

JOURNALISM 2003 CONFERENCE

FRIDAY 17 OCTOBER 2003

ABC IWAKI AUDITORIUM

11am Keynote: John M Barrie (US)

Plagiarism: John Barrie, one of the world's leading experts on plagiarism, discusses the Jayson Blair affair at the New York Times, the extent of plagiarism in the media and the legal and technological issues surrounding detection and deterrence.

Moderator: Mike Smith, MPC Committee

Mike Smith (MS): Ladies and gentlemen, if I could have your attention, please. My name is Mike Smith from the Press Club committee and welcome to this keynote section of Journalism 2003.

If the Iraq war was the biggest news story of the year, then plagiarism was perhaps the year's biggest professional issue. The Jayson Blair affair, involving plagiarism and journalistic fraud, led to the demise of editors at the *New York Times* and a whole lot of nervousness and soul-searching throughout the world of journalism.

Plagiarism: it's an ugly word. It's a creepy word. In a way it's a bit like the word, homicide. Just as homicide can mean everything from an accidental killing to a mass murder you can be accused of plagiarism for anything from a wholesale lifting of a story to an unconscious or even sub-conscious borrowing of a fact or a turn of phrase. How do we define it? How do we police it? How much does it matter? How do we punish it?

One thing we do know is that in recent years the Internet makes it easier for plagiarists. It also makes it much easier to detect. For instance, this year a reporter on the other side of the world did a *Google* search on a key phrase in one of his own articles to find out to what extent his piece had been syndicated around the world. And what he found is that slabs of his work had been picked up by a reporter in Australia and put under the Australian reporter's name.

To answer some of these questions and these ugly issues raised by plagiarism, we're privileged this morning to have as our first keynote speaker of the day one of the world's experts on plagiarism, Mr John Barrie, from Oakland, California. At this stage to introduce John I will steal – sorry, borrow – from your program notes but that's OK because I wrote them, having earlier pinched them from his website – with permission of course.

John graduated in rhetoric and neurobiology – what a mix that is - at UC Berkeley and then did a PhD in biophysics. He founded companies that developed plagiarism detection systems that track, index, archive and compare literally billions of documents in a short period of time – a frightening thought for plagiarists. More recently, he has taken an interest in plagiarism in the media and that will be one of his focuses this morning. Ladies and gentlemen, please welcome the original John Barrie.

John Barrie (JB): Thank you very much. I'd like to thank the Press Club for the gracious invitation, for bringing me here to speak today and I want to tell you - this is hopefully the only disclaimer I'll make before I begin – I am jetlagged, nervous and highly caffeinated and I have a long presentation so I will try to keep my smart arse

comments to a minimum and hopefully with the microphone in this hand and the laser in this hand everything will go very smoothly.

So today I'm going to talk about plagiarism: a new look at an old problem. As Mike said, I did my studies at Berkeley in rhetoric and neurobiology. I've a doctorate in biophysics and my study had to do with theoretical computational electroneurophysiology. I study spatial patterns of brain waves and how that can tell you about how the brain works and how the brain transduces the sensory world into the neuro world. That's important because I don't have a PhD in catching plagiarists. I don't have a PhD in education. I couldn't have cared less about catching plagiarists. It wasn't even anything I intended ever to go on to do.

So how did I get here? While I was at Berkeley, I was a teaching assistant in some classes and I thought it would be kind of cool to develop some technology to allow students to read and peer review each other's papers. I think it's a little ironic that peer review is pretty much the foundation of all professional academics across the spectrum, but you talk to your typical undergrad and probably 98% of them have never even engaged in peer review. So I thought you know what? At Berkeley we're going to do this. And we're going to allow these students to answer a couple of fundamental questions. That is : I wrote this term paper. What did all these other students write about? And finally I got a "B, check your spelling" or "A, interesting paper" or "D, did you ever come to this class?" but, I mean, that's not a lot of feedback for the students. So I thought what we're going to do is have the students read and peer review each other's papers and have the students give each other critical feedback. And that was published in an article in 1996 in *Science Magazine* about how to use the Internet to extend traditional models of education.

