## TurnItIn.com

July 26, 2016 Jessica Cox

## That speech...



Melania Trump RNC speech | July 18, 2016

66

From a young age, my parents impressed on me the values that you work hard for what you want in life, that your word is your bond and you do what you say and keep your promise, that you treat people with respect. They taught and showed me values and morals in their daily lives. That is a lesson that I continue to pass along to our son," Trump said.

And we need to pass those lessons on to the many generations to follow. Because we want our children in this nation to know that the only limit to your achievements is the strength of your dreams and your willingness to work for them."



Michelle
Obama

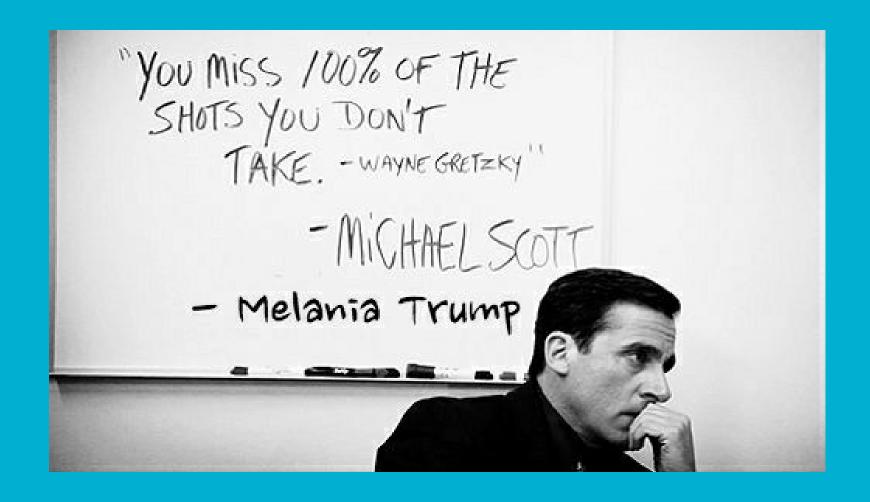
DNC speech | August 25, 2008

66

And Barack and I were raised with so many of the same values: that you work hard for what you want in life; that your word is your bond and you do what you say you're going to do; that you treat people with dignity and respect, even if you don't know them, and even if you don't agree with them.

And Barack and I set out to build lives guided by these values, and to pass them on to the next generation. Because we want our children -- and all children in this nation -- to know that the only limit to the height of your achievements is the reach of your dreams and your willingness to work for them."





#### From recode.net:

POLITICS SOFTWARE

# Here's the software that could have saved Melania Trump from her plagiarism debacle

Donald Trump's campaign co-chair mentioned TurnItln.com in an interview today.

BY AMY KEYISHIAN · @MADFOOT · JUL 19, 2016, 5:22P

#### TurnItIn.com

- Founded in the late 90's by students at Berkley
- Used by more than 120,000 higher ed institutions
- Compares content against:
  - Websites
  - other/previous submissions
  - Books
  - Journals
  - Magazines
- Produces an originality score

Originality GradeMark PeerMark

My Submission

turnitin (7)

98%

65

This first bit is different. Brilliant example of a new originality report.

Anademic writing in American institutions is filled with rules that writers often don't know how to follow. A working knowledge of these rules, however, is critically important; insidervenet missistes can lead to changes of plagarism or the unacknowledged use of somebody else's words or ideas. While other cultures may not insist so heavily on documenting sources. American institutions do. A change of plagarism can have severe consequences, including explains from an university. This handout, which does not reflect any official university policy, is designed to help written develop strategies for knowing how to avoid accidental plagarism.

In college courses, we are continually engaged with other people's ideas: we read them in texts, hear them in lecture, discuss them in class, and incorporate them into our own writing. As a result, it is very important that we give credit where it is due. Plagarism is using others ideas and words without clearly acknowledging the source of that information.

Pagarism Thans using another's work without giving credit. You must put others' words in quotation marks and doe your source(s) and must give citations when using others' ideas, even if those ideas are paraphrased in your own words!

It's an unfortunate fact of life for University lecturers that the pressure on students leads some of them to copy other sudernis satignments or at least to obtain more assistance from their Fiends than is appropriate. Aper from discrediting the use of assignments for assessment, the copying of assignments also writes the assignments educational aim. The typical institutional response is to require that assignments only form a small part of a student's assessment. However, such a response is inappropriate because it other results in trivial assignments or in assignments which do not adequately repay students in marks for the effort that they have invested.

is trucial phrase in this definition is "as one's own." In all learning at University it is completely acceptable to use associated person's Woodysts, writings or inventions to aid our own learning and understanding. Indeed, this is a primary method of learning. We all need sextbooks, released papers, insmalls and many other documents, and male use of the material constanted in them. This is perfectly normal and acceptable?

