Self-Plagiarism in Computer Science

Christian Collberg **Stephen Kobourov**
Department of Computer Science,
University of Arizona, Tucson, AZ 85721.
{collberg,kobourov}@cs.arizona.edu

We are all too aware of the ravages of scientific misconduct in the academic community. Students submit assignments inherited from their friends who took the course the year before, on-line paper-mills allow students to browse for term-papers on popular topics, and occasionally researchers are found out when falsifying data or publishing the work of others as their own.

Recent cases of falsified data include the Bell Labs nanotechnology researcher J. Hendrik Schön who was found guilty of 16 counts of scientific misconduct [4] and the Lawrence Berkeley National Laboratory's retraction of the discovery of superheavy element 118 [17]. The New York times recently reported [6] that:

> *In the last decade, there have been about 50 cases of misconduct among basic science research sponsored by the National Science Foundation and 137 cases of misconduct among biological and medical research financed by the National Institutes of Health.*

Recently, controversies surrounding books by historian Stephen Ambrose [13] and Pulitzer Prize winner Doris Kearns Goodwin [14] have drawn the attention of the general public to the issue of plagiarism. A specific instance of plagiarism in academia is documented in a CACM article by N. Kock [15].

The purpose of this paper is to revive a discussion about a very specific form of scientific misconduct, namely self-plagiarism. Apart from Samuelson's [18] 1994 CACM article about the legal and ethical implications of self-plagiarism, Louie [16] and Bird's [2] recent repartee on the topic, and Brogan's [3] detailed study of the systematic self-plagiarism practices of political scientist Irving Louis Horowitz, self-plagiarism has received little public attention. This is unfortunate since we believe self-plagiarism to be much more prevalent than other forms of scientific misconduct. We also believe that rampant self-plagiarism can have very detrimental effects on our research community:

- It can prevent worthy and novel ideas from being published;
- It can give the public the idea that their research dollars are spent on rehashing old results rather than on original research, simply to further the careers of researchers;
- It can indicate to our students that we do not take our own unethical behavior as seriously as we do theirs;
- It can indicate to our colleagues that academic dishonesty is not a big deal. In the worst case this could lead to more serious forms of academic dishonesty becoming more acceptable.

What is Self-Plagiarism?

Self-plagiarism occurs when an author reuses portions of their previous writings in subsequent research papers. Occasionally, the derived paper is simply a re-titled and reformatted version of the original one, but more frequently it is assembled from bits and pieces of previous work.

It is our belief that self-plagiarism is detrimental to scientific progress and bad for our academic community. Flooding conferences and journals with near-identical papers makes searching for information relevant to a particular topic harder than it has to be. It also rewards those authors who are able to break down their results into overlapping *least-publishable-units* over those who publish each result only once. Finally, whenever a self-plagiarized paper is allowed to be published, another, more deserving paper, is not.

Many of us have probably been guilty of some form of self-plagiarism. Maybe we have recycled an introduction or a related work section from one paper to the next. Maybe we have failed to include one of our own related papers in the list of references because it was a bit *too* related to the topic we were writing on. Maybe we have submitted very similar work to two different communities in order to "advertise" a new important result.

What we will attempt to do in this paper is to discuss different instances of what could be deemed self-plagiarism and continue a discussion about which are acceptable and which represent academic dishonesty. We believe that this is a problem that can be potentially damaging to the community and deserves an in-depth and open discussion. We also believe that formal (or informal) guidelines for avoiding self-plagiarism can supplement existing policies outlined in publications by the ACM and IEEE. Finally, such a discussion may be of help to junior faculty who are struggling to keep up with the "publish or perish" way of academic institutions.

This is not a scientific study of self-plagiarism. Rather, we will relate some personal anecdotes, show the results of an automated search for self-plagiarism among Computer Science academics, relate what some of our colleagues think about the issue, and discuss possible ways of addressing the problem.

Three Anecdotes

Our interest in self-plagiarism first arose when A_a[1] was on the program committee of C_x. One of the papers assigned to A_a was long, complicated, and clearly outside his area of expertise. He therefore searched the web for earlier papers by the same authors that might help to give some background to the area. One paper looked particularly relevant. It had first appeared in a regional conference whose proceedings were later published by a major publisher. On closer scrutiny A_a found that – except for the title and formatting – the earlier paper was word-for-word identical to the one submitted to C_x. No mention of the original paper was made in the submitted one.