The editors at *Science Magazine* said OK fine we're going to publish your manuscript but we'd like you to make a couple of predictions about how the Internet may not be so good for education. And it was a no-brainer because I had a parade of students coming into my office telling me : well you know all those papers you put on the Internet? These other students are downloading the papers and turning them in to other classes or there's a student out there selling them for \$10 bucks a pop. And I couldn't believe it. I thought OK, what we're going to do here is take some of this pattern recognition technology I was working on in neurobiology and apply it to pattern recognition of large databases of texts: that is to detect plagiarism. And my transformation to the plagiarism expert was complete I think in November 1999 when an AAP story went out regarding *Turnitin* and the journalist said well she ended her first sentence with "John Barrie could be your worst nightmare." So I went from random neurobiology researcher, Berkeley, to worst nightmare.

The unattributed use of a person's ideas seems to be a general problem in our society. Now I'm going to read from some of these slides because typically I used to have the computer right here but OK, bear with me. In our society and let's just say that a lot of these slides apply to what goes on in American society. Whether they go on here, I don't know. Many of our politicians and corporate leaders rarely write their own speeches, their policy positions or books. Some researchers routinely take credit for work done with their grant money. Some famous journalists have research teams to write their articles for them. Many legal opinions written by distinguished judges and justices are really written by their judicial clerks. And plagiarism in academia is rampant. According to the largest study of plagiarism that just came out a few months ago, more than 40 percent of the students have admitted to plagiarising from the Internet. And actually this is a pretty good article right here talking about the culture of plagiarism.

OK, Jayson Blair: Blair plagiarised and fabricated over 36 articles at *The New York Times*. Apart from the last 15 years of plagiarism in the media ...lexus nexus. Here's just the most recent example of a long line of other high profile incidents of plagiarism in the media. And therefore I don't think it's that instructive to focus on Blair to gain insight into this problem. But I do think there are important similarities between journalism and academia. Both institutions consider plagiarism to be a capital crime. Both institutions suffer sporadic plagiarism scandals brought to light mainly by chance. Both institutions rely on a code of honour and integrity to deter plagiarism and fraud. Both institutions typically react to plagiarism scandals by placing the offending party's head on a stick for others to see to further deter unethical activity. In both cases, each institution is being reactive instead of pro-active and their digital plagiarism problem continues despite all efforts to deter it. This is resulting in my humble opinion in a decrease in public confidence.

Probably the most striking aspect, I think, of the Blair episode concerns how the public related to the scandal and that's with a lot of indifference. By placing enough heads on sticks over the years, public confidence in journalism is eroding. In an October 1998 survey by the Media Studies Centre 76 percent of respondents believed that journalists often or sometimes plagiarised and 66 percent said they make stuff up. According to a January 15, 2002 Gallup poll in the United States, 14 per cent of the American public has no confidence in the media and only 32 percent said journalists get the facts straight and that was the lowest level since the Gallup organisation began tracking that number. Well, I think plagiarism in the media and academia is a digital problem. Attitudes among students, that is your future journalists, regarding digital and intellectual property theft are enlightening. Now this is from a September 30, 2003 Gallup poll. When asked is it morally acceptable to download music from the Internet, 83 percent of student said yes, no problem doing it and only 15 percent said it's morally wrong. And I believe that only a digital solution can address this particular digital problem and from here on out I would like to talk about some of our experiences in academia since I'm not a journalist. I'm not really an expert in this field.

Turnitin is our effort at providing a digital solution to this digital problem and after the last couple of years we now have thousands of institutional clients in 51 countries representing over nine million students. And we receive about 15,000 to 20,000 student term papers every day to be checked for plagiarism, to see if they're original or not. Here are some of our clients: every university in the United Kingdom, a lot of the prestigious universities in the United States. And we've gone on to provide actually a commercial version of that service designed at checking other documents for originality: books, magazine articles, things written by hired writers like technical journals. For example, we've done some checking of certain books for large publishers and found that there were some best selling books in their eighth editions that were just plagiarised pieces of crap.

