Turnitin and iThenticate
How Education Now Works At All Levels
Dr. John Barrie – Turnitin, Creator
My Background

- Undergraduate: UC Berkeley
  Rhetoric & Neurobiology

- Doctorate: UC Berkeley
  Biophysics Multidisciplinary Graduate Group, Neurobiology

- Dissertation: Theoretical and computational electro-neurophysiology
  Spatiotemporal dynamics of the neocortical EEG
  (in other words, the physiology of perception)
Problem

- Internet has allowed the public to access massive amounts of information
  This access has been accompanied by a manifold increase in the theft and misappropriation of digital intellectual property

- This problem has become especially pronounced in academia where it takes the form of plagiarism
  - Students are using the Internet like a multi-billion page searchable, cut-and-paste(able) encyclopedia
  - The problem jeopardizes the integrity of the entire system and the addressable academic market is in the billions
Problem

• According to one of the largest studies of plagiarism in the world, Donald McCabe found that almost 40% of students surveyed admitted to plagiarizing information from the Internet - and that percentage may be low because many students did not consider ‘borrowing’ from the Internet without attribution plagiarism.
Problem - Components

- Increasingly IT literate student population (Wikipedia is replacing the library)
- Emergence and proliferation of cheat sites
- End justifies the means philosophy towards education (parents)
- Lack of enforcement by institution
- Competitive pressure
- Bad examples in society (cheaters win)
- Things on the internet are not like their material counterparts
Social Epidemic

- Politicians and leaders rarely write their own material
- Researchers take credit for work done with their grant money
- Some journalists have research teams write their articles
- Legal opinions may be written by judicial clerks not judges
- Plagiarism is rife in professional publications
- Plagiarism in academia is widespread
Major science publishers are gearing up to fight plagiarism. The publishers, including Elsevier and Springer, are set to roll out software across their journals that will scan submitted papers for identical or paraphrased chunks of text that appear in previously published articles. The move follows pilot tests of the software that have confirmed high levels of plagiarism in articles submitted to some journals, according to an informal survey by Nature of nine science publishers. Incredibly, one journal reported rejecting 23% of accepted submissions after checking for plagiarism.

Over the past two years, many publishers (including Nature Publishing Group) have been trialling CrossCheck, a plagiarism checking service launched in June 2008 by CrossRef, a non-profit collaboration of 3,108 commercial and learned society publishers. The power of the service — which uses the iThenticate plagiarism software produced by iParadigms, a company in Oakland, California — is the size of its database of full-text articles, against which other articles can be compared. Publishers subscribing to CrossCheck must agree to share their own databases of manuscripts with it. So far, 83 publishers have joined the database, which has grown to include 25.5 million articles from 48,517 journals and books.

As publishers have expanded their testing of CrossCheck in the past few months, some have discovered staggering levels of plagiarism, from self-plagiarism, to copying of a few paragraphs or the wholesale copying of other articles.

"We are using CrossCheck on about a dozen journals, and it has spotted things that we would otherwise have published," says Aldo de Pape, manager of science and business publishing operations at Springer in Rotterdam, the Netherlands. "Some were very blatant unethical cases of plagiarism. It has saved us a lot of embarrassment and trouble."
NSF Plagiarism Cases

- 2004: 8 cases
- 2005: 9 cases
- 2006: 8 cases
- Subscribed to iThenticate
- 2007: 19 cases
- 2008: 14 cases
- 2009: 25 cases

Most of the cases pertain to more than one proposal. For example, one professor submits multiple proposals all of which have some form of research misconduct in them. Another example is when a professor is found suspect of some sort of research misconduct and then his past proposals are reviewed and it is revealed that those were also plagiarized.
Even In China....

  More than 60 scholars from various universities, research centers and media organizations have jointly issued an open letter urging the Chinese Academy of Social Sciences (CASS) and Tsinghua University to investigate if a well-known scholar and professor is guilty of plagiarism in his new book.

- Plagiarism rocks China's academia again. English.news.cn, July 09, 2010

- Professor Faces Plagiarism Charges. People’s Daily Online, July 12, 2010
Technology is the Key

“Warning students [or journalists or authors or researchers or anyone else] not to plagiarize, even in the strongest terms, appears not to have had any effect whatsoever. Revealing the use of plagiarism-detection software [Turnitin] to the students prior to completion of an assignment, on the other hand, proved to be a remarkably strong deterrent.”