A_a alerted the conference chair who preemptively withdrew the paper from further reviewing. A_a was also able to submit the shortest referee report of his life: "*The following paper should also be referenced: [...].*"

In a similar incident A_b was asked to review a paper for C_y. The paper assigned to him was very much in his own area of expertise. A_b's review reads as follows:

> *Assume the following definitions:*
>
> P_0 *is the article [...], published in* C_w.
>
> P_1 *is the article [...], published in* C_z.
>
> P_2 *is the paper submitted for review.*
>
> P_0 *and* P_2 *and* P_1 *and* P_2 *have intersecting author lists.*
>
> *Four pages in* P_2 *occur identically in* P_1. *Two figures (one of them containing the major algorithms of the paper) in* P_2 *also occur in* P_1. *The bottom of a figure in* P_2 *(containing a major algorithm) also occurs in* P_0.
>
> *Neither* P_0 *nor* P_1 *are referenced in* P_2*'s bibliography. In fact, only a few of the authors' many papers on the same topic are referenced. I have included a more complete list below for the authors' benefit.*
>
> *Unless the authors have cleared the republication of (parts of)* P_0 *and* P_1 *then this appears to be a violation of copyright.*

Such incidents made the first author of this paper (henceforth, *CC*) think about his own previous papers. Had he ever engaged in self-plagiarism himself? As it turns out, he had once submitted a paper

[1] In an attempt to preserve anonymity we make use of the following abbreviations. A_a and A_b refer to the authors of the present paper. Similarly, *Conference on* $x/y/\cdots$ (abbreviated as $C_x, C_y, \cdots$) refer to the conferences where the incidents described took place.

to the *ACM 1996 Conference on Programming Languages Design and Implementation*, PLDI'96. It was rejected, and then resubmitted to a regional conference (the 1997 *Australasian Computer Science Conference*) where it was accepted [7]. He then completely rewrote the paper, added some major new results, and submitted the new paper to PLDI'97. This time it was accepted [10]. However, CC conveniently "forgot" to include a reference to [7] in his submission to PLDI'97. Why? Most likely he was afraid that the previous publication would prevent him from getting the publication in a prestigious international conference he felt he desperately needed.

So, What Happened?

So, what happened to the authors in these cases? Apparently, not much. The program chair of C_x writes that:

> *I reported the case to [the publisher of the original paper] and got a reply about how they would look into it, but never heard anything more.*
>
> *I reported it to the [steering committee of the superstructure to which C_x belongs] but they did not show much interest.*
>
> *I complained to the two authors about their unethical behaviour but got no reply. (Not surprising.)*
>
> *I toyed with the thought of complaining to the head of the department where the authors work but didn't do it. The action struck me as being a little vindictive and would achieve little.*

Most likely, the program chair of C_x was more diligent than most but still the authors were not punished for their actions.

The program chair of C_y writes that:

> *The paper was rejected and I just sent the review to the authors. Nothing more I'm afraid.*

CC later saw the errors of his ways and in the journal version [8] of his paper both [7] and [10] are cited.

Current Policies

Conference call-for-paper announcements and journal submission guidelines usually have a short statement about use of previously published results. Consider the current ACM and IEEE publication policies. The ACM policy refers to the issue of self-plagiarism as follows:

> *Authors may submit to ACM conferences and journals revised versions of papers that appeared previously in refereed or formally reviewed publications or under consideration for such publication elsewhere if:*
>
> - *the paper has been substantially revised (this generally means that at least 25% of the paper is material not previously published; however, this is a somewhat subjective requirement that is left up to each publication to interpret);*
> - *upon submission, the author notifies the Editor-in-Chief(s) or Program Chair(s) that the paper has been previously published or simultaneously submitted; and*
> - *the published policies of the publications or conferences involved do not prohibit this.*

Similarly, the IEEE policy expressly states that plagiarism, self-plagiarism, fabrication and falsification are "unacceptable":

> **f)** *Plagiarism is unacceptable. The verbatim copying or reuse of one's own research which results in subsequent journal or conference papers without appropriate reference (as indicated in paragraph "h" below) is considered another form of plagiarism, or self-plagiarism; it is also unacceptable.*
>
> **g)** *Fabrication and falsification are unacceptable.*

h) *Except as indicated in Section 6.3.4 (Multiple Publication of Original Technical Material in IEEE Periodicals), authors should only submit original work that has neither appeared elsewhere for publication, nor which is under review for another refereed publication. If authors have used their own previously published work(s) as a basis for a new submission, they are required to cite the previous work(s) and very briefly indicate how the new submission offers substantively novel contributions beyond those of the previously published work(s).*

Both policies give substantial leeway to the journal editor or program chair to decide when a submitted work meets minimum novelty standards. Both policies emphasize novelty of the new result as important criteria and ACM puts a number to it: "at least 25% of the paper is material not previously published".

