciplinary accounts for a percentage of only 6.145% or equivalent to n=1,438,466. The total publications of all disciplinary (n=23,153,925).

Retraction causes harm to authors and a publisher's reputation [12]. On this point, R. Grant Steen [13] states that retracted articles could reflect on medical misinformation [13]. Thus, it can cause misfortune to the world that referred and trusted that medical misinformation. Also, retraction has harmful acts during clinical practices [14]. Adding to the seriousness of this issue, Tianwei He [15] has pinpointed that the highest rates of misconduct are found in biomedical research (i.e., clinical, medical and pharmacological) rather than in other disciplines.

The intention here is to reflect on the reason for the importance of understanding retraction. It is for preparing junior researchers better for conducting their research and improving their writing processes along with assisting them from being retracted through having a better understanding of the retraction reasons, from other disciplines not only biomedical studies [5]. Ravi Murugesan [16] indicates that any extent retraction does not impair or diminish the journal; it is conserved in the journal contained by retraction [16]. The retraction notices consider being a crucial issue in the integrity of the scientific record. Therefore, they must be exact and coherent [17]. A written Guideline on Retractions from the Committee on Publication Ethics (COPE) says: '*The intention of the retraction is to correct the literature and ensure its integrity rather than to punish authors who misbehave*'[18].

This paper classifies retraction reasons into random and non-random and attempts to spotlight the different forms of errors and misconduct [2]. As Steen [1] reasons which based on honest error, misconduct, frauds and ethical matters. Furthermore, the main goal of the present paper follows a precise classification in this regard for the intention to distinguish honest error (unintentional) from misconduct and frauds

(intentional), particularly in computer science publications. The review also includes the poor of adherence to ethical /legal issues that were discussed and classified as intentionally unethical acts.

The present study will be organized as follows: In Section 2, the reasons found in the relevant literature will be described along with the terms widely used in this paper. This is followed by the review process that was applied in Section 3. All the results are then discussed in Section 4, and finally, Section 5 would provide the conclusion.

II. REASONS ON RETRACTION

There are literature reviews associated with the reasons of retraction[19]–[21]. Over and above, the causes of retraction have been extensively investigated in details by R. G. Steen [1]; M. Zhang and M. L. Grieneisen [5]; Susan Feng Lu , Ginger Zhe Jin , Brian Uzzi and Benjamin Jones [10]; R. Nambiar, P. Tilak, and C. Cerejo [12]; and Tianwei He [15] in the publications indexed in WoS. Presently, two science journalists that had spent seven years in writing on retraction have launched a beta version of retraction database. The intention of this database is to keep track of researchers that failed or committed misconduct in their publications. This has also aimed at raising community awareness of paper retraction.

Retraction is caused by human errors; it includes intentional and/or unintentional misbehaviour [2]. In the two studies of Andrew B. Rosenkrantz [8] and Yan et al. [19], honest error or misconduct have been identified as contributing factors to the large proportion of retraction reasons [8], [19]. Moreover, there are also different forms of misconduct defined by N. A. Trikalinos, E. Evangelou, and J. P. A. Ioannidis [22], indicating that retraction from misconduct is also due to falsification [22]. Image manipulation, which was discussed by Debra Parrish, and Bridget Noonan [23], and faked emails by Xingshun Qi, Han Deng, and Xiaozhong Guo [14] are examples for fallen into the category of misconduct. No matter what type of reason there is , it is the authority of author and publisher that is involved [17]. There is a high proportion of retraction reasons. All the reasons discussed under this paper are classified into two categories; random and non-random.

A. Random Reasons

Whereas random reasons commonly refer to an honest error and improper data, non-random reasons have dissimilar forms of random/unintentional errors [24]. An honest error is that error performed unintentionally due to human error. As stated by Roig [25], 'unintentional errors' may exist in the process of

writing committed by humans to extent that it may reach to the violation of agreement [25]. Tianwei He [15] defines improper data as any part of a publication of inaccurate data that is published unintentionally [15]. Tianwei He [15] added that inadvertent performance appears if an author has incorrectly done miscalculation or made an experimental error [15]. For instance, Andrew B. Rosenkrantz [8] comments that 16 of 48 (33.3%) of examined honest errors retractions are due to following incorrect methods or arriving at unreliable results [8]. R. Grant Steen [13] endorses the comments made by Andrew B. Rosenkrantz [8], and says that publication bias occurs whenever errors of improper data could appear as experimental errors, data collection, errors in bias, and non-replicable results [13].

Evidence provided by recent studies indicates that retraction notices engage multiple parties [20]. In their lengthy discussion, M. L. Grieneisen and M. Zhang [4]; and B. K. Redman, H. N. Yarandi, and Jon F. Merz [26] underline that a good number of authorities usually constituted the source of retractions (i.e., authors, editors, peer-reviewers, and publishers) [4] [26]. B. K. Redman, H. N. Yarandi, and Jon F. Merz [26] clarify that the reasons for retractions in a high proportion of many cases involve different parties [26]. In an analytical study by Wager and Williams [2] indicate that the percentage of 65 retractions commonly issued by authors accounts for (63%), while those by other parties amount to (21%) by editors, (6%) by journalist, (2%) by publishers, (1%) by institutions [2], and unclear statements reach (7%)[2]. Errors made by authors could be mistakes appearing in data, samples, or the methods and results. Supporting this point, E Carafoli [7] pinpoints that in a case of preclinical research, the impact on faulty statistical analysis which is made by the author can lead to arriving at serious and wrong conclusions [7]. A study by M. Hosseini, M. Hilhorst, I. de Beaufort, and D. Fanelli [27] argues that fourteen authors tried to correct their honest errors, despite that the journal editors have treated all their cases as retractions made by authors [27]. Wager and Williams [2] agree to what has been stated by M. Hosseini, M. Hilhorst, I. de Beaufort, and D. Fanelli [27], that, occasionally, editors and publishers feel uncomfortable to retract an article [2] [27]. And, the journal's retraction notices are not clear enough to ascertain whether such retractions are due to honest errors or misconduct [2]. R. G. Steen [24] states that it is not only the reason when editors feel embarrassed, they will write unclear statements by concurrent retraction notices [24]. Undoubtedly, this ascertains that editors and publishers make authors bear the whole responsibility of retraction. Harvey Marcovitch [11] disputes that the responsibility of an honest error as improper data lies more upon the shoulder of editors rather than on authors [11], [15]. However, editors still require from authors to attach a copy of their related papers to

