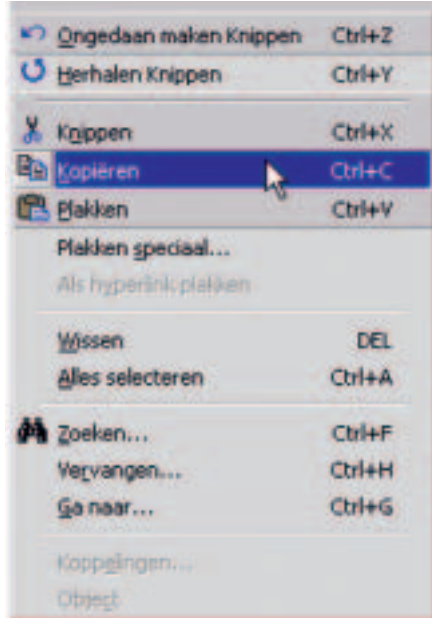


Intellectuele normen en waarden

Steeds meer onderwijsinstellingen halen de werkstukken van hun studenten door een plagiaatscanner. Het massale jatten van andermans werk via internet is anders niet meer in te dammen. Studenten reageren betrappt of verontwaardigd. Toch is de controle nog lang niet waterdicht, blijkt uit een test

George van Hal



Een recent onderzoek onder studenten aan de universiteit van Utrecht laat zien dat één op de twaalf studenten hele lapen tekst ongewijzigd en zonder bronvermelding van het internet kopieert. Na het verschijnen van het rapport was in veel kranten te lezen dat maar liefst zestig procent van de Utrechtse studenten plagieert. Dat getal blijkt echter onjuist. “*De Volkskrant*, die als eerste over ons schreef, heeft percentages opgeteld waar dat niet mocht,” vertelt onderzoeker Femke Jacobs van stichting Onderwijs Evaluatie Rapport, “daarna is dat door iedereen overgenomen.” Klakkeloos overnemen uit een dubieuze bron is dus ook de media niet vreemd.

In wetenschappelijke kring is het plegen van plagiaat een doodzonde. Daarom ondernemen steeds meer universiteiten actie tegen het knippen en plakken. De plagiaatscanner lijkt daarbij aan terrein te winnen. Deze software vergelijkt het ingeleverde werkstuk met teksten in een internetdatabase, eventueel uitgebreid met een privé-database van op de universiteit zelf ingeleverde stukken.

Een kleine zoektocht op internet leert dat er een ware wildgroei aan scanners is. Er zijn algemene plagiaatscanners die internationaal veel klanten hebben, maar ook kleinere, specifieke scanners die bijvoorbeeld regels code van computerprogramma's vergelijken. Het bekendst zijn de programma's Ithenticate en Turnitin en hun grootste concurrent, Mydropbox.

Een populair programma van Nederlandse makelij is Ephorus. De scanner wordt inmiddels door meer dan

dertig onderwijsinstellingen gebruikt. “In het begin moesten we hen uitleggen waarom anti-plagiaatsoftware handig is, maar inmiddels stijgt de interesse duidelijk”, vertelt Joep Chappin van Ephorus.

Vermoedelijk is ook de aandacht die staatssecretaris Rutte van onderwijs de afgelopen maanden aan het probleem schonk een belangrijke stimulans. Hij noemde in *De Telegraaf* de resultaten van het Utrechtse onderzoek ‘zeer verontwaardig’ en overweegt om de maximale straf (een jaar uitsluiting van onderwijs en tentamens) te verhogen. Zo stelde hij voor om studenten die plagieën bij hun eindschrijving definitief uit te sluiten van onderwijs in dat vakgebied. De staatssecretaris werkt op dit moment aan een voorstel voor een nieuwe wet op het hoger onderwijs en onderzoek, waarbij de sanctie op het plegen van plagiaat waarschijnlijk aan de orde zal komen.

