Fachhochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences

Fremde Federn Finden

Test von Plagiatssoftware 2007

Wann ist ein Plagiat ein Plagiat? Wie kann man es entdecken?

Plagiate hat es auch in der Wissenschaft schon immer gegeben. Das Thema hat in Zeiten des Internets allerdings eine neue Dimension angenommen. Es ist einfacher geworden, Plagiate in kurzer Zeit und mit überschaubarem Aufwand herzustellen. Copy, Shake, Paste, fertig ist der „Remix", ohne lästiges Recherchieren, umständliches Zitieren, Nachlesen und Formulieren.

Was versteht man unter einem Plagiat?

„Plagiat ist die aus freier Entschließung eines Autors oder Künstlers betätigte Entnahme eines nicht unbeträchtlichen Gedankeninhalts eines anderen für sein Werk in der Absicht, solche Zwangsanleihe nach ihrer Herkunft durch entsprechende Umgestaltung zu verwischen und den Anschein eigenen Schaffens damit beim Leser oder Beschauer zu erwecken."

Definition von Dr. Paul Englisch, aus:
„Meister des Plagiats oder die Kunst der Abschriftstellerei"

Was ist Software zur Plagiatserkennung und wie funktioniert sie?

Zur leichteren Erkennung von Plagiaten haben verschiedene Unternehmen Computerprogramme entwickelt. Sie werden zu unterschiedlichen Konditionen wahlweise online oder offline angeboten. Die Programme vergleichen digital vorliegende Inhalte mit dem Internet oder mit internen Datenbanken bzw. leisten beides. Dabei suchen sie nach Übereinstimmungen, melden Verdächtiges und versuchen am Ende eine Antwort auf die Frage zu geben, ob es sich bei dem vorliegenden Text um ein Plagiat handelt oder nicht.

Welche Systeme sind getestet worden?

Im Sommersemester 2007 haben Prof. Dr. Debora Weber-Wulff und ihr studentischer Mitarbeiter Martin Pomerenke von der FHTW Berlin siebzehn Softwaresysteme getestet, die vorgeben, Plagiate zu erkennen.

Folgende Softwareprodukte wurden getestet:

- AntiCutAndPaste

- ArticleChecker
- CatchItFirst
- Copyscape (Free und Premium Versionen)
- DocCop
- Docoloc
- Ephorus
- iPlagiarismCheck
- JPlag
- picapica
- PlagAware
- StrikePlagiarism
- TextGuard
- turnitin
- Urkund
- WCopyFind

Überblick – Top, Flop, Absteiger und Aufsteiger

Top
Nummer eins im Feld war Ephorus, ein niederländisches System: einfach, praktisch, zielorientiert. Ephorus zeigte eine sehr gute Nase für Plagiate und erreichte 40 von 60 Punkten. Verbesserungsbedürftig ist lediglich die Handhabung. (http://plagiat.fhtw-berlin.de/software/s01-ephorus/)

Flop
Das kanadische System CatchItFirst fuhr mit 0 von 60 möglichen Punkten einen Rekord ein: Kein einziges Plagiat wurde erkannt. Für einen Dollar pro Test mussten die Wissenschaftler bis zu 32 Stunden auf ein Ergebnis warten. (http://plagiat.fhtw-berlin.de/software/s18-catchitfirst)

Absteiger: Das US-amerikanische Produkt turnitin gilt zwar weithin als sehr effektiv, war jedoch oft nicht einmal in der Lage, Wikipedia als Quelle zu finden, obwohl dies die erste Adresse für viele Plagiatoren ist. Turnitin erreichte mit 26 von 60 Punkten nur Platz 8 und erhielt die Note „befriedigend". (http://plagiat.fhtw-berlin.de/software/s02-turnitin/)

Aufsteiger
Das deutsche System PlagAware war ursprünglich gar nicht auf der Testliste enthalten, weil sein eigentliches Ziel nicht die Plagiatserkennung ist, sondern die Überwachung einer Internetpräsenz und das Auffinden von