prevent their papers from being revoked [28]. On the other hand, Philippa J. Benson [29] ascertains that a sin of sloth exists whenever an author disregards modifying the manuscript meticulously. It is important that authors review the manuscripts many times before the camera is ready, in order to avoid occurrence of visible errors such as formatting manuscript improperly or missing the name of co-authors [29]. In fact, editors feel annoyed at receiving several drafts of one author, however, publishers sometimes are the ones who ask the authors to keep editing their papers. Therefore, it is reasonable to state that errors made by the author have been aware of editing readability before the final submission [29]. In contrast, publisher-based errors refer to all typing errors during the manuscript submission stage. Michael L. Grieneisen, and Minghua Zhang's study classified five types of publisher errors; 1) accidental duplicate publication, 2) publishing author's paper without final correction, 3) publishing in a wrong journal, 4) publishing a special or regular issue, and 5) publishing paper with errors after a paper was rejected [4]. Michael L. Grieneisen, and Minghua Zhang [4] give an example of publisher error, forty-nine papers are retracted in the entire issue of *Gene Express Pattern* Journal, at whatever time it had to be published in the Journal of *Mech Dev*. Publishing a paper in the wrong journal recorded accidentally as a type of publisher error.

Another example of honest error issued by publishers happens to an author who attempts to clarify this situation by different publishers. An example is when Bohannon [30] sent a fake paper to see which publisher could detect the former's mistakes. When Bohannon contacted more than fifty percent of open access Journals to run a test, the vast majority of those journals accepted Bohannon 's paper. Surprisingly, few referred to Bohannon for fundamental reviews or did that by themselves. One of the feedbacks was made by PLOS ONE which rejected the work due to ethical problems. The work lacks literature 'about the treatment of animals used to generate cells for the experiment'. In the case, it is possible to draw that half of the publishers treated this situation of dishonesty as a normal act. John Bohannon's case indicates that reviewing any paper or the acceptance of a number of papers can happen mistakenly due to negative decisions made by reviewers or open access journals/publishers [30]. Andrew B. Rosenkrantz [8] and Felicitas Hesselmann [20] contend that retraction may be found in ambiguous words categorized in inadvertent or deliberate acts [8], [20]. In a study conducted by M. L. Grieneisen and M. Zhang [4], errors on retraction have been classified as random reasons appertaining to a number of factors such as authors and publishers along with unspecified ones, including honest errors, misconduct, and all sorts of fraud practices [4]. It is, therefore, safe to argue that retraction caused by an author or publisher is viewed in some cases as an intentional error.

To sum up, this paper asserts that errors are also found in works produced by an editor/publisher with (or) without the cooperation of an author [2]. Table I below is intended to put this point in a more explicable manner as it provides the Glossary of Random Reasons.

TABLE I. GLOSSARY FOR RANDOM REASONS

| Terms	Description
Improper Data	Publish article with incorrect information due to behaviour (e.g., errors in samples or data, skewed statistical analysis, inaccuracies or unverifiable information, irreproducibility).
Errors of Author/ Publisher	Errors by authors appear in data samples or due to sin of sloth. However, the publisher error is the errors in typing after a submission stage.
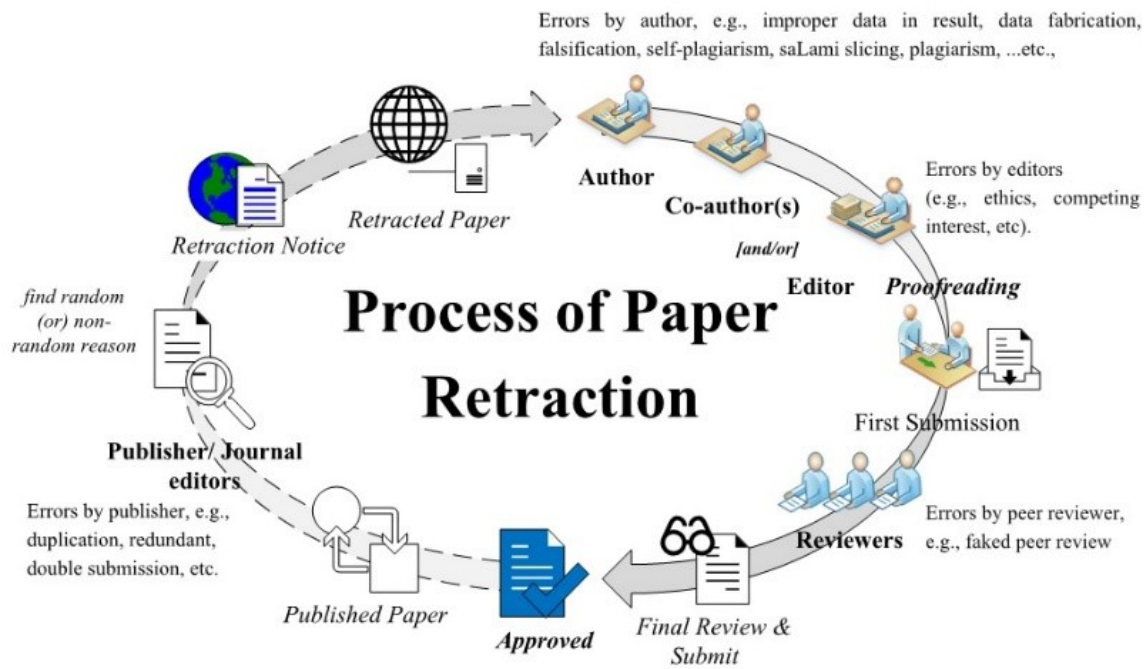
Fig. 1. Process of Paper Retraction

People involved in reporting an improper work or misconduct are often the readers and/or co-authors. However, their 'failure to report misconduct on the part of others' is still of a lower level. As a result, a lack of comprehension by readers and/ or co-authors is a forefront to the unintentionally acts which refer to random errors. Fig. 1 shows the Process of the Paper Retraction which includes errors that are normally caused by the main author, and sometimes by other parties.