Let me tell you a little bit about how it works. There are two primary ways that somebody can go ahead and take from another source and try to conceal that they have taken from that source. One is to substitute words. In this situation I've shown a paragraph about *Macbeth* and right here I have replaced just about every other word with a synonym. And what this graph shows is that you're going to have to replace about 45 percent of every word in that paper or the paper you've plagiarised from to escape detection by a computer. In example number two, I've taken the same paragraph about *Macbeth* and I've begun to add information from a second source. And what this graph right here is showing you is that not only is this even a worse

way to plagiarise material because we can still find the first source as you add more and more of the red source we can begin to detect that source as well.

OK. Let me just briefly go over how this works. Documents are digitally up-loaded to our system so when an academic client licences our service, the faculty of that institution will add their class sources and the faculty will go to their class: OK, class, give me the paper version of your paper. Go to the Internet, find my class and up-load the digital version of your paper, the version you wrote in Microsoft Word. So the student goes, browses for their paper, up-loads it into our system. At that point our computers take the student's manuscript and transform that manuscript from a collection of characters and words and into a collection of very long numbers. We then compare that collection of very long numbers to a copy of the Internet that's more than three billion pages large and that we update to the tune of about 40 million pages per day, a group of electronic books including tens of thousands of books including all the classics and a group of journals including the last five years of more than 95 percent of all the academic journals out there. This is material that you won't find on the Internet. Finally, and I'm not showing this here but for our academic clients, we also compare the student's paper to every single student term paper ever submitted to us from day one.

After we extract all those documents, we compare all the documents from the Internet, books and journals to the original student paper and produce a report that is essentially an exact copy of the student paper except that every single phrase, paragraph or page that came from any other source is underlined and colour coded. And that takes place in about less than 10 seconds for a 10-page paper. At that point, all the manuscripts that have a problem are highlighted. So if there is a class of 300 hundred students, the instructor doesn't have to take a look at every single manuscript that was submitted to us. They can just go to certain manuscripts that are highlighted as being well maybe less than original. And here's a very important point I'd like to make. I think it's probably going to be a cold day in hell before any computer can determine whether a document, whether it's a news article or a book or a student term paper is plagiarised or not plagiarised. All we say to our clients is that that underlined material came from these other sources verbatim. You're the professional. You make the call.

Before we talk about this report, and again it's not really important what those words are just that there are some that are colour-coded and some that aren't colour coded. The point I'd like to make is that maybe for some certain specific topics like *Macbeth*, there are not so many new ideas about that topic but there are probably an infinite number of ways to express those ideas. And what we're talking about here is just verbatim copying. For example, the probability that 16 words are strung together in the same order as somebody else has strung them in the same order.... for example these 16 words appeared in this paper and they also appeared in this journal article.... the probability of that happening is on the order of one in a trillion. It's just not going to happen. And the probability that a whole paragraph was copied by chance from another source is about as small as our sun exploding. It's just not going to happen. So what we see here is that here's a copy of the student paper and every single red word was taken from this red source and all the blue words were taken from the blue source and so on and so forth and the words that were changed are not underlined. So it allows the person analysing this to very rapidly determine whether or not there is a problem. After that, if the person wants to take a look at exactly the source that we highlighted versus the paper that was submitted to us they can take a look at the paper

and then our computers also underline all the words from that source that were also found in this paper.

Finally, publishers or non-academic clients have one additional problem that our academic clients don't have. And that is if the documents that they produced is original, for example if a book is original, once that book goes out into our society, the additional problem is that people begin to steal that intellectual property and re-post it somewhere else. As of mid-2001, there were more than 7200 different pirated books on the Internet. So what we do in this case is also show every single place some of these intellectual properties ended up, whether it's in a journal or whether it's somewhere on the Internet.

I think this is a pretty good glimpse of the future. What I've shown here just up until midday 13 October is the number of submissions that we receive from our academic clients and we're talking – again it's bumping around 16,000-20,000 student papers - every day. And we're seeing that submissions are doubling year over year. Essentially what this means is that academic institutions are beginning to take plagiarism at their institution extremely seriously. They're going to do something about it.