–Actions Do Speak Louder than Words: Deterring Plagiarism with the Use of Plagiarism-Detection Software by Bear F. Braumoeller, Harvard University and Brian J. Gaines, University of Illinois at Urbana-Champaign
Technology is the Key

“Warning students or anyone else not to plagiarize, even in the strongest terms, appears not to have had any effect whatsoever. Revealing the use of plagiarism-detection software [Turnitin] to the students prior to completion of an assignment, on the other hand, proved to be a remarkably strong deterrent.”

“This is a digital problem, and it must be addressed with a digital solution. The status quo is not working big time.”

—Actions Do Speak Louder than Words: Deterring Plagiarism with the Use of Plagiarism-Detection Software by Bear F. Braumoeller, Harvard University and Brian J. Gaines, University of Illinois at Urbana-Champaign
Plagiarism Detection

- An internet based service (no software)
- Accepts student papers
- Produces originality reports
- Identifies content copied from:
  - The internet
  - ...and the archived internet
  - Nearly 200 million submitted student works from nearly 10,000 institutions around the world
  - GALE OneFile
  - The COMPLETE EBSCO database
  - ALL of the content from the world’s STM and academic publishers
  - All Yahoo web and current news content
  - Content from over 31 languages
New Content - Partnerships

- JISC Collections and e-Book project
- British Library EThOS project
- EBSCO
- Pearson
- McGraw Hill
- PubMed Central
- Open Archives
- arXiv.org
- University content (e.g., University of Malaysia; CHE, Thailand)
New Content - CrossRef

- 28+ million journal articles already
  ACM, BMJ Publishing Group, Elsevier, IEEE, Nature Publishing Group,
  Oxford University Press, Sage, Informa UK (Taylor & Francis), Wiley
  Blackwell, Springer, Sage, Am. Institute of Physics, Am. Physical Society,
  Am. Psychological Society, AAAS

- Publisher participation is growing quickly
Key Turnitin Metrics

- 150+ million student papers submitted to Turnitin — projected to grow to over 200 million by the end of 2010
- 180,000-250,000 new student papers daily
- 108 countries
- Integrations with major LMS’s
- Full support for 30+ languages
  Arabic, Catalan, Chinese (simplified and traditional), Croatian, Czech, Danish, Dutch, Finnish, French, German, Greek, Hebrew, Hungarian, Italian, Japanese, Malay Norwegian (Bokmal, Nynorsk), Farsi, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swedish, Thai, Turkish
Largest Online Academic Community

- Nearly 10,000 institutional clients in 108+ countries
- In the UK:
  - Over 99% of Universities
  - Every high stakes awarding body
  - Over 100 schools and colleges
  - 2007 NCC Group Report ranked Turnitin #1 in the world
- We process 250,000 student papers a day at peaks
Largest Online Academic Community

We process 250,000 student papers a day at peaks

THE POWER OF TURNITIN AND ITHENTICATE IS IN THE NETWORK EFFECT
Core Technology

1. A document is uploaded to iParadigms' ASP service.

2. It is then compared to a database of millions of manuscripts, books, periodicals, and billions of web pages.

3. iParadigms adds millions of web pages per day to its Internet node.

4. The results are returned in seconds, highlighted with hyperlinks to the original text.
Needle in the Haystack
Searching the Entire Document

- Manuscript or article submitted to iParadigms
- Computer transforms manuscript to digital fingerprint
- Copy of internet
- Books, journals, newspapers
- Extract matching documents
- Student papers or client node
- Next slide
Digital fingerprint of the manuscript/article is remapped into a high dimensional space and test for clustering.
Originality Report

Matching passages from 12+ billion Internet web pages: updated at a rate of 60-240 million pages/day

Matching passages from millions of Books, Journals, Newspapers

Matching passages compared to original manuscript/article

Matching passages from millions of Student Papers or Client Node

Create Originality Report Entire process: seconds-minutes
To Speak or Sign

The Parental Pressure to Choose between Cochlear Implants, Sign Language, or Both

Sign language is a beautifully expressive, visual language using hand, body and facial expressions to convey words and meaning, however, as more and more parents are learning even as early as birth that their child may be deaf, more children are getting cochlear implants then ever before and there is fear that fewer and fewer children will learn to sign. In fact, as of May 2002, 69 percent of babies born in the United States had been screened for hearing loss as part of universal newborn hearing screening programs. Such early identification of infant hearing loss represents a dramatic change from the situation that existed only six years ago when hearing loss was not identified in children until they were 2-1/2 years of age (Sorkin 2003). It is likely that the implementation of these early screenings have lead to the growth of cochlear implants within the United States to 20 percent or more per year (Sorkin 2003). Many parents view deafness as a disability that with cochlear implants can be cured. They desire that their child learn language and function "normally" in society.
To Speak or Sign