Other points include that notifications are made during the publication stages where editors and reviewers detect some other misconducts such as duplication, redundancy, and self-plagiarism.

B. Non-Random Reasons

Non-random reasons refer to the presence of deliberate fraudulent data onto different nature and high levels [13]. M. Zhang and M. L. Grieneisen [5] classifies the misconduct and frauds into two categories; publishing misconduct and research misconduct [5]. plagiarism and duplicate publications considered to be publishing misconduct [4], [5]. Henning Schulzrinne [31] asserts that publishing misconduct refers to double submission of plagiarism, and falsification [31]. Philippe Mongeon and others claimed that retraction acts with misconduct and fraud have gradually grown compared to that with honest errors [3], [9], [19]. A study by Ankur khajuria [32] clarifies the rise of research misconduct including fraud, plagiarism, and duplication [32]. As to emphasize by Tianwei He [15], a retraction is happening quite extensively nowadays, it has commonly been due to duplicate publications acts, and plagiarism [15], in that case, the impacting factor of those retractions has become lower compared to the journals who have been retracted due to errors and frauds [15]. Table II considers the errors below as part of non-random reasons and presents the Glossary for Non-Random Reasons. There are three classifications of the main forms for Non-Random Retraction Reasons; research misconduct, publishing misconduct, and ethical/ legal issues.

TABLE II. GLOSSARY FOR NON-RANDOM REASONS

	Terms	Description
Research Misconduct	Data fabrication	An act of making up a fake data intentionally.
	Data Falsification	An act of misleading or manipulating manuscript components as a false data or result intentionally.
	Plagiarism	An act of submitting a manuscript using a certain work (idea, data, or image) lacking any references to copyrights to original authors, or without any extent permission.
	Data or Image Manipulation	An act which may include adding and/or changing an data, displaying different data/images without referring to the original one along with the modification.
	Fraudulent (Fake) Review	An act when someone uses a reviewer's email account illegally to send fake emails.
Publishing	Duplicate (Dual) Publication	An act of submitting the same paper to multiple conferences or journals.

	Duplication Submission	An act of having authors produce multiple papers, in different venues, arriving at the same results.
	Redundancy	An act of providing unnecessary information to be duplicated at other times.
	Salami Slicing	An act of using the same data set to publish multiple papers.
	Self-Plagiarism	An act of copying or reusing a fully or piece of published work for someone else and submitting it as if it the first work.
Ethical/ legal issues	Authorship Dispute	An act of violation that affects all authors in a published paper and leads to conflict of interest.
	Copyright Violation	An act of copying materials without getting a permission from copyright holder.

The total number of those misconducts and frauds inclusion is twelve terms in all, namely; data fabrication, data falsification, plagiarism, and image manipulation are forms of frauds that happen intentionally. As well as, fake peer review. And, the forms of publishing misconduct are duplication of several types such as duplicate publication, a duplicate submission, redundancy, salami slicing, and self-plagiarism. Furthermore, authorship and copyright violation are the numbers of ethical/legal research. All the numerous infractions are parts of non-random reasons.

Philippe Mongeon [9] indicates that fraudulent data onto biomedical research does only not harm merely science but also affects the people who are working with fraudulent authors [9]. As a result, fraudulent data is a type of the non-random errors. Philippa J. Benson [29] claimed that one of the extremely worst immoral acts is data fabrication [29]. For instance, several retractions appearing in researchers' experiment when using fake data. Elizabeth C Moylan and Maria K Kowalczuk [34] described data falsification and data fabrication as acts that come under misconduct; Moyan and Kowalczuk [34] refer to when an author manipulates data or result(s) [34]. Philippe Mongeon [9] postulates that errors on retraction due to frauds may influence an author's profession [9]. Philippa J. Benson [29] advise authors to keep away from fabrication and once authors confess that, they have to be ready to correct the manuscripts [29]. Data falsification pointed out by N. A. Trikalinos, E. Evangelou, and J. P. A. Ioannidis [22] as it has more negative impacts than plagiarism and duplication on ethics.

Plagiarism is defined by Oxford Dictionary as "presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into your work without full acknowledgement". Plagiarism is an act of stealing the data or text as if the writing of others without making any reference at all to the source [24]. M. Mohan, D. Shetty, T. Shetty, and K. Pandya [35] classify Plagiarism as a form of misconduct [35], consisting of Plagiarism of ideas, Plagiarism of text, Plagiarism of the source, Mosaic Plagiarism, Self-Plagiarism, and Ghost-Writing [35]. Most authors that commit an act of plagiarism are non-native English Language who commonly came from lower-income countries as pointed out by S. Stretton et al. [21]. Plagiarism is considered the greatest significant reason and widely appears in scientific writing [25]. Philippe Mongeon [9] reports that fabrication, falsification and plagiarism (FFP) are three serious forms of intentional infraction which have been treated as frauds [9]. Wager and Williams [2] consider FFP as forms of misconduct [2]. Additionally, data manipulation is similar to data falsification particularly if it appears as digital figures [29]. Debra Parrish, and Bridget Noonan notify that data manipulation increased sharply [23]. This form of frauds seems to be difficult to explore compared to FFP [20].