Kopien derselben. Dennoch machte die Software einen überraschend guten Job und überzeugte außerdem mit seinem User-Interface. Das System schaffte 34 von 60 mögliche Punkte (befriedigend) und teilt sich damit mit dem schwedischen Urkund und Indigo Streams Copyscape Premium den dritten Platz.
(http://plagiat.fhtw-berlin.de/software/s22-plagaware/)

Das Gesamtergebnis finden Sie unter
http://plagiat.fhtw-berlin.de/software/

Testmethodik

Um Plagiatserkennungssoftware zu testen, haben wir eine Sammlung von Testfällen konstruiert. Diese sind kurze Aufsätze, von denen wir wissen wie plagiiert wurde, wie viel plagiiert ist, und welches die Quellen sind. Wir haben verschiedene Plagiatstechniken eingesetzt und haben dabei versucht, die ganze Bandbreite anzuwenden. Es wurden auch Plagiate von Büchern und CDs beigemischt sowie Übersetzungsplagiate angefertigt, aber auch Originale hergestellt. Weiter unten auf dieser Seite ist eine genaue Auflistung der Plagiate. Da wir die Testfälle als Übungen für die Lerneinheit "Fremde Federn Finden" einsetzen, war es notwendig, die Erlaubnis von allen Urhebern einzuholen, damit die Ergebnisse publiziert werden können.

Wir haben dann Namen von potentiellen Testkandidaten ermittelt und um Zugänge für einen Test beim Hersteller nachgefragt. Viele haben sofort geantwortet, andere haben zwar versprochen, sich darum zu kümmern, aber nicht mal auf Nachfrage gelang es uns, dort einen Testkonto zu bekommen. Die kostenlosen Dienste haben wir ohne Anmeldung verwendet. Insgesamt gab es 25 Kandidaten, von denen zwei Code-Vergleichs-Systeme und eines ein Collusion-Erkennungs–System war. Von den 22 identifizierten Plagiatserkennungssystemen waren zwei identisch und nur unter anderen Namen vermarktet (turnitin und iThenticate). Es blieben 21 Systeme, von denen wir aus verschiedenen Gründen 7 nicht testen konnten - die Gründe sind im Testbericht vermerkt.

Wir hatten die Testfälle als .doc, .html, .pdf und als .txt vorzuliegen, damit jedes System genau das Material so bekommen konnte, wie es es brauchte. Alle 20 Testfälle wurden dem System gleichzeitig zum Testen angeboten. Ein Fall hatte im Dateinamen einen Umlaut, um zu sehen, ob die Systeme damit zu Recht kommen können. Eine Kopie der Datei ohne Umlaut im Namen wurde verwendet, wenn es nicht mit dem Umlaut-Namen klappte.

Wir haben die Zeit gemessen, die benötigt wurde um die Suche durchzuführen (inkl. der Zeit, um die Dateien hochzuladen) und unser subjektives Empfinden bei der Verwendung der Sites notiert. Als die Ergebnisse vorlagen, wurden die einzelnen Berichte analysiert. Wir haben versucht, die Position einer Nicht-Informatik-Lehrkraft einzunehmen, um zu entscheiden, ob wir die Ergebnisse als korrekt (Erkennung ob Plagiat oder Nicht-Plagiat) werten konnten. Die Kriterien für die Bewertung sind auf einer getrennten Seite beschrieben.

Nach Abschluss der Analyse gab es einen zweiten Durchgang, da viele Systeme "aktualisierte" Versionen für Ende August angekündigt hatten. Wir haben diese erneut getestet und als Gesamtnote den Durchschnitt aus beiden Durchgängen gebildet. Auch fanden sich neue Systeme, die ebenfalls mitgetestet wurden.