Definitions of Self-Plagiarism

While student plagiarism has received a great deal of attention by the Computer Science community, as witnessed by the dozens of papers about plagiarism-detection, there is comparatively little written about self-plagiarism. There also appears to be little agreement among academics as to what should be regarded as self-plagiarism and what is acceptable republication. This is even more so across disciplines and even between different computing subfields. We will therefore introduce terminology to allow us to express the actions performed by authors that might be referred to as self-plagiarism. We will adopt the neutral word *reuse* to refer to texts or ideas that are published multiple times. When appropriate, we have tried to integrate the terminology of others (for example, Evans [11]).

We introduce the following terms:

textual reuse: incorporating text/images/etc. from previously published work[2].

semantic reuse: incorporating ideas from previously published work.

blatant reuse: incorporating texts or ideas from previously published work such that the two works are virtually indistinguishable.

selective reuse: incorporating bits-and-pieces from previously published work.

incidental reuse: incorporating texts or ideas not directly related to the new ideas presented in the paper (such as related work sections, motivating examples, etc.).

reuse by cryptomnesia [5]: incorporating texts or ideas from previously published work while unaware of the existence of that work.

opaque reuse: incorporating texts or ideas from previously published work without acknowledging the existence of this work.

advocacy reuse: incorporating texts or ideas from previously published work when writing to a community different from that in which the original work was published.

When these actions pertain to ones own work we talk about **textual *self*-reuse**, etc. When it is believed that the actions are ethically or legally questionable we replace *reuse* by *plagiarism*, as in **blatant semantic opaque self-*plagiarism*** (reusing ones own previously published ideas in a new publication without adequate attribution.)

It is not clear what may be considered as self-plagiarism and what is clearly not. Most would agree that *blatant textual self-plagiarism* (as exhibited in the submission to C_x) is wrong. Just as we do not condone publishing someone else's work as our own we should not publish our own work more than once.

Many academics appear to differ on their views of *advocacy plagiarism*. Certainly, some would argue, it is important to make the public aware of new results relevant to their field, and if a particular result applies to more than one community multiple publications of the same idea is perfectly reasonable.

[2] By *published work* we mean *articles published in refereed conferences and journals where copyright is assigned to someone different from the author.*

Some make a distinction between *horizontal* and *vertical* advocacy reuse. Vertical reuse (republishing a suitably restructured scholarly article in a more "popular" forum) is often deemed acceptable, whereas horizontal reuse (republishing a scholarly article in a similar research forum) is not. The reasoning is that vertical reuse typically does not earn the author academic credit whereas horizontal reuse does.

Incidental plagiarism is also contentious. For example, once we have written the perfect introduction to a problem on a particular topic, can we reuse it for other papers on the same topic or do we need to reword it every time? Some authors do rewrite their related work section for every paper whereas others simply cut-and-paste.

And where should the line be drawn between *blatant* and *selective* self-plagiarism? The authors of the submission to C_y reused more than four pages from their previously published work. It is clearly not as bad as trying to republish an entire paper, but clearly not aboveboard either.

While *textual self-plagiarism* is easier to detect than *semantic self-plagiarism*, is it also less ethical? Copyright law covers the expression of an idea, not the idea itself, so rewording and republishing a paper – which many would regard unacceptable – may be perfectly legal.

Searching for Self-Plagiarism

It is difficult to know how common self-plagiarism is. Anecdotally, we hear of colleagues who publish the same result with minor modifications over and over again, and occasionally we come across a paper whose content we feel is too close to previous publications.

Over the last few months we conducted an experiment whereby we examined the publications found on the web sites of Computer Science professors from 50 schools. Our system, `SPlaT` (*Self-Plagiarism Tool*) consists of a specialized web spider and a text-similarity analyzer. `SPlaT` downloaded the publications of each author from each institution, converted the articles to text and compared them pair-wise for instances of textual reuse. Works that exceeded a certain threshold were examined manually. `SPlaT` can be downloaded from `splat.cs.arizona.edu`. For details on the system see [9].