There is a kind of fraud categorised in retraction notice as 'faked peer review (FPR)' or 'faked emails (FE)'. Faked peer review or faked emails involve an author getting help from a third party as an outside expert on reviewing a real research and making suggestions on it [36]. Xingshun Qi, Han Deng, and Xiaozhong Guo [14] explain such a matter in terms of how a given author may review his own paper and persuasively makes the

concerned institution accept it. The fundamentals of reviewer's process included quality, integrity, and reproducibility. However, FPR has led to the production of a prominent number of retracted papers [14]. Lots of samples have been retracted due to fake peer review. Xingshun Qi, Han Deng, and Xiaozhong Guo [14] have spotted five publishers and 48 journals out of 250 retracted papers with regards to faked peer reviews. China is behind the utmost immense number of faked reviewers' accounts; its top three provinces of faked peer reviews are Taiwan, Shanghai, and Liaoning [14]. This form of fraud is intentional errors and may bring damage to an author and editor's reputation. One solution to refrain faked peer review of fake peer reviewer is by expanding the number of a reviewer prior to training them [36]. Xingshun Qi, Han Deng, and Xiaozhong Guo [14] conclude that there is a work recently on the issue of fake reviewer emails and this type of misconduct will disappear in the near future [14].

Elizabeth C Moylan and Maria K Kowalczuk [34] refer Duplication as an action when a scientific paper is published twice in different journals or may be published in different languages [34]. Nowadays, to discover any duplication practices, electronic searching is the base for any authors or reviewers [11]. Duplicate publication is considered as a result of misconduct made by an author [34]. M. Castillo [37] study indicates that the most retraction due to duplications comes from China and India and the reason is that of non-native speaking [37]. Another term is a duplication of submission. The retraction notices may appear as due to "Double submission" or "multiple submission". In other words, Henning Schulzrinne [31] explains that this could be referred to as re-publication of a conference (or) journal paper when a corresponding paper has been published in another publication or the paper has been submitted to the editor for checking in another conference or a journal and still under review [31]. Sometimes retracted papers appear due to double submission when they are sent to different journals and are examined by the same reviewer(s). Henning Schulzrinne [31] claims that this form of duplications is annoying because conference organizers are still scarce [31]. Although a number of publications were retracted due to duplicate publication (or dual), they covered other terms such as redundant publication, salami slicing, and self-plagiarism [38]. A redundant publication reported by E. Wager, V. Barbour, S. Yentis, S. Kleinert [18] that appears when an author duplicates his/her paper in different publications without specific explanation, authorization or citation [18]. Further, Naseem Shah [39] refers to redundancy as the act of republishing of an existing work with additional information [39]. Ernesto Carafoli [7] has clarified the possible usage of redundancy could be found in translation, following the acquisitions of permission from the original publisher [28]. As a Guidance of the

Committee on Publication Ethics Retractions (COPE) [18], redundant would be happened if an author publishes the same paper in different journals without permission, therefore the first published paper may be notified having redundant. Also, that paper will not be retracted unless the journal have checked the findings [18]. And in this regard, Felicitas Hesselmann [20] claims that redundancy and salami slicing are one of the extremely significant reasons followed by conflict of interest and plagiarism[20]. Redundancy and actual duplication are further described as a salami-slicing which leads to deform the work [25]. 'Salami Publication' is defined by Miguel Roig [25] as a segmented publication which is often referred to as a case of self-plagiarism, it overlaps with the previous work of the same author [25]. Miguel Roig [25] determines that it is possible if an author(s) republishes the dissertation or thesis in separated papers, whereas it is acceptable the citation and quotation are written correctly [25].

Another term related to duplication is Self-plagiarism that further refers to submitting multiple papers in different journals with the same results but with some changes on the papers' titles [7]. In the studies of Zhang and Grieneisen [5], it is indicated that self-plagiarism is when authors re-publish previous works without any extent permission from other authors [4], [5]. S. Stretton et al. [21] calls self-plagiarism as a duplication of re-using a previous work from the same author[21]. A similar point to what was written by M. Roig [25] is that it may happen to the event that an author copies and pastes the sentences from their previously published articles [25]. Self-plagiarism covers the issue of misconduct [38]. R. G. Steen [24] argues that self-plagiarism is a type of plagiarism usually caused by an author's misconduct [24]. And, S. Stretton [21] concludes that self-plagiarism and plagiarism have the same meaning [21]. This is so because both terms involve reiterating data in a published article. Moreover, both self-plagiarism and plagiarism connote to the act of cheating [38].

Ethical criteria involve honesty, integrity, and social responsibility [39], however, publishing misconduct or research misconduct is an act of dishonesty intentional behaviour or fraud affects research integrity. The number of suspected cases appeared due to lack of author's awareness of the ethical writing and authorship. However, it is still argued by Elizabeth Wager [40] that ethical publishing is responsible for all parties [40]. Naseem Shah [39] observes that in the post-submission, all journals send the manuscript to a number of reviewers [39]. Hence, all authors should abide by the ethical regulations [32]. E. Wager [28] states that authors may prevent their manuscripts from publication bias by following the journal regulations[28]. Besides

that, M. Mandal, D. Bagchi, and S. R. Basu [41] added that authors must do efforts to enhance their first draft, by using detection software to enhance their manuscript [41]. In the latter, it would be a great step if authors and editors set up international regulations that will attempt to improve integrity and transparency.