In the next graph, I'm showing an example from UCLA. UCLA has been our client for over three years. In the first year of their service, the papers that we received from UCLA peaked out about 30 to 40 percent level of unoriginality of every paper that we received. In the second term, that level decreased and into the third term the levels of unoriginal work in student papers decreased to almost negligible levels. And this I believe is the deterrent effect of using this type of digital solution to address this problem. Once the institution wrote a few articles about using *Turnitin*.... once word got out among the students that well, it's really a pretty bad idea to plagiarise work, after a while their plagiarism problem went to nearly zero.

So why not just use *Google* to search for these things? Of the papers we receive, about 70 percent of those papers pretty much have minimal plagiarism. Thirty percent of those papers - again 30 percent of 15,000-20,000 papers a day - contain significant amount of materials from other sources. And less than one percent of the papers that we receive were entirely copied from another source. What I'm trying to say here is that the plagiarism we see is not so much lifting from one source and being put right into the student paper. It's going to *Google*, searching on a topic, ah Gothic architecture, finding 20,000 web pages, taking a blurb here and a blurb there, adding some transitional sentences, putting an introduction and conclusion on it and turn it in. That's the problem we're talking about. The sources of this material include the Internet, published works in non-academics setting client works - for academia, other student papers. So if a student pulled something out of a journal or pulled something out of another student's paper you can search until the cows come home and you are never going to find that. Let's say the information is on the Internet. At that point this is what you're going to be up against: for a typical 15-page paper, if the paper was put through some type of computer detection, you're talking about 15 seconds regardless of how many papers you submit. If that paper were completely copied from another web-site, you're talking about you know a little over an hour to search for about 120 papers. If those papers were taken from multiple sources, you're up to about two hours for you know for in the order of 10 papers. The take-home message is it's just a complete and massive waste of time to try and go to the Internet and hunt down people who are cheating. The value that our clients get from this type of service actually has nothing to do with hunting down students for cheating. The value that academic institutions get from this type of service is the deterrent effect. The ability to tell their students, OK, look at this institution, there is an expectation of originality in

your work. We're going to use this type of system to let you to peer review each other's papers, to allow you to exchange information about your work and to make sure that everybody is playing by the same set of rules. As a matter of fact, I don't think it's an unfair comparison to say this type of technology is almost like a referee on a football field. The referee is not there because the institution thinks all of their athletes are cheating bastards. The referee is there to make sure everybody is playing by the same set of rules and to call the foul when it occurs. This is the same type of thing.

Now I would like to spend just my last few minutes discussing some of the intellectual property issues regarding using this type of service to determine if something is original or not. As a matter of fact, right before I came over here when I was sitting in the LAX airport somebody sent me an e-mail saying well, hey, there was this article in *The Australian* talking about how plagiarism detection services have hit a glitch in terms of us potentially violating a student's intellectual property rights. Before I go on to talk about all this legal stuff, let me say this. As a practical matter, the thousands of institutions that have signed our contracts would never have signed our contracts if what we were doing was illegal. Those contracts are vetted by the legal departments of all these institutions.