Some highly-correlated pairs of papers represented what is generally though of as acceptable forms of re-publication, such as technical reports published in conferences, conference articles recast as journal papers, etc. These were all weeded out manually.

However, we found a number of instances of papers with questionable originality. In particular we ran across cases such as:

- pairs of conference publications with common introduction and/or related work sections that do not reference each other.
- pairs of conference publications with over 50% common text, that do not reference each other.[3]
- pairs of nearly identical conference and journal versions of the same paper, where the journal version does not reference the conference version.
- pairs of nearly identical conference and journal versions of the same paper, with different titles, where the journal version does not reference the conference version.

This is in many ways an unsatisfactory study. We only looked at papers downloaded from the web pages of the authors and we did not attempt to collect all papers by a given author (for example, by using services such as `http://citeseer.nj.nec.com`). Authors who deliberately engage in self-plagiarism are unlikely to put both the original and the derived works on their web site. Furthermore, if we believe that in general most referees and editors thoroughly vet submitted works then many attempted cases of self-plagiarism will be caught and will never make it to publication. However, our study confirms that textual reuse does occur.

[3] Note that when measuring textual reuse we look for sentences and paragraphs that co-occur rather than for words or phrases.

What do our Colleagues Think?

To get a feel for how others in the Computer Science community feel about the issue of self-plagiarism we conducted an informal survey among our colleagues. We sent out thirty questionnaires and received ten responses. Although no statistical significance can be attributed to this survey the answers illustrate the range of opinions on this topic.

Our first question asked if the respondent had encountered cases of self-plagiarism:

a) *I've seen many cases of papers that served up the same basic ideas in different ways [· · ·].*

b) *I have encountered submissions where the authors provide the same (word-by-word) introduction, background, and parts of the paper.*

c) *I encountered bad self-plagiarism once. I considered it to be a violation of the submission requirements for a major conference. [· · ·] Had I known about the previous publication of the similar work—which I'd guess was 80% identical to submission—I'm sure it would have been killed in committee.*[4]

d) *I'm an associate editor for [· · ·] and we frequently see papers that are more or less complete resubmissions of conference papers. I throw anything out that doesn't make the 30% rule but I believe some of the editors are more "flexible."*

e) *I have identified three such cases, bad ones. Case 1) [was a] paper submitted to good ACM conference by flourishing research group [· · ·]. Contains 2 pages of introduction and two pages of "future work" identical to that in a conference I chaired six months before, and some other. The paper was rejected on these grounds [· · ·]. [I wrote to the chair but his] reply was disappointing, and I felt it was a cop out. "Nevertheless, please review the paper anyway. I agree it is very irritating [· · ·]. However, I think we must judge the paper primarily based on its actual contribution, [· · ·]. To update the introduction and related work sections to be more appropriate to the paper contents could be an advice to the authors (or a condition for accept)."*

Case 2) I got a paper to review, [· · ·] looked up some references and found that the paper in hand was more than half a copy of one of [the author's] own references. I wrote this in my review. There was no PC meeting, and I was astonished when the paper was accepted. It was not till I arrived at the conference that I realized that the person in question was very established and famous, and that the angle of the other PC members had been that they could not turn him down because of a review by an unknown like me [· · ·] Are there different rules for the famous? [· · ·]

Our second question asked if the respondents themselves worry about reusing material from previous papers:

f) *I always rewrite every paper from scratch.*

g) *No, I don't really worry too much about this.*

h) *For sure. In systems work it's always rather hard because two or three sections of your papers (motivation, architecture, implementation) often tend to be rather similar. [· · ·] I think there's a strong sense that CS papers should be largely self-contained and that inevitably means duplication. I [· · ·] try to make very sure that each new paper has a very definite purpose/message.*

i) *Indeed yes, and I have to drill younger colleagues and students in how to avoid it.*

We next sketched a few reuse scenarios and asked whether the respondent thought they would be cause for concern. The first scenario asked about "two conference papers sharing word-for-word introductions and/or related work sections:"

j) *I think this is very disturbing. As a reviewer, this gives me a very negative impression of the paper and makes me suspicious about the content as well.*

[4]Private correspondence with the respondents who wrote anecdotes c and q indicate that these, in fact, refer to the same incident.