According to a recent article in Retraction Watch, Luann ZanZola detects three retraction papers using plagiarism detection service, iThenticate. Zanzola declares that some journals are shy to inform authors of the order of their plagiarism, and categorised a reason of "citation and attribution errors" in retraction notice. Furthermore, there are various pitfalls of misconduct causing retraction. Elizabeth C Moylan and Maria K Kowalczuk [34]; and D. M. Markowitz and J. T. Hancock [42] indicate that ethical and authorship dispute are types of misconduct [34], [42]. A number of studies by E. Wager, V. Barbour, S. Yentis, S. Kleinert [18]; Philippa J. Benson [29]; and M. Mandal, D. Bagchi, S. R. Basu [41] study authorship improprieties. Wager, V. Barbour, S. Yentis, S. Kleinert [18] state that a number of authors demand journals to retract an article due to authorship violations [18]. However, the acts of other authors that did not accept retraction endanger the journals and editors. [18]. The point here belongs to different types of authorship disputes. Authorship disputes practices consist of Gift authorship, and Ghost-writing, made by authors. Gift authorship being so-called 'Honorary', happens when individual author include author's name without any significant contribution [29]. In addition, Philippa J. Benson [29] defines Ghost-writing as a form of authorship dispute that happens when a co-author contributes in writing a part of research without mentioned him/her as an author; this author is called a Ghost-writer instead of co-author[29]. E. Wager [28] adds that practices of a supervisor who do not contribute to the manuscript cause authorship dispute [28]. E. Wager [28] explains the intentional ethical matter which includes a Ghost-writer name without a notification or permission. Sometimes there will be a multi-authors from a number of institutions, it is better to declare the statement of competing of interest before the submission [28]. Additionally, E. Wager [28] states that authorship violation tends to be a form of frauds [28], while Philippa J. Benson [29] assures that Guest and gift authorship are forms of improper authorship violations [29]. M. Mandal, D. Bagchi, S. R. Basu [41] explain that in terms of submission, a student who is an author should include all contributors in the authorship list [41]. However, Harvey Marcovitch [11] added that to settle authorship dispute, editors have to conceal the consent until the conflict of authors being solved, also editors must be conscious when dealing with groups of authors [11]. The problem derived here is competing for interest that will cause retraction. E. Wager [28] stated that authorship dispute brings a conflict of interest which is compulsory for authors who must declare all the contributors to the journal. Also,

E. Wager [28] added that it is essential that authors must declare not only who contribute, but also the research funder[28]. A study by Philippa J. Benson [29] declares that competing of interest is not a misconduct, however, practising this term improperly will lead to misconduct [29]. Philippa J. Benson [29] describes another ethical and legal issue which is copyright infringement. This infringement is caused by author, or editor/publisher or other parties in the research community [29]. M. Liviu Andreescu [38] associates copyright infringement with self-plagiarism, a fact which refers to a matter between author and journal editor [38]. Therefore, a copyright law prevents the author's work from being self- plagiarized. Mandal, D. Bagchi, S. R. Basu [41] mentions that the first author is recognized as a corresponding author that should identify any responsibility of all co-authors [41]. In addition, Philippe Mongeon [9] asserts that the legal consequences on retraction are influenced by all authors, however, the last author comes after the first author who is affected strongly, while middle authors (co-author) may have a lesser significant impact [9]. To improve the scientific community, Yongyan Li [33] recommends senior authors to educate novices in collaboration with the scientific publishing [33]. Naseem Shah [39] also states 'Poor supervision of junior researchers' (postgraduate students), will drive into non-random reasons [39], however, a clear understanding of the retraction will reduce unethical behaviours.

Despite the presence of a high proportion of reasons on retraction, Wager and Williams [2]; Ferric C Fang [3]; M. Zhang and M. L. Grieneisen [5]; S. Stretton et al. [21]; and Philippa J. Benson [29] have shown that there is 'No reason/unclear/unknown' statement, stating that retractions are related to unspecified errors. This is so due to vague reasons for retraction notices [2], [3], [5], [21], [29]. In addition, Elizabeth C Moylan and Maria K Kowalczuk [34] state that 'No reason/unclear' is difficult to identify if the statement is 'honest error' or 'misconduct'[34]. At the end of this paper, random and non-random reasons on retraction have set the overall titles as shown in the diagram below (Fig. 2).
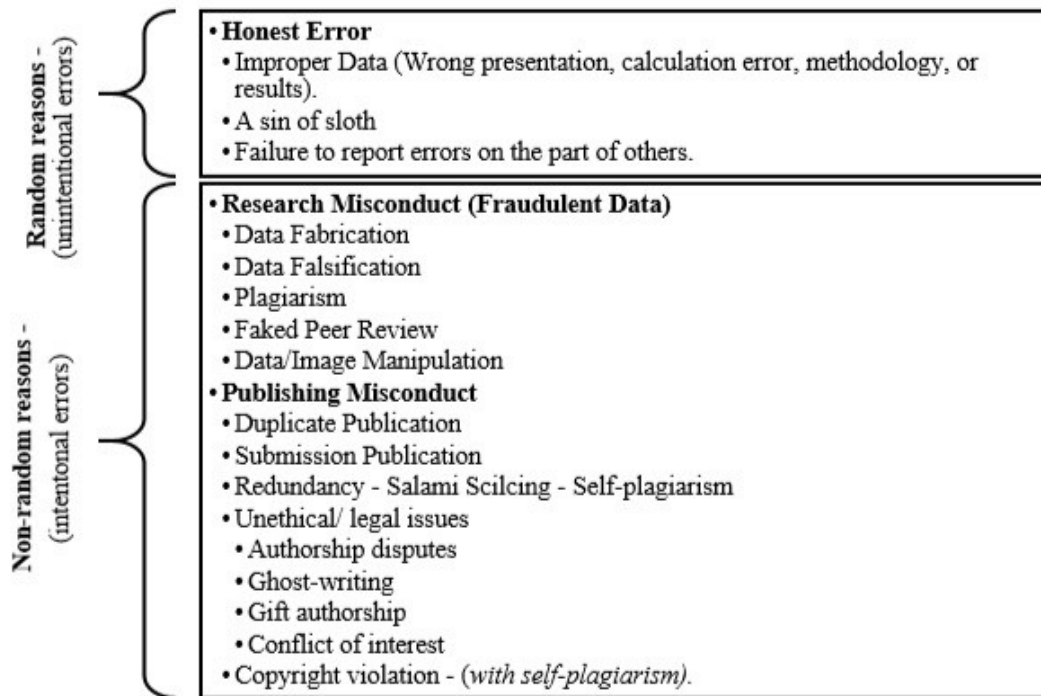
Fig. 2. Random and Non-random Reasons on Retraction

III. RESEARCH METHOD

The present research paper has been done via three phases (Fig. 3).