Vingerafdruk Een plagiaatscanner in z'n simpelste vorm zet een tekst om in een rij bits, nullen en enen (bijvoorbeeld door elk karakter te vervangen door z'n ascii-code). Het programma inventariseert vervolgens welke rijen nullen en enen van een zekere lengte in een tekst voorkomen. “Hebben twee teksten meer dan een bepaalde grenswaarde aan dezelfde bitrijen, dan krijgt de gebruiker een melding”, vertelt Chappin. Overigens zullen, afhankelijk van het vak, specifieke woorden vaker overeenkomen zonder dat van plagiaat sprake hoeft te zijn. “Daarom kan de gebruiker de grenswaarde met de hand instellen”, aldus Chappin. Sommige scanners gaan ook na hoe ‘verwant’ teksten zijn door op veel voorkomende combinaties van woorden te letten. Zo ontstaat van een tekst een digitale vingerafdruk met statistische gegevens.

De vingerafdruk versnelt het zoeken in een grote database, een belangrijk aspect bij het scannen. Alle tek-

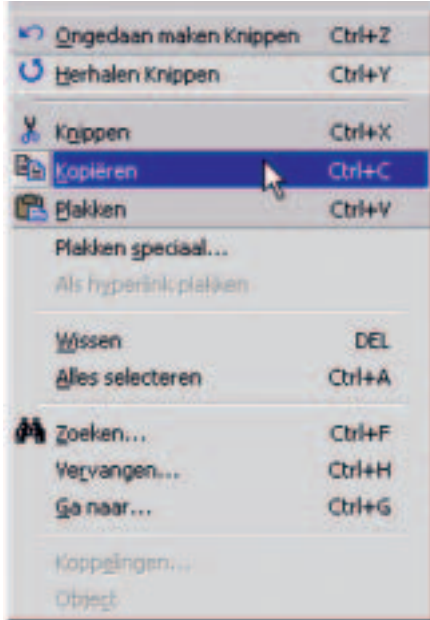
sten met een sterk afwijkende vingerafdruk ten opzichte van het te controleren artikel zijn niet relevant.

“In de toekomst zal het met Ephorus ook mogelijk zijn om op stijl te controleren”, vertelt Chappin. “We kijken dan naar de gemiddelde woord- en zinslengte en het aantal gebruikte lidwoorden.” Van de schrijver wordt dan een stijlprofiel aangemaakt. Dat profiel wordt natuurlijk nauwkeuriger naarmate meer tekst beschikbaar is. Onderwijsinstellingen doen er dus goed aan zoveel mogelijk teksten van een student te bewaren.

Ephorus wil ook compressietechnieken, vergelijkbaar met programma's als Winzip, inzetten bij het plagiaatscannen. Dergelijke programma's analyseren een tekst als een verzameling karakters die met een bepaalde regelmaat voorkomen, aangevuld met volledig willekeurige tekens. De truc is om die regelmaat te ontdekken, want dat is een eigenschap waardoor een tekst korter weer te geven is. “Die eigenschappen gebruiken we om iets over de tekst en de schrijver te weten te komen”. Heeft een deel van de tekst ineens andere eigenschappen, dan wordt een stijlbreuk geconstateerd en is waarschijnlijk sprake van plagiaat.

Intellectueel eigendom “De scanner heeft vooral een preventief effect”, vertelt Joop Kerkhof, aanspreekpunt voor docenten die met Ephorus aan de slag gaan bij de faculteit letteren van de Nijmeegse Radboud Universiteit. Daar gebruikt men de plagiaatscanner al sinds 2004. “Veel docenten vertellen me dat het aantal gevallen van plagiaat terugloopt.” Bijna alle docenten zijn positief over het systeem, ondanks de beperkingen: de scanner detecteert geen plagiaat als een student overschrijft uit een boek dat elektronisch niet beschikbaar is.

Uiteraard is de nieuwe waakzaamheid niet onder alle studenten even



populair. “Tijdens de invoering ontstond hier paniek”, vertelt Kerkhof. “Mischien is dat een teken dat er veel geplagieerd werd”.

Op het webforum van studentenpartij AKKURAad uit Nijmegen ontstaat een levendige discussie naar aanleiding van het in gebruik nemen van Ephorus. Albrecht (achternaam onbekend) opent met de mededeling: “Jongens, hier gaat iets gigantisch fout!” Sommigen denken dat het scannen en bewaren van werkstukken niet mag omdat ze intellectueel eigendom zijn van de student. Ilja Faber, van de facultaire studentenraad, weerlegt dat idee echter in een berichtje. “Van alles wat een student aflevert in de studen-

tentijd gaan de rechten automatisch naar de universiteit. Daar kun je je vraagtekens bij zetten, maar aanvechten zal niet lukken”, schrijft hij. Toch verstomt het protest niet.

“Het komt allemaal zo wantrouwend over”, schrijft een student. Docent Toon merkt op dat Ephorus ook goede kanten heeft: “Kun je je voorstellen dat Ephorus ook een manier is om jullie auteursrechten te beschermen?” Toch plaatst ook hij kanttekeningen. Vertalen van artikelen uit buitenlandse tijdschriften en parafaseren is volgens hem voldoende om de scanner te misleiden.

Ook vindt Toon plagiaat meer een kwestie van eigen verantwoordelijkheid. “Wie op deze manier zijn of haar bul behaalt, valt later vanzelf wel door de mand”, schrijft hij. “De enige die hier echt beter van wordt, is de meneer die Ephorus te gelde maakt”.

Inmiddels is het protest weer verstomd. “Studenten zijn het nu gewend,

en de universiteit is erg open in het gebruik van de scanner”, vertelt Kerkhof. Op de website van Ephorus wordt een gratis demoaccount aangeboden. “Studenten maken daar gebruik van om eens te kijken wat die scanner eigenlijk kan.”

Overigens vindt Kerkhof dat de oorzaak van het probleem rond plagiaat niet alleen bij de student ligt. “Het begint al op de middelbare scholen. Daar worden scholieren door docenten al snel verwezen naar internet. Hoe moeten ze dan weten wat wel en niet mag? Dat moeten wij ze hier leren.” ■

Onderzoek onder studenten Universiteit Utrecht: www.oeronline.nl
 Plagiaatscanners Ithenticate en Turnitin: www.plagiarism.org
 Mydropbox www.mydropbox.com
 Ephorus www.ephorus.nl

Plagiaatfest

Hoe volledig is de database van een plagiaatscanner? Het updaten van de database gebeurt door ‘bots’, computerprogramma’s die het internet afstruinen op zoek naar nieuwe informatie, net zoals grote zoekmachines doen. We gaan aan de slag met de gratis demoversie van Ephorus, dus zonder privé-database. Eerst uploaden we een gekopieerd nieuwsbericht van de *Natuurwetenschap & Techniek*-site www.natutech.nl. De plagiaatscanner vangt bot, hoewel het bericht met zoekmachine Google makkelijk te vinden is. Een oud nieuwsbericht van de website van de Leidse studentenkrant Mare, uit 2003, wordt wel als plagiaat herkend.

Vervolgens controleren we een tekst die gekopieerd is van www.tweedekamer.nl. Binnen enkele seconden herkent Ephorus dit als plagiaat. Ook een artikel van de voorpagina van onze eigen website valt door de mand.

Joep Chappin van Ephorus: “We concentreren ons niet op nieuwsberichten. Die kunnen morgen al weer veranderd zijn. Het zou dus heel goed kunnen dat onze database die gewoon nog niet bevat.”

Later blijkt dat Ephorus ook nog kinderziekten vertoont: “De berichten worden wel gevonden en zitten dus in onze database, maar ons programma ‘snapt’ ze niet”, legt Chappin uit. De nieuwsberichten staan op de website van Natuurwetenschap & Techniek in een database en zijn niet als aparte links beschikbaar. Zoekmachines als Google zijn daar op ingesteld, maar blijkbaar gaat er in de technologische keten bij Ephorus wat mis. Chappin: “Ik ga het meteen doorgeven aan onze helpdesk.”

Uiteraard is het voor studenten nu ook mogelijk om bij Ephorus eerst zelf hun stukken na te kijken en eventueel nog aan te passen. ‘Veilig’ plagiëren kost dan al gauw net zo veel tijd als het schrijven van een eigen tekst.