k) *Not a big deal.*

l) *This is something I've done to some extent [· · ·] but the way I deal with it is by thinking that I'll rewrite it if the paper is accepted.*

The second scenario asked about "two papers that are essentially the same but sent to different relevant communities (to advertise a single result):"

m) *This is not acceptable.*

n) *I think this [deserves] public flogging.*

o) *Probably OK, provided the papers are substantially rewritten for the two different communities and reference is made in one paper to the other. For example, a paper on a topic in AI might appear in both an AI journal and a philosophy journal, but substantially rewritten for the different communities.*

p) *If I consider the purpose of publication to be the dissemination of results (and that's my primary mode), I don't consider this unethical. From the point of view of the load on the reviewers, this is problematic. From the point of view of using publication counting to evaluate performance, such practices probably hurt this system, but since publication counting is idiotic anyway, I don't see this as unethical.*

The third scenario asked about "two very related papers (80% or more) and neither of the two cites the other." All respondents agreed that this was unacceptable. One respondent wrote:

q) *I wrote a paper for [a regional conference] then added essentially a results section, and submitted it to [a major conference] where it was accepted. A few years later [a colleague who had been on the conference program committee] asked me about this; I had not cited the [first] paper, nor made the PC aware of this in some other way. I could not remember my motives for this then and I cannot now, except that I am sure that during the submission I was not aware that I was doing anything wrong. Nowadays, with a little more distance, I can just explain it with being young and dumb* `:-/`.

The fourth scenario asked about "a conference paper and a journal version of the same paper that are virtually identical (95% or more)." Most respondents agreed that this was acceptable, or should even be encouraged:

> *We need conferences to advertise results quickly and journals to archive those results for longer periods of time [· · ·].*

Our final question asked if the respondents thought self-plagiarism in the computing community is a problem that deserves more attention:

r) *I think it's a problem, yes, but mainly as a symptom of a deeper problem: the superficiality of the methods used to evaluate academic contribution. The self-plagiarism behavior arises because the success of an academic is often measured primarily by counting the number of publications without any attempt to evaluate their value.*

s) *It is a problem that any program chair of journal editor must be aware of. It would be nice to have some automated way of checking for similar publications when reviewing papers.*

t) *I don't think this is a big problem. Real plagiarism is much worse, as is pre-plagarism, where someone hears about a new result, either from the inventor or second-hand, and then goes on to reproduce and publish that result himself.*

u) *Not really. [· · ·] I think it's an inevitable consequence of too much pressure to publish too much material. [· · ·] I'm really more for reducing the number of publication forums and then I think the problem will go away.*

What can be done?

Missing from the ACM and IEEE policy documents is any discussion of what the consequences of ignoring the rules and guidelines might be and whose responsibility it is to prevent plagiarized and self-plagiarized papers from being published. In contrast, most (if not all) university course syllabi address the definition of plagiarism, who will look for it, how it will be looked for, as well as the potential consequences of plagiarism.

The program chair of C_x made a valiant effort to alert the parties involved to the attempted transgression, but received little or no response. We believe this is indicative of how we as a community view the self-plagiarism issue. We know that it occurs, we deplore that it occurs, but we are not willing to expend the energy to deal with it when it does occur.

N. Kock [15] described one story involving a plagiarized paper, the difficulties he ran across, and suggestions of steps that could be taken in the future. In response to the article, Saul I. Gass, a CACM reader, wrote [12] that:

> *There are worrying implications stemming from Kock's article. The first is his suggestion that professional societies establish ethics communities "[$\cdots$] to bring some measure of justice to the people involved." Such volunteer committees would, I believe, require Solomonic decisions from ordinary people and could lead to false results and the possible ruination of careers. I say, leave it to the people involved and the courts.*

The reader goes on to say that:

> *The second implication is that referees may be asked to become plagiarism police, endangering the refereeing process and the structure of peer review. We must enter into the peer review process under the assumption that all involved work under personal and professional codes of ethics that advance the objectives of the profession as a whole.*

Whereas replacing one ethical failure with another would hardly be an improvement, should these issues be discussed more, or taken seriously into account? It is our belief that self-plagiarism is detrimental to scientific progress. But what should be done to combat it? We pose the following questions to the community:

1. Should conferences and journals provide guidelines describing in more detail what practices will be considered self-plagiarism and detailing the process of dealing with such practices?

2. Should the burden of detecting and dealing with plagiarism and self-plagiarism be on professional organizations such as ACM, IEEE, and/or with the companies that publish the articles?