A. Research Process

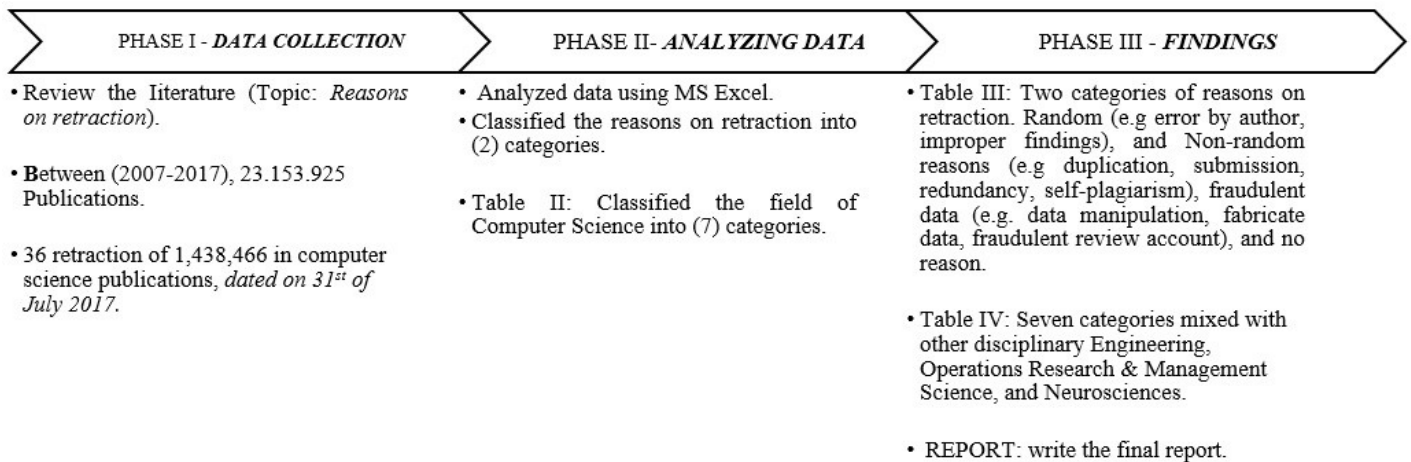| PHASE I - *DATA COLLECTION*	PHASE II- *ANALYZING DATA*	PHASE III - *FINDINGS*
• Review the literature (Topic: *Reasons on retraction*). • **B**etween (2007-2017), 23.153.925 Publications. • 36 retraction of 1,438,466 in computer science publications, *dated on 31ˢᵗ of July 2017*.	• Analyzed data using MS Excel. • Classified the reasons on retraction into (2) categories. • Table II: Classified the field of Computer Science into (7) categories.	• Table III: Two categories of reasons on retraction. Random (e.g error by author, improper findings), and Non-random reasons (e.g duplication, submission, redundancy, self-plagiarism), fraudulent data (e.g. data manipulation, fabricate data, fraudulent review account), and no reason. • Table IV: Seven categories mixed with other disciplinary Engineering, Operations Research & Management Science, and Neurosciences. • REPORT: write the final report.

Fig. 3. Research Process Chart

1) Data Collection: At the initial phase of the study, a comprehensive literature review is performed summarising the "reasons on retraction". All the reviewed papers are collected from reputable impact factor journals and databases such as ACM, Elsevier, PLoS ONE, Springer, SAGE, Wiley, and other Open Access

Journals (OAJ) such as BMJ Open, MDPI, and PNAS.

The period of retraction was conducted on Jan 2007 through July 2017. The latest retraction took place on July 31th, 2017. A total number of 23,153,925 publications were successfully indexed in WoS. Also, a total number of retraction publications being equal to the total number of computer science "disciplinary" accounted for 1,438,466. Recently, there are 2,248 retracted publications in Computer Science and 1,098 retractions indexed in the WoS (dated July 31st, 2017). Out of the numbers above, thirty-six retraction articles in Computer Science among all other types of were downloaded from WoS, constituted seven categories under Computer Science. The present research was limited to the selection of retracted publication written in English. They were further set into seven categories. By so doing, consideration was given to that they mixed with other related disciplines such as Engineering, Operations Research & Management Science, and Neurosciences & Neurology. Not only that, every notice for every single retracted paper was sought to identify the reason(s) for retraction.

2) Data Analysis: The research filtered the retracted publication. Regarding the WoS, retraction processes can be implemented by using two types of documents: "Retracted Publications" and "Retraction". Retracted Publication is defined as a withdrawal document. However, Retraction is a document that shows the retraction notice.

3)

Doing a search for articles in through WoS is more secure and trustworthy than in websites for sciences, social sciences, arts, and humanities disciplines [10]. The Web of Science allows authors to examine and scrutinize papers with close similarities with a specific field of knowledge and know dealt with papers through limited databases [10]. The WoS includes 1.3 billion, indexed and cited references. Moreover, there are more than 18000 high-impact journals with other databases [2]. All information on those (36) retracted papers were examined and analysed by using MS Excel; besides that, the reasons for retraction (Random and Non-Random reasons) were analysed as well.

---

[2] https://clarivate.com/products/web-of-science/web-science-form/web-science-core-collection/

TABLE III. COMPUTER SCIENCE CATEGORIES

#	Computer Science Categories
1	Computer Science, Software Engineering (CSSE)
2	Computer Science, Theory & Methods (CSTM).
3	Computer Science, Hardware & Architecture (CSHA)
4	Computer Science, Information Systems (CSIS)
5	Computer Science, Cybernetics (CSC)
6	Computer Science, Interdisciplinary Applications (CSIA)
7	Computer Science, Artificial Intelligence (CSAI)

4) The Findings: Based on the relevant retraction notices, the reasons on retraction were identified and summarised. They had further classified them into two categories of errors. We also set the discipline of computer science into seven categories according to WoS in Table III. It was important to decide if what has been discussed is still relevant to the related work which focused on studying different cases of retractions through WoS. Then, a consideration is given to the significant question. Next section will figure out the significant results to arrive at a better understanding of retraction along with relevant reasons.