Es folgte eine Gesamtdurchsicht der Bewertungen, für einen Testfall (Nr. 6) wurde die Bewertung umgestellt und alle Ergebnisse entsprechend korrigiert. Dann wurde geschaut, dass alle Systeme an einigermaßen gleichen Maßstäben gemessen wurden, bevor die Rangliste mit der Einteilung in Güteklassen aufgestellt wurde.

Testfälle

Zwanzig Testfälle wurden bei dem Test der Plagiatssoftware 2007 verwendet. Die Testfälle 0-9 wurden bereits für den Test 2004 erstellt und befanden sich daher gelegentlich in Datenbanken von Plagiatserkennungssoftware - ohne Erlaubnis, versteht sich, denn es wurde damals explizit gebeten, die Testfälle wieder zu tilgen. Durch eine Missachtung der Suchmaschinendirektiven wurden ausserdem teilweise diese Texte in Suchmaschinen-Datenbanken vorrätig gehalten, sie sind inzwischen durch persönliche Intervention wieder bereinigt worden.

Die verwendeten Testfälle lassen sich wie folgt klassifizieren

- Originalaufsätze
- Originalaufsatz, von der Autorin in die Wikipedia eingestellt
- Übersetzungsplagiate
- Copy & Paste mit Shake & Paste
- Copy & Paste
- Shake & Paste
- Halbsatzflickerei
- Gekauft bei Hausaufgabenbörse

- Copy & Paste eines offline Mediums (Buch, bzw. CD-ROM)
- PDF Quellen für Copy & Paste.

Die Testfälle sind als Übungsmaterial der E-Learningeinheit "Fremde Federn Finden" zu finden (http://plagiat.fhtw-berlin.de/ff/)

Bewertung

Bei der Bewertung des Plagaitstests 2007 gab es 4 Stufen, die erreichbar waren, mit den Punktzahlen 0-3 versehen. Für jede Arbeit waren die Kriterien leicht anders. Wir haben versucht, im Nachgang der Punktevergabe noch mal alle miteinander zu vergleichen und hoffen, dass jetzt alles gleichmäßig bewertet wurde. Wird eine Diskrepanz festgestellt, erbitten wir Nachricht unter plagiat@fhtw-berlin.de.

	3	2	1	0
00-schaltjahr	wenn nichts gefunden wurde	bis 10% Plagiat gemeldet	bis 25% Plagiat gemeldet	Große Mengen Plagiat gemeldet und/oder Warnfarbe vergeben
01-djembe	Englische Quelle gefunden	Plagiat der Seite (auf Englisch) gefunden	Plagiat der Seite (auf Deutsch) gefunden	Nichts gefunden
02-atwood	Amazon.de Quelle gefunden	Plagiatsseiten gefunden ohne Amazon	unter 20% gemeldet	Nichts gefunden
03-IETF	WZ Berlin Quelle gefunden	Quelle da, aber verwirrender Bericht	Nur Plagiat gefunden	Nur wenig oder irrelevantes gefunden
04-döner	Alle drei Quellen gefunden, Humboldt-gesellschaft, tk-logo, Wikipedia	Zwei von drei Quellen gefunden	Eine Quelle gefunden	Nichts gefunden
05-telent	Hacker's BlackBook (pdf) gefunden	Auszug aus BlackBook gefunden	Quelle gefunden, aber viel Blödsinn dabei	Nichts gefunden
06-ftf	Wikipedia gefunden, Hinweis auf Autorin	Wikipedia gefunden und prominent (also wenige Mirrors)	Nur Mirror gefunden	Nichts gefunden
07-ahorn	Beide Quellen gefunden	Eine Quelle gefunden	Nur Wikipedia Mirror	Nichts gefunden
08-lettau	Da ich hier nicht prominente Autorin bin, haben	Nur Wikipedia Mirrors gefunden	Kein Wikipedia, keine Mirrors, nur Unsinn ge-	Nichts gefunden

Fachhochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences

Fremde Federn Finden

Test von Plagiatssoftware 2007

08-lettau	Da ich hier nicht prominente Autorin bin, haben wir grün dafür gegeben, wenn nur die Wikipedia gefunden wurde	Nur Wikipedia Mirrors gefunden	Kein Wikipedia, keine Mirrors, nur Unsinn gemeldet	Nichts gefunden
09-frosch	Hinweis auf schoolunity prominent, ggf. mit weitere Quellen			Nichts gefunden
10-fraktur	Wikipedia, PDF und Buch gefunden	Wikipedia und PDF gefunden	Wikipedia oder PDF gefunden	Nichts gefunden
11-mankell	Quelle e-script gefunden	Plagiat gefunden	Unsinn gemeldet	Nichts gefunden
12-mikrobrauerei*	Englische Quelle gefunden (Wikipedia)			Nichts gefunden
13-allspice	Hinweis auf Buch gefunden			Nichts gefunden
14-schmeling	Nichts gefunden	Kleinigkeiten gemeldet		SEO-Page mit mehr als 10% Plagiat gemeldet
15-bedürfnisanstalt	Buch oder DVD gefunden			Nichts gefunden
16-jelenik	Alle drei Quellen: dieterwunderlich, Hamburger Abendblatt und englische Quelle	dieterwunderlich und Abendblatt	nur eins, oder Plagiate davon	Nichts gefunden
17-squaredance	Quelle eaasdc gefunden	Quelle gefunden, aber viele unsinnige Stelle dazu	Quelle mit extremer Anstrengung zwar zu sehen, aber sehr viel anderes	Nichts gefunden
18-vikinger	Quelle Reinhold Wulff/HU gefunden	Quelle nur abschnittsweise gefunden	Quelle zwar genannt, aber erscheint im Bericht nicht direkt	Nichts gefunden
19-blogs	PDF Quelle Erik	Quelle ange-ge-	Quelle gefunden,	Nichts gefunden

* Es wurde hier einmal Schwarz = -1 Punkt vergeben für absoluten Blödsinn, was als Plagiat angekündigt wurde

Ergebnisse

Sehr gute Software

Leider ist diese Kategorie überhaupt nicht besetzt. Das beste Softwaresystem im Test erreichte lediglich 40 von 60 Punkten. 51 Punkte (17/20 korrekt oder 85%) wären notwendig, um in diese Kategorie aufgenommen zu werden. Studenten brauchen in ihre Hausarbeiten oder Klausuren 90% der erreichbaren Punkte für ein sehr gut…

Gute Software

Folgende Systeme sind für gut befunden worden: (haben 60-85% der erreichbaren Punkte bekommen, also 36 Punkte oder mehr)

- Platz 1 und alleiniges System in dieser Kategorie ist **Ephorus**, das wir in zwei Versionen getestet haben, eine alte (mit 42 Punkten) und eine neue (mit gerade 36 Punkten), die im Durchschnitt 38 Punkte erreichten. Das System hat einige Usability-Probleme, man kann leicht in Zustände hinein kommen, in dem die Berichte nicht scrollbar sind, und die Bedienung ist nicht intuitiv. Es gibt auch Probleme mit Umlauten, und es stimmt uns recht bedenklich, dass die neue Fassung schlechter geworden ist. Das System bietet drei verschiedene "Stärken" der Überprüfung an, die aber relativ sinnlos sind. Beim Auffinden von Quellen hat dieses Softwaresystem die meisten Quellen aufgedeckt. Wenn für das System die Oberfläche überdacht wird, sich mit PDF-Inhalten auseinander gesetzt wird, und das Umlautproblem gelöst wird, sowie herauskommt, wieso der zweite Test nicht so effektiv war, wird das System immerhin brauchbar sein, um auf Plagiats-Verdachte hinweisen zu können.

Befriedigende Software

Folgende Software wurde als befriedigend empfunden, die wir als alle Systeme definieren, die mindestens 40% der Punkte erreichten, also 24 Punkte oder mehr. Man sollte aber beachten, dass bei allem, was unter 30 Punkten ist (50%) man genau so gut eine Münze werfen könnte, um zu entscheiden, ob eine Arbeit plagiiert ist oder nicht.