3. Would self-plagiarism-detection programs or services be useful to authors, reviewers, and program committee members? Or, does the potential for abuse outweigh the benefits? For example, false positive detection results could lead to accusations of academic dishonesty that could do irreparable harm to an author's career.

4. What should the consequences be when we do find a paper we feel has been self-plagiarized? The chair of C_y simply rejected the paper although several pages had already appeared in print. The chair of C_x, on the other hand, felt that the unethical nature of the incident warranted further actions, but stopped short of reporting to the author's department chair.

5. As reviewers and program committee members, should we become plagiarism police as is indeed the case in some fields?[5]

It is our belief that we should hold ourselves to the same high standards as we do our students. Many professors use tools such as `moss` [1] or *Glatt Plagiarism Services* (`http://www.plagiarism.com`) to detect plagiarism among students. Similar tools would be useful to detect self-plagiarism among academics. We are currently modifying `SPlaT` to act as a *reviewer's workbench*. The program will compare a paper under review to a record of the author's previously published articles extracted from their web site and online article repositories (such as `http://portal.acm.org` and `http://citeseer.nj.nec.com`). Given our own personal experiences we naturally would welcome the advent of such a tool. `SPlaT` will also be

[5] For example, on their paper review form the *Journal of Advanced Nursing* asks: *Consider here whether 'salami slicing' of publications, or plagiarism (including self-plagiarism) are possibilities. Has this paper (or parts of it) been published before? Can you identify any potential copyright problems?*

useful to authors weary of committing *textual self-plagiarism by cryptomnesia* (reusing ones own previously published text while unaware of its existence).

We also believe that — just as university codes of ethics specify the consequences of academic dishonesty and the procedures by which it arrives at these consequences — our professional organizations should actively develop guidelines for what constitutes self-plagiarism and how cases of self-plagiarism should be pursued.

References

[1] Alex Aiken. A system for detecting software plagiarism.

[2] Stephanie J. Bird. Self-plagiarism and dual and redundant publications: What is the progblem? Commentary on 'Seven ways to plagiarize: Handling real allegations of research misconduct'. *Science and Engineering Ethics*, 8(4), 2002.

[3] Martha Brogan. Recycling ideas. *College and Research Libraries*, 52(5):453–464, September 1992.

[4] Geoff Brumfiel. Physicist found guilty of misconduct. *Nature*, September 2002.

[5] Siri Carpenter. Plagiarism or memory glitch? Inadvertent plagiarism complicates efforts to end cheating. *Monitor on Psychology*, 33(2), February 2002.

[6] Kenneth Chang. On scientific fakery and the systems to catch it. *New York Times*, 2002.

[7] Christian Collberg. Automatic derivation of machine descriptions. In *Australasian Computer Science Conference 1997*, February 1997.

[8] Christian Collberg. Automatic derivation of compiler machine descriptions. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 24(4):369–408, 2002.

[9] Christian Collberg, Stephen Kobourov, Joshua Louie, and Thomas Slattery. A study of self-plagiarism in computer science. Technical Report TR03-02, University of Arizona, February 2003.

[10] Christian S. Collberg. Reverse interpretation + mutation analysis = automatic retargeting. In *Conference on Programming Language Design and Implementation*, Las Vegas, Nevada, June 1997. SIGPLAN, ACM.

[11] Jim Evans. The new plagiarism in higher education: From selection to reflection. *Interactions*, 4(2), 2000. `http://www.warwick.ac.uk/ETS/interactions/vol4no2/index.htm`.

[12] Saul Gass. Letter to Forum, October 1999.

[13] David D. Kirkpatrick. 2 accuse Stephen Ambrose, popular historian, of plagiarism. *New York Times*, January 5th 2002.

[14] David D. Kirkpatrick. Historian says borrowing was wider than known. *New York Times*, February 23rd 2002.

[15] Ned Kock. A case of academic plagiarism. *Communications of the ACM*, 42(7):96–104, July 1999.

[16] Michael C. Loui. Seven ways to plagiarize: Handling real allegations of research misconduct. *Science and Engineering Ethics*, 8(4), 2002.

[17] Zeeya Merali. Element 118 dropped from periodic table. *Scientific American*, July 2002.

[18] Pamela Samuelson. Self-plagiarism or fair use. *Communications of the ACM*, 37(8):21–25, 1994.