IV. RESULTS AND DISCUSSION

A. *Results*

Based on the retraction notices, Table IV proves that the 36 retracted papers are classified into two categories; random and non-random reasons of paper retraction. Table V also highlights that seven

categories from the papers of retraction in Computer Science mixed with other areas such as Engineering (n=5), Operations Research & Management Science (N=5), and Neurosciences (n=1).

B. Discussion

The present study paper shows that during the last 126 months (i.e. Jan 2007 – July 2017), the total number of 36 retraction reasons gives ten common reasons. Yet, it is important to take note that there are significantly fewer random reasons than non-random.

TABLE IV. FREQUENT AND PERCENTAGE OF REASONS ON RETRACTION (JAN 2007 TO JULY 2017) *

#	Reasons on Retraction	N	Errors Category	Total	%
1	Improper findings	3	Random	4	11
2	Error-by-author	1			
3	Duplicate	9	Non-random	31	86
4	Redundancy	3			
5	Duplicate	6			
6	Self-Plagiarism	2			
7	Data manipulation	2			
8	Fraudulent	3			
9	Plagiarism	6			
10	No reason	1	No reason	1	3
Total				36	100

(all were searched on 31.07.2017)

TABLE V. PERCENTAGE OF COMPUTER SCIENCE CATEGORIES

#	Computer Science Categories	N, %
1	Computer Science, Software Engineering (CSSE).	1, 2.7%
2	Computer Science, Theory & Methods (CSTM).	1, 2.7%
3	Computer Science, Hardware & Architecture (CSHA) and etc.	2, 5.6%

4	Computer Science, Information Systems (CSIS) and etc.	4, 11%
5	Computer Science, Cybernetics (CSC) and etc.	6, 6.6%
6	Computer Science, Interdisciplinary Applications (CSIA) and etc.	7, 19.4%
7	Computer Science, Artificial Intelligence (CSAI) and etc.	15, 41.6%

Non-random reasons have scored the highest (n=31/36, 86%) out of other paper retraction with regards to fraud and misconduct. Specifically, duplication has scored the first highest number of reason for retraction (n=20/36, 55.5%), to include the reasons for duplication of publication; while (n=9/36, 25%) is for Redundancy in data accounts for (n=3/36, 8.3%) and Duplicate of submission amounts to (n=6/36, 16.6%). The least is Self-plagiarism, accounting for n=2/36, 5.5%), followed by the result Data manipulation, amounting to (n=2, 5.5%) while Fraudulent reviewer for (n=3, 8.3%), and Plagiarism for (n=6, 16.6%). The statement "No reason" has only one case (n=1, 2.7%). It is clear that Computer Science, Artificial Intelligence (CSAI), etc. has the highest rate (N=15/36) among all computer science categories. Given that our findings are based on a limited number of retraction papers, the result from such analysis should therefore be treated with considerable the utmost caution. Conclusively, the paper ends with a conclusion and future research.

V.   CONCLUSION

Eventually, this paper has attempted to classify in detail the available information relevant to the reasons on retraction. Until July 2017, thirty-six studies were retracted in the discipline of computer science. We have found that there is no retraction notice regarding authorship, violation of copyright rules, or ethical issues in Computer Science. The evidence from this study indicates that Non-random reasons such as duplicate publication, submission, and plagiarism have the highest number in total compared to other random reasons. In addition, the majority of findings show that there is a significant relationship between the duplicate of publications from the category of Computer Science Artificial Intelligence (CSAI). The study has some limitations; however, the results so far have been very promising and should be validated by a larger sample size. To further our research, we are planning to expand the retractions to involve various disciplines such as linguistics and psychology. Also, the research will involve additional database (i.e., SCOPUS). Ultimately, we

are confident that our results may improve knowledge about the reasons on retraction. And, future work should be attention to distinguish random reasons from non-random ones towards increasing awareness of retraction notices.

REFERENCES

[1] R. G. Steen, "Retractions in the scientific literature: is the incidence of research fraud increasing?," J. Med. Ethics, vol. 37, no. 4, pp. 249–253, 2011.

[2] E. Wager and P. Williams, "Why and how do journals retract articles? An analysis of Medline retractions 1988-2008.," J. Med. Ethics, vol. 37, no. 9, pp. 567–570, 2011.

[3] F. C. Fang, R. G. Steen, and A. Cadadevall, "Misconduct accounts for the majority of retracted scientific publications," Proc. Natl. Acad. Sci., vol. 109, no. 42, pp. 17028–17033, 2012.

[4] M. L. Grieneisen and M. Zhang, "A Comprehensive Survey of Retracted Articles from the Scholarly Literature," PLoS One, vol. 7, no. 10, 2012.

[5] M. Zhang and M. L. Grieneisen, "The impact of misconduct on the published medical and non-medical literature, and the news media," Scientometrics, vol. 96, no. 2, pp. 573–587, 2013.

[6] D. Fanelli, "Why Growing Retractions Are (Mostly) a Good Sign," PLoS Med., vol. 10, no. 12, pp. 1–6, 2013.

[7] E. Carafoli, "Scientific misconduct: the dark side of science," Rend. Lincei, vol. 26, pp. 369–382, 2015.

[8] A. B. Rosenkrantz, "Retracted Publications Within Radiology Journals," Am. J. Roentgenol., vol. 206, no. 2, pp. 231–235, 2016.

[9] P. Mongeon and V. Lariviere, "Costly Collaborations: The Impact of Scientific Fraud on Co-Authors' Careers," J. Assoc. Inf. Sci. Technol., vol. 67, no. 3, 2016.

[10] S. F. Lu, G. Z. Jin, B. Uzzi, and B. Jones, "The retraction penalty: evidence from the Web of Science.," Sci. Rep., vol. 3, p. 3146, 2013.

[11] H. Marcovitch, "Misconduct by researchers and authors.," Gac. Sanit., vol. 21, no. 6, pp. 492–499, 2007.