- Platz 2 mit 35 Punkten: **docoloc**, ein System der Technischen Universität Braunschweig. Die Berichte sind etwas gewöhnungsbedürftig und die Benutzerführung könnte von einer Überarbeitung profitieren (Symbolik und Namen überdenken, Layout verbessern,

Hochladen von ZIP-Archiven zulassen). Aber bei der Erkennung von Plagiatsquellen war dieses System vom Mittelfeld das Beste, sicherlich weil es auch Quellen in PDF-Dokumenten finden konnte.

- Platz 3 mit 34 Punkten ist dreifach besetzt:

  - **Urkund**, das schwedische System, haben wir zweimal getestet, einmal mit 33 Punkten und einmal mit 35 Punkten. Das neue System ist beim Finden von Plagiaten etwas besser geworden, es ignoriert kleinere Bearbeitungsversuche - aber nicht ganze Sätze, die eingeschoben oder entfernt werden. Aber die neue Oberfläche zeigt sehr stark, dass es noch in der Entwicklung ist und es bedarf noch viel Arbeit, bevor es produktiv eingesetzt werden kann.

  - **Copyscape Premium**. Mit geringfügig besserem Komfort und ohne Werbung bekommt man fast so gute Ergebnisse wie mit der kostenlosen Version. Die Preise sind mit 5 US Cent pro Test für Bezahltests extrem günstig.

  - **PlagAware**, eigentlich nur zum Auffinden von Plagiaten von Web-Sites gedacht, schneidet recht gut ab. Wir haben für den Test unseren Aufsätzen online gestellt und mit dem Logo präpariert, das die Site verlinkt. Nur solche vorbereiteten Seiten werden auch überprüft, was sicherlich nebenbei viel dazu tun wird, den Google PageRank der Site zu erhöhen.

- Platz 6, mit 32 Punkten ist das System **Copyscape free**. Wenn man nur wenige Texte testen will (maximal 10 im Internet erreichbare Dokumenten pro Monat), dann kann man sehr einfach mit befriedigenden Resultaten Copyscape einsetzen. Es funktioniert schnell und ohne viele Umstände.

- Platz 7 mit 29 Punkten ist das System **TextGuard**. Das System hat die einfachen Plagiate erkannt, aber auch viele unsinnige Stellen markiert, die teilweise mitten im Wort begannen oder aufhörten. Das System kann aber mit PDF-Quellen umgehen. Allerdings ist die Handhabung sehr schwerfällig für mehr als einen Einzeltest. Die Ergebnisfenster sind extrem schwer zu lesen und zu vergleichen, daher nicht unbedingt tauglich für den Hochschulalltag.
- Platz 8 mit 26 Punkten ist ebenfalls doppelt besetzt:

- **turnitin**, von vielen als "das beste Plagiatserkennungssystem" angesehen, weil das System eine sehr gute Bedienoberfläche hat, die sich nahtlos in den Hochschulbetrieb einpassen lässt. Wie im letzten Test 2004, kränkelt turnitin an vielen Fronten: Wikipedia wird nicht als Quelle erkannt, es gibt immer noch Probleme mit Umlauten in Dateinamen und im Text (die Übereinstimmung bricht an der Stelle, wo ein Wort ein Umlaut hat, ab), nur 1 von 3 PDF-Quellen wurden gefunden. Dafür sind neue Probleme hinzugekommen: Spamseiten dominieren die Suchresultate und die vorgenommene "Eindeutschung" hätte gerne von einer Fachkraft durchgesehen werden können ("veranschlagen ausschließen" für "Zitate ignorieren" ist ein besonders problematischer Fall). Nicht erklärbar ist die extrem starke Übereinstimmung zwischen turnitin und iPlagiarismCheck, siehe hierzu die besondere Seite im Langbericht.
- **ArticleChecker** (es gibt zwei unterschiedliche Produkte, die so heißen, gemeint ist hier das unter articlechecker.com zu findende) hat eine grauenhafte Oberfläche. Man kann bis zu 5 Dateien angeben, es wird sehr schnell nachgeschaut, auf Wunsch in Google, Yahoo und MSN. Die Ergebnisseite ist selbst für erfahrene "Plagiatsjäger" eine Zumutung. Man muss die Quelltexte genau kennen (was in der Regel nicht gegeben ist), weil nicht gekennzeichnet wird, welches Ergebnis zu welcher Datei passt. Man muss selbst auf einen Link klicken, der mit schwer lesbaren Zahlen 0 bis 8+ gekennzeichnet ist und sich aus dem Ergebnis die Übereinstimmung zusammensetzen. Wir haben uns durchgebissen und waren erstaunt über die recht guten Ergebnisse. Es hat also die einfachen Plagiate gefunden, ist aber nicht für den Alltagseinsatz tauglich.