That practical point aside, we've gone out and obtained legal opinions from law firms in the United States, in Australia, in Canada and in the United Kingdom. And in every single instance, the opinion of the law firm is that using this type of technology to vet a paper for originality does not violate the intellectual property rights of the author. And let me go over some of those points right now, using just our legal opinion from the United States. I'm sorry I don't have a separate set of Australian slides but I think you'll get the point here. In the United States, the purpose of our copyright law is to promote the progress of science and useful arts. And from... one of the main thinkers about intellectual property law, "the aim of copyright is to give an author an exclusive right sufficient to create an incentive to produce but not so great right as to undermine the public domain. These rights must be for a limited term and they must promote the progress of science." From our copyright law: fair use: "The fair use of a copyrighted work for purposes such as criticisms, comment, news reporting, teaching, scholarship or review is not an infringement of copyright." Now onto our legal opinion, I'll read this as fast as I can. "The company – our company's – activity may be fairly characterised as criticism as that term is used in the preamble of section 107 of *The Copyright Act* according to Webster's dictionary. 'Criticism, use, fault, finding or censure or the act of judging the merits of something' by this definition the company by investigating the integrity of an author's work is engaged in a form of criticism. The company is judging the merits of the author's work. As a result we concluded the company's criticism of written works constitutes the type of activity that the courts have traditionally characterised in the legislature as recognises fair use of copyrighted material. Our opinion is further supported by a closer look at whether the company has sufficiently transformed the original author's work...and that transformation they are talking about is the transformation of the collection of characters into the collection of numbers "...certainly the company's use involves the complete transformation of the raw material when the fingerprint is created. Further, the purpose of the company's fingerprint creation and analysis is to identify potential plagiarism, which has absolutely nothing to do with the purpose of the original work. Further, the company's originality report, by identifying potential plagiarists, provides new insights and understandings about the original. We believe that the identification of plagiarists is the type of activity that the fair use doctrine is intended to protect for

the enrichment of society.” Finally, “In short, for the same reason delineated above, with regard to our plagiarism analysis, we are of the opinion that storing a copy of an author’s work in the database to be used solely for the purpose of comparing the work to other works constitutes a fair use of the work. In view of the foregoing it is our opinion that under copyright law both types of use by the company of written works despite the lack of expressed consent of the author fall within the ambit of the fair use doctrine and accordingly the company should not be liable for the claims of copyright infringement. A court would find in favour of a defence of fair use.”

Essentially what they are saying is that to take a written work, transform it into a group of numbers, to do the type of analysis we do... and let me tangentially add, again it would take until the cows come home to try to compare a written work to a few billion other things just by looking at similarities in text. You just can’t do it. So the value of a work to us is not so much the collection of characters and words. It’s that transformation into the realm of numbers. And I can tell you right now we have some publishing clients who will not publish a manuscript before it’s vetted by this sort of technology. They’re just not going to put their legal arse out on a line for that. So the case could somehow be made that in that instance value is actually added to the work by the publisher feeling more confident to put it out there in the public domain. Finally, let me give you an example. Leslie A Kelly was a photographer who posted his pictures on a website. *Ariba* own a search engine for images. They copied all of Kelly’s pictures without his permission. They copy the entire picture. They transform the pictures into small thumbnails. They stored the thumbnails in a database they owned. They are a commercial venture and they profited by this type of service but they didn’t harm the market value of Kelly’s work and the court in the United States found that *Ariba* was making a fair use of Kelly’s pictures. You just couldn’t go from that small fingerprint on a website and blow it up into anything that even remotely resembled the original work of Kelly.

Again, in my humble opinion I see sort of two scenarios for the future. Scenario number one is go with the status quo. It’s not working. It hasn’t been working for a while. In that scenario, I believe we produce students, journalism future leaders, with a shaky ethical foundation - for example, Enron, Arthur Anderson; fill in whatever company you like from Australia – and poor critical thinking skills. Trust in our institution continues to erode. I think it’s naive to believe that once a student lays a foundation of shaky ethics in secondary school that when they receive their diploma that all of a sudden they become an ethical person. And you can extrapolate that through college, through professional school right out into our society and into whatever profession they are going into.

Just to give you a good example of that when we were profiled in *Sixty Minutes*, it’s not because we were such great guys even though we are such great guys. It was because at the University of Virginia, in two classes ... let me back up a little bit. The University of Virginia, the University in the United States that has one of the oldest honour codes in our country and a zero tolerance policy. You’re caught cheating one time you’re gone. They catch you after the fact, they call you up and say well, you know that degree we gave you, you might as well shred it because we’re taking it back. Well in two classes without comparing anything to the Internet they found about 130 students plagiarising from each other’s term papers. And they were able to do that because the classes were divided up into sections and the rest of the students knew that their papers in the different sections would be graded by different teaching assistants. The only reason that plagiarism scandal came to light is because one of the students in one of those classes went up to faculty and said OK, I may have gotten a C

on my paper but at least I wrote the damn thing, which is more than I can say for a lot of your other students. And when that physics professor developed a program to compare the student papers to each other, 130 papers dropped out. Now that's going on in two classes to the University with one of the oldest honour codes in our country. Number one, it's naive to believe it's not going on in other classes and it's really naive to believe that it's not going on just about every place.