Used Nightingsle, the deputy vice-chancellor of Kent University said he would not commert on individual cases because the external examiners' meeting for the School of English would take place next week, when exam entries would be considered.

But he said: "I would stress that throughout their time at Kent, all students are given clear guidelines a well as practical advice and support as so what constitutes plagarism. These spell is out that it is not acceptable under any circumstances.

"For example, in the School of English this information is provided in the fisculty handbook and in the department's own handbook, both of which are issued so all students.

"All students are also encouraged to attend the regular workshops on study-skills run by the university which also provides on-line advice."

Match Overview					
30		1/2			
1	Submitted to University Student paper	30%			
2	www.hilpers.org Internet source	28%			
3	www.cs.su.oz.au Internet source	20%			
4	www.cs.bham.ac.uk Internet source	13%			
5	WWW.UWO.CB Internet source	7%			

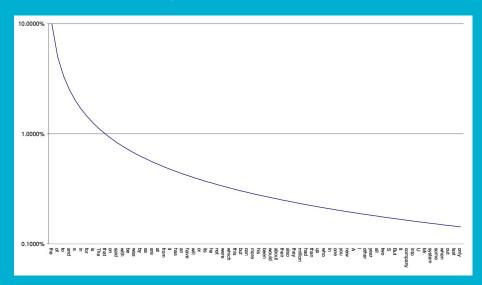
## Possible TurnItIn Algorithms

#### Hits and Misses

- Make a list of all unique words in the document
- For each 'hit' in paper being evaluated, similarity score increases
- Does not take word frequency into account

#### Word Frequency

Zipf's Law



#### **Hamming Distance**

- Defined as the number of bit positions in which a pair of words differ in binary form.
- Example:
  - o ASCII(dog) = (100, 111, 103)
  - ASCII(cat) = (99, 97, 116)

As ASCII limits the number of characters to  $2^8$ =256, we can write each word:  $dog = 100 \cdot 256^2 + 111 \cdot 256^1 + 103 \cdot 256^0 = 6553600 + 28416 + 103 = 6582119$ =  $(11001000110111101100111)_2$ 

$$cat = 99 \cdot 256^{2} + 97 \cdot 256^{1} + 116 \cdot 256^{0} = 6488064 + 24832 + 116 = 6513012$$
$$= (11000110110000101110100)_{2}$$

Those two binary strings

11001000110111101100111

1100**011**0110**000**1011**1**01**0**0

24 - 9 / 24 = 62.5% similar

## Hamming distance

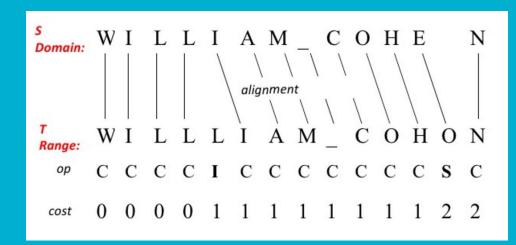
The	dog	ran	Home	because	it	was	raining
The	cat	ran	home	because	it	was	sunny

And calculate the Hamming distance to be 2 out of 8 words (atoms). We say, the two

sentences are 
$$\frac{8-2}{8} = 75\%$$
 similar.

#### Levenshtein Distance

- Measure of the similarity of two strings
  - distance is the number of deletions,
     insertions, or substitutions required to
     transform
  - Penalty assigned for every change need to match the string being compared to



## **TurnItIn.com** Experiment

A researcher turned in 3 different papers:

- 1. An original research report
- 2. An entirely copied report
- 3. A report cobbled together from several different internet sources, but completely plagiarized

## **TurnItIn.com** Experiment

A researcher turned in 3 different papers:

- 1. An original research report 3%
- 2. An entirely copied report 90%
- 3. A report cobbled together from several different internet sources, but completely plagiarized 82%

#### Tricks- Problems

I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character.

I hav a dr am that my four littl childr n will on day liv in a nation whr thy will not b judg d by th color of thir skin but by th cont nt of thir charactr.

#### **Ethical Concerns**

# Turnitin: if you're a student, all your intellectual property might belong to us

To most students, it seems like Turnitin is just another case of "The Man" being out to get them.



By David Gewirtz for ZDNet Government | April 14, 2010 -- 08:22 GMT (01:22 PDT) | Topic: Legal

TurnItIn creates a database of your papers for other companies to check against, and in the end uses it to sell their product

# Thank you!