- Platz 10 mit 25 Punkten: **picapica**, ein experimentelles System, das angeblich Textanalysen anfertigt und Brüche erkennt. Da es sich noch in der Experimentierphase findet, ist das System extrem schlecht bedienbar und die Berichte schwierig zu lesen. Einige einfache Plagiate wurden erkannt.

Nicht zweckmässige Software

Folgende Software-Systeme eignen sich nicht, Plagiate zuverlässig zu erkennen:

- Platz 11 mit 17 Punkten ist **DocCop**. Das System macht eine unglaublich rechenintensive Operation, in dem es alle Teilzeichenreihen untersucht, also ein Textfenster immer um ein Zeichen weiter verschiebt. Die Berichte dauern ewig, man kann nicht mehrere gleichzeitig starten und es ist eine Zumutung, die Berichte zu empfangen. Es sind riesig große E-Mails mit sehr wenig Inhalt. Das System findet nicht mal alle einfachen Plagiate. Es gibt keine Gegenüberstellung, man muss selbst in einer Suchmaschine nachschlagen. Da kann man gleich die Suchmaschine bedienen und sich das Warten auf den Bericht ersparen.
- Platz 12 mit 12 Punkten ist doppelt besetzt:

  - **iPlagiarismCheck**. Dieser Software liefert sehr ähnliche Berichte und extrem ähnliche Ergebnisse wie turnitin. Da man jedoch keine Quellen wegklicken kann, konnte man bei den ersten 10 Aufsätzen den Link auf unsere eigene Seiten nicht löschen, die natürlich so um 100% Plagiat waren. Eine längere Diskussion der Ähnlichkeit turnitin/iPlagiarismCheck ist zusammengestellt worden.

  - Die polnische Firma **StrikePlagiarism**. Das Ergebnis müßte eigentlich noch schlechter sein, weil so oft plagiierte Seiten gemeldet wurden, weil keine Quellen gefunden wurden und weil man die "korrekten" Originale nicht werten sollte. Überhaupt wurden nur 2 korrekte Quellen im Test gefunden. Dass sie für diese Leistung 2 € pro Test verlangen, ist recht verwunderlich.

- Platz 14 und letzter mit 0 Punkten ist der Online-Dienst **CatchIt-First**. Für viel Geld und viel Geduld bekommen die Kunden nichts geboten - man bekommt immer 100%ige Originalität versichert, auch wenn man mit solcher Software so etwas nie beweisen kann, sonder höchtens auf nicht-originale Passagen hinweisen kann. Man sollte sich das Geld sparen und Münzwürfe einsetzen - das ist effektiver.

Es gibt von einigen Herstellern Stellungnahmen zur Test, die auf der Website bei den einzelnen Systemen aufgenommen wurden.

Für mehr Information kontaktieren Sie
Prof. Dr. Debora Weber-Wulff
FHTW Berlin *weberwu@fhtw-berlin.de*