Scenario number two is that something happens to change what is going on: to apply some type of digital solution to the very digital plagiarism problemfor example, when people are out there downloading music from the Internet the RIAA goes out there, sues a handful of students or a handful of people to try and make examples of those people. Hopefully, those people's heads on sticks begin to deter other people from downloading music. It just doesn't happen the result of this would be more originality and creativity in work, increased sharing among students and within our society and better critical thinking skills.

Well, that was it. Thank you very much.

MS: Thank you, John. We have some time for questions. Who wants to quiz your worst nightmare? Just one quick one from me first. A lot of evidence that universities and other parts of the education system are applying systematic approaches to detecting and educating about plagiarism. What do you know about what's happening in news organisations particularly in the wake of the Jayson Blair affair? Is there any such systematic approach and what can your technology tell us about the incidence of plagiarism in the media?

JB: It's a good question. Actually, I see plagiarism in the media as very similar to what was going on in universities just a few years ago and that is denial that it is a big problem. One of the biggest problems we had in the beginning was to convince universities ...was that they should take a stand against the plagiarism among their students. A lot of things we heard was that well we have an honour code here. Our students aren't like that. Our students don't cheat. Our faculty can spot plagiarism. It's not a problem. It was only after a few brave and high profile institutions stood up and said OK, well we do have a problem. And what we've been doing isn't solving this problem and I think that we're going to have to take another look at it. And that's when those universities began to apply digital solutions in addition to everything else they were trying to do to solve this problem.

And I think it's going to be the same way with media. The first few publishing clients we had, I can remember when one very large publisher saw some of the reports that we produced with some of their books. All she said was, oh my God, oh God. Well, so and so get over there and start paraphrasing all that stuff as fast as you can to get the eighth edition out and oh, and by the way, Mr Barrie, I'd like to remind you of that non-disclosure agreement that your company signed with us. We wouldn't like anybody to talk about this.

Again, I think it's going to be a few brave publishers who initially stand up and say OK, you're right. This is a problem. We're going to do something about it. We're going to take a stand. Just to address how prevalent the situation may be. We were contracted from one large northern Californian TV station to look at the large (San Francisco) Bay Area newspapers and it took us literally less than one hour to find major cases of plagiarism in every single one of the large San Francisco Bay area newspapers. And if it took us an hour to do that, then I believe the problem is a lot more massive than you hear about right now. I believe the problems you hear about

are again just by pure chance. Somebody ran across something and you know called the journalist and they did a story about it. And that publisher goes well, OK, we'll weather this storm and hopefully it'll go away or next time it happens, hopefully it'll happen to somebody else.

MS: Question from the floor.

Richard Connor: I recently completed an MBA at RMIT without plagiarising. However, a friend of mine who was a fellow student had the embarrassment of having one of his final research papers run through software to check for plagiarism and it returned a false positive and it caused him much inconvenience and embarrassment to clear his name. So I just wanted fellow students to know this sort of software can get it wrong.

JB: Well, you know I think it's only fair to allow me to address that. My comment would be that for our type of service, it's impossible to have a false positive. I mean literally, it's impossible. We may say that OK, that text came from these sources. Again we're not saying that that's plagiarism. It's up to the instructor and sit down with the student and say OK, well these computers said that this text came from these sources. Could you please explain why that is? We make no judgment about anything that a student has done. It's just bringing information to the table.