[12] R. Nambiar, P. Tilak, and C. Cerejo, "Quality of author guidelines of journals in the biomedical and physical sciences," Learn. Publ. -Association Learn. Prof. Soc., vol. 27, no. 3, pp. 201–206, 2014.

[13] R. G. Steen, "Misinformation in the medical literature: what role do error and fraud play?," J. Med. Ethics, vol. 37, pp. 498–503, 2011.

[14] X. Qi, H. Deng, and X. Guo, "Characteristics of retractions related to faked peer reviews: an overview," Postgrad. Med. J., vol. 93, no. 1102, p. 499, 2016.

[15] T. He, "Retraction of global scientific publications from 2001 to 2010," Scientometrics, vol. 96, no. 2, pp. 555–561, 2013.

[16] R. Murugesan, "What Happens When Ethical Violations are Detected in Research?," 13-Dec-2014.

[17]   E. Bilbrey, N. O'Dell, and J. Creamer, "A Novel Rubric for Rating the Quality of Retraction Notices," Publications, vol. 2, no. 1, pp. 14–26, 2014.

[18]   E. Wager, V. Barbour, S. Yentis, and S. Kleinert, "Retractions: Guidance from the Committee on Publication Ethics (COPE)," Maturitas, vol. 64, no. 4, pp. 201–203, 2009.

[19]   J. Yan, A. MacDonald, L.-P. Baisi, N. Evaniew, M. Bhandari, and M. Ghert, "Retractions in Orthopaedic research - A Systematic Review," Bone Joint Res., vol. 5, no. 6, pp. 263–268, 2016.

[20]   F. Hesselmann, V. Graf, M. Schmidt, and M. Reinhart, "The visibility of scientific misconduct: A review of the literature on retracted journal articles," Curr. Sociol., no. Published online before print October 13, 2016, pp. 1–32, 2016.

[21]   S. Stretton et al., "Publication misconduct and plagiarism retractions: a systematic, retrospective study.," Curr. Med. Res. Opin., vol. 28, no. 10, pp. 1575–83, 2012.

[22]   N. A. Trikalinos, E. Evangelou, and J. P. A. Ioannidis, "Falsified papers in high-impact journals were slow to retract and indistinguishable from nonfraudulent papers," J. Clin. Epidemiol., vol. 61, no. 5, pp. 464–470, 2008.

[23]   D. Parrish and B. Noonan, "Image Manipulation as Research Misconduct," Sci. Eng. Ethics, vol. 15, no. 2, pp. 161–167, 2009.

[24]   R. G. Steen, "Retractions in the scientific literature: do authors deliberately commit research fraud?," J. Med. Ethics, vol. 37, no. 2, pp. 113–117, 2011.

[25]   M. Roig, "Avoiding plagiarism, self-plagiarism, and other questionable writing practices: A guide to ethical writing," pp. 1–71, 2015.

[26]   B. K. Redman, H. N. Yarandi, and J. F. Merz, "Empirical Developments in Retraction," J. Med. Ethics, vol. 34, no. 11, pp. 807–809, 2008.

[27]   M. Hosseini, M. Hilhorst, I. de Beaufort, and D. Fanelli, "Doing the Right Thing: A Qualitative Investigation of Retractions Due to Unintentional Error," Sci. Eng. Ethics, no. office 321, 2017.

[28]   E. Wager, "Ethical publishing: the innocent author's guide to avoiding misconduct," Menopause Int., vol. 13, no. 3, pp. 98–102, 2007.

[29]   P. J. Benson, "Seven sins in publishing (but who's counting...)," Ann. R. Coll. Surg. Engl., vol. 98, no. 1, pp. 1–5, 2016.

[30]   John Bohannon, "Who's Afraid of Peer Review?," Science (80-.)., vol. 342, no. 6154, pp. 60–65, 2013.

[31]   H. Schulzrinne, "Double Submissions - Publishing Misconduct or Just Effective Dissemination?," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 3, pp. 40–42, 2009.

[32]   A. Khajuria and R. Agha, "Fraud in scientific research - birth of the Concordat to uphold research integrity in the United Kingdom.," J. R. Soc. Med., vol. 107, no. 2, pp. 61–65, 2014.

[33]   Y. Li, "Text-Based Plagiarism in Scientific Publishing: Issues, Developments and Education," Sci. Eng. Ethics, vol. 19, no. 3, pp. 1241–1254, 2013.

[34]   E. C. Moylan and M. K. Kowalczuk, "Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central," Br. Med. J. Publ. Gr., vol. 6, no. 11, 2016.

[35]   M. Mohan, D. Shetty, T. Shetty, and K. Pandya, "Rising from plagiarising," J. Oral Maxillofac. Surg., vol. 14, no. 3, pp. 538–540, 2014.

[36]   Retraction Watch, "Can a tracking system for peer reviewers help stop fakes?," 2017. [Online]. Available: http://retractionwatch.com/2017/06/23/can-tracking-system-peer-reviewers-help-stop-fakes/.

[37]   M. Castillo, "The fraud and retraction epidemic," Am. J. Neuroradiol., vol. 35, no. 9, pp. 1653–1654, 2014.

[38]   L. Andreescu, "Self-Plagiarism in Academic Publishing: The Anatomy of a Misnomer," Sci. Eng. Ethics, Springer J., vol. 19, no. 3, pp. 775–797, 2013.

[39]   N. Shah, "Ethical issues in biomedical research and publication," J. Conserv. Dent., vol. 14, no. 3, pp. 205–208, 2011.

[40]   E. Wager, "Publication ethics: whose problem is it?," Insights, vol. 25, no. 3, pp. 294–299, 2012.

[41]   M. Mandal, D. Bagchi, and S. R. Basu, "Scientific misconducts and authorship conflicts : Indian perspective," Indian J. Anaesth., vol. 59, no. 7, pp. 400–405, 2015.

[42]   D. M. Markowitz and J. T. Hancock, "Linguistic Obfuscation in Fraudulent Science," J. Lang. Soc. Psychol., vol. 35, no. 4, pp. 435–445, 2016.