The Parental Pressure to Choose between Cochlear Implants, Sign Language, or Both

On the Rise

Sign language is a nonverbal, expressive, visual language that is used to convey words and meaning, however, as more and more parents are learning even as early as birth that their child may be deaf, more children are getting cochlear implants then ever before and there is fear that fewer and fewer children will learn to sign. In fact, as of May 2002, 69 percent of babies born in the United States had been screened for hearing loss as part of universal newborn hearing screening programs. Such early identification of infant hearing loss represents a dramatic change from the situation that existed only six months before.

(Corkin 2003) It is likely that the implementation of these early screenings have lead to the growth of cochlear implants within the United States to 20 percent or more per year (Corkin 2003). Many parents view deafness as a disability that with cochlear implants can be cured. They desire that their child learn language and function "normally" in society.

Supplemented by sign language, the cochlear implant has been around for 25 years.
To Speak or Sign

The Parental Pressure to Choose between Cochlear Implants, Sign Language, or Both

On the rise

Sign language is a beautifully expressive, visual language using hand, body and facial expressions to convey words and meaning; however, as more and more parents are learning even as early as birth that their child may be deaf, more children are getting cochlear implants than ever before and there is fear that fewer and fewer children will learn to sign. In fact, as of May 2002, 69 percent of babies born in the United States had been screened for hearing loss as part of universal newborn hearing screening programs. Such early identification of infant hearing loss represents a dramatic change from the situation that existed only six

with hearing loss were not identified until they were 2-1/2 years of age (Sorkin 2003). It is likely that the implementation of these early screenings have lead to the growth of Cochlear implants within the United States to 20 percent or more per year (Sorkin 2003). Many parents view deafness as a disability that with cochlear implants can be cured. They desire that their child learn language and function "normally" in society.
Chinese Culinary Arts
BY JOHN SMITH

Unbeknownst to many Westerners, dishes and chopsticks are part of China's unique culinary culture during Chinese New Year and various other holidays.

culinaryculture.com

Cluster View
1. culinaryculture.com 36%
2. Lincoln Land Commity Colle 22%
3. Norwich University 18%
4. JAMA Publication 13%
5. about.com 10%
6. Medical Examiner Publication 8%
7. Diablo Valley College 5%

4. 1949年後，由于国共内战的结果，使得两个继承“中国”之名的国名但互不隶属的政治实体并存，中华人民共和国和中华民国均声称对整个中国拥有主权。中国是世界上最早产生文明并延续至今的国家之一，因其文化传播的广泛性，独特性和成熟性而对周边国家和民族的文化产生深远影响，同时不断吸收各个民族的文明成果而演变为今天的中华文化。
To Speak or Sign: Choosing Cochlear Implants

The Parental Pressure to Choose between Cochlear Implants, Sign Language, or Both

On the Rise

Sign language is a beautifully expressive, visual language using hand, body, and facial expressions to convey words and meaning, however, as more and more parents are learning even as early as birth that their child may be deaf, more children are getting cochlear implants than ever before and there is fear that fewer and fewer children will learn to sign. In fact, as of May 2002, 69 percent of babies born in the United States had been screened for hearing loss as part of universal newborn hearing screening programs. Such early identification of infant hearing loss represents a dramatic change.

Mike Jones

with hearing loss were not identified until they were 2-1/2 years of age (Sorkin 2003). It is likely that the implementation of these early screenings have lead to the growth of

Figure 1 Infant Hearing Test

Cochlear implants within the United States to 20 percent or more per year (Sorkin 2003). Many parents view deafness as a disability that with cochlear implants can be cured. They desire that their child learn language and function "normally" in society.
“...if anything, iParadigms’ use of the students’ works fostered the development of original and creative works ‘by detecting any efforts at plagiarism by other students.’”

–Judge Traxler, U.S. 4th Circuit Court of Appeals
April 16th, 2009

Replace ‘students’ with ‘researchers’ and you have the ultimate benefit that iThenticate holds for the STM world:

more originality & more creativity

iThenticate and Turnitin are the next-generation spell-checkers.
Thank You