Bruce Williams, Northern Times, Northern Victoria, Kerang: In this country we have our hosts here, the ABC, have a program called *Media Watch*. And often it has brought to the attention of newspapers and television issues of where plagiarism has been rife whether it's been lifting of news footage or articles in newspapers. Now I returned home (Melbourne is my home) from six years living in your country. Yet I found nowhere in the United States and I did quite a bit of travelling there of any sort of similar things where a watch is actually kept on the media other than from an organisation like yours. And I was just wondering if you wanted to comment.

JB: I couldn't agree with you more and as a matter of fact that's why I said in my presentation that typically when these sort of scandals break it's by pure chance. when Jayson Blair was caught plagiarising it was almost by chance. And in Jayson Blair's case, it was inevitable because it was so egregious and it involved so many articles and such a major newspaper. It was bound to happen. But yes, I don't know of anything even similar to the *Media Watch* you have in Australia.

Julie Zilco: I'm doing a doctorate of Business at Monash. I'm interested in what you think the role of an Internet engine like *Google* have to play because obviously when you're doing a literature review for a thesis for instance you have to cover all bases and make sure the theories are touched on. Google's a fantastic way of doing this but even a good researcher and a good, honest researcher can't often cite those sources because those sources are not cited within itself. Do you think that Google is going to change the way it's doing things or any of these Internet search engines to make it easy for somebody to cite.

JB: Right, I think I know what you're talking about. I mean, let me say that as far as the Internet's concerned I think that's probably one of the biggest boons to education since the Guttenberg press in ,what 1440, or something like that? And you're right it

is a problem that sources on the Internet come and go and it's hard to nail them down. *Google* does have their cash diversion of the source, so if it's not there they do show you at least something that could have been there. I don't think that problem is going to be solved anytime soon and I think that's one reason why people will go straight to the academic journals more than they probably will go to the Internet. I think that in the final analysis, to make a good-faith effort at citing something from the Internet is probably the best you can do.

Jonathon Kendall, RMIT: You talked earlier on about students plagiarising in their student career and then maybe going on to do unethical things. Do you think it's fair to make that assumption that they will continue to be unethical after maybe one mistake?

JB: Well, again I wasn't saying that that is necessarily the case. I was saying that I believe it's the height of naivety to believe that's not the case. From what we see, our clients range from elementary school all the way through to professional schools. We get papers across the curriculum. We see about a level of 30 per cent plagiarism across the board. So why is it that somebody that who would cheat their way to a degree would suddenly stop that type of behaviour once they graduated? Maybe, maybe it's the case. Maybe when you get into University of Virginia you say, well there's an honour code here. Of course, I wouldn't cheat at this institution. Maybe that's the case. I don't believe it is.

Listen the final thing that I would like to say is that we tell every single one of our clients to use this type of technology to try to catch students cheating would be a misuse of this technology. To analyse papers and to bring a student in and say hah, I got you. Finally I got you, you cheating bastard would actually be a misuse of the service that we're providing. We recommend to every single one of our clients that ...well first of all, we say you're going to find a lot of people cheating. You just will. But we recommend in every single instance, at least for the first instance, that you bring the student in and say OK, look there is a problem here. Maybe you didn't know how to cite a source. Maybe you didn't know how to write a term paper. Here's the proper way to do it. Let's do it again and this time get it right. It's an educational moment for the student. I believe that's the proper use of something like this. And something I didn't mention is that there are a handful of very well known journalists who had this similar type of problem early on in their career. They were able to overcome it, get over it and go on to have a prestigious career in journalism. I don't think it should be sort of one fatal mistake.

MS: Thank you, John. On that note I'm afraid we're out of time. In my experience, often when the subject of plagiarism comes up in groups of communications professionals, the conversation is stopped when somebody uses the epithet "When you steal from one author it's plagiarism. If you steal from many it's research." They never attribute it and for your information it is from the American writer and dramatist, William Mittner 1876-1937, as quoted in Collins Quotation Finder on page 506, Harper Collins revised edition 2001. John, thank you very much for your time. We could have quizzed you for much longer but John will be around in the corridors for a little while during the day if you want to catch up with him. Please accept these gifts from the Press Club to mark your visit and once again